



**POLITECHNIKA
GDAŃSKA**

Wydział Elektroniki, Telekomunikacji i Informatyki



Imię i nazwisko autora rozprawy: Magdalena Mazur-Milecka
Dyscyplina naukowa: Inżynieria biomedyczna

ROZPRAWA DOKTORSKA

Tytuł rozprawy w języku polskim: Analiza możliwości automatycznego rozpoznawania zachowań zwierząt doświadczalnych

Tytuł rozprawy w języku angielskim: Analysis of the possibility of automatic recognition behavior of the experimental animals

Promotor

podpis

dr hab. inż. Jacek Rumiński, prof. PG

Gdańsk, rok 2021

Streszczenie

Automatyzacja analizy zachowań zwierząt laboratoryjnych umożliwia szybką i obiektywną ocenę obserwacji. Rozwiązania w istniejących systemach skupiają się na detekcji prostych aspektów zachowań, która jest możliwa dzięki pomiarowi parametrów motorycznych zwierząt, ich punktów charakterystycznych czy kształtu ciała. Jednym z istotnych problemów w dziedzinie rozpoznania zachowań są zachowania złożone, dla których trudno jest zdefiniować opisujące je parametry nawet doświadczonemu obserwatorowi. Kolejnym aspektem są testy socjalne lub grupowe (stadne) podczas których często występuje kontakt między osobnikami skutkujący złączeniem się obserwowanych obiektów na obrazie lub nawet ich częściowym zasłonięciem. W ramach rozprawy zaproponowano 3-etapową metodę segmentacji zwierząt zależną od stopnia złączenia, dla której zaadaptowano i przetestowano algorytmy uczenia maszynowego. Opracowano również metody symulacyjne testów socjalnych. W przeprowadzonych badaniach zweryfikowano także możliwość zdefiniowania deskryptorów zachowań nietypowych wspomagających automatyzację analizy zachowań socjalnych na podstawie obrazów termograficznych.

Abstract

Automating the analysis of laboratory animals' behavior enables a quick and objective evaluation of observations. Solutions in existing systems focus on the detection of simple aspects of behavior, which is possible due to the measurement of the animals' motor parameters, characteristic points or body posture. One of the significant problems in the field of behavior recognition are complex behaviors, for which it is difficult to define describing parameters, even for an experienced observer. Another aspect is social or group test, when contacts between individuals often appear resulting in objects occlusions or even partially covering in the image. In the presented dissertation a 3-step method of animal segmentation have been proposed, together with machine learning algorithms adaptation and testing. Simulation methods for social testing have also been developed. The conducted research also verified the possibility of defining descriptors of atypical behaviors that can support the automation of social behavior analysis based on thermal images.

Podziękowania

Pragnę serdecznie podziękować promotorowi pracy, prof. Jackowi Rumińskiemu za całą okazaną mi pomoc, wsparcie i nadzór nad pracą. Dziękuję za wszystkie cenne wskazówki, pomogły mi one doprowadzić pracę do końca. Dziękuję także prof. Antoniemu Nowakowskiemu, który również był promotorem pracy, za jego wsparcie i motywację.

Dr. Wojciechowi Glac dziękuję za wsparcie badań od strony biologicznej.

Osobne podziękowania składam mojej rodzinie, Mężowi oraz Mamie, za cierpliwość i pomoc. W szczególności dziękuję mojemu Tacie, który zawsze wierzył w moje możliwości.

Serdecznie dziękuję wszystkim osobom, które przyczyniły się do powstania tej pracy.

Spis treści

1. Wstęp	1
1.1. Systemy analizy zachowań zwierząt doświadczalnych	3
1.1.1. Systemy komercyjne	6
1.1.2. Dostępne bazy danych nagrań zachowań gryzoni	9
1.2. Podsumowanie	10
2. Tezy i zakres pracy	13
3. Zachowania socjalne gryzoni	17
3.1. Detekcja zachowań prostych	20
3.2. Zachowania agresywne	21
3.3. Automatyczna detekcja prostych zachowań	22
3.3.1. Wybór cech	23
3.3.2. Wyniki	24
3.4. System dedykowany analizie określonych funkcji motorycznych myszy	27
3.5. Podsumowanie	29
4. Projekt eksperymentów i akwizycji danych	33
4.1. Założenia w zakresie przeprowadzanych eksperymentów	33
4.2. Stanowisko pomiarowe	34
4.3. Zwierzęta uczestniczące w badaniach	34
4.3.1. Szczury	35
4.3.2. Myszy	35
4.4. Przebieg doświadczeń	35
4.5. Baza danych	36
4.6. Podsumowanie	38
5. Analiza obrazów dla potrzeb oceny zachowań zwierząt	39
5.1. Przetwarzanie wstępne	40
5.1.1. Przekształcenie danych termicznych na obrazy	40
5.1.2. Segmentacja obiekt-tło	42
5.1.3. Metoda połączonych komponentów	43
5.1.4. Podsumowanie przetwarzania wstępnego	45
5.2. Algorytmy śledzenia obiektów	45
5.2.1. Metoda CamShift	45
5.2.2. Przepływ optyczny	45
5.2.3. Filtracja cząstkowa	45
5.2.4. Metoda aktywnych konturów	46
5.2.5. Wyniki zastosowania algorytmów śledzenia obiektów - analiza porównawcza	46
5.2.6. Podsumowanie analizy algorytmów śledzenia	47
5.3. Algorytmy ekstrakcji obiektów połączonych - segmentacja obiekt-obiekt	47
5.3.1. Algorytm działów wodnych (wododziałów)	48



5.3.2.	Segmentacja semantyczna	50
5.3.3.	Segmentacja instancji	51
5.4.	Zaproponowany algorytm ekstrakcji obiektów połączonych	52
5.5.	Segmentacja przy użyciu uczenia głębokiego	60
5.5.1.	Segmentacja semantyczna	60
5.5.2.	Segmentacja instancji	66
5.5.3.	Podsumowanie zastosowania metod uczenia głębokiego	74
5.6.	Podsumowanie	75
6.	Badania symulacyjne	77
6.1.	Symulacja kształtu obiektu	77
6.2.	Symulacja tekstury	79
6.2.1.	Punkty charakterystyczne ciała	81
6.3.	Symulacja ruchu	82
6.3.1.	Detekcja granic obrazu	82
6.4.	Symulacja śladów śliny	84
6.4.1.	Symulacja ruchu śladu śliny	85
6.4.2.	Symulacja zmian temperatury śladu w czasie	89
6.5.	Podsumowanie	90
7.	Parametry w detekcji zachowań socjalnych	93
7.1.	Detektory punktów charakterystycznych	93
7.2.	Porównanie detektorów punktów charakterystycznych	94
7.3.	Analiza detekcji śladów śliny	96
7.3.1.	Ogólna analiza śladów śliny	98
7.3.2.	Porównanie detekcji różnych punktów na tym samym obrazie	98
7.3.3.	Porównanie cech śladów pozostawionych przez ugryzienie oraz grooming	101
7.3.4.	Analiza wpływu detektora na detekcję - badania symulacyjne	103
7.4.	Dynamiczna analiza detekcji śladów śliny	106
7.4.1.	Analiza zmian temperatury na obrazach rzeczywistych	107
7.4.2.	Analiza zmian temperatury na symulacjach	114
7.5.	Podsumowanie	117
8.	Podsumowanie	121
8.1.	Osiągnięcia i znaczenie pracy	123
Dodatek A	136
A.	Publikacje związane z tematem pracy	136

Spis rysunków

1.1. EthoVisionXT	7
1.2. Oznaczenie interesujących przedmiotów - VideoMot2	7
1.3. Oznaczenia zwierząt - VideoMot2	8
1.4. Podział obszaru sceny w VideoMot2	8
1.5. Oświetlenie labiryntu VideoMot2	9
1.6. Automatic Rat Behavior Recognition – EthoVision XT	9
1.7. SocialScan firmy Clever Sys	10
2.1. Schemat elementów pracy	16
3.1. Ofensywna pozycja boczna	22
3.2. Wyniki zastosowania klasyfikatorów	24
3.3. Czas działania	25
3.4. Wyniki przydatności cech dla binarnych drzew decyzyjnych i lasu losowego	26
3.5. Labirynt bez przeszkód	27
3.6. Obraz labiryntu ze zwierzętami	28
5.1. Obrazy 8-bitowe i ich histogramy stworzone z różnych zakresów danych termicznych	41
5.2. Wykres wartości progu obliczonego metodą Otsu dla 19-minutowego nagrania	43
5.3. Histogram obrazu z zaznaczonymi progami	43
5.4. Histogramy dla a) tła oraz b) obiektu	44
5.5. Wyniki progowania dla progu obliczonego metodą Otsu	44
5.6. Przykład złączenia obiektów	48
5.7. Działanie algorytmu działów wodnych	48
5.8. Działanie filtracji oraz algorytmu działów wodnych	49
5.9. Wyniki segmentacji przy użyciu algorytmu działów wodnych	50
5.10. Architektura U-Net, na podstawie [120]	51
5.11. Różnice pomiędzy segmentacją a) semantyczną a b) segmentacją instancji	52
5.12. Przykłady złączenia obiektów	53
5.13. Algorytm wyboru stopnia złączenia.	53
5.14. Zastosowanie binaryzacji, szkieletyzacji i erozji.	54
5.15. Wykres średnich procentowych liczby różnych rodzajów wzajemnego ułożenia obiektów w trakcie 20 minutowego nagrania	54
5.16. Algorytm rozdzielania obiektów dla złączenia minimalnego.	55
5.17. Wyniki działania algorytmu rozdzielania obiektów dla złączenia minimalnego.	56
5.18. Algorytm segmentacji obiektów dla złączenia małego	57
5.19. Wyniki pomiaru odległości linii	58
5.20. Wyniki błędnego wyboru linii rozdzielającej	59



5.21. Wyniki poprawnego wyboru linii rozdzielającej	60
5.22. Przykładowe obrazy oraz ich maska	61
5.23. Przykładowe wyniki segmentacji dla sieci U-Net	64
5.24. Wyniki dokładności segmentacji obiektów (J_{segm})	65
5.25. Przykłady segmentacji referencyjnej	71
5.26. Porównanie wyników segmentacji dla różnych typów uczenia	72
5.27. Porównanie wyników segmentacji dla różnych architektur	73
6.1. Algorytm tworzenia symulowanego kształtu gryzonia	78
6.2. Przykłady stworzonych kształtów ciała	79
6.3. Przykład rzeczywistej tekstury powierzchni ciała zwierząt na obrazach termograficznych	79
6.4. Działanie algorytmu szkieletyzacji	81
6.5. Wpływ zmian wartości wariancji rozkładu na symulowaną teksturę	81
6.6. Obliczanie miejsc położenia punktów charakterystycznych ciała	82
6.7. Przykłady symulacji ruchu	83
6.8. Proces tworzenia dwóch obiektów na obrazie	83
6.9. Ślad okrągły o największym rozmiarze o różnych temperaturach	84
6.10. Ślad okrągły o temperaturze 297 K i różnych rozmiarach	86
6.11. Ślad podłużny o temperaturze 297 K i różnych rozmiarach	87
6.12. Przykłady nałożenia śladów śliny o różnych temperaturach na symulacje zwierząt	88
6.13. Symulacja przesuwania śladu na obiekcie	88
6.14. Wykres zmian temperatury dla przykładowej obserwacji oraz symulacji	90
7.1. Wynik detekcji punktów charakterystycznych detektorem Harris'a	96
7.2. Rozkład punktów odnalezionych przez detektor	97
7.3. Obraz trzech różnych analizowanych śladów	99
7.4. Temperatura T_{cLow} śladów oraz średnia temperatura obiektu	99
7.5. Wartości R śladów	100
7.6. Rankingi śladów	100
7.7. Średni gradient temperatury w obszarze śladów	101
7.8. Ślady śliny pozostawione po groomingu	101
7.9. Zmiany rankingu śladów w czasie	102
7.10. Zmiany temperatury wykrytych punktów śladów oraz średniej temperatury ciała w czasie	103
7.11. Wyniki analizy detektorem FAST	104
7.12. Wyniki analizy detektorem Harris	105
7.13. Zmiany temperatury T_{cp} , T_{cLow} i temperatury średniej w czasie	107
7.14. Zmiany rankingu i wartości R w czasie	108
7.15. Obrazy przykładowych punktów śladów śliny wybranych do analizy	108
7.16. Obrazy przykładowych punktów referencyjnych wybranych do analizy	109
7.17. Algorytm półautomatycznego śledzenia wybranego narożnika	110
7.18. Średnia temperatura śladu $Trace_2$ w sąsiedztwie N_4	111
7.19. Średnia temperatura punktów referencyjnych	112
7.20. Dopasowanie krzywych dla procesu chłodzenia	113
7.21. Dopasowanie krzywych dla procesu ogrzewania	113
7.22. Wykresy rozrzutu dla wpływu długości śladu (LR) na współczynnik dopasowania R^2	116
7.23. Przykłady dopasowania obserwacji do modelu	118

7.24. Wykresy średnich a) wartości R i b) rankingów dla wszystkich 27 klatek symulacji.	119
----------------------------------------------------------------------------------------------------	-----

Spis tabel

3.1. Niektóre kategorie zachowań socjalnych wg [19] oraz [52]	18
3.2. Podstawowe parametry pozwalające na detekcję określonych zachowań	20
3.3. Wyniki dwuczynnikowego testu ANOVA w postaci poziomu prawdopodobieństwa	29
4.1. Bazy danych nagrań	37
5.1. Wyniki algorytmu Otsu dla 19-minutowego nagrania	42
5.2. Średnia liczba różnych rodzajów zachowań w trakcie nagrań	54
5.3. Porównanie miar M_l odległości linii otrzymanych w wyniku działania algorytmu do linii wybranej przez obserwatora	58
5.4. Parametry trenowania w eksperymentach segmentacji semantycznej	61
5.5. Wyniki treningu architektury U-Net i V-Net	63
5.6. Wyniki parametrów Intersection over Union (IoU) oraz F_1 dla testowania architektur U-Net i V-Net	64
5.7. Parametry trenowania, dla których osiągnięto najlepsze wyniki	67
5.8. Wyniki testowania dla uczenia transferowego Mask R-CNN dla 200 obrazów	68
5.9. Wyniki testowania dla trenowania od podstaw Mask R-CNN dla 500 obrazów	69
5.10. Wyniki testowania dla uczenia transferowego TensorMask dla 200 obrazów	70
5.11. Wyniki testowania dla trenowania od podstaw TensorMask dla 500 obrazów	71
5.12. Wyniki segmentacji obrazów $ch0$ dla modeli wytrenowanych jedynie na obrazach z popularnych baz danych	73
5.13. Wyniki mAP dla 3-krotnej cross-walidacji dla uczenia transferowego oraz trenowania od podstaw	74
7.1. Średnie rankingi dla punktów charakterystycznych	95
7.2. Analiza statystyczna śladów śliny	98
7.3. Tabela porównania cech śladów po groomingu i ugryzieniu	103
7.4. Wyniki trójczynnikowej analizy ANOVA	106
7.5. Wyniki obliczonych współczynników dopasowanej funkcji oraz miara dopasowania R^2	114



Lista skrótów

CNN	ang. Convolutional Neural Network - splotowa sieć neuronowa
SVM	ang. Support Vector Machine - maszyna wektorów nośnych
HMM	ang. Hidden Markov Model - statystyczny model Markova
LSTM	ang. Long short-term memory - rodzaj sieci rekurencyjnej
AdaBoost	klasyfikator składający się z dużej liczby słabych klasyfikatorów
VGG	rodzaj splotowej sieci neuronowej wprowadzonej przez Visual Geometry Group
CRIM13	Caltech Resident-Intruder Mouse - baza danych interakcji myszy
MBADA	baza danych Mice Behavior Analysis dataset
RatSI	baza danych zachowań szczurów
SABN	System Analizy Badań Na Zwierzętach
CC	Metoda połączonych komponentów (ang. Connected Components)
T_{cp}	Temperatura narożnika wyznaczonego przez detektor punktów charakterystycznych
$Corner_{LOW}$	Punkt znajdujący się w miejscu piksela o najniższej temperaturze w otoczeniu N4 od narożnika wyznaczonego przez detektor
N4	śasiadztwo piksela składające się z 4 innych najbliższych pikseli połączonych z punktem w pionie i poziomie
N8	śasiadztwo piksela składające się z 8 innych najbliższych pikseli połączonych z punktem w pionie i poziomie oraz ukośnie
T_{cLow}	Temperatura punktu $Corner_{LOW}$
$DB_{T1}, DB_{T2}, DB_{T3}$	Nazwy stworzonych baz danych termograficznych nagrań testów socjalnych szczurów
DB_{W1}, DB_{W2}	Nazwy stworzonych baz danych nagrań testów socjalnych szczurów kamerą z zakresu promieniowania widzialnego
DB_{W3}	Nazwa bazy danych testów motorycznych myszy
U-Net	Sieć typu enkoder-dekoder opracowana do segmentacji obrazów biomedycznych
V-Net	Sieć typu enkoder-dekoder oparta na sieci U-Net
Mask R-CNN	Dwuetaapowa architektura segmentacji instancji
TensorMask	Jednoetaapowa sieć segmentująca instancje
ReLU	ang. Rectified Linear Unit - rektyfikowana jednostka liniowa, nielin-iowa funkcja aktywacji
IoU	ang. Intersection over Union - miara podobieństwa zbiorów danych
F₁ lub dice	współczynnik dice - miara podobieństwa zbiorów danych
TP	ang. True Positive - wynik (lub obraz) prawdziwie dodatni
TN	ang. True Negative - wynik (lub obraz) prawdziwie ujemny
FP	ang. False Positive - wynik (lub obraz) fałszywie dodatni
FN	ang. False Negatice - wynik (lub obraz) fałszywie ujemny
LR	parametr długości czasu detekcji śladu



Lista symboli

W_{AGRES}	współczynnik agresywności osobnika
A	powierzchnia obiektu
L	obwód obiektu
Q	współczynnik cyrkularności
$linia_{prop}, linia_{wybr}$	zaproponowana i wybrana linia separacyjna rozdzielająca 2 obiekty
w_1, w_2	wagi składowych miary $linia_{prop}$
w_3, w_4	wagi składowych miary $linia_{wybr}$
d_{lin}	długość zaproponowanej linii separacyjnej
$Area_1, Area_2$	powierzchnie podzielonych obszarów
$\overline{\delta}_{lin}$	średnia odległość punktów linii otrzymanych w wyniku algorytmu podziału wodnego od zaproponowanej linii separacyjnej
M_l	miara odległości dwóch linii
P	obrazy (dane) będące wynikiem predykcji
GT	ang. Ground Truth - obrazy (dane) referencyjne
J_{segm}	parametr wprowadzony do celów tej pracy, zdefiniowany jako stosunek liczby obrazów poprawnie rozseparowanych do liczby wszystkich obrazów testowych
n	liczba obrazów poprawnie rozseparowanych
N_{test}	liczba wszystkich obrazów testowych
AP	średnia precyzja
mAP	mean average precision - uśredniona średnia precyzja
AP_{50}	średnia precyzja dla IoU wynoszącego 50
AP_{75}	średnia precyzja dla IoU wynoszącego 75
$p(r)$	krzywa precyzja-czułość
R	wartość R - miara jakości detekcji narożnika
$ranking$	miara jakości detekcji narożnika
I_p	intensywność potencjalnego narożnika
I_x	intensywność jednego z szesnastu pikseli dookoła potencjalnego narożnika
t	wartość progu
a, b, c	współczynniki wielomianu drugiego stopnia
R^2	współczynnik determinacji
y_t	rzeczywista wartość zmiennej (pomiaru) w czasie t
\bar{y}	wartość średnia rzeczywistej zmiennej (pomiaru)
f_t	estymowana (wynikająca z modelu) wartość zmiennej w czasie t
\bar{f}	wartość średnia estymowanej (zamodelowanej) zmiennej
ρ	współczynnik korelacji Pearsona
σ_y	odchylenie standardowe pomiarów
σ_f	odchylenie standardowe wartości modelowanych



Rozdział 1

Wstęp

Badania laboratoryjne przeprowadzane na zwierzętach mają swój niekwestionowany wkład w rozwój medycyny, farmakologii czy biologii. Stanowią wstępny etap badań nad wszelkiego rodzaju lekami, poprzedzają badania kliniczne z udziałem ludzi. Gryznie wykorzystywane są w badaniach nad fizjologią, toksykologią związków a także badaniach nad komórkami nowotworowymi, badaniach biochemicznych, mikrobiologicznych i immunologicznych. Wraz z pojawieniem się technologii inżynierii genetycznej, która umożliwia zmianę genomu np. myszy, zwierzęta te stały się jednym z popularniejszych źródeł materiału do badań nad chorobami ludzkimi. Automatyzacja tych badań na przestrzeni ostatnich dziesięcioleci umożliwiła obiektywną, długotrwałą i wiarygodną obserwację, a także łatwiejszą, szybszą i bardziej zaawansowaną analizę. Automatyczne systemy monitorujące scenę doświadczalną oraz klatki wyposażone w czujniki są już w powszechnym użyciu w laboratoriach oraz chętnie wprowadzane do nowych doświadczeń. Za stosowaniem zautomatyzowanych systemów rejestracji przemawia przede wszystkim ich obiektywność, powtarzalność oraz standaryzacja wyników. Częstym problemem podczas wykonywania badań metodą tradycyjną jest zmiana osoby obserwującej przebieg doświadczenia skutkująca znaczną zmianą rezultatów obserwacji, wynikająca z odmiennej percepcji każdego człowieka. Sposób obserwacji jednej osoby może także zmieniać się pod wpływem zmęczenia, stopnia skupienia czy też warunków otoczenia. Trudności w osiągnięciu powtarzalności i wiarygodności danych behawioralnych gryzoni oraz konieczność standaryzacji doświadczeń są przedstawione w pracach [124], [109], [145], [148] oraz [10]. Systemy skomputeryzowane zapewniają zawsze takie same wyniki doświadczenia dla takich samych (idealnych) warunków, niezależnie od miejsca oraz osoby wykonującej badanie. Cecha ta pozwala porównywać wyniki badań różnych ośrodków badawczych jeśli tylko do rejestracji i analizy użyty został ten sam system. Komputer jest także dużo bardziej precyzyjny niż człowiek, jego dokładność zależy od użytego sprzętu (rozdzielczości kamery czy sieci czujników), który w dzisiejszych czasach ma praktycznie nieograniczone możliwości.

Inną zaletą systemów skomputeryzowanych jest możliwość pracy w różnym czasie i różnych warunkach. Automatyczny system jest w stanie pracować tak długo, na ile wystarczy mu zasobów sprzętowych. Ta cecha jest bardzo ważna podczas całodobowych testów, których jeden człowiek nie jest w stanie przeprowadzić z jednakowym stopniem skupienia, a zmiana obserwatora skutkuje odmiennymi wynikami. Także obecność sprzętu ma prawdopodobnie mniejszy wpływ na zachowanie zwierząt niż obecność człowieka, dlatego zachowania podczas doświadczenia przeprowadzonego przy użyciu komputera są bardziej zbliżone do naturalnych zachowań obiektów doświadczalnych [150].

Kolejną istotną cechą automatyzacji badań jest szybkość działania. Człowiek nie

zawsze jest w stanie zauważyć wszystkie zdarzenia jakie zachodzą podczas badania. Przykładem może być tu test interakcji socjalnych dwóch agresywnych szczurów. Kiedy osobniki zaczynają walkę, nie jest możliwe zauważenie konkretnych zachowań każdego gryzonia przez człowieka, ponieważ ludzki narząd wzroku nie jest tak szybki. Także w przypadku wolniejszych zachowań dających się zauważyć, problemem jest zanotowanie i opisanie ich, a jednocześnie ciągła obserwacja dalszego przebiegu doświadczenia. Jedynym wyjściem w takich sytuacjach jest nagrywanie sceny doświadczenia, a następnie przeglądanie jej klatka po klatce.

Analiza nagrania, obliczanie parametrów pośrednich wynikających z parametrów otrzymanych w procesie rejestracji czy też analiza statystyczna wyników są także szybciej wykonywane przez komputer niż przez człowieka. Mało tego, bardzo często system rejestrujący jest zintegrowany z oprogramowaniem analizującym oraz oceną statystyczną, co eliminuje procedurę wprowadzania wyników doświadczenia do komputera oraz ewentualne pomyłki człowieka podczas tego procesu. Automatyczny system bez wątplenia przyspiesza przeprowadzanie badań doświadczalnych, ale także potrafi w prosty sposób wykonać pomiary, z którymi człowiek nie radzi sobie tak dobrze. Mowa jest tu np. o pomiarze prędkości. Dla oprogramowania jest to prosta zależność między pokonaną drogą, a czasem jej pokonania – dwie wartości łatwo mierzalne. Dla człowieka obliczenie prędkości wiąże się z zastosowaniem przyrządów dodatkowych, które pomierzą czas oraz drogę, nierzadko z małą dokładnością. Również złożone długoterminowe zależności przy czynowo skutkowe między zdarzeniami czy schematy zdarzeń są możliwe do określenia przez komputer na podstawie analizy dużej ilości danych [24], [25], co znacznie przewyższa ludzkie możliwości.

Kolejną zaletą systemów automatycznych jest możliwość wprowadzenia dodatkowych technologii i tym samym uzyskanie danych wcześniej niemożliwych do zaobserwowania. Najczęściej używanymi gatunkami zwierząt laboratoryjnych są gryzoni, które charakteryzują się dużą dostępnością, niskimi kosztami utrzymania oraz występowaniem wielu genetycznie modyfikowanych szczepów. Światło niezbędne standardowym kamerom czy obserwatorom jest dla zwierząt nocnych źródłem strachu i stresu, a co za tym idzie – wpływa na zmianę ich zachowań [118], [7]. Obrazowanie w zakresie promieniowania podczerwonego wydaje się być najlepszą technologią w przypadku analizy zachowań gryzoni ze względu na ich nocny charakter życia. Dane o rozkładzie powierzchniowym temperatury ciała są z kolei ważną informacją na temat stanu zwierzęcia, mogą być pomocne w analizie stresu lub strachu [68]. Ciągły pomiar temperatury pozwala również na obserwację termoregulacji, która, jak zostało udowodnione, jest wynikiem zmian zachowania zwierząt [138].

Zalety systemów automatycznych oraz znaczna przewaga nad standardową metodą rejestracji badań zwaną „stoper, kartka, ołówek” [116] zachęca biologów i lekarzy do stosowania takich systemów. Jednak to, wydawałoby się łatwe rozwiązanie, ma także swoje wady. Jedną z nich jest konieczność spełnienia wielu założeń. Jeśli system działa na zasadzie analizy nagrań wideo, użytkownik musi zapewnić odpowiednie warunki nagrania zgodne ze specyfikacją systemu, np. właściwe oświetlenie czy też brak innych, podobnych do obiektów doświadczalnych przedmiotów w tle lub scenie badania.

Każdy sprzęt ulega awariom oraz degradacji jakości na skutek zużycia czy też wadliwego użytkowania. Badanie przeprowadzone wadliwym sprzętem może wprowadzić błędne dane do bazy wyników doświadczeń, a to z kolei może prowadzić do wysnucia błędnych wniosków. Nie do końca wiadomo również jak obecność



urządzeń wchodzących w skład systemu wpływa na zachowanie zwierząt. Przyпуска się, że jest to mniejszy wpływ niż w przypadku obecności ludzi, ponieważ człowiek jest naturalnym wrogiem większości zwierząt doświadczalnych i czują się one zagrożone w jego towarzystwie [150].

Ostatnią, ale najpoważniejszą wadą komputeryzacji badań doświadczalnych jest trudność w identyfikacji zdarzeń stosunkowo prostych do rozpoznania przez człowieka. Maszyna nie ma problemów z rejestracją dokładnych parametrów zdarzeń, jednak określenie kontekstu zdarzenia jest dla niej trudnym zadaniem. Z tego powodu rozpoznanie zdarzenia czy też ostateczna decyzja bardzo często jest wciąż dokonywana przez użytkownika. System może dostarczyć dane wspomagające decyzję lub ją zasugerować, jednak ostatecznie to użytkownik decyduje o klasyfikacji zdarzeń.

Systemy automatycznego śledzenia i rozpoznania zachowań zwierząt doświadczalnych są powszechnie rozwijane. Bardzo często dostosowywane są one do potrzeb danego doświadczenia: mierzą określony parametr w określonych warunkach.

1.1. Systemy analizy zachowań zwierząt doświadczalnych

Wiele jednostek badawczych na świecie rozwija różnego rodzaju systemy wspomagające wykonanie i analizę badań laboratoryjnych z udziałem zwierząt doświadczalnych. Niektóre z nich oferują jedynie śledzenie obiektu, inne wykonują także detekcję dodatkowych parametrów lub nawet zdarzeń.

Jednym z takich projektów jest „Monitoring Animal Behavior” prowadzony przez Kristin Branson [22], [21], [20], [72], mający na celu stworzenie inteligentnego terrarium do automatycznego monitoringu gryzoni. Wynikiem pracy jest algorytm śledzenia do trzech myszy w klatce z kamery ustawionej z boku. System wykorzystuje algorytm przepływów optycznych (ang. Optical Flow) i oferuje poprawną identyfikację obiektów po zasłonięciu. Jednak, ani podczas nałożenia się obiektów, ani podczas rozłączenia nie są obliczane żadne inne parametry oprócz położenia. Ułożenie kamery bardzo ogranicza możliwość analizy bardziej złożonych zachowań niż analiza przemieszczenia, powoduje obserwację złączenia obiektów podczas gdy nie musi faktycznie zachodzić kontakt fizyczny gryzoni, analiza jest także zależna od ułożenia się osobnika względem kamery.

Kolejnym przykładem jest praca [105] przedstawiająca system śledzenia trajektorii wielu osobników w ciągu wielu dni. Rozpoznanie gryzoni odbywa się na podstawie wzoru wyfarbowanego na futrze, którego system jest nauczony. Pozwala to na dokładną i pewną identyfikację osobników, nawet po zaginięciu (zakopaniu się w trocinach) czy zasłonięciu wzoru. Nie jest to jednak zgodne z zasadą preferowaną podczas obserwacji zwierząt laboratoryjnych - jak najmniejszej ingerencji w przebieg doświadczenia. W tym przypadku nie jest znany wpływ dodatkowych czynników, np. zapachu farby, na przebieg doświadczenia.

Pionierami w dziedzinie analizy zachowań gryzoni byli Rousseau [123] oraz Heeren [60]. Pierwsza praca przedstawia zastosowanie sieci neuronowych do rozpoznania dziewięciu prostych zachowań szczurów na podstawie ułożenia ciała oraz położenia osobnika. Nagrania wykonane były kamerą usytuowaną z boku klatki. Poprawność rozpoznania wyniosła 63,7% wszystkich klatek. Drugi artykuł klasyfikuje postawy gryzoni na podstawie obrysu ich sylwetek przy pomocy sieci Kohonena z 96,9% poprawnością.

Kolejnym przykładem jest system rozpoznający zachowanie myszy z nagrań kamery usytuowanej z boku klatki [38] na podstawie klasyfikacji cech czaso-przestrzennych otrzymanych po zastosowaniu stworzonego klasyfikatora. Dzięki analizie kilku regionów charakterystycznych obiektu (oczy, uszy, łapy, grzbiet itp.), zwanych w pracy sześcianami, algorytm rozróżnia zachowania, takie jak jedzenie, picie, spanie, grooming oraz eksploracja. Całkowita poprawność detekcji wyniosła 72%. Do poprawnego działania system wymaga niezastłoniętego widoku kluczowych regionów ciała. Zastosowanie kamery z boku, co prawda umożliwia lepszą detekcję tych regionów niż użycie kamery znajdującej się nad klatką, jednak liczba scen nadających się do analizy jest dużo mniejsza. Spowodowane jest to możliwością ustawienia się obiektu tyłem do kamery lub zastąpieniem się za innym obiektem.

W 2010 r. w pracy [69] przedstawiono system obserwacji zwierząt, który na podstawie porównania parametrów ośmiu automatycznie wykrytych zachowań rozpoznał je z 76% dokładnością. System składa się z modułu ekstrakcji ponad 300 cech ruchu i położenia obiektu oraz modułu klasyfikacji, który został stworzony przy użyciu algorytmów Support Vector Machine i Hidden Markov Model (SVM-HMM) oraz nauczony na ponad 10 godzinnej bazie filmów oznaczonych przez kilku niezależnych ekspertów. Kamera podczas nagrań do bazy ustawiona była z boku pojemnika, w którym znajdowało się tylko jedno zwierzę. Korzystając z tej bazy, powstały później także inne systemy, które rozpoznawały te same, zdefiniowane wcześniej osiem prostych zachowań: picie, jedzenie, mikro-ruchy, chodzenie, odpoczywanie, grooming, stawanie na tylnych łapach i zwisanie z górnej kratki pojemnika. W pracy [80] użyto sieci głębokiej będącej kombinacją sieci 3D-CNN i LSTM. Jiang i inni w [70] zastąpili zbiór punktów czaso-przestrzennych wektorem Fischera a klasyfikator SVM siecią neuronową. Z kolei Nguyen [102] zastosował modele I3D i R(2+1)D wstępnie przetrenowane na pozach ciała człowieka.

W artykułach [28] oraz [27] zaprezentowano system monitorowania dwóch myszy na podstawie współrzędnych geometrycznych oraz modeli fizycznych obiektów, interakcje zaś zostały określone na podstawie deterministycznych zestawów zasad opartych na położeniu względnym zwierząt oraz ich części ciała (głowy, karku, nasady ogona), ich przyspieszeniu i innych danych numerycznych ustawionych na stałe na zasadzie prób i błędów.

W pracy [35] poddano analizie zachowanie jednej myszy. Obserwowano zdarzenia takie jak: jedzenie, picie, skoki, stanie na tylnych łapach, opieranie się przednimi łapami na ścianach, odpoczynek, wachanie, chodzenie oraz grooming używając współrzędnych położenia różnych części ciała, konturu obiektu, informacji o rozmieszczeniu dodatkowych elementów w klatce (poidło, jedzenie) oraz statystyki ruchu w celu śledzenia zwierzęcia. Klasyfikacja odbywała się przy użyciu kwadratowej funkcji klasyfikacyjnej, gdzie argumentami były wyżej wymienione parametry ruchu i ułożenia ciała gryzonia. Zgodność między wynikami systemu a obserwacją przez człowieka plasowała się w przedziale 70-85%.

W pracy [23] do śledzenia obiektów użyte zostały detektory czaso-przestrzenne, a algorytm klasyfikacji oparty na AdaBoost oraz metodzie auto-context został nauczony zachowań na podstawie przykładów. Podczas detekcji zdarzenia w określonym momencie brane są pod uwagę poprzednie i następne klatki nagrania oraz wcześniej zaobserwowane schematy występowania zdarzeń. Klasyfikowanych jest 12 typów zachowań oraz jedno dodatkowe zachowanie określone jako „inne” – mniej ważne. Średnia zgodność rozpoznania 13 klas zachowań między systemem a obserwatorami wyniosła 61%, a zgodność pomiędzy samymi obserwatorami 70%. Mała zgodność rozpoznania między obserwatorami wynika z dużej rozbieżności dotyczącej obserwacji zachowań określonych jako „inne”.

Dodatkową technologię w postaci nagrań termograficznych wykorzystano w oprogramowaniu [50], [131], które jest w stanie śledzić i analizować zachowania więcej niż 2 myszy umieszczonych w tym samym pomieszczeniu. Analiza zachowania jest uzyskiwana dzięki dwóm komponentom: modułowi śledzącemu oraz klasyfikatorowi. Moduł śledzący składa się z trzech składowych:

- algorytmu detekcji blobów do wstępnej detekcji zwierząt,
- algorytmu działań wodnych używanego w przypadku, gdy wykrytych obiektów jest mniej niż rzeczywistych (złączenie się obiektów),
- oraz algorytmu dopasowania osobników opartego na wcześniejszych wynikach algorytmów, momentach H_u oraz wspieranego danymi termograficznymi.

Klasyfikator zachowań otrzymuje kilka określonych cech czaso-przestrzennych takich jak: położenie względne, kształt, przemieszczenie itp. z kilku kolejnych i poprzednich klatek, a następnie na podstawie wcześniej nauczonego algorytmu Random Forest klasyfikuje rodzaj zachowania. Średnia zgodność rozpoznania wszystkich zachowań klatka po klatce między systemem a obserwatorami wyniosła 74%-83%, a zgodność pomiędzy samymi obserwatorami 75%.

W ostatnich latach, wraz ze wzrostem popularności sieci neuronowych, powstało coraz więcej systemów wykorzystujących właśnie te architektury do detekcji zachowań gryzoni. Kramida i inni [77] zastosowali model LSTM i VGG w celu rozpoznania czterech zachowań dla pojedynczej myszy. Praca [48] przedstawia system śledzący myszy o różnych kolorach i na różnym tle, przy użyciu sieci typu enkoder-dekoder. System nie zapewnia segmentacji obiektów, ani śledzenia więcej niż jednego osobnika. Projekt przedstawiony w [143] wykorzystuje augmentację zbioru uczącego oraz sieć MF-Net do klasyfikacji 9 podstawowych zachowań gryzoni laboratoryjnych.

Powstały także systemy wyspecjalizowane w detekcji konkretnych regionów lub zdarzeń: groomingu [81], chodzenia [5], drapania się [2] czy śledzenia nosa [82].

Stworzono także systemy do śledzenia, detekcji lub rozpoznania pozy ciała różnych obiektów, które można pobrać i zainstalować bez żadnych opłat:

- **idtracker.ai** - [119] śledzi do 100 różnych kształtów na podstawie kolorów pikseli, wykorzystuje różne sieci splotowe do detekcji tego, czy wystąpiło złączenie obiektów oraz do śledzenia każdego osobnika osobno. Uczenie następuje w 3 etapach (protokołach) i jest wykonywane przez użytkownika.
- **DeepLabCut** - [101] oferuje rozpoznanie pozy ciała dla różnych gatunków. Uczenie transferowe sieci ResNet odbywa się w nawet 10 etapach. Użytkownik musi nie tylko zapewnić obrazy referencyjne dla każdej pozy ciała, ale także bardzo dokładnie oznaczyć punkty zainteresowania na ciele obiektów, jak np. oczy, pysk, łapy wraz z miejscami zgięć, ogon (nasadę, środek oraz jego koniec).
- **Tractor** - [132] prosty program do śledzenia ruchu oraz pomiaru podstawowych jego wielkości (prędkość, przyspieszenie, czasy i odległości) jednego zwierzęcia w zróżnicowanym środowisku lub wielu zwierząt w środowisku jednorodnym. Nie wymaga uczenia, ani obrazów wzorcowych, rozróżnia zwierzęta od tła dzięki progowaniu adaptacyjnemu, a od siebie nawzajem dzięki algorytmowi k-średnich. Nie działa podczas zasłonięcia obiektu.

- **ToxTrack** - [117] system do śledzenia różnych organizmów na zróżnicowanym tle, wg autorów utrzymuje identyfikację nawet po złączeniach, nie segmentuje obiektów.
- **ThermoLab** - [45] to proste oprogramowanie do analizy obrazów termograficznych, umożliwia automatyczną segmentację tła używając algorytmu Otsu, detekcję obiektów, o ile nie są połączone oraz pomiar temperatury i tworzenie profili termicznych nagrań.

Wiele z wyżej wymienionych prac wykorzystuje metody uczenia maszynowego, które skutkują większą elastycznością systemu oraz łatwością jego rozszerzenia o nowe, bardziej zaawansowane zachowania. Kosztem takiego rozwiązania jest większa nieprzewidywalność wyników: mniejsza powtarzalność i poprawność, co do tej pory skutecznie zniechęcało firmy do ich stosowania w systemach komercyjnych. Istniejące na rynku produkty stosują dużo prostsze algorytmy, które nie są w stanie wykonać analizy złożonego zachowania, ale charakteryzują się wysoką powtarzalnością i precyzją.

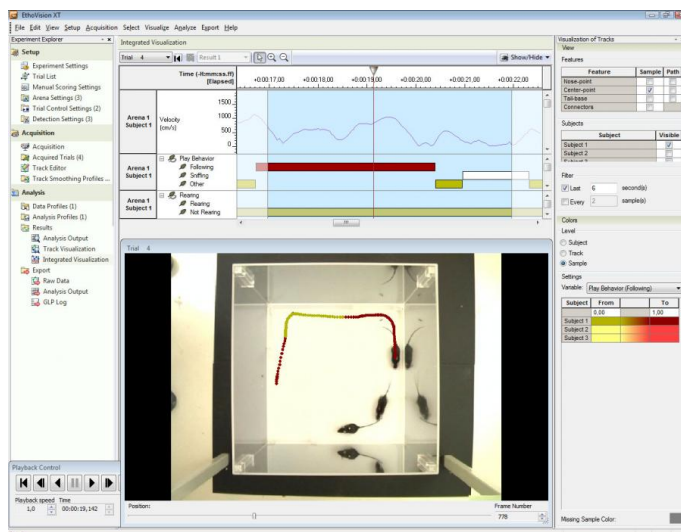
1.1.1. Systemy komercyjne

Istnieje wiele różnych rozwiązań pomocnych przy analizowaniu ruchu, aktywności oraz zachowania zwierząt. Klatki z czujnikami podczerwieni pozwalają obliczać parametry motoryczne gryzoni na podstawie zarejestrowanego położenia. Klatki wyposażone w czujniki drgań umożliwiają obserwację typu zachowania na podstawie amplitudy i częstotliwości drgań, m.in. zachowania nazwanego „zamarciem”, które jest wynikiem np. strachu. Urządzenia ultrasonograficzne analizują odgłosy wydawane przez myszy czy szczury w badaniach neurologicznych lub podczas testu leków. Rotometry z kolei zliczają liczbę i prędkość obrotów gryzoni wywołanych zmianami patologicznymi mózgu [134]. Dostępne są także specjalne środowiska monitorujące procesy uczenia [65] czy siłę uchwytu [135]. Przykłady takich systemów dedykowanych do konkretnych badań i mierzących ograniczoną liczbę parametrów są liczne, jednak prawdopodobnie najbardziej uniwersalnym i rozpowszechnionym sposobem automatycznej analizy badań doświadczalnych jest analiza obrazu z nagrań wideo.

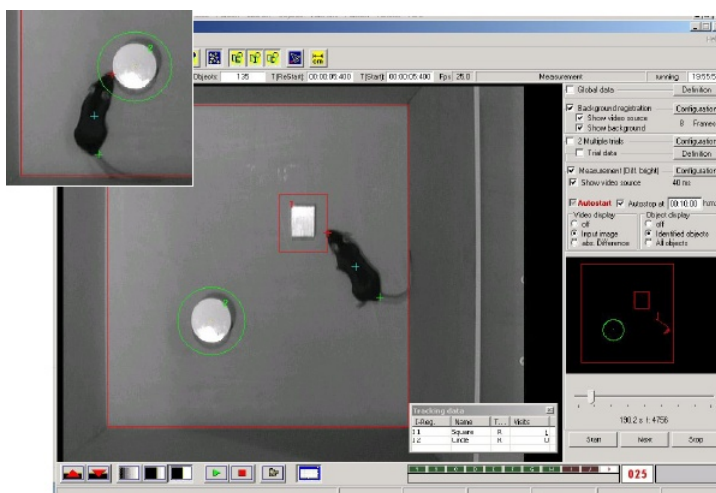
Wśród firm oferujących komercyjne systemy rejestrujące i analizujące doświadczenia przeprowadzane na zwierzętach znajdują się m.in.

- Noldus z produktem EthoVisionXT (rys. 1.1) [104],
- TSE Systems i VideoMot2 (rys. 1.2, 1.4) [136],
- Med. Associates z Video Tracking Software [64],
- San Diego Instruments oferujące system ANY-Maze [66],
- Clever Sys z pakietami oprogramowań: SocialScan, GroupScan, HomeCageScan czy TopScan [63].

Wszystkie te systemy składają się z jednej lub więcej kamer, komputera oraz oprogramowania nagrywającego i analizującego przebieg doświadczenia. Kamery umieszczone są nad klatkami ze zwierzętami. System ma możliwość analizy obrazu przechwytywanego przez kamerę w czasie rzeczywistym lub wcześniej nagranych i zapisanego filmu. Analizowany może być jeden osobnik w jednym pomieszczeniu, wiele osobników w wielu pomieszczeniach lub wiele osobników w jednym



RYS. 1.1. Przykład zastosowania systemu EthoVisionXT, źródło: [104]



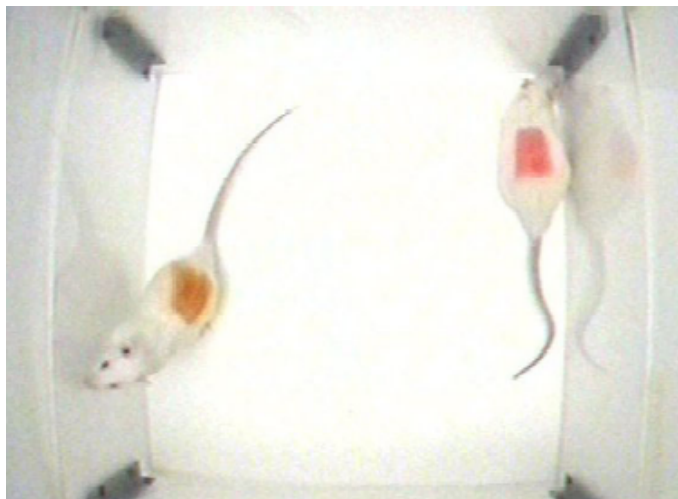
RYS. 1.2. Oznaczenie interesujących przedmiotów znajdujących się wewnątrz klatki – VideoMot2, źródło: [136]

pomieszczeniu. Niektóre systemy wymagają wcześniejszego oznakowania znacznikami różnych kolorów zwierząt badanych w jednym pomieszczeniu w celu rozróżnienia osobników od siebie (rys. 1.3)

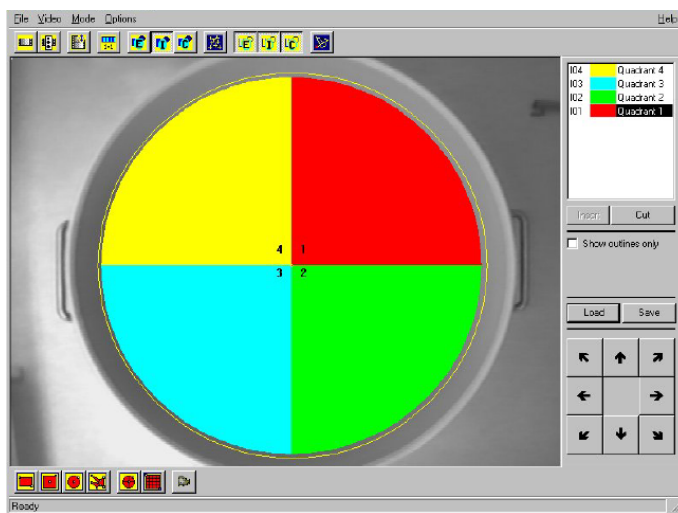
Kolejną wspólną cechą wymienionych wcześniej systemów jest możliwość definiowania własnych obszarów w polu widzenia kamery. Obszary te mogą określać kształt zastosowanego labiryntu i jego podział na podobszary (rys. 1.4) lub ruchome obiekty umieszczone wewnątrz klatki (rys. 1.2).

Zwierzęta odnajdowane są na obrazie na podstawie różnicy kolorów między przedmiotem analizy a tłem lub na podstawie różnicy kolejnych klatek obrazującej ruch. Większość systemów używa obu metod jednocześnie dla lepszej dokładności. W przypadku zbyt małego kontrastu koloru pomiędzy zwierzęciem a otoczeniem, VideoMot2 używa oświetlenia w zakresie podczerwieni zapewniającego kontrast dzięki odpowiedniemu materiałowi, z którego wykonany został labirynt (rys. 1.5).

W większości przypadków systemy te dostosowane są do powszechnie używanych przez biologów czy behawiorystów zwierzęcych scen testowych, takich jak:



RYS. 1.3. Zwierzęta oznaczone kolorowymi markerami – VideoMot2, źródło: [136]

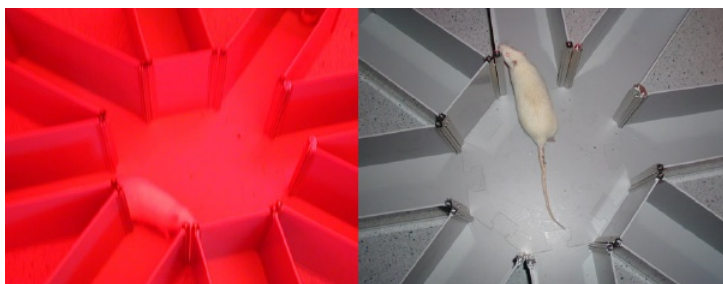


RYS. 1.4. Podział obszaru analizy na cztery podobszary – VideoMot2, źródło: [136]

- basen Morris'a [100] – basen wypełniony wodą z zanurzoną platformą wykorzystywany jako narzędzie do badania uczenia przestrzennego i pamięci,
- labirynt Barnes'a [9] - okrągła podwyższona arena,
- różnego rodzaju labirynty (np. krzyżowy, 0, T) stosowane m.in. w ocenie pracy pamięci roboczej, badaniach lęku i eksploracji czy rozróżniania bodźców akustycznych.

Parametrami mierzonymi przez te systemy są przede wszystkim: prędkość poruszania, przebyta droga, przemieszczenie środka ciężkości ciała, nosa oraz nasady ogona, czasy przebywania we wcześniej określonych strefach, liczba wizyt, częstotliwość eksploracji oraz przebyta droga w każdej strefie, czas spędzony w określonych strefach w porównaniu do całkowitego czasu itp. Jeżeli chodzi o detekcję zachowań, to systemy te oferują dodatkową możliwość zaznaczenia przez użytkownika różnych rodzajów zdarzeń na nagraniu poprzez wciśnięcie odpowiedniego przycisku po zaobserwowaniu zdarzenia. Odpowiednie klatki nagrania zostaną





RYS. 1.5. Oświetlenie labiryntu światłem podczerwonym – Video-Mot2, źródło: [136]

wtedy opatrzone etykietą o rodzaju zdarzenia zanotowanego przez użytkownika. Niektóre z systemów umożliwiają także detekcję podstawowych pozycji ciała (często nazywanymi zachowaniami). Dodatek do produktu EthoVision XT firmy Noldus - Automatic Rat Behavior Recognition [35] - rozpoznaje 10 zachowań gryzonia przedstawionych na rys. 1.6. System analizuje zachowanie tylko jednego osobnika.

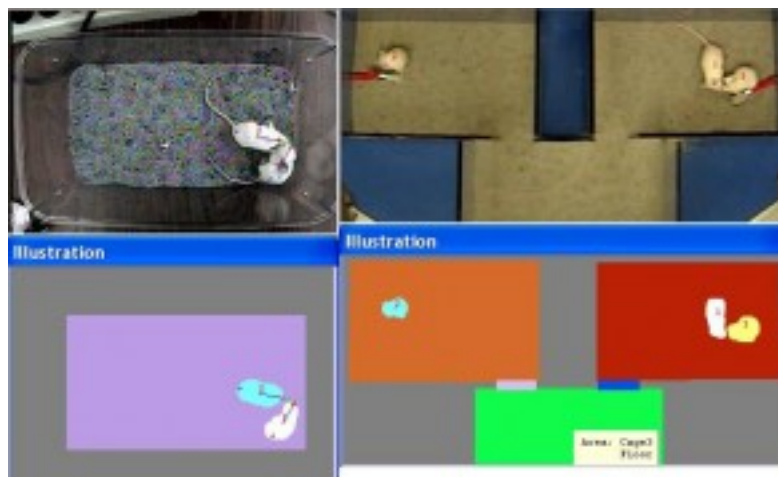


RYS. 1.6. Automatic Rat Behavior Recognition – EthoVision XT, na podstawie [104]

SocialScan firmy Clever Sys (rys. 1.7) - oferuje ilościowy i jakościowy pomiar zachowań socjalnych takich jak: aktywny i pasywny kontakt, kontakt części ciała, zbliżanie się, odchodzenie, śledzenie, obwąchiwanie czy wspinanie się. Warunkiem identyfikacji osobników oraz wykonania analizy podczas kontaktu jest widoczna różnica pomiędzy osobnikami - różnica koloru futra lub wielkości ciała. W przeciwnym razie konieczne jest dodatkowe oznaczenie zwierząt.

1.1.2. Dostępne bazy danych nagrań zachowań gryzoni

Wraz ze wzrostem popularności algorytmów uczenia maszynowego wzrosło zapotrzebowanie na obszerne bazy danych. Istnieje kilka baz przedstawiających zachowania gryzoni.



Rys. 1.7. SocialScan firmy Clever Sys, źródło: [63]

Pierwszą z nich jest baza przedstawiona w pracy [69], często wykorzystywana w systemach używających metod uczenia głębokiego. Składa się ona z ponad 10 godzin nagrań pojedynczych myszy znajdujących się w klatkach będących miejscem ich przechowywania (ang. home-cage). Osiem typów zachowań (picie, jedzenie, mikro-ruchy, chodzenie, odpoczywanie, grooming, stawanie na tylnych łapach i zwisanie z górnej kratki pojemnika) zostało oznaczonych na nagraniu przez kilku niezależnych ekspertów.

Druga baza - Caltech Resident-Intruder Mouse (CRIM13) [23] - składa się z 88 godzin nagrań dwóch myszy kamerami umieszczonymi nad i z boku sceny. Adnotacje rozróżniają 13 zachowań: zbliżanie się, atak, krążenie, węszenie, odchodzenie, kopulacja, pogoń, czyszczenie, picie, jedzenie, wspinanie się oraz pozostałe.

Baza MBADA (Mice Behavior Analysis dataset) przedstawiona w pracy [50] to 2,5 godziny nagrań zachowań socjalnych trzech myszy, 1 godzina nagrań dwóch myszy oraz 9 godzin nagrań dwóch lub czterech myszy pochodzących z różnych szczepów (C57BL/6J i BTBR T+tf/J). Wszystkie dane pozyskiwane są przez kamerę termograficzną, nagrania nie zawierają adnotacji.

Kolejna baza - RatSI [86] - stworzona została przez firmę Noldus. Składa się z dziewięciu 15-minutowych nagrań 2 szczurów z oznaczonymi zachowaniami: groomingu, zbliżania się, śledzenia, odchodzenia, ataku, odosobnienia, kontaktu nos-nos, przygniecenia oraz pozostałe.

W porównaniu do baz danych innych obiektów (np. obrazów twarzy, gestów czy przedmiotów codziennego użytku), baz danych przedstawiających zachowania gryzoni laboratoryjnych jest niewiele i nie są one obszerne. Nie przedstawiają także rzadszych typów zachowań jak np. strach, agresja, ustalanie hierarchii stadnej lub zachowania macierzyńskie.

1.2. Podsumowanie

Zarówno komercyjne jak i niekomercyjne rozwiązania nie pozwalają na automatyzację detekcji wielu form interakcji socjalnych. Oprogramowanie najczęściej jest w stanie rozpoznać moment kontaktu badanych osobników i rozróżnić czy był to kontakt nos-nos, czy nos-ogon. Program może pomierzyć średnią odległość dwóch osobników, czas spędzony w niewielkiej odległości od siebie, liczbę i częstotliwość

kontaktów, średni czas spędzony podczas kontaktu lub w znacznej odległości, czas do pierwszego kontaktu itp. Parametry te są, co prawda bardzo przydatne do późniejszej analizy i na ich podstawie można wiele wywnioskować na temat wzajemnych relacji gryzoni, jednak nie opisują one wielu ważnych zdarzeń jakie mogą zajść podczas przeprowadzania tego typu testów. System sam nie jest w stanie stwierdzić, czy podczas interakcji nastąpiło gryzienie, obwąchanie czy przewrócenie i przytrzymanie przeciwnika na grzbiecie. Niektóre z prac przedstawionych powyżej analizują zachowania pojedynczych zwierząt, jednak wciąż są to zachowania proste, jak czyszczenie się, węszenie czy skakanie. Pomimo zastosowania metod uczenia głębokiego, przedstawione wyżej rozwiązania nie klasyfikują zachowań w sposób jaki jest powszechnie używany przez behawiorystów, wykrywają raczej pozycje ciała zwierzęcia lub określone konfiguracje ułożeń obiektów. Co więcej, w większości przypadków moment, kiedy to zwierzęta są w bliskim kontakcie - a co za tym idzie zachodzi nałożenie się dwóch obiektów na obrazie otrzymanym z kamery - jest wykluczony z analizy.

Wciąż brakuje rozwiązań dotyczących detekcji chociażby agresji oraz metod jej klasyfikacji. Żadna z prac nie analizuje sytuacji walki pomiędzy osobnikami. Testy interakcji socjalnych, agresji, zachowań stadnych czy też testy zachowań związanych ze stresem prenatalnym wykonywane są przez obserwatorów. Powodem tego jest m.in. brak odpowiednich danych. Nie jest łatwe zdefiniowanie parametrów i ich wartości niezbędnych do automatycznej analizy bardziej złożonych zachowań. Kolejnym problemem są częste zasłonięcia się obiektów, podczas których większość prezentowanych systemów rezygnuje z analizy [89].

Praca ta ma na celu przedstawienie metod i parametrów pomocnych w automatyzacji testów socjalnych gryzoni poprzez:

1. definicję parametrów dodatkowych wspomagających analizę zachowań, w szczególności tych agresywnych;
2. opracowanie algorytmu segmentacji zwierząt podczas złączenia w celu ich poprawnej identyfikacji i możliwości analizy standardowych parametrów bez względu na wzajemne położenie zwierząt.

W dalszej części pracy przedstawione zostaną tezy i zakres pracy.

Rozdział 2

Tezy i zakres pracy

Rozprawa ma na celu analizę możliwości automatycznego rozpoznawania złożonych zachowań zwierząt w badaniach laboratoryjnych. Celem pracy jest automatyzacja analizy złożonych zachowań socjalnych zwierząt doświadczalnych, na drodze budowy scenariuszy testowych i cyfrowej rejestracji obrazu sceny eksperymentu, automatyczna segmentacja obiektów złączonych oraz detekcja i pomiar parametrów umożliwiających indywidualizację cech niezbędnych w detekcji złożonych zachowań socjalnych. Określono również przydatność opracowanego systemu w badaniach zachowań socjalnych gryzoni.

Przyjęto następujące założenia:

- system przeznaczony jest do badania interakcji socjalnych dwóch szczurów,
- nagrania wykonywane są kamerą termograficzną usytuowaną nad sceną testu,
- w polu widzenia nie znajdują się inne przedmioty o temperaturze zbliżonej do temperatury ciała zwierząt,
- nie jest możliwe rozróżnienie zwierząt na podstawie ich wyglądu.

Zakres pracy obejmuje:

- opracowanie środowiska do badań testowych,
- zebranie potrzebnej bazy danych,
- adaptację i przetestowanie algorytmów rozpoznawania i śledzenia obiektów,
- zaproponowanie algorytmu rozdzielania obiektów złączonych,
- opracowanie detektorów określonych zachowań lub zdarzeń,
- opracowanie metod symulacji zachowań szczurów doświadczalnych,
- opracowanie metod symulacji i analizy zdarzeń ugryzienia jednego osobnika przez drugiego,
- ocenę przydatności wykonanego systemu w warunkach eksperymentu.

Przystępując do realizacji pracy postawiono następujące tezy:

1. Możliwe jest zautomatyzowanie procesu analizy złożonych zachowań socjalnych zwierząt doświadczalnych, na drodze budowy scenariuszy testowych i cyfrowej rejestracji obrazu sceny eksperymentu oraz zastosowanie procedur cyfrowej analizy sekwencji obrazów termograficznych do rozpoznawania określonych klas zachowań socjalnych.

2. Stosując metody uczenia maszynowego, możliwa jest automatyczna segmentacja obiektów złączonych oraz detekcja i pomiar parametrów umożliwiających indywidualizację cech niezbędnych w detekcji zachowań socjalnych.

Podział zachowań socjalnych w pracy na proste i złożone jest podziałem wynikającym z możliwości automatycznej detekcji tych zachowań. Zachowania, dla których możliwe jest zdefiniowanie mierzalnych parametrów względnie prostych w detekcji i pomiarze przez systemy i algorytmy widzenia komputerowego uznawane są za zachowania proste. Przykładami takich zachowań są np. śledzenie, pogoń, obwąchiwanie czy eksploracja. Zachowania te występują często podczas interakcji zwierząt. Parametrami detekcji zachowań prostych mogą być: wzajemne położenie obiektów (ich środków ciężkości), odległość pomiędzy obiektami, wektory przemieszczeń obu osobników w kolejnych klatkach, ilość czasu spędzonego w bliskiej odległości itp.

Za złożone zachowania socjalne zwierząt doświadczalnych w pracy uważane są zachowania, które nie posiadają prostych w detekcji parametrów i bardzo często wymagają dodatkowej wiedzy lub doświadczenia obserwatora. Są to między innymi: agresywne czyszczenie, wspinanie się lub gryzienie. Bardzo często zachowania te różnią się między sobą w sposób mało zauważalny, zależą od wcześniejszych zdarzeń lub mają swoje własne sekwencje. Detekcja takich zdarzeń wykonywana jest przez doświadczonego obserwatora.

Podsumowując, detekcja zdarzeń prostych polega na detekcji i pomiarze zdefiniowanych mierzalnych parametrów (lub, w przypadku niektórych systemów wykorzystujących uczenie głębokie, trenowaniu sieci na obrazach referencyjnych), detekcja zdarzeń złożonych oparta jest na subiektywnej ocenie człowieka. W pracy skupiono się jedynie na agresywnych zachowaniach złożonych: agresywne czyszczenie, gryzienie.

Za zdarzenie nietypowe uważa się zdarzenie (akcję) wskazujące na wystąpienie złożonego zachowania socjalnego. W pracy skupiono się na zdarzeniach pozostawienia śladów śliny na futrze jednego osobnika przez drugiego. Ślady te są skutkiem wystąpienia zachowań socjalnych takich jak: ugryzienie czy agresywne czyszczenie. Są to zdarzenia rzadkie.

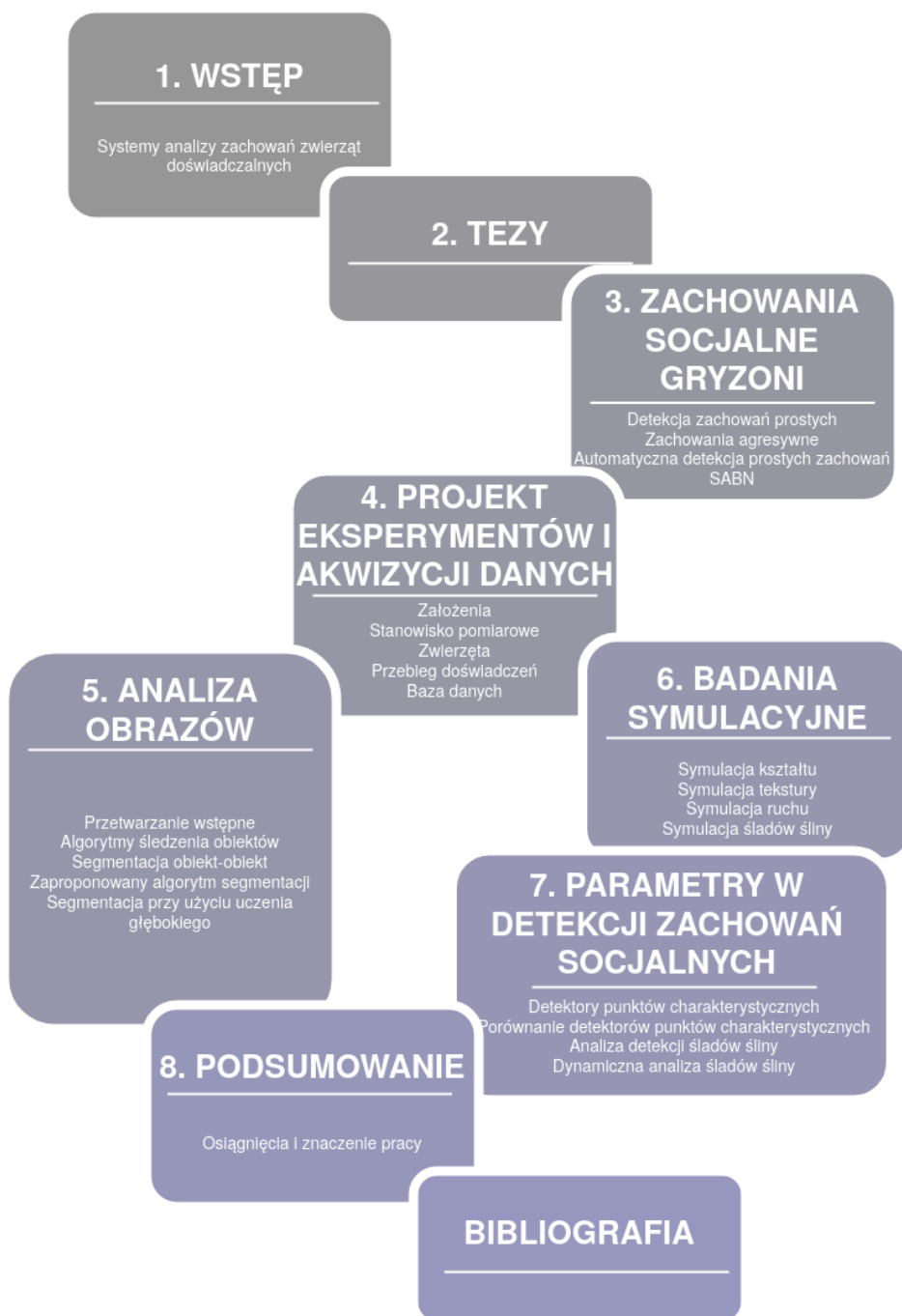
Na potrzeby pracy obiekt złączony zdefiniowano jako obraz przedstawiający dwa osobniki podczas kontaktu, który skutkuje brakiem możliwości poprawnej identyfikacji obiektu oraz pomiarów jego parametrów bez procesu rozdzielenia obszarów każdego osobnika.

Obszarem pracy są metody przetwarzania i analizy obrazów z zastosowaniem uczenia maszynowego dla potrzeb detekcji zwierząt oraz określonych zachowań socjalnych podczas testów zwierząt laboratoryjnych. Analiza zachowań gryzoni jest pomocna w nieinwazyjnej diagnostyce zaburzeń nastroju, w tym depresji, zaburzeń lękowych, zmian neurodegeneracyjnych i wielu innych [83]. Pozwala również na jakościową ocenę stopnia dominacji każdego osobnika poprzez parametryzację jej zachowań społecznych. Tego rodzaju ocena dominacji, agresywności, podatności na stres czy strach jest bardzo często wymagana przed różnego rodzaju badaniami laboratoryjnymi. Zachowania społeczne są także bardzo ważnym parametrem samopoczucia i zdrowia zwierząt. Zmiany, nieprawidłowości lub odchylenia w zachowaniu społecznym mogą wskazywać na chorobę psychiczną, stres lub strach, ale także mogą być wynikiem stosowania produktów leczniczych.

Realizacja celów pracy stanowi treść następujących rozdziałów.

Rozdział 1. przedstawia wstęp pracy, rozdział 2 tezy i zakres prac. W rozdziale 3 opisane zostały typowe testy interakcji socjalnej. Rozdział 4 opisuje metodę pomiarową oraz dokładny sposób przeprowadzania doświadczeń. Rozdział 5 zawiera opis przetwarzania wstępnego oraz analizy obrazu wykonanych nagrań w celu detekcji, śledzenia i segmentacji obiektów. Rozdział 6 przedstawia opis badań symulacyjnych kształtu, ruchu, agresywnej interakcji i zmian temperatury śladu w czasie. Rozdział 7 to prezentacja metod detekcji punktów charakterystycznych oraz ich analizy. Rozdział 8 zawiera podsumowanie rozprawy.

Schemat pracy został przedstawiony na rys. 2.1.



RYS. 2.1. Schemat elementów pracy

Rozdział 3

Zachowania socjalne gryzoni

W rozdziale tym przedstawione zostaną typowe zachowania socjalne gryzoni. Parametry mierzone podczas detekcji zachowań prostych opisane zostały w rozdziale 3.1. Rozdział 3.2 zawiera opis zachowań agresywnych. Przykład automatycznej detekcji zachowań prostych przedstawiono w rozdziale 3.3. Rozdział 3.4 prezentuje system stworzony w celu analizy zmian motorycznych myszy.

Zachowanie socjalne jest częścią osobowości każdego stworzenia. Od niego w znacznej mierze zależy przetrwanie osobnika czy gatunku. U zwierząt żyjących w stadach lub koloniach, takich jak szczury czy myszy, wykształca się hierarchia socjalna, która decyduje m.in. o pierwszeństwie dostępu do pokarmu. Wiele chorób klinicznych przejawia się problemami w utrzymywaniu kontaktów socjalnych. Opracowanie modeli zachowań badanych zwierząt oraz wiarygodnych testów to kluczowy element procesu zrozumienia mechanizmów lęku, stresu lub chorób i zaburzeń psychicznych takich jak m.in. autyzm, depresja czy schizofrenia. Istnieje kilka typów testów zachowań socjalnych [12]:

- testy interakcji socjalnych w środowisku neutralnym lub awersyjnym – polegają na 10-cio minutowej obserwacji zachowań socjalnych oraz lokomocji w środowiskach odpowiednio: neutralnym (słabe oświetlenie, pomieszczenie identyczne jak klatka, w której zwierzęta były wcześniej przetrzymywane) oraz awersyjnym (mocniejsze oświetlenie, nowe pomieszczenie);
- test mieszkaniac-intruz to test badający zachowania agresywne w sytuacji, gdy do pomieszczenia z osobnikiem przetrzymywanym osobno przez 3 tygodnie został wprowadzony na 10 minut osobnik przetrzymywany wcześniej w grupie;
- test długotrwałego stresu – zwierzęta przez pewien okres czasu poddawane są różnym czynnikom stresującym. Stres wywoływany jest między innymi przez rotacyjne rozmieszczanie osobników obojga płci w klatkach [139], unieruchomienie lub zanurzenie w zimnej wodzie [115].

Często wykonanie testu interakcji socjalnych lub testu mieszkaniac-intruz określających stopień dominacji każdego osobnika, jest wymagane przed różnego rodzaju testami biologicznymi. Stopień dominacji jest jednym z parametrów charakteryzujących stan organizmu gryzoni doświadczalnych, jak ciężar ciała czy ciśnienie krwi. Testy te są czułe na liczne środowiskowe i fizjologiczne czynniki wpływające na stopień lęku zwierząt. Niewątpliwą zaletą testów interakcji socjalnych jest ich prostota, nie wymagają specjalnych warunków, pozbawiania pokarmu lub wody, czy też treningów zwierząt.

Pierwsze testy socjalne oraz typowe zachowania wraz z ich klasyfikacją zostały opracowane około 40 lat temu [3], [41], [49]. W ciągu tych lat testy socjalne wykryły



działanie wywoływania lęku lub działanie przeciwlękowe niektórych leków, peptydów, hormonów, składników diety oraz warunków zewnętrznych takich jak np. sposób przetrzymywania zwierząt, separacja z matką w okresie poporodowym i zapachy naturalnych wrogów. Przyczyniły się także do poznania neuronowych podstaw powstawania lęku [43] oraz wpływu stresu na organizm.

Aby wywołać zachowania socjalne między badanymi osobnikami prowadzące do ustalenia dominacji, muszą zostać spełnione pewne warunki otoczenia. Jeśli zwierzęta zostaną umieszczone w nowym pomieszczeniu podczas testów, większość ich zachowań będzie związana z eksploracją otoczenia (wacanie, spacerowanie, śledzenie, przechodzenie przez drugiego osobnika), a zachowania agresywne czy seksualne będą sporadyczne [79]. Na rodzaj behawioru mają wpływ także takie czynniki jak oświetlenie, izolacja socjalna przed testem, przyzwyczajanie do miejsca testowania oraz czas trwania testu. Niekorzystne warunki (np. zbyt mocne oświetlenie i nowe otoczenie) wywołują w zwierzętach strach, co powstrzymuje chęć kontaktu z drugim osobnikiem, ale jednocześnie dostarcza wygodne narzędzie do badań nad lękiem. Test oparty na metodzie File i Hyde [42] zakłada, że gryzoni przestraszone zachowują się w towarzystwie drugiego osobnika w pewien określony sposób: nie dążą do konfrontacji i kontaktu, większość zachowań jest obojętna (eksploracja, czyszczenie itp.). Zwierzęta, które czują się pewnie w danym otoczeniu wykazują większą chęć kontaktu.

Test interakcji socjalnych oparty na metodzie Albonetti i Farabollini [4] ma na celu ustalenie pozycji socjalnej każdego badanego osobnika. Powinien być wykonywany w warunkach sprzyjających kontaktom socjalnym gryzoni. Obserwacje muszą być wykonane w systemie „każdy z każdym”, czyli każdy osobnik w grupie musi zostać porównany ze wszystkimi pozostałymi osobno. Osoba prowadząca test obserwuje zachowanie dwóch szczurów w klatce przez 10 minut i zlicza typy zachowań. Zachowania zwierząt można podzielić na cztery kategorie (Tabela 3.1), na podstawie których wyznaczany jest współczynnik agresywności W_{AGRES} według wzoru 3.1

$$W_{AGRES} = \frac{\sum zach.agresywne}{\sum zach.agresywne + \sum zach.obronne} \quad (3.1)$$

gdzie:





$\sum zach.agresywne$ - liczba zachowań agresywnych łącznie we wszystkich pojedynkach,


$\sum zach.obronne$ - liczba zachowań obronnych łącznie we wszystkich pojedynkach.

TABELA 3.1. Niektóre kategorie zachowań socjalnych wg [19] oraz [52]

kategoria	zachowanie	charakterystyka
eksploracja na odległość	zbliżanie się	zbliżanie się do przeciwnika z uwagą skierowaną ku niemu
	zauważenie	uwaga skierowana na przeciwnika, z oddali
	krażenie	krażenie wokół przeciwnika zbliżając się i oddalając bez dodatkowej aktywności
	przyglądanie się	głowa odwrócona w stronę przeciwnika



kategoria	zachowanie	charakterystyka
poznawanie socjalne	napięte zauważenie 	podobnie jak w „zauważeniu”, ciało dodatkowo jest wydłużone, a głowa wyciągnięta do przodu
	wspinanie się (crawl over) 	obie przednie łapy umieszczone na przeciwniku
	wspinanie się (crawl under)	głowa i przednia część ciała ułożona na przeciwniku
	śledzenie	poruszanie się w bliskiej odległości od przeciwnika, gdy ten przemieszcza się w klatce
	czyszczenie	czyszczenie ciała przeciwnika (z wyjątkiem głowy) przy użyciu pyska
	czyszczenie głowy	czyszczenie głowy przeciwnika
	rozpoznawanie	obwąchiwanie ciała lub ogona przeciwnika
	obwąchiwanie pyska	obwąchiwanie głowy lub pyska
	obwąchiwanie	obwąchiwanie odbytowo-płciowych obszarów ciała przeciwnika
	przepychanie	dochodzi do bliskiego bocznego kontaktu dwóch osobników w trakcie poruszania się w przeciwnych kierunkach
grożenie	chodzenie dookoła	chodzenie co najmniej częściowo wokół przeciwnika w niewielkiej odległości i z głową skierowaną w jego kierunku
	agresywne czyszczenie 	energiczne czyszczenie nieruchomego przeciwnika w pozycji bocznej przy użyciu zębów i kończyn przednich
	ofensywna pozycja boczna 	postawa na trzech łapach w kierunku przeciwnika, łapa uniesiona w górę blisko przeciwnika, oczy zmrużone, uszy położone

kategoria	zachowanie	charakterystyka
	ofensywna pozycja pionowa	
		postawa na dwóch łapach w kierunku przeciwnika, oczy zmrużone, uszy położone
	potrząsanie ogonem	gwałtowne drżenie lub uderzanie ogonem w pozycji bocznej
atak	atakowanie	szybkie zbliżanie się do przeciwnika
	skok do przodu	gwałtowny ruch wykonany w stronę przeciwnika całą przednią częścią ciała lub tylko głową, pozornie w celu ugryzienia, jednak bez użycia zębów
	gryzienie	ugryzienie przeciwnika
	pogoń	pogoń za uciekającym przeciwnikiem

3.1. Detekcja zachowań prostych

Istniejące systemy oferujące analizę zachowań oparte są na dwóch podejściach:

1. pomiarze czasów, przemieszczeń, prędkości i wzajemnych odległości,
2. detekcji postury.

Parametry motoryczne zwierząt są cenną informacją o zachowaniach socjalnych. Szczególnie istotne są wzajemne zależności odległości i prędkości. Co więcej, takie pomiary są łatwe w implementacji, ponieważ wymagają jedynie współrzędnych położenia obiektów.

Postura osobnika, czyli geometryczne ułożenie jego ciała, jest trudniejsza do automatycznej analizy. W większości przypadków prezentowanych w systemach opisanych w rozdziałach 1.1 i 1.1.1 bazują one na detekcjach punktów charakterystycznych ciała (łapy, pysk, ogon) oraz jego konturach lub sieciach neuronowych, jednak tylko w przypadku pojedynczych osobników. Analiza postury podczas testów socjalnych, kiedy to bardzo często dochodzi do kontaktu i złączenia się zwierząt, jest trudna ze względu na konieczność wcześniejszego rozdzielenia obu obiektów. Dodatkowo, zasłonięcie prowadzi do braku widoczności całego ciała, a co za tym idzie, także części istotnych punktów charakterystycznych. Dlatego właśnie zagadnienie automatycznej analizy złożonych zachowań socjalnych wciąż nie jest dostatecznie poznane.

Tabela 3.2 przedstawia spis parametrów, które mogą być stosowane jako miara konkretnych zachowań. Ostatnia kolumna zawiera przykłady niektórych zastosowań.

TABELA 3.2. Podstawowe parametry pozwalające na detekcję określonych zachowań

zachowanie	parametry	przykłady zastosowań
zbliżanie się	wzajemna odległość obiektów	[72], [23], [35]

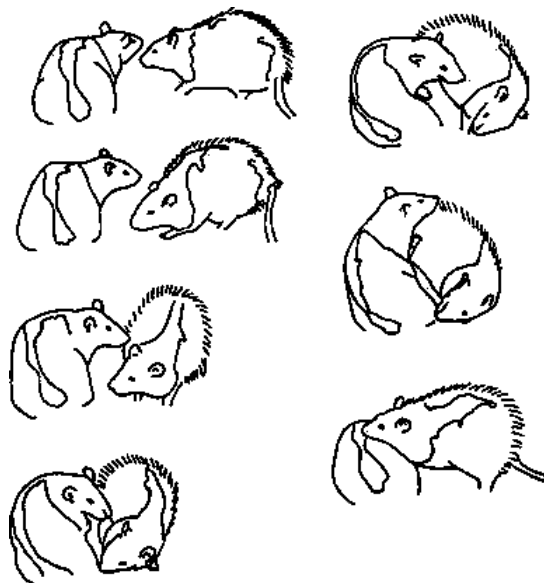
zachowanie	parametry	przykłady zastosowań
	parametry motoryczne	[72], [23], [35]
śledzenie	przemieszczenia	[72], [23], [35]
	parametry motoryczne	[72], [23], [35]
pogoń	przemieszczenie	[72], [23], [35]
	parametry motoryczne	[72], [23], [35]
eksploracja	przemieszczenie	[72], [23], [35]
	parametry motoryczne	[72], [23], [35]
	położenie nosa	
czyszczenie się	zmiany położenia środka ciężkości	[69], [70], [72], [23], [35]
	kształt ciała (kolistość)	[35]
	punkty charakterystyczne ciała	[69], [70]
chodzenie	przemieszczenie	[69], [70], [72], [23], [35]
	prędkość	
	punkty charakterystyczne ciała	[69], [70]
odpoczynek	przemieszczenie	[69], [70], [72], [23], [35]
	kształt ciała (kolistość)	
	punkty charakterystyczne ciała	[69], [70]
mikro-ruchy	zmiany położenia środka ciężkości	[69], [72], [23], [35]
	kształt ciała (kolistość)	
	punkty charakterystyczne ciała	[69], [70]
obwąchiwanie	wzajemne położenia	[63]
	parametry motoryczne	[63]
skok do przodu	przemieszczenie	
	parametry motoryczne	
stanie na tylnych łapach	punkty charakterystyczne ciała	[69], [70], [86]
	kształt ciała	[86]
	przemieszczenie	[86]

Większość podstawowych zachowań może zostać zaobserwowana dzięki bardzo prostym i łatwo obliczalnym parametrom, jakimi są np. wartość przemieszczenia, parametry motoryczne lub kształtu ciała a także położenie punktów charakterystycznych ciała. Zachowania przedstawione w Tabeli 3.2 w większości przypadków są zachowaniami neutralnymi. Zachowania, które jednak najlepiej definiują stopień dominacji osobnika są zachowania agresywne.

3.2. Zachowania agresywne

Celem niektórych zachowań agresywnych jest ugryzienie podporządkowanego osobnika najczęściej w kłęb (zad) tułowia, a celem broniącego się jest uniknięcie tego ugryzienia. Wśród innych zachowań agresywnych znajduje się agresywne czyszczenie (grooming) karku, kiedy to osobnik atakujący chwytą zębami fałdy skóry osobnika czyszczonego. Osobnik podporządkowany w tym czasie pozostaje nieruchomy, może wydawać ciche piski, jego ruch może spowodować ugryzienie lub kopnięcie przez osobnika czyszczącego. Innym przykładem agresywności jest pogoń trwająca kilka sekund. Pogoń może zakończyć się ucieczką osobnika gonionego,

zaprzestaniem pościgu lub złapaniem. Złapanie zwykle skutkuje podgryzaniem lub ugryzieniem ściganego szczura, zachęcaniem go do starcia, wspinaniem się na niego lub agresywnym czyszczeniem karku. W przypadku, gdy osobnik ścigany nie ustępuje, dochodzi do starcia. Atakowany szczur odwraca się do agresora, oba osobniki pozostają nieruchome stojąc naprzeciwko siebie (na czterech lub dwóch łapach), ich sierść jest podniesiona. Zdarza się, że szczur uległy otwiera pysk pokazując zęby. Jeśli atakowany szczur nie ucieknie lub podda się, agresor przystępuje do kontaktu fizycznego. Zaczyna boksowanie lub grożenie (zbliżanie się) ustawiając się w pozycji z boku drugiego osobnika popychając go przy tym lub kopiąc (rys. 3.1) [14], [15], [16], [108], [56].



RYS. 3.1. Ofensywna pozycja boczna w celu ataku bocznego z ugryzieniem. Rysunek stworzony na podstawie [108]

Testy interakcji socjalnych opierają się na detekcji określonych zachowań. Detekcja tych zachowań jest złożonym zagadnieniem, także dla obserwatora-człowieka. Kluczową rolę pełni tu detekcja ugryzienia, które jest ważną miarą stopnia dominacji a także strachu każdego osobnika. Nie istnieją systemy klasyfikacji czy nawet detekcji zachowań agresywnych, co tym samym uniemożliwia wykonanie prawidłowego automatycznego testu interakcji socjalnych. Jednak, oprócz samej detekcji ugryzienia niezbędna do pełnej analizy zachowania jest także detekcja innych, mniej złożonych zachowań. Istotną informacją są m.in. długość i rodzaj kontaktu przed ugryzieniem, kształty ciał obu zwierząt czy ich prędkości poruszania się, które uzupełniają pełną charakterystykę sytuacji. Dlatego, rozpoznanie prostych zachowań jest konieczna także w przypadku analizy zachowań złożonych.

W kolejnym rozdziale przeanalizowano możliwości zastosowania pięciu popularnych klasyfikatorów do rozróżnienia czterech typowych zachowań gryzoni na podstawie podstawowych miar najczęściej stosowanych w pracach opisanych w rozdziałach 1.1 i 3.1.

3.3. Automatyczna detekcja prostych zachowań

W celu klasyfikacji zachowań prostych zastosowano wektor cech określający przemieszczenie oraz kształt ciała w dwóch kolejnych klatkach. Zastosowano algorytmy:



- binarne drzewa decyzyjne,
- las losowy (ang. Random forest),
- AdaBoost,
- Naiwny klasyfikator Bayes'a,
- wielowarstwowy perceptron.

Dla każdego z algorytmów wybrane zostały najbardziej optymalne ustawienia dobrane na drodze prób i błędów.

Analizowane były nagrania dwóch samców szczurów podczas testów interakcji społecznych przy użyciu kamery w zakresie promieniowania widzialnego. Z nagrań wyodrębniono pięć różnych sekwencji przedstawiających określone pary zachowań oraz jeden plik zawierający wszystkie analizowane zachowania:

- bieg i eksploracja,
- eksploracja i spoczynek,
- grooming i bieg,
- grooming i eksploracja,
- grooming i spoczynek,
- grooming, bieg, eksploracja i spoczynek.

Analiza przeprowadzana była na różnych wielkościach zbiorów treningowych i testowych tych samych plików, aby sprawdzić również zależności efektywności działania algorytmów od ilości danych. Zastosowano także krosvalidację i metody samosporne (ang. bootstrapping [40]) w celu lepszego doboru danych.

3.3.1. Wybór cech

Efektywność działania klasycznych algorytmów uczenia maszynowego silnie zależy od takiego wyboru cech tworzących wektor wejściowy, aby prezentowały one różne wartości dla różnych grup klasyfikacji. Aby dobrze wybrać wektor cech, należy przyjrzeć się zdarzeniom, które mają być klasyfikowane. Każde z czterech analizowanych zachowań posiada swoje własne charakterystyczne akcje:

- **Grooming** to czyszczenie całego ciała i futra, najczęściej jednak pyska, przednimi łapami. Zachowanie to jest charakterystyczne dla wszystkich gryzoni. Podczas groomingu można zaobserwować niewielkie zmiany położenia środka ciężkości, powierzchni oraz konturu ciała bez przemieszczenia. Kształt ciała przypomina koło.
- **Spoczynek** generalnie charakteryzuje się brakiem ruchu i kształtem ciała także zbliżonym do koła. Mogą zdarzyć się nagłe zmiany kształtu, położenia środka ciężkości ciała i przemieszczenie podczas np. zmiany pozycji leżenia.
- **Bieg** to duże przemieszczenia i prędkości obiektu, kształt ciała ulega ciągłym zmianom bez wyraźnej tendencji.
- Podczas **eksploracji** zwierzę raczej nie przemieszcza się, ale wykonuje dużo ruchów w miejscu. W czasie węszenia ciało jest rozciągnięte.

Wybrane zachowania, to typowe i często występujące zachowania gryzoni. Z przedstawionych obserwacji można wstępnie określić parametry obiektów różne dla różnych zdarzeń: długość wektora przemieszczenia, powierzchnia i współczynnik cyrkularności ciała. Współczynnik cyrkularności jest miarą zwartości kształtu obliczany ze wzoru 3.2

$$Q = \frac{4\pi A}{L^2} \quad (3.2)$$

gdzie:

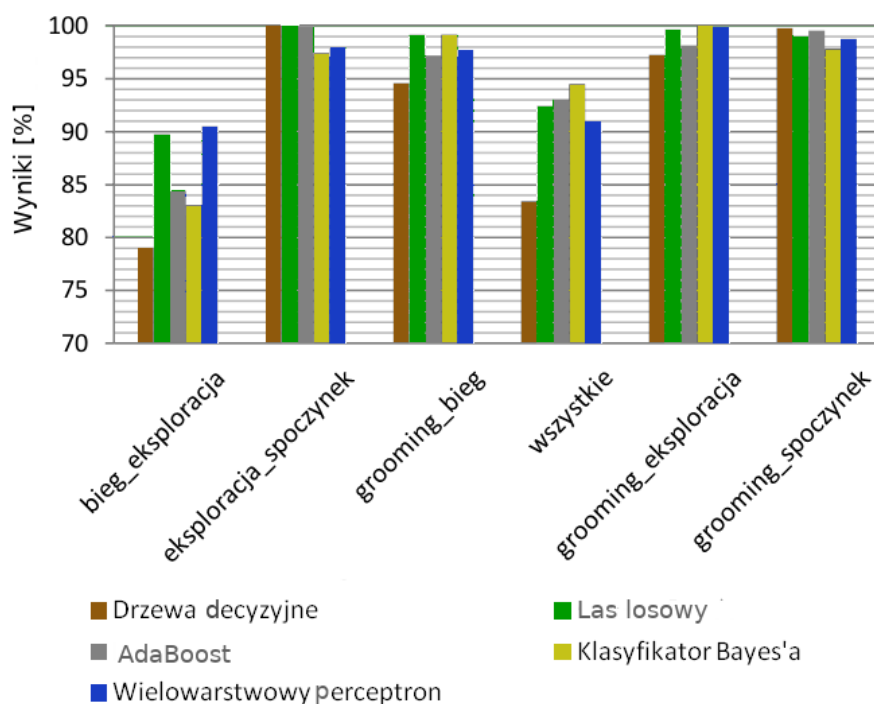
A - powierzchnia obiektu,

L - obwód obiektu.

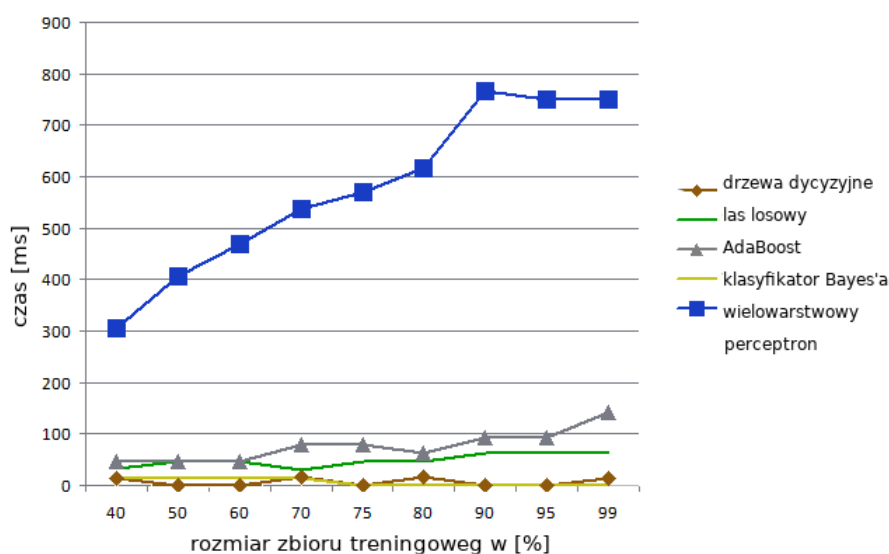
Ostatecznie, wektor cech został zdefiniowany z sześciu parametrów - trzech cech obiektu dla aktualnej i poprzedniej klatki nagrania: współczynnik cyrkularności ciała dla klaki aktualnej, współczynnik cyrkularności ciała dla klaki poprzedniej, długość wektora przemieszczenia dla klatki aktualnej, długość wektora przemieszczenia dla klatki poprzedniej, powierzchnia ciała dla klatki aktualnej oraz powierzchnia ciała dla klatki poprzedniej. Podczas doboru parametrów kierowano się wnioskami prac opisanych w rozdziale 1.1.

3.3.2. Wyniki

Wszystkie algorytmy osiągnęły zadowalające wyniki rozpoznania pomiędzy czterema zachowaniami gryzoni. Najlepszym wyborem okazał się algorytm lasu losowego oraz wielowarstwowy perceptron, które bardzo dobrze rozpoznały wszystkie zachowania (rys. 3.2).



RYS. 3.2. Wyniki zastosowania klasyfikatorów w celu rozpoznania zachowań prostych jako procentowa część liczby poprawnie wykonanych klasyfikacji

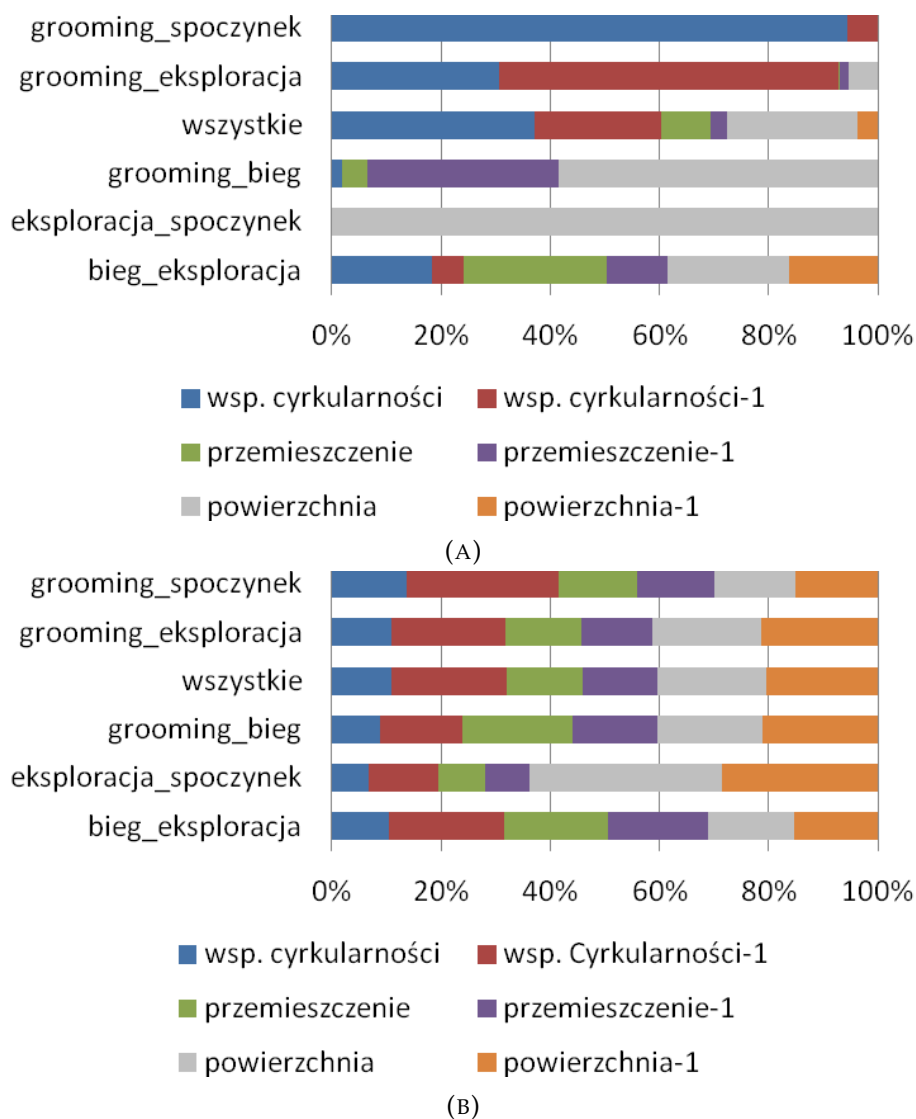


RYS. 3.3. Czas trenowania dla wszystkich algorytmów w zależności od wielkości zbioru treningowego

Jednocześnie, wielowarstwowy perceptron uzyskał najdłuższy czas działania (rys. 3.3) ściśle zależny od ilości danych. Pozostałe algorytmy działały dużo szybciej, jednak czas ten także zależny był od indywidualnych parametrów każdego algorytmu. Badania dowiodły, że efektywność rozpoznania przez wielowarstwowy perceptron nieznacznie większa od rozpoznania przez inne algorytmy wymaga dłuższego czasu pracy. Najtrudniejszymi zachowaniami do rozpoznania okazały się bieg kontra eksploracja, co widać na rys. 3.2. Dla lepszej klasyfikacji tych dwóch zdarzeń potrzebne są dodatkowe cechy. Również rozróżnienie pomiędzy wszystkimi cechami osiągnęło gorsze wyniki, czego przyczyną może być zbyt mała ilość cech w wektorze lub próbek podczas etapu uczenia.

Obliczono również liczbą ważność cech dla drzew decyzyjnych (rys. 3.4a) i lasu losowego (rys. 3.4b). Rysunki obrazują jakie cechy w najlepszym stopniu rozróżniają odpowiednie zachowania. Parametry *wsp. cyrkularności-1*, *przemieszczenie-1* i *powierzchnia-1* oznaczają wartości tych parametrów w klatce $t-1$. Dla każdej pary zachowań obliczono stopień poprawnego rozróżnienia przez kolejne cechy i przedstawiono jako procentowy udział. Wyniki te ukazują znaczne różnice w działaniu tych algorytmów. Oprócz kilku przypadków, gdzie wszystkie cechy były używane przez oba algorytmy, las losowy przeważnie używał wszystkich cech w zbliżonym stopniu, podczas gdy drzewa decyzyjne potrzebowały tylko kilku wybranych cech (np. *wsp. cyrkularności* lub *powierzchnia obiektu*) osiągając przy tym tylko trochę gorsze wyniki efektywności klasyfikacji.

Wyniki powyższej analizy pokazują, że proste do pomiaru parametry kształtu i przemieszczenia są w stanie rozróżnić podstawowe zachowania gryzoni przy użyciu nieskomplikowanych algorytmów. Szereg badań dotyczących automatycznej detekcji zachowań prostych opublikowano w pracy [93].

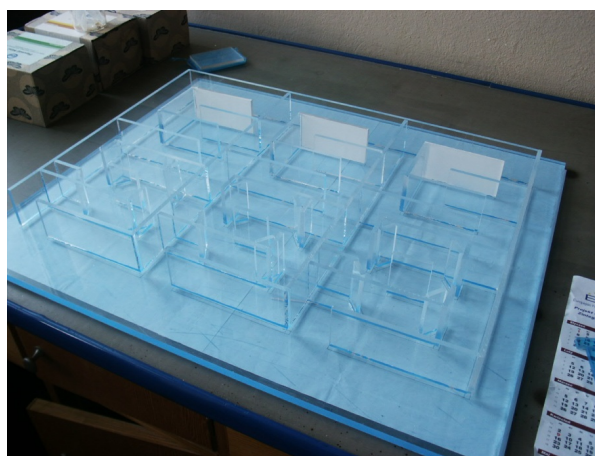


RYS. 3.4. Wyniki przydatności cech dla a) binarnych drzew decyzyjnych, b) lasu losowego. Współczynniki oznaczone są różnymi kolorami

3.4. System dedykowany analizie określonych funkcji motorycznych myszy

Na podstawie pomiaru wyłącznie parametrów motoryki zwierząt możliwa jest także detekcja zachowań nietypowych. Przykładem jest SABN (System Analizy Badań Na zwierzętach) [92] - system dedykowany do badania metod leczenia udaru mózgu, analizujący funkcje motoryczne myszy. Wykonuje on 24-godzinną obserwację myszy i porównuje stopień poprawności funkcji motorycznych, które są jednym z objawów uszkodzenia mózgu.

System składa się z labiryntu wykonanego z pleksi (rys. 3.5) i kamery internetowej umieszczonej nad labiryntem połączonej z komputerem. Labirynt pozwala na obserwację trzech czarnych osobników jednocześnie, w jednakowych pomieszczeniach. Rozkład labiryntu przedstawiono na rys. 3.6. Czerwone linie oznaczają ściany z pleksi niewidoczne na nagraniach. Wszystkie 3 części labiryntu są osobne, myszy nie mają możliwości wyjścia ze swojej części, ani przejścia do innego osobnika. Dodatkowo, ściany pomiędzy jednakowymi pomieszczeniami nie są transparentne, aby zwierzęta się nie widziały, nie wyklucza to jednak rejestracji obecności innych osobników poprzez inne zmysły. Każde z trzech pomieszczeń labiryntu zbudowane jest z wąskich korytarzy (17 stref rozróżnianych przez oprogramowanie), w których można umieszczać przeszkody (stopnie, rampa, kratka). Dodatkowo uwzględniono także korytarz z rozwidleniem przeznaczony do wykrycia potencjalnych preferencji jednej połowy ciała, mogących sugerować uraz mózgu. Oprogramowanie rozpoz-

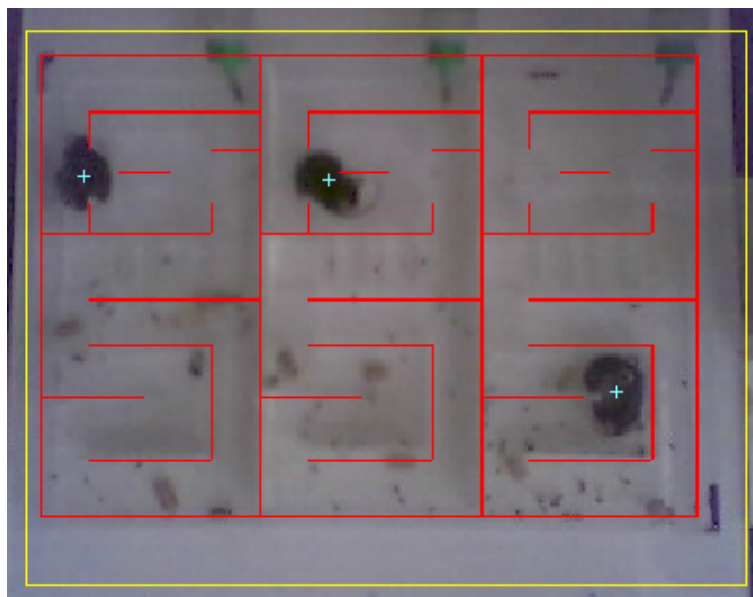


RYS. 3.5. Labirynt bez przeszkód, opracowanie własne

naje zwierzęta na podstawie koloru oraz ruchu, oblicza ich środek ciężkości (oznaczony zielonym krzyżykiem na rys. 3.6) oraz mierzy parametry takie jak:

- całkowita prędkość średnia - parametr często używany w analizie zachowań zwierząt[13] [cm/s];
- prędkość średnia obliczana tylko w momencie ruchu – ruch jest zdefiniowany jako przemieszczenie. Zdarzenia takie jak obwaciwanie, kręcenie się w miejscu czy czyszczenie nie są uznawane za ruch, mimo że powodują zmiany położenia środka ciężkości [cm/s];
- czas spoczynku jako procent całkowitego czasu trwania doświadczenia (24 godziny) [%];

- czas aktywności jako procent całkowitego czasu trwania doświadczenia (24 godziny) [%],
- prędkość średnia podczas pokonywania stopni [cm/s],
- prędkość średnia podczas pokonywania rampy [cm/s],
- wybór drogi w miejscu rozgałęzienia,
- czasy przebywania w poszczególnych strefach [%],
- czas przebywania w pobliżu pojemników z wodą [%],
- prędkości średnie w poszczególnych strefach [cm/s].



RYS. 3.6. Obraz po analizie ze ścianami labiryntu zaznaczonymi kolorem czerwonym oraz środkami ciężkości myszy - kolor zielony

Wszystkie parametry obliczane są dla całego doświadczenia (24 godziny) oraz dla sześciu 4-godzinnych przedziałów. Takie rozwiązanie zostało wprowadzone z powodu zmiennej aktywności dobowej gryzoni.

Pierwszym doświadczeniem przeprowadzonym przy użyciu systemu było sprawdzenie, czy myszy zwyczajne i myszy pozbawione genu białka rozprzęgającego 2 (UCP-2) otrzymane z laboratorium Jackson'a (Bar Harbor, Maine, USA) są zróżnicowane pod względem motorycznym. Białko kodowane przez gen UCP-2 odgrywa kluczową rolę w predyspozycjach do rozwoju chorób naczyniowych zarówno w modelach zwierzęcych jak i u ludzi, szczególnie w odniesieniu do otyłości, cukrzycy i nadciśnienia [111].

W badaniu wzięło udział 13 myszy pozbawionych genu białka (dalej zwane UCP) oraz 11 myszy kontrolnych. Każda trójka rejestrowanych myszy dobierana była w sposób losowy z obu grup, następnie zamykana w labiryncie znajdującym się w osobnym pomieszczeniu na 24 godziny. Podczas nagrań w pomieszczeniu nie znajdował się człowiek, światło zapalone było przez cały czas nagrania. Dokładny opis bazy danych nagrań przedstawiony został w rozdziale 4.5.

W procesie analizy badano następujące parametry: całkowitą prędkość średnią, prędkość podczas ruchu, czas spoczynku, prędkość na stopniach oraz prędkość na rampie.

Wykonano dwuczynnikową analizę wariancji ANOVA dla poziomu istotności $<0,095$, dla czynników: typ myszy (UCP i kontrolne) oraz pora dnia (6 4-godzinnych przedziałów czasu). Wyniki testu przedstawione w tabeli 3.3 ujawniły, że pora dnia ma wpływ na wszystkie wyniki eksperymentu oprócz prędkości na rampie. Natomiast typ myszy wpływa na wyniki prędkość podczas ruchu, czas spoczynku oraz prędkość na stopniach. Pora dnia badania nie wpłynęła na różnice w mobilności między dwiema grupami zwierząt. Warto jednak zauważyć, że zwykły cykl dnia i nocy badanych myszy był zakłócony przez ciągłe sztuczne oświetlenie niezbędne do nagrań.

TABELA 3.3. Wyniki dwuczynnikowego testu ANOVA w postaci poziomu prawdopodobieństwa

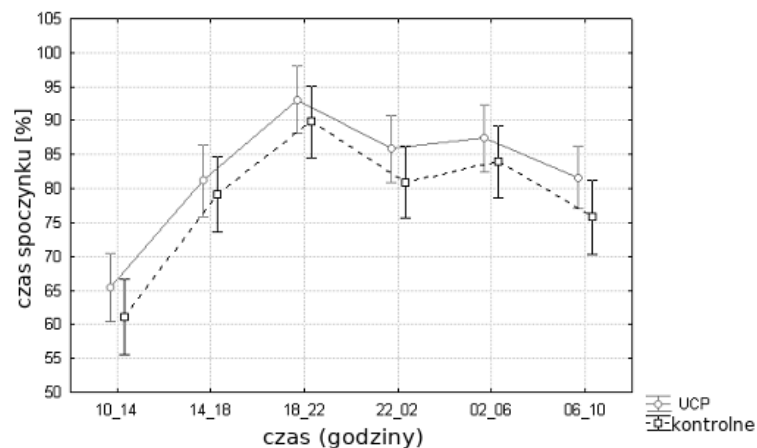
p	całkowita prędkość średnia	prędkość podczas ruchu	czas spoczynku	prędkość na stopniach	prędkość na rampie
pora dnia	0,0	0,0	0,0	0,023	0,322
typ myszy	0,085	0,037	0,008	0,005	0,350
pora dnia *	0,982	0,839	0,970	0,991	0,989
typ myszy					

Rys. 3.7 przedstawia wyniki czasu spoczynku (typ myszy $p=0,008$, pora dnia*typ myszy $p=0,97$) oraz prędkości na stopniach (typ myszy $p=0,005$, pora dnia*typ myszy $p=0,991$) dla każdej grupy w zależności od pory dnia. Wyniki nie wykazują statystycznych różnic pomiędzy grupami typów myszy oraz pory dnia, w przeciwieństwie do analizy każdego z czynników osobno (tabela 3.3). Niespodziewanie, jak na zwierzęta nocne, wysoka aktywność między godziną 9 a 12 związana była z pojawianiem się badacza rozpoczynającego eksperyment a następnie eksploracją nowego miejsca.

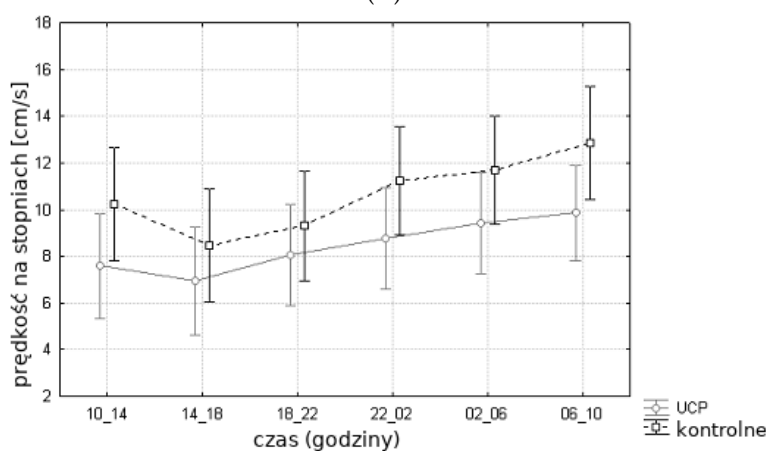
Korzystając z 24-godzinnego automatycznego systemu monitorowania, udowodniono, że myszy pozbawione genu białka rozprzegającego 2 wykazują upośledzenia w pewnych zachowaniach motorycznych. Różnicę tę należy wziąć pod uwagę przy interpretacji wyników badań, w których występuje regulacja genu UCP-2. Szereg badań dotyczących systemu opisano w publikacji [92].

3.5. Podsumowanie

Zachowania socjalne gryzoni są bardzo różnorodne. Niektóre polegają na zachowywaniu dystansu między osobnikami lub wręcz przeciwnie. Inne z kolei na subtelnych zmianach postawy ciała. Jeszcze inne charakteryzują się szybkim ruchem i bliskim kontaktem fizycznym. Z tego powodu automatyzacja zachowań zwierząt laboratoryjnych wymaga pomiaru różnorodnych parametrów. Wartości niektórych parametrów są mniej lub bardziej łatwe do uzyskania, np. przemieszczenie czy prędkość poruszania lub położenie punktów charakterystycznych ciała. Możliwa jest detekcja prostych zachowań na podstawie łatwych do pomiarzenia parametrów, co zostało przedstawione w rozdziale 3.3. Systemy oparte na pomiarach



(A)



(B)

Rys. 3.7. Wyniki dwuczynnikowej analizy wariancji dla: (a) czasu spoczynku oraz (b) prędkości na stopniach. Pionowe wąsy oznaczają 0,95 przedziału ufności. Oś pozioma przedstawia zakresy czasu (godziny) analizowanych danych

takich parametrów mogą także poprawnie rozpoznać niektóre nietypowe zachowania. Zostało to przedstawione w rozdziale 3.4, gdzie system dobowej kontroli motoryki dwóch grup myszy wykazał nieznaczne upośledzenie niektórych funkcji motorycznych zwierząt o zmodyfikowanym genie UCP-2. W pełni zautomatyzowany proces analizy zachowań prostych na podstawie rejestracji obrazu sceny eksperymentu zdolny do detekcji nietypowych zachowań może także być przydatny podczas opracowywania systemów detekcji bardziej złożonych zachowań (np. sekwencji zachowań).

Wyniki przedstawione w tym rozdziale potwierdzają tezy w odniesieniu do zachowań prostych. Rozdział 3.3 odnosi się do drugiej części tezy 2.: stosując metody uczenia maszynowego możliwa jest detekcja i pomiar parametrów umożliwiających indywidualizację cech niezbędnych w detekcji zachowań. Badania przedstawione w rozdziale 3.4 potwierdzają tezę 1: możliwe jest zautomatyzowanie procesu analizy zachowań zwierząt doświadczalnych, na drodze cyfrowej rejestracji obrazu sceny eksperymentu oraz zastosowanie procedur cyfrowej analizy sekwencji obrazów do rozpoznawania określonych klas zachowań.

Proste parametry nie są jednak wystarczające do detekcji bardziej złożonych zachowań socjalnych, takich jak gryzienie, atak czy agresywne czyszczenie. Co więcej, detekcja tych zdarzeń na obrazach rejestrowanych w zakresie promieniowania widzialnego jest zagadnieniem skomplikowanym. Dużym problemem jest również pomiar tych parametrów podczas kontaktu zwierząt. Dlatego w dalszej części pracy przedstawione zostaną metody zarówno segmentacji umożliwiającej identyfikację zwierząt podczas kontaktów, jak i wspomaganie detekcji zachowań agresywnych, zbadane zostaną również możliwości analizy obrazów termograficznych.

Rozdział 4

Projekt eksperymentów i akwizycji danych

Do wykonania analizy złożonych zachowań socjalnych zwierząt niezbędne są odpowiednie nagrania zawierające testy obrazujące takie zachowania. W ramach pracy wykonano nagrania testów interakcji socjalnych, gdzie występują zachowania charakteryzujące ustalanie hierarchii pomiędzy dwoma osobnikami. Założenia oraz stanowisko pomiarowe zostały opisane w rozdziałach 4.1 i 4.2. Rozdział 4.3 przedstawia charakterystykę zwierząt, a 4.4 - przebieg doświadczeń. Baza danych składa się z nagrań w różnych warunkach i na różnych obiektach. Przeprowadzono również dodatkowe testy analizy motorycznej na myszach przedstawione w rozdziale 3.4. Różnorodność bazy zapewnia lepsze możliwości walidacji zaproponowanych algorytmów. Wszystkie rodzaje nagrań przedstawione są w rozdziale 4.5 oraz tabeli 4.1.

4.1. Założenia w zakresie przeprowadzanych eksperymentów

Założenia przeprowadzenia poprawnych nagrań do analizy testów interakcji socjalnych:

- minimalizacja dodatkowego oświetlenia w celu uniknięcia jego wpływu na badane zwierzęta,
- brak ruchomych obiektów innych niż gryzonie o temperaturze lub kolorze zbliżonym do temperatury lub koloru ciała zwierząt (w zależności od zastosowanej kamery) podczas nagrania,
- brak obecności człowieka w pokoju nagrań,
- kamera usytuowana nad sceną nagrania,
- pomieszczenia do przetrzymywania i nagrywania zwierząt wykonane z materiałów transparentnych dla danej kamery - w szczególności mowa jest tu o pokrywie, która w przypadku kamer działających w zakresie promieniowania widzialnego powinna być wykonana z pleksi, a w przypadku kamer termograficznych powinno nie być jej wcale (ściany powinny być tak wysokie, aby uniemożliwić zwierzętom ucieczkę).

Proces przeprowadzania eksperymentów był konsultowany z ekspertami w dziedzinach biologii i medycyny z Uniwersytetu Gdańskiego oraz Gdańskiego Uniwersytetu Medycznego.



4.2. Stanowisko pomiarowe

Stanowisko pomiarowe składa się z klatki badań i kamery termograficznej lub zwykłej umieszczonej 120 cm nad klatką. W większości przypadków kamery zostały ustawione centralnie nad polem doświadczenia, ponieważ taki widok dostarcza najwięcej danych na temat poruszania i kształtu obiektu. Poza tym, jest to najczęściej używane położenie w istniejących komercyjnych systemach analizy prostych zachowań i motoryki zwierząt. Istnieją rozwiązania analizujące zachowanie gryzoni, które bazują na innym niż górne ustawieniu kamery. Jednym z takich rozwiązań jest ułożenie kamery z boku klatki opisane np. w [21], [22]. Jednak takie podejście musi poradzić sobie z zasłonięciem jednego osobnika przez drugiego, prawidłową oceną położenia względnego obiektów i innymi problemami, od których wolne jest górne ułożenie kamery. W celu przetestowania właściwości nagrań z boku, baza BD_{W2} (tabela 4.1) zawiera taki rodzaj nagrań wraz ze standardowymi nagraniami z góry.

Do badań wykorzystano klatki o wymiarach:

- klatka duża - 35 x 45 x 46 cm (długość x szerokość x wysokość) - dla bazy BD_{T1} , BD_{T2} , BD_{T3} oraz BD_{W1} , nie posiada górnej pokrywy, dzięki czemu możliwe było wykonanie nagrań kamerami termograficznymi;
- klatka mała - 22 x 35 x 45 cm - dla bazy BD_{W2} ;
- klatka potrójna - 36 x 53 x 6 cm - dla bazy BD_{W3} , przeznaczona wyłącznie dla myszy.

Wszystkie klatki wykonane były z pleksi specjalnie do celów nagrań zwierząt. Przeznaczeniem dwóch pierwszych klatek były krótkie testy socjalne szczurów. Ostatnia klatka zaprojektowana była dla długotrwałych (dobowych) testów zdolności motorycznych myszy. Klatka ta zawiera 3 osobne, jednakowe labirynty z przeszkodami w formie rozwidleń, schodów i ramp. Posiada przykrywą z wywierconymi otworami oraz miejsce na pokarm i pojemnik z wodą. Wszystkie klatki były dla zwierząt obcym środowiskiem, ponieważ czas osuwania w nich był krótki lub żaden. Możliwe jest, że materiał wykonania miał wpływ na zachowanie się gryzoni, ponieważ przy pewnym oświetleniu zauważalne są odbicie w ścianach z pleksi. Wrażliwość większej liczby osobników w pomieszczeniu może mieć bardzo istotny wpływ na jakość testów interakcji socjalnej.

Kamery używane w doświadczeniach to:

- FLIR A320G, maksymalna rozdzielczość przestrzenna - 320 x 240 pikseli, maksymalna częstotliwość nagrywania obrazu - 60 klatek/s - kamera używana we wszystkich nagraniach termograficznych,
- Iqin Vision IQEye, maksymalna rozdzielczość przestrzenna - 640 x 480 pikseli, maksymalna częstotliwość nagrywania obrazu - 30 klatek/s - kamera używana w dodatkowych nagraniach w zakresie promieniowania widzialnego.

4.3. Zwierzęta uczestniczące w badaniach

Wszystkie procedury eksperymentalne zostały zatwierdzone przez Lokalną Komisję Etyczną ds. doświadczeń na zwierzętach i zostały przeprowadzone zgodnie z dyrektywą Rady Wspólnoty Europejskiej.

4.3.1. Szczury

W badaniach rejestrowanych w bazach: BD_{T1} , BD_{T2} , BD_{T3} , BD_{W1} oraz BD_{W2} brały udział samce szczurów szczepu Wistar w wieku od 12 do 20 tygodni. Wykonane zostało również jedno nagranie porównawcze z udziałem 17-miesięcznego osobnika. Łączna liczba zwierząt to 23 osobniki. Nagrania wykonywane były w Katedrze Fizjologii Zwierząt Uniwersytetu Gdańskiego. Szczegółowa charakterystyka zwierząt została przedstawiona w tabeli 4.1. Osobniki z bazy BD_{T1} i BD_{W1} pochodzili od innego hodowcy niż te z bazy BD_{T2} , BD_{T3} i BD_{W2} . Wyraźnie różniły się podatnością na agresywne zachowanie, co prawdopodobnie spowodowane było różnym pochodzeniem genetycznym. Szczury BD_{T1} i BD_{W1} wykazywały dużą większą skłonność do walki.

Każde nagranie wykonywane było na parze zwierząt. Przed testami zwierzęta przetrzymywane były w pięcioosobowych grupach z nieograniczonym dostępem do wody i pożywienia. Zwierzęta w większości przypadków nie były przyzwyczajane do klatki. Wyjątek stanowią nagrania z bazy BD_{T3} , w których próbowano wywołać zachowania agresywne u badanych osobników. poprzez:

- wymuszenie społecznej konfrontacji wg modelu rezydent-intruz poprzez wcześniejsze przystosowanie niektórych osobników do pomieszczenia nagrań w formie kilkudziesięciominutowego przebywania w nim kilka dni przez wykonaniem nagrań [76, 55],
- podskórne podanie amfetaminy w dawce 3 mg/kg na 15-75 minut przed doświadczeniem [130, 133].

W dniu nagrań do bazy BD_{T3} zwierzęta były bardzo aktywne i niespokojne od wczesnych godzin porannych (z nieznanych przyczyn). Skutkowało to wieloma mokrymi śladami na futrze widocznymi na nagraniach, które prawdopodobnie powstały w wyniku przypadkowego dotknięcia dystrybutora wody.

4.3.2. Myszy

Do nagrań z bazy BD_{W3} wykorzystano 13 myszy UCP2 KO (szczep z inaktywowanym genem UCP2) i 11 myszy kontrolnych. Wszystkie zwierzęta były w wieku 12 miesięcy, pochodziły z laboratorium Jackson (Bar Harbor, Maine, USA). Badania przeprowadzane były w Katedrze Neurologii oraz Katedrze Histologii i Immunologii Gdańskiego Uniwersytetu Medycznego.

4.4. Przebieg doświadczeń

Każde pojedyncze nagranie testów socjalnych kamerą termowizyjną pary szczurów (bazy BD_{T1} , BD_{T2} , BD_{T3} w tabeli 4.1) trwało około 17 minut. Miejscem nagrania była klatka (duża lub mała), która nie służyła do przetrzymywania zwierząt. W pomieszczeniu nagrania wyłączono światło, oświetlony był pokój obok przy otwartych drzwiach. Podczas trwania całego nagrania w pomieszczeniu panowała cisza, nie przebywał w nim człowiek. Zwierzęta nie miały dostępu do pokarmu ani wody.

Nagrania testów socjalnych kamerą z zakresu promieniowania widzialnego (bazy BD_{W1} , BD_{W2}) trwały od 7 do 20 minut. Miejscem nagrania była klatka (duża lub mała), która nie służyła do przetrzymywania zwierząt (szczurów). W pomieszczeniu nagrania było włączone światło, co mogło wpłynąć na zachowania zwierząt.

Zwierzęta nie miały dostępu do pokarmu i wody.

Każde z nagrań testów motorycznych (baza BD_{W3}) trwało 24 godziny przy zapalonym świetle w klatce potrójnej. Jednocześnie rejestrowano 3 osobniki (myszy), które miały dostęp do pokarmu i wody. W pomieszczeniu nagrania nie przebywał człowiek.

4.5. Baza danych

Do celów pracy stworzono bazę danych przedstawioną w tabeli 4.1 składającą się z nagrań różnych osobników w różnym czasie i przy różnych warunkach.

Dane oznaczone jako BD_{T1} składają się z 2 nagrań o łącznej długości 40 minut. Bazy BD_{T2} i BD_{T3} są obszerniejsze, zawierają odpowiednio 110 i 125 minut nagrań zapisanych jako sekwencje danych termicznych (.SEQ). Baza BD_{W1} to 60 minut nagrań, z kolei w skład bazy BD_{W2} wchodzi dwa 90-cio minutowe nagrania (z kamer górnej i bocznej). Największą ilość danych zawiera baza BD_{W3} przedstawiająca 8 24-godzinnych nagrań myszy.

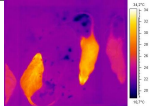
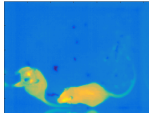
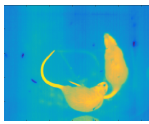



Warunki standardowe wymienione w tabeli 4.1, w jakich nagrywane były bazy BD_{T1} i BD_{T2} , oznaczają typowe warunki testów socjalnych, czyli:

- udział 2 samców szczurów,
- zwierzęta nie były wcześniej przyzwyczajane do klatki,
- brak dostępu do pokarmu i wody w trakcie trwania testu,
- widok kamery z góry,

Nagrania do bazy BD_{T1} zostały wykonane przy standardowych warunkach na tych samych osobnikach co do bazy BD_{W1} przy użyciu innej technologii kamery. Nie przedstawiają jednak tych samych testów. W pierwszej kolejności wykonano nagrania do bazy BD_{W1} , następnie do bazy BD_{T1} . Zwierzętom wykonano oznaczenia markerem na ogonach w celu identyfikacji. Ślady te widoczne były tylko na nagraniach w zakresie promieniowania widzialnego. Nagrania bazy BD_{T2} oraz BD_{T3} wykonane są również na jednakowej grupie zwierząt przy użyciu kamer termowizyjnych w miesięcznym odstępnie czasu, różnią się warunkami wcześniejszego przygotowania zwierząt. Warunki nagrań bazy BD_{T2} to warunki standardowe. Podczas rejestracji bazy BD_{T3} zwierzęta także nie miały dostępu do pokarmu i wody, kamera usytuowana była nad klatką w której jednorazowo znajdowały się wyłącznie 2 samce. Zwierzęta jednak zostały wcześniej przygotowane wg opisu zawartego w rozdziale 4.3.1 w celu zwiększenia agresji między osobnikami, która była niespodziewanie mała podczas wcześniejszych obserwacji zwierząt. Baza BD_{W3} została stworzona do analizy motoryki myszy w badaniach opisanych w pracy [92] oraz rozdziale 3.4. Nie zawiera informacji o jakichkolwiek zachowaniach socjalnych gryzoni.

Dane otrzymane z kamery termowizyjnej zostały przekształcone z pliku .SEQ do zestawu plików .MAT (Matlab) za pomocą oprogramowania FLIR, w wyniku czego każda klatka mogła być prezentowana jako macierz (240 wierszy x 320 kolumn) temperatury w stopniach Kelvina.

TABELA 4.1. Bazy danych nagrań

Na- zwa	Roz- miar [min]	Liczba testów	Osobniki	Rozdziel- czość obrazu	fps	Warunki	Przykład
Testy socjalne - nagrania termograficzne							
BD_{T1}	40	2	4 w wieku około 16 tyg.	320 x 240	1	warunki standardowe	
BD_{T2}	110	6	10 w wieku 12 tyg.+1 w wieku 17 miesiący	320 x 240	60	warunki standardowe	
BD_{T3}	125	7	10 w wieku 15 tyg.	320 x 240	60	zwierzęta wcześniej przyzwycza- jone do klatki, prze- głodzone, po podaniu amfetaminy	
Testy socjalne - dodatkowe nagrania w zakresie promieniowania widzialnego							
BD_{W1}	60	7	4	640 x 480	10	zwierzęta oznaczone markerem na ogonach w celu identy- fikacji	
BD_{W2}	90 x 2	6	5	640 x 480	10	zwierzęta oznaczone markerem na ogonach oraz futrze, klatka o wymiarach 22 x 35 x 45 cm, nagrania z góry i z boku	
Testy motoryczne - dodatkowe nagrania w zakresie promieniowania widzialnego							
BD_{W3}	11520 (192h)	8	24 myszy w wieku 12 mies.	640 x 480	30	myszy	

W nagraniach bazy BD_{T1} oraz BD_{W1} można zaobserwować liczne zachowania agresywne, częste podgryzanie się osobników, zastraszanie czy śledzenie. Są to typowe reakcje socjalne szczurów w nowym otoczeniu w stosunku do nieznanego osobnika - pożądane ze względu na badania prowadzone w ramach tej pracy. Stopień agresji zwierząt uczestniczących w nagraniach do bazy BD_{T2} był znacznie mniejszy. Wynikało to prawdopodobnie z tego, że zwierzęta pochodziły z innego źródła - hodowli zwierząt tylko laboratoryjnych, u których zanikają pierwotne instynkty.

4.6. Podsumowanie

W celu ewaluacji postawionych w pracy tez stworzono stanowisko oraz procedurę nagrań zachowań socjalnych przy użyciu różnych technologii i zwierząt. Zebrano bazę różnorodnych zachowań składającą się z ponad 500 minut nagrań socjalnych i prawie 200 godzin nagrań motorycznych. Wszystkie bazy przedstawione w tym rozdziale stanowią razem zbiór zachowań gryzoni. Były one tworzone dla różnych celów, w różnych warunkach i przy użyciu różnych technologii przedstawiając tym samym różne modele zachowań. Ze względu na niewielką liczbę zebranych nagrań ilustrujących różnorodność zachowań agresywnych, spowodowaną nietypowym brakiem cech agresywnych u niektórych zwierząt, zaproponowano w pracy metody symulacyjne przedstawione w rozdziale 6.1. Nagrania zebrane w opisanych bazach danych zostały wykorzystane do analizy obrazów opisanej w rozdziale 5 oraz analizy parametrów zachowań agresywnych w rozdziale 7.

Rozdział 5

Analiza obrazów dla potrzeb oceny zachowań zwierząt

Podstawowym etapem analizy zachowań socjalnych gryzoni jest analiza obrazu, w której skład wchodzi przetwarzanie wstępne, detekcja obiektu, jego śledzenie w kolejnych klatkach nagrania oraz ekstrakcja określonych cech charakterystycznych. Wynikiem zastosowania tych algorytmów są informacje o obiekcie, takie jak: położenie wewnątrz klatki, wzajemne położenie obiektów, parametry kształtu ciała (powierzchnia, kontur itp.) oraz, w przypadku obrazowania termicznego, rozkład temperatury ciała, położenie i temperatura punktów charakterystycznych (nos, oczy, podstawa ogona, ślady na futrze).

Rozdział 5 zorganizowany jest w następujący sposób:

- Rozdział 5.1 opisuje przetwarzanie wstępne: proces tworzenia obrazów na podstawie danych termicznych oraz zastosowanej segmentacji obiekt-tło,
- Rozdział 5.2 przedstawia popularne metody śledzenia obiektu oraz wyniki ich zastosowania na obrazie gryzoni,
- Rozdział 5.3 to przegląd popularnych algorytmów segmentacji, które mogą być zastosowane do rozdzielania obszaru każdego z osobników (segmentacji obiekt-obiekt),
- Rozdział 5.4 prezentuje propozycje algorytmu rozdzielania zwierząt podczas złączenia (segmentacja obiekt-obiekt),
- Rozdział 5.5 opisuje wyniki adaptacji algorytmów uczenia głębokiego do segmentacji obiektów,
- Rozdział 5.6 podsumowuje rozdział.

W pracy wyróżnione zostały pojęcia dwóch segmentacji: (i) segmentacja obiekt-tło opisana w rozdziale 5.1 oddzielająca analizowane obiekty (zwierzęta) od tła (klatki) oraz (ii) segmentacja obiekt-obiekt przedstawiona w rozdziałach 5.3 i 5.4. Podczas kontaktów fizycznych zwierzęta znajdują się w małej odległości od siebie, a często nawet ich ciała na siebie zachodzą. Powoduje to powstanie na obrazie złączenia dwóch podobnych do siebie obiektów lub zasłonięcia jednego przez drugi. Z tego powodu segmentacja obiekt-obiekt jest zagadnieniem trudniejszym niż odseparowanie obiektu od otaczającego go tła.

5.1. Przetwarzanie wstępne

Pierwszym krokiem przetwarzania wstępnego jest przekształcenie danych termicznych otrzymanych z kamer na jednokanałowe obrazy 8 bitowe. W celu zbadania wpływu dokładności wartości temperatury na wyniki segmentacji stworzono także jednokanałowy obraz 16-bitowy. Następnie wykonane zostało progowanie z określoną wartością temperatury w celu odseparowania tła od obiektów. Ta metoda progowania mogła zostać użyta dzięki stałej i dużej różnicy temperatury ciała gryzonia i otoczenia (około 6 K). Ostatnim etapem przetwarzania wstępnego było zastosowanie algorytmu połączonych komponentów w celu odróżnienia obiektów od siebie i nadania im osobnych etykiet. Algorytm połączonych komponentów był stosowany tylko w przypadkach, gdy ciała zwierząt nie stykały się, w przeciwnym razie stosowano metody segmentacji obiekt-obiekt opisane w sekcji 5.3, 5.4 i 5.5.

5.1.1. Przekształcenie danych termicznych na obrazy

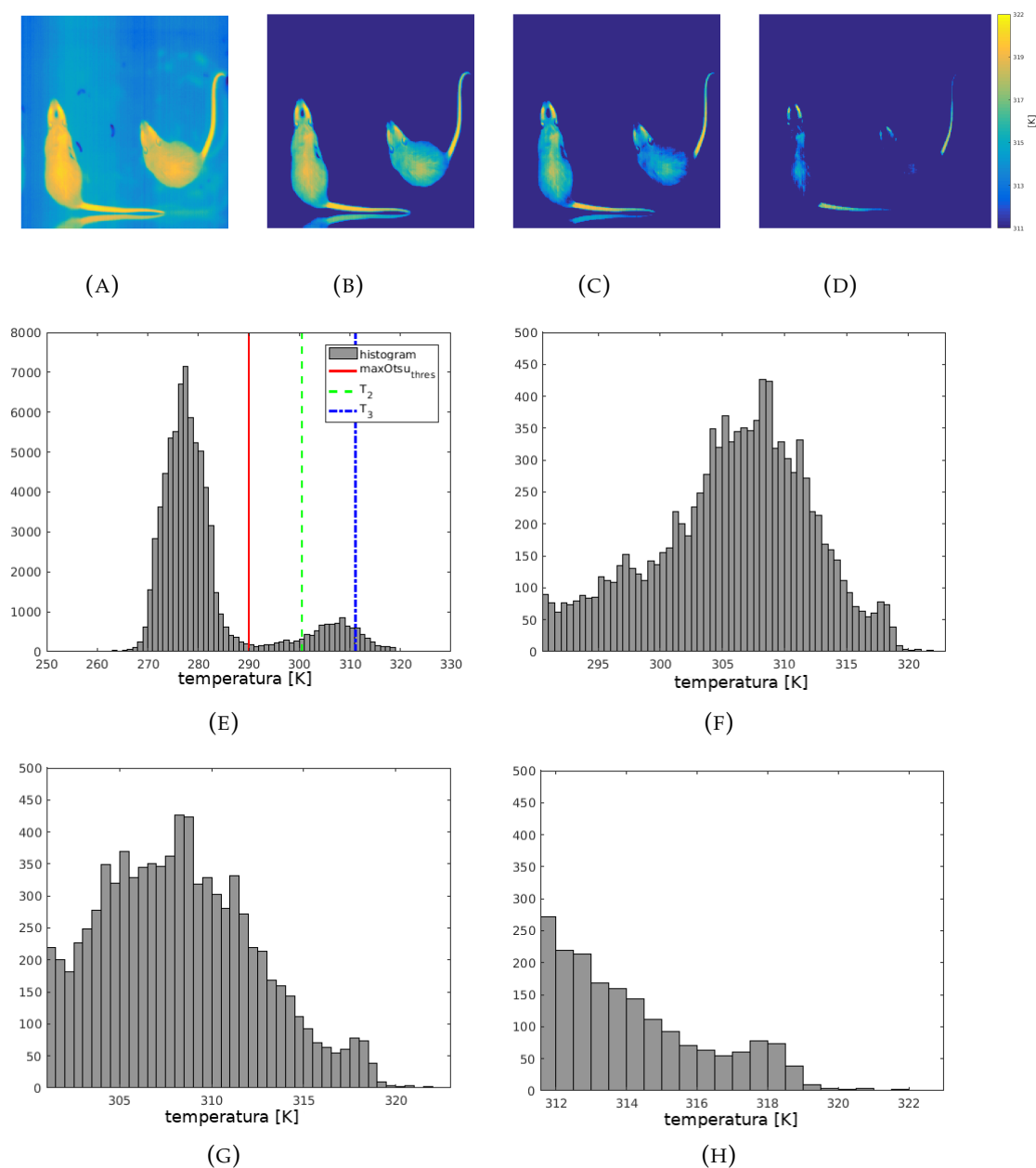
Przeskalowanie 65536 (2^{16}) poziomów do 256 (2^8) poziomów powoduje utratę dokładności. W zastosowaniach medycznych okna obserwacyjne są zwykle definiowane w celu wybrania zakresu wartości z oryginalnego zestawu danych i odwzorowania tego zakresu na dostępny zakres odcieni szarości. Histogram przykładowych wartości temperatury zarejestrowanych dla jednej klatki pokazano na rys. 5.1e. Przeskalowanie całego zakresu temperatur do jednego kanału standardowego obrazu (256 wartości) powoduje utratę dokładności.

Badanie niniejszej pracy miało na celu m.in. zbadanie wpływu doboru zakresu danych na wyniki segmentacji obiektów. W tym celu z każdej klatki zarejestrowanych wartości temperatury utworzono 4 różne obrazy 8-bitowe i jeden 16-bitowy. Dla pierwszego obrazu (zwanego dalej "ch0"), cały zakres temperatur, od wartości tła do najcieplejszych obszarów ciała zwierzęcia (T_{max}), został przeskalowany na standardowy obraz w 8-bitowej skali szarości (rys. 5.1a). Pozostałe 3 8-bitowe obrazy powstały tylko na podstawie temperatury zwierząt. Aby określić temperaturę progową między tłem a obiektem, zastosowano metodę Otsu [107]. Średnia wartość progu Otsu dla przedstawionego przykładu wynosiła $295,7 \pm 1,2$ K; 291,4 i 300,1 K to odpowiednio wartości minimalne i maksymalne. W dalszej analizie uwzględniono jedynie minimalną wartość progu Otsu, oznaczoną jako $minOtsu_{thres} = 290$ K (rys. 5.1e). Powody wyboru tych wartości przedstawione są w kolejnym rozdziale. Drugi obraz (dalej określany jako "ch1") został stworzony na podstawie wartości temperatur pomiędzy $minOtsu_{thres}$ a maksymalną wartością obrazu (T_{max}) (rys. 5.1b, 5.1f). Zakres trzeciego obrazu (zwanego "ch2") wynosił od T_2 do T_{max} (rys. 5.1c, 5.1g). Zakres czwartego obrazu ("ch3") został ustawiony między T_3 a T_{max} (rys. 5.1d, 5.1h) gdzie:

$$T_2 = minOtsu_{thres} + \frac{T_{max} - minOtsu_{thres}}{3} \quad (5.1)$$

$$T_3 = minOtsu_{thres} + 2 \cdot \frac{T_{max} - minOtsu_{thres}}{3} \quad (5.2)$$

Obrazy 16-bitowe zachowały oryginalne wartości danych termicznych.



RYS. 5.1. Obrazy 8-bitowe stworzone z różnych zakresów danych termicznych wraz z ich histogramami: a) obraz uzyskany z całego zakresu - $ch0$, e) histogram przykładowej klatki danych termicznych z zaznaczonymi wartościami: $minOtsu_{thres}$ oznaczona kolorem czerwonym, T_2 - zielonym i T_3 - niebieskim, b) obraz uzyskany z pierwszego zakresu: od $minOtsu_{thres}$ do T_{max} : $ch1$, f) histogram pierwszego zakresu, c) obraz uzyskany z drugiego zakresu: od T_2 do T_{max} : $ch2$, g) histogram drugiego zakresu, d) obraz uzyskany z trzeciego zakresu: od T_3 do T_{max} : $ch3$, h) histogram trzeciego zakresu.

5.1.2. Segmentacja obiekt-tło

Segmentacja obrazu jest jednym z etapów przetwarzania wstępnego używanego do podzielenia obrazu na podobszary. Podstawowym etapem segmentacji jest oddzielenie pierwszego planu od tła. Istnieje kilka technik segmentacji sklasyfikowanych wg [151], [26] w skład których wchodzi m.in.:

- metody gradientowe i oparte na histogramach,
- progowanie i metody działania na regionach,
- metody oparte na sieciach neuronowych.

Metody stosowane dla obrazów termograficznych obejmują: algorytm level-set [113], progowanie [8], detekcję krawędzi [36] lub algorytm K-średnich [142, 99]. Bardzo często progowanie Otsu wraz z odpowiednimi operacjami morfologicznymi są wystarczające do segmentacji obiekt-tło w przypadku obrazów termograficznych [46, 44].

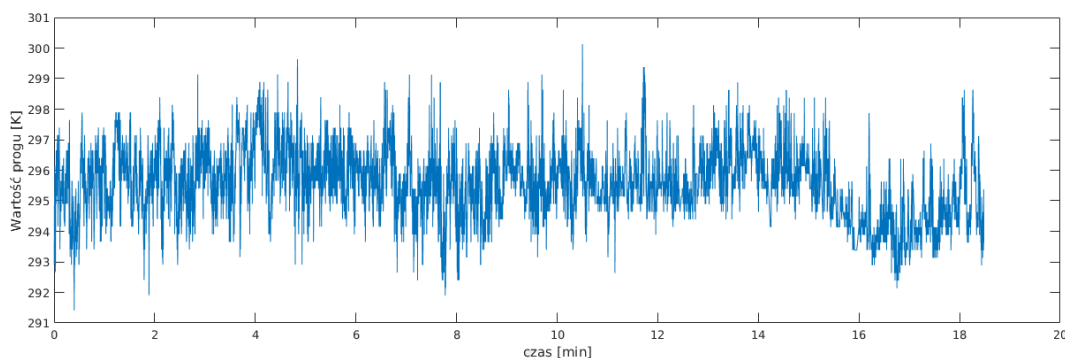
Stopień skomplikowania problemu oddzielenia tła od pierwszego planu zależy od obrazu. Projektując system do analizy badań można nałożyć pewne ograniczenia i założenia sposobu wykonywania doświadczeń oraz warunki otoczenia tak, aby jak najbardziej uprościć pierwszy etap analizy obrazu. Z tego powodu w wielu systemach wymagany jest znaczny kontrast pomiędzy obiektem a tłem. Umożliwia to zastosowanie najprostszej metody detekcji zwierzęcia jaką jest progowanie. Obraz zamieniany jest na obraz bitowy (binarny) w taki sposób, że wszystkie piksele o kolorze mniejszym niż zadany próg są zamieniane na kolor czarny, a większym na kolor biały, lub odwrotnie. Użytkownik stosujący tę metodę musi zabezpieczyć nagranie przed wystąpieniem innych przedmiotów o kolorze (lub temperaturze) zbliżonym do koloru obiektu na scenie doświadczenia.

W zaproponowanym systemie dla obrazów, które zostały przetworzone z danych termicznych w całym swoim zakresie, razem z tłem (obrazy 16-bitowe oraz *ch0*) wykonano progowanie, ze względu na dużą różnicę temperatury obiektu i otoczenia oraz brak innych obiektów o zbliżonej temperaturze w polu widzenia kamery. Poniżej przedstawiono sposób wyboru progu dla przykładowego nagrania. Dla każdego nagrania wykonano osobne obliczenia ze względu na potencjalne niedokładne pomiary kamery.

Tabela 5.1 oraz rys. 5.2 obrazują zmiany progu obliczanego przez metodę Otsu dla każdej pojedynczej klatki w ciągu jednego przykładowego 19-minutowego nagrania. Wartość średnia progu to 295,71 K o małym odchyleniu standardowym równym 1,17 K, co wskazuje, że wyniki dla różnych klatek są do siebie zbliżone. Ostatnia wartość w tabeli oznaczona jako EM to metryka efektywności, która przyjmuje wartość 1 dla obrazów o dwóch rozdzielnych histogramach, a wartość 0 dla histogramów unimodalnych lub bez wyraźnej granicy. Wartość $1 \pm 8 * 10^{-15}$ wskazuje na istnienie wyraźnej różnicy pomiędzy temperaturami obiektów i tła dla całego nagrania.

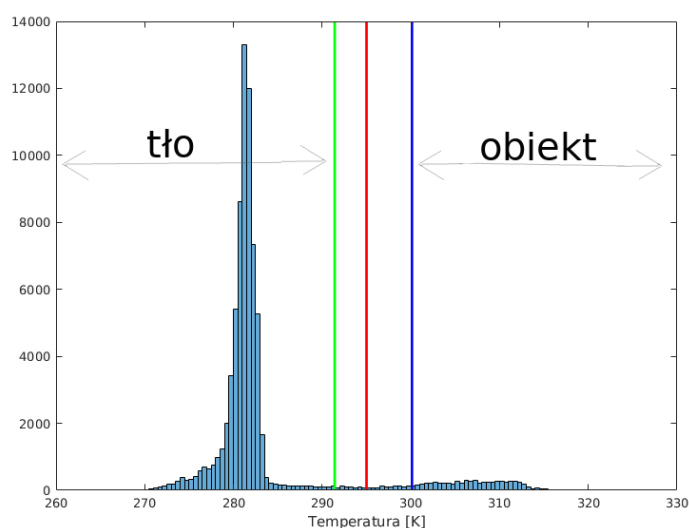
TABELA 5.1. Wyniki algorytmu Otsu dla 19-minutowego nagrania

próg [K] ± odchylenie standardowe	295,71 ± 1,17
minimum [K]	291,41
maksimum [K]	300,13
EM ± odchylenie standardowe	$1 \pm 8 * 10^{-15}$



RYS. 5.2. Wykres wartości progu obliczonego metodą Otsu dla 19-minutowego nagrania

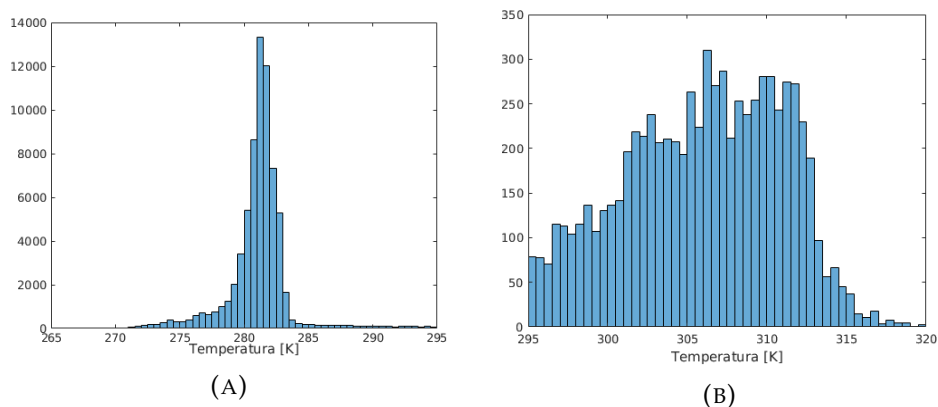
Histogram przykładowego obrazu z zaznaczonymi wartościami minimalną, maksymalną oraz średnią obliczonego progu przedstawiono na rys. 5.3. Wyraźnie widoczna jest granica pomiędzy temperaturą ciała zwierzęcia (wartości od 295 do 320 K) oraz tła (poniżej 295 K). Wszystkie zaznaczone wartości znajdują się w obszarze, który jest granicą obu histogramów. Osobne histogramy dla tła i obiektu zostały zaprezentowane na rys. 5.4. Rys. 5.5 obrazuje zastosowanie progowania o obliczonych wartościach progu odpowiednio: minimalnej, maksymalnej oraz średniej. Jak widać dobór dowolnego progu z rozpatrywanego zakresu zapewnia poprawne progowanie obrazu. Minimalne różnice występują jedynie w przypadku ogona, który został pominięty dla progowania wartością maksymalną oraz odbić obiektów w ścianach klatki, które wystąpiły podczas progowania wartością minimalną. Obraz dla progu średniego jest pozbawiony tych błędów.



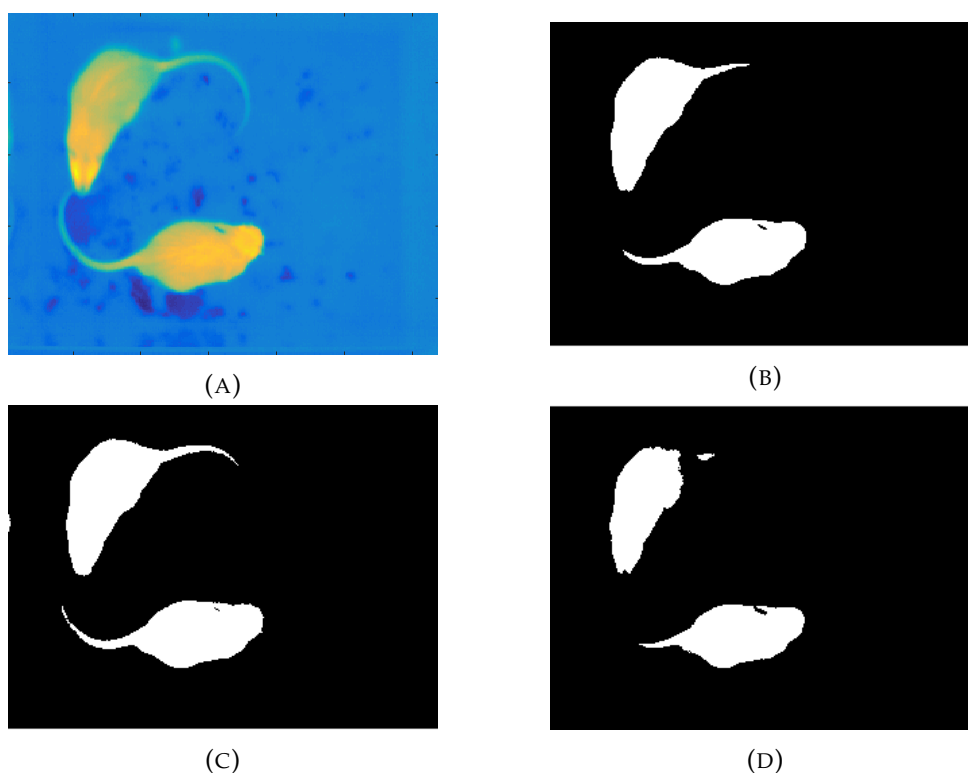
RYS. 5.3. Histogram obrazu z zaznaczonymi obliczonymi progami: próg minimalny - kolor zielony, próg maksymalny - kolor niebieski, próg średni - kolor czerwony

5.1.3. Metoda połączonych komponentów

Metoda połączonych komponentów (ang. Connected Component Analysis) to metoda detekcji obiektów polegająca na przyporządkowywaniu etykiet do elementów obrazu [121], [37]. Algorytm skanuje macierz obrazu linia po linii w poszukiwaniu pikseli



RYS. 5.4. Histogramy dla a) tła oraz b) obiektu



RYS. 5.5. Wyniki progowania obrazu (a) dla wartości progu równego b) obliczonej wartości średniej, c) obliczonej wartości minimalnej, d) obliczonej wartości maksymalnej

spełniających zadany warunek (np. o odpowiednim kolorze). Jeśli natrafi na pierwszy taki piksel, nadaje mu etykietę. Przy odnalezieniu kolejnych pikseli algorytm sprawdza czy sąsiedzi tych pikseli także posiadają jakąś etykietę: jeśli nie – nadaje nową etykietę pikselowi, jeśli tak – nadaje pikselowi etykietę o najmniejszej wartości z etykiet sąsiadów. Po zakończeniu sprawdzania całego obrazu algorytm przystępuje do powtórnego skanowania w celu skorygowania etykiet obszarów, których różnie etykietowane podobszary połączyły się w pewnym momencie. Metoda połączonych komponentów została zastosowana wyłącznie dla obrazów z dwoma rozdzielnymi obiektami.

5.1.4. Podsumowanie przetwarzania wstępnego

Przedstawione powyżej przetwarzanie wstępne ma na celu: po pierwsze, przekształcenie surowych danych na obrazy; po drugie, przygotowanie obrazów do dalszej analizy. Dla wszystkich obrazów wykonano segmentację obiekt-tło. Dla przypadków, dla których było to możliwe zastosowano wstępną etykietyzację obiektów.

5.2. Algorytmy śledzenia obiektów

Śledzenie jest jednym z elementów analizy zachowań. Poniżej przedstawiono kilka popularnych prostych algorytmów dotyczących śledzenia oraz ich porównanie wykonane dla obrazów w zakresie promieniowania widzialnego ze zgromadzonej bazy danych.

5.2.1. Metoda CamShift

Metoda CamShift (ang. Continously Adaptive Mean-SHIFT) oparta jest na przesuwaniu wcześniej określonego okna o adaptacyjnym rozmiarze zgodnie z obliczonym środkiem ciężkości [31], [47]. Algorytm ten działa na obrazie rozkładu gęstości koloru (ang. histogram backprojection) w następujący sposób [17]:

1. wybierane jest okno poszukiwań – jego początkowa lokalizacja, typ, kształt, rozmiar,
2. obliczany jest środek ciężkości dla pikseli zawartych w oknie,
3. okno zostaje wyśrodkowane do obliczonego środka ciężkości,
4. powrót do kroku 2. aż do momentu, gdy okno przestanie się przemieszczać (zgodnie z kryteriami określającymi maksymalną liczbę iteracji lub przemieszczenie środka okna między kolejnymi iteracjami).

W swojej pracy Bradski [18] przedstawił CamShift jako algorytm użyteczny w śledzeniu ludzkiej twarzy na podstawie koloru skóry.

5.2.2. Przepływ optyczny

Przepływ optyczny (ang. Optical Flow) jest to pole wektorowe, które opisuje zmianę jednego obrazu z sekwencji w drugi. Metoda ta została przedstawiona m. in. w pracach [61], [88]. Umożliwia ona śledzenie określonych punktów obrazu (intensywność pikseli) na kolejnych klatkach filmu określając wektor przemieszczenia każdego punktu pomiędzy poszczególnymi klatkami. Algorytm zakłada niezmienną jasność oraz małe przemieszczenie między kolejnymi obrazami.

5.2.3. Filtracja cząstkowa

Algorytm filtracji cząstkowej (ang. Particle Filtering) został stworzony w celu śledzenia częściowo zasłanianych obiektów [67], [6]. Ma on na celu oszacowanie sekwencji ukrytych parametrów na podstawie obserwacji. Rozkład jest aproksymowany przez pomiary (cząstki - particles) oraz przypisane im wagi, co powoduje uproszczenie obliczeń. Cząstki oraz ich wagi uaktualniane są po każdej obserwacji (resampling). Jest on powszechnie używany w śledzeniu wielu obiektów: mrówek [74], myszy z kamery ułożonej z boku klatki [20], [21] lub z góry [112].

5.2.4. Metoda aktywnych konturów

Metoda ta (ang. Active Contours) polega na minimalizacji funkcji energii dla każdego punktu konturu. Pozycja i kształt konturów zmieniają się iteracyjnie pod wpływem przyłożonych do nich sił nadającym tym ciałom pewną energię. Energia konturu jest zależna od kształtu (ciągłości i krzywizny) oraz położenia konturu na obrazie [73].

Wyżej wymienione algorytmy zostały zastosowane do śledzenia zwierząt doświadczalnych, ich wyniki przedstawiono w kolejnym rozdziale.

5.2.5. Wyniki zastosowania algorytmów śledzenia obiektów - analiza porównawcza

Przeprowadzono analizy porównawcze czterech algorytmów śledzenia (CamShift, przepływ optyczny, filtracja cząstkowa i metoda aktywnych konturów) podczas analizy ruchu dwóch szczurów znajdujących się w klatce wykonanej z plexi przy użyciu kamery z zakresu promieniowania widzialnego. Proces śledzenia wykonywany był tylko dla klatek z obiektami w pozycji rozłącznej. Badania przedstawione w wyżej wymienionej pracy sprawdzają, które algorytmy i w jakich warunkach, działają najlepiej oraz określają wpływ zaburzeń podczas nagrania na poprawność działania. Mierzonych jest kilka parametrów, w tym:

- optymalne warunki oświetlenia,
- średni czas analizy jednej klatki (o określonych rozmiarach i formacie),
- jakość śledzenia podczas zmian natężenia światła,
- jakościowa analiza śledzenia podczas zaburzeń (3 typy zaburzeń: duże obiekty w polu zainteresowania, małe obiekty obok pola zainteresowania, zmiany wartości pikseli lub odbicia w ścianach klatki)
- czas i stopień powrotu poprawnego śledzenia po zasłonięciu obiektu i przerwaniu śledzenia.

Rezultaty przeprowadzonych badań na obrazach widzialnych:

- Obraz niedoświetlony jest łatwiejszy w analizie niż obraz prześwietlony (ma to związek z zastosowanymi materiałami klatek, które powodują odbicia często rozpoznawane jako obiekty), jasne światło negatywnie wpływa również na zachowanie się gryzoni - jest to ważny powód do zastosowania alternatywnej technologii jaką jest termografia.
- Szybkość działania każdego z algorytmów zależy od kilku zmiennych (np. liczby śledzonych punktów dla przepływu optycznego, liczby obiektów dla filtracji cząstkowej lub wielkości konturu dla metody aktywnych konturów), jednak metoda CamShift jest najszybciej działającym algorytmem, a najdłużej działają metoda aktywnych konturów oraz filtracja cząstkowa dla więcej niż jednego obiektu.
- Zmiany częstotliwości klatkowania pomiędzy 5 a 30 klatek/s nie wpływają znacznie na jakość śledzenia obiektów nawet dla przepływu optycznego, dla którego jednym z założeń jest mały wektor przemieszczenia określonego piksela pomiędzy kolejnymi klatkami nagrania; jedynie częstotliwość 2,5 klatek/s powoduje gubienie obiektu przez prawie wszystkie metody.

- Zwiększenie natężenia światła (wzrost średniej wartości pikseli o 50% w skali szarości) nie miało wpływu na działanie wyłącznie algorytmu CamShift, wszystkie pozostałe metody zgubiły obiekt lub wykrywały ich więcej niż powinny (wynika to bezpośrednio z ich sposobów działania).
- Algorytm CamShift ma najlepsze wyniki działania podczas wystąpienia zaburzeń; chociaż gubi obiekt po zasłonięciu go przez inny przedmiot, bardzo szybko do niego powraca po zniknięciu zaburzenia (po około 0,3s).
- Przepływ optyczny jest bardzo wrażliwy na zaburzenia w polu zainteresowania, jakikolwiek obiekt występujący w bliskim sąsiedztwie śledzonego zwierzęcia wprowadza błędne rozpoznanie. Nie znajduje raz utraconego obiektu. Nie jest czuły na zaburzenia poza polem zainteresowania.
- Filtracja cząstkowa oraz metoda aktywnych konturów bardzo dobrze radzą sobie z odróżnieniem dużego zaburzenia w polu zainteresowania od obiektu, potrzebują natomiast znacznie więcej (4,7 s) czasu na całkowite odnalezienie zagubionego obiektu (odnalezienie części obiektu – około 50% - następuje po mniej więcej 0,4s).

Szereg badań dotyczących zagadnienia zastosowania algorytmów śledzenia zawarto w pracy [93].

5.2.6. Podsumowanie analizy algorytmów śledzenia

Z powyższych badań wynika, że proste algorytmy mogą być stosowane do śledzenia zwierząt laboratoryjnych. Jeżeli chodzi o szybkość oraz jakość śledzenia, najlepsze wyniki osiągnęła metoda CamShift. Wykazała także najlepszy czas powrotu po wystąpieniu zaburzenia oraz najmniejszą wrażliwość na zwiększenie natężenia światła. Prześwietlenie obrazów okazało się stanowić największe wyzwanie dla wszystkich algorytmów.

Powyższe metody były używane jedynie w przypadku obiektów rozłącznych. Nie są odpowiednie do śledzenia zwierząt znajdujących się w bliskim kontakcie. W przypadku złączenia się obiektów, należy wykonać segmentację obiekt-obiekt.

5.3. Algorytmy ekstrakcji obiektów połączonych - segmentacja obiekt-obiekt

Wyraźna różnica temperatur w obrazowaniu termicznym znacznie upraszcza segmentację obiekt-tło. Bardzo często jednak istnieje potrzeba dokładniejszej segmentacji, np. oddzielenie obiektu od innych części pierwszego planu lub rozróżnienie poszczególnych osobników. W przypadku analizy nagrań testów socjalnych zwierząt doświadczalnych jednym z największych problemów, z którymi musi się zmierzyć algorytm detekcji i śledzenia obiektu jest zasłonięcie obiektu przez inny lub „złączenie” się obiektów podobnych. Zasłonięcie obiektu występuje w momencie, gdy widok śledzonego obiektu zostanie przysłonięty przez inny obiekt. Złączenie natomiast jest wynikiem wystąpienia podobnego (strukturalnie, kolorystycznie itp.) obiektu w bardzo bliskiej odległości. Ułożenie kamery nad sceną obserwacji pozwala wyeliminować przypadki całkowitego zasłonięcia obiektów. Dużym problemem wciąż jednak pozostaje moment połączenia się obiektów często występujący w tym przypadku z częściowym zasłonięciem wynikającym z faktu, iż gryzonie wchodzi jeden na drugiego (rys. 5.6).



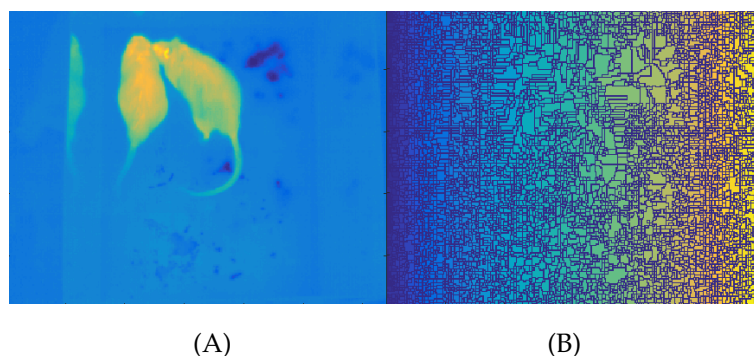
RYS. 5.6. Przykład złączenia obiektów

Moment kontaktu osobników jest w przypadku analizy zachowań socjalnych bardzo ważnym elementem, w którym to można zaobserwować charakterystyczne zachowania agresywne lub defensywne. Z tego powodu pominięcie go w analizie pozbawia wielu cennych danych.

Poniżej przedstawiono popularne algorytmy segmentacji, jakie mogą być przydatne podczas segmentacji typu obiekt-obiekt.

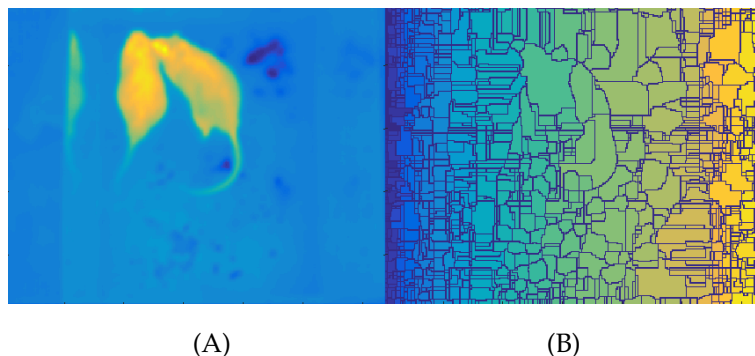
5.3.1. Algorytm działów wodnych (wododziałów)

Algorytm działów wodnych (tzw. watershed) jest metodą segmentacji obrazu tworzącą zamknięte regiony reprezentujące obszary [127], [144], [97]. Działa na zasadzie analogicznej do tworzenia się zlewni terenu geograficznego. Jako wysokość obrazu przyjmuje się najczęściej jego intensywność lub amplitudę gradientu intensywności. Obszary o małej wysokości stanowią lokalne doliny, a obszary o dużej wysokości - wzniesienia. Obszary o największej wysokości przedstawiają krawędzie - tzw. wododziały. Każdy region jest „zalewany” do pewnej granicy (wododziału) - piksele znajdujące się wewnątrz pikseli granicznych tworzą region. Często wadą tego algorytmu jest zbyt szczegółowa segmentacja, co widoczne jest na rys. 5.7. Obraz oryginalny został podzielony na podobszary, z którego każdy został oznaczony inną etykietą (kolorem).



RYS. 5.7. Działanie algorytmu działów wodnych: a) obraz oryginalny, b) algorytm działów wodnych wykonany na obrazie oryginalnym

W celu ograniczenia liczby obszarów stosowane są różne techniki wstępnego przetwarzania obrazu, m. in. rozmycie, uśrednianie, zastosowanie innych punktów początkowych (tzw. seed) czy transformacja odległościowa. Wpływ rozmycia na wynik algorytmu działów wodnych jest przedstawiony na rys. 5.8.



RYS. 5.8. Działanie filtracji oraz algorytmu działów wodnych: a) oryginalny obraz rozmyty poprzez zastosowanie filtru o masce o rozmiarze 5x5 pikseli i wartości każdego jej elementu wynoszącej 1/25, b) algorytm działów wodnych wykonany na obrazie rozmytym

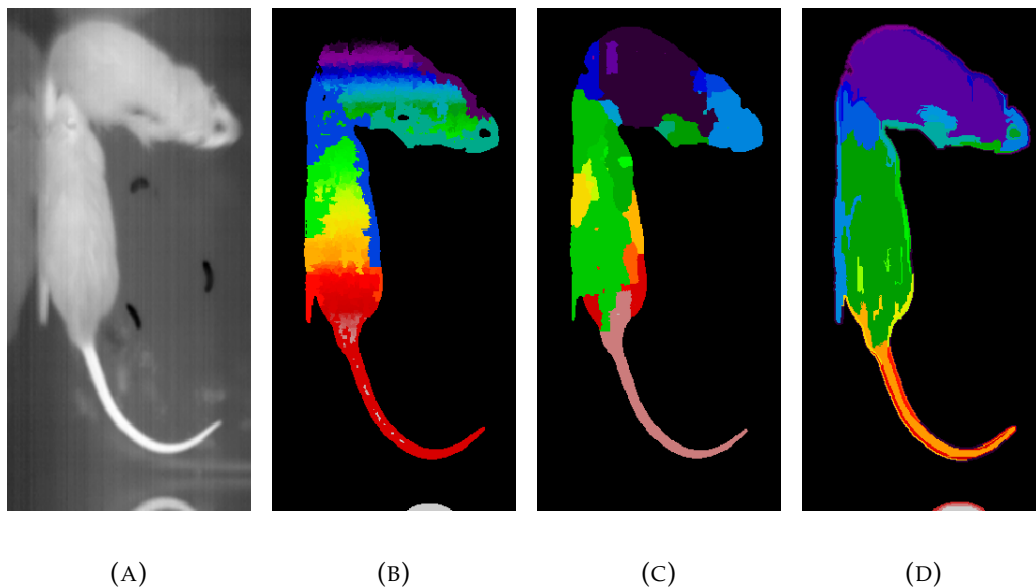
Rys. 5.7 i 5.8 przedstawiają obszary otrzymane w wyniku zastosowania algorytmu Fernanda Meyera, opisanego w pracy [97], dla sąsiedztwa N8. Każdy z otrzymanych regionów otrzymał etykietę przedstawioną kolorem na rysunkach. Region pierwszy (lewy górny róg obrazu) otrzymał etykietę = 1, przedstawioną kolorem ciemno-niebieskim, ostatni region (prawy dolny róg) - etykietę = 256, przedstawioną kolorem żółtym.

Porównując wyniki z rys. 5.7 oraz 5.8 można zaobserwować zmniejszenie liczby obszarów wykrytych przez algorytm działów wodnych. Nie wyodrębnione zostały jednak granice obiektów.

Inną metodą zmniejszenia liczby segmentów jest połączenie metody działów wodnych z bardziej ogólnymi algorytmami, wykrywającymi mniej szczegółów. Zhao i Zhuang w swojej pracy [153] połączyli algorytm aktywnych konturów (Snake) z algorytmem działów wodnych. Podobny algorytm nazwany WaterBalloons zaproponował w 2008 roku Dagher i Tom [33]. Natomiast Rafał Petryniak w swojej pracy doktorskiej [110] stworzył algorytm detekcji niejednorodnych obiektów w oparciu o algorytm działów wodnych oraz algorytm Canny. W pracy [75] zaproponowano dodanie do algorytmu działów wodnych markerów oznaczających minima transformacji odległościowej. Wynik segmentacji obrazu termograficznego szczurów dla takiego algorytmu przedstawiony jest na rys. 5.9b. Rys. 5.9c i 5.9d prezentują odpowiednio działanie metody działów wodnych z markerami opartymi na regionach najniższych gradientów oraz segmentację Felzenszwalb'a. Żadna z metod nie wyznacza jednoznacznie granic zwierząt.

Segmentacja obiektów, które w znacznym stopniu nachodzą na siebie oraz są do siebie podobne, nie jest łatwa i większość standardowych metod analizy obrazu zawodzi. Dlatego w ostatnim czasie prace badawcze dotyczące segmentacji skupiły się na metodach opartych na głębokim uczeniu, które zapewniają lepsze, niż metody standardowe, wyniki w wielu zagadnieniach widzenia komputerowego. Techniki te można podzielić na dwa podejścia:

1. segmentację semantyczną oraz
2. segmentację instancji.



RYS. 5.9. Wyniki segmentacji (a) obrazu oryginalnego przy użyciu: b) algorytmu działów wodnych z markerami opartymi na regionach najniższych gradientów, c) algorytmu działów wodnych z markerami opartymi na transformacji odległościowej d) segmentacja Felzenszwalb'a

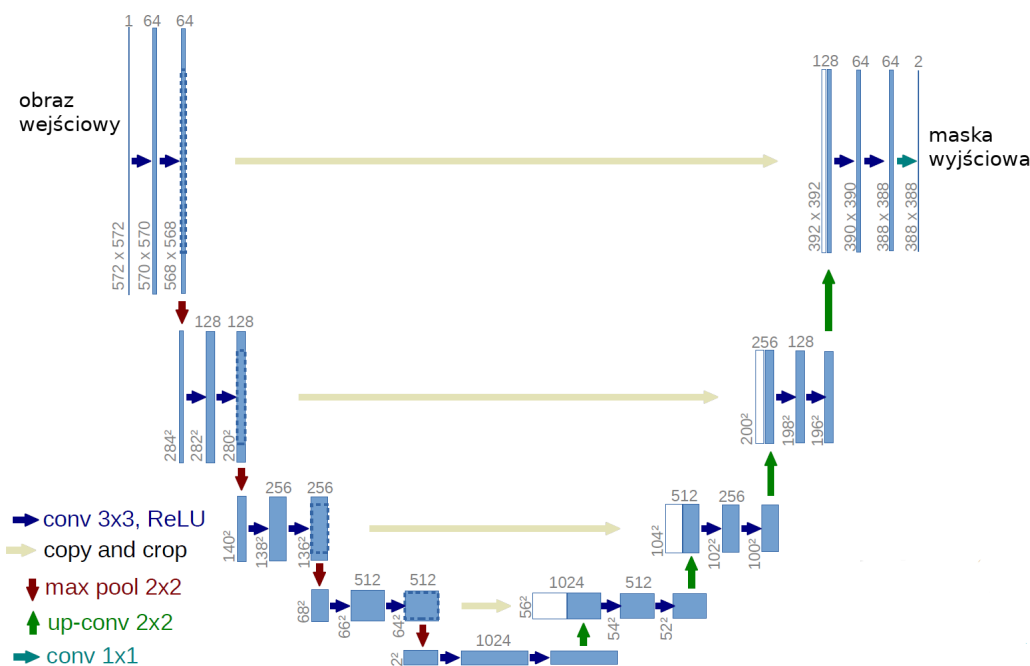
5.3.2. Segmentacja semantyczna

Segmentacja semantyczna to taki podział obrazu na podobszary, w którym każdy piksel przypisany jest do określonej klasy na podstawie swojej etykiety. Określenie etykiety każdego piksela jest wynikiem zastosowania metody segmentacji. Piksele mające takie same etykiety tworzą segmenty określonej klasy bez rozróżnienia instancji. Oznacza to, że wszystkie obiekty (instancje) tej samej klasy będą posiadały jednakową etykietę. Jednym z najczęściej stosowanych oraz najbardziej wydajnych algorytmów segmentacji semantycznej jest U-Net.

Został on wprowadzony w 2015 roku przez F. Ronnenberga, P. Fishera i T. Broxa w pracy [120] do segmentacji obrazów medycznych. Sieć U-Net jest oparta na modelu enkoder-dekoder (rys. 5.10), gdzie ścieżki kompresji (zbocze opadające) i dekompresji (zbocze narastające) są mniej lub bardziej symetryczne i tworzą architekturę w kształcie litery U. Ścieżka kompresji jest zgodna z typową architekturą sieci spłotowej z powtarzanymi: dwoma spłotami i ReLU (rektyfikowana jednostka liniowa -strzałki granatowe na rys. 5.10), następnie operacją max pooling (strzałka czerwona). Część dekompresji składa się ze spłotu (strzałka zielona), konkatencji z odpowiednią mapą cech ze ścieżki kompresji (strzałka szara), dwóch spłotów i ReLU. W ostatniej warstwie zastosowano spłot 1x1 (strzałka niebieska), aby odwzorować wektor cech na żadaną liczbę klas. Architektura auto-enkodera, który na każdym etapie ścieżki kompresji wyodrębnia cechy, a następnie przekazuje je do odpowiedniej ścieżki dekompresji, poprawia jakość końcowego wyniku [34]. Kolejną zaletą U-Net jest wymóg niewielkiej liczby obrazów treningowych dzięki zastosowaniu augmentacji danych przed treningiem. Augmentacja to proces generacji przekształconych obrazów na podstawie posiadanej bazy danych, czyli tworzenie obrazów poprzez np. translacje lub rotacje zaobserwowanych przypadków.

Sieć jest również w stanie nauczyć się rozdziałania połączonych obiektów tej samej klasy. Funkcja optymalizowana podczas treningu jest oparta na entropii krzyżowej.





Rys. 5.10. Architektura U-Net, na podstawie [120]

V-Net jest jedną z wielu architektur opartą na architekturze U-Net. Został przedstawiony w 2016 roku w artykule [98]. Inspiracją dla autorów był nie tylko model U-Net, ale także architektura sieci koniunkcyjnej VGG i jej lustrzanego odpowiednika [103], a także trzy w pełni splotowe głębokie sieci neuronowe wytrenowane do segmentacji semantycznej [128]. V-Net jest również w pełni splotową siecią neuronową, w której operacje łączenia (ang. pooling) zostały zastąpione przez splot. Proponowana funkcja kosztu optymalizowana podczas treningu jest oparta na współczynniku dice.

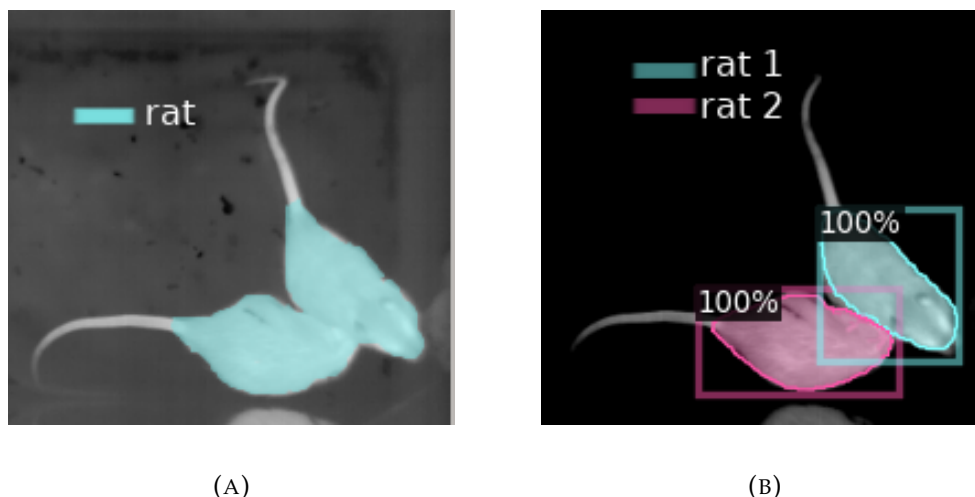
5.3.3. Segmentacja instancji

Segmentacja instancji jest jednym z najtrudniejszych zadań przetwarzania obrazu. W odróżnieniu od segmentacji semantycznej, pozwala na rozróżnienie obiektów tej samej klasy. Rys. 5.11 przedstawia różnice pomiędzy segmentacją semantyczną (rys. 5.11a) a segmentacją instancji (rys. 5.11b) na obrazach gryzoni.

Algorytmy segmentacji instancji klasyfikują i lokalizują obwiednię każdego obiektu, jednocześnie precyzyjnie segmentując każde wystąpienie. Oba te podzadania można wykonać jako jedno- lub dwuetapowy proces wykrywania.

Większość dwuetapowych metod przyjmuje paradygmat "znajdź-potem-segmentuj" (detect-then-segment). Na początku wykonywana jest detekcja obiektów a ich położenie oznaczane jest poprzez prostokąty (box). Następnym krokiem jest segmentacja każdego z wykrytych obiektów dzięki wcześniej określonym prostokątom. Chociaż metody dwuetapowe wykazują bardzo dobre wyniki, osiągają to kosztem czasu i zasobów sprzętowych.

Jednoetapowe metody upraszczają całą procedurę poprzez usunięcie prostokątów



RYS. 5.11. Różnice pomiędzy segmentacją a) semantyczną a b) segmentacją instancji

definiowanych w pierwszym etapie, jednak nie są w stanie uzyskać tak dokładnych masek segmentacji jak metody dwuetapowe. Przykładami architektur jednoetapowych mogą być FCOS [141], EmbedMask [152] lub TensorMask [30]. TensorMask to przykład metody segmentacji instancji wykorzystującej przesuwalne okno. Wykonuje klasyfikację wieloklasową i generowanie masek równoległe. Maski reprezentowane są jako strukturalne tensory 4D. Autorzy, twierdzą, że architektura TensorMask osiąga wyniki zbliżone do najpopularniejszej architektury dwuetapowej - Mask R-CNN.

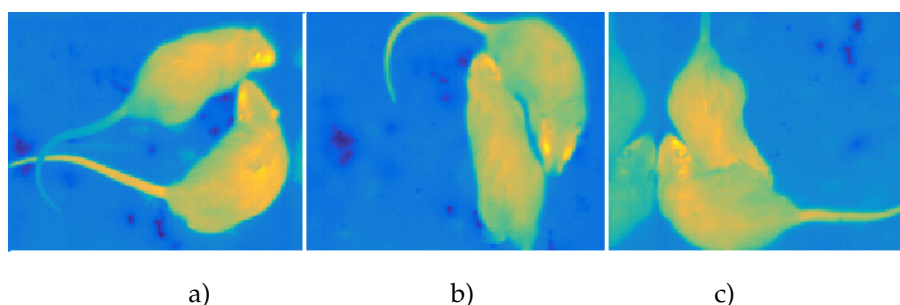
Mask R-CNN [59] oparta jest na metodach Fast [51] i Faster R-CNN [114], dodając element równoległy do rozpoznawania obszaru obiektu, który określa jego maskę. Ta metoda wykrywa obszar obiektu i generuje dla niego prostokąt, a następnie przycina go i dokładniej segmentuje. Wiele podejść opartych na Mask R-CNN zdominowało rankingi ostatnich wyzwań związanych z segmentacją: Mask Scoring R-CNN [62], PANet [85] czy Hybrid Task Cascade (HTC) [29].

5.4. Zaproponowany algorytm ekstrakcji obiektów połączonych

W celu analizy zachowań zwierząt podczas bliskich kontaktów fizycznych konieczne jest wykonanie segmentacji obiekt-obiekt. W pracy opracowano i zaadaptowano szereg algorytmów segmentacji. W rozdziale 5.4 zaproponowano algorytmy segmentacji wykorzystujące podstawowe metody widzenia komputerowego w celu rozdzielania złączonych obiektów. W kolejnym rozdziale 5.5 przedstawiono adaptację algorytmów segmentacji semantycznej i instancji do problemu rozróżnienia obiektów częściowo zasłoniętych w analizie zachowań obserwowanych zwierząt.

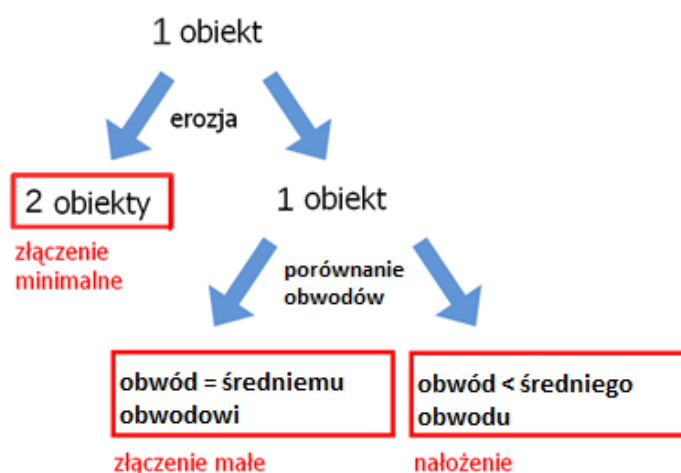
Zaproponowany algorytm rozdzielania dwóch połączonych obiektów dostosowany jest do wcześniej wykonanych nagrań termograficznych szczurów doświadczalnych. Zostaje on automatycznie uruchomiony w sytuacji, gdy liczba obiektów wykrytych przez metodę połączonych komponentów jest równa 1. Algorytm składa się z trzech etapów, wybór używanej metody zależy od stopnia złączenia. Wyróżniono następujące stopnie złączeń:

1. złączenie minimalne – występuje, gdy zwierzęta są blisko siebie lub się mijają, (rys. 5.12a),
2. złączenia małe – występują w sytuacjach takich samych jak złączenia minimalne, ale część wspólna obwodów obiektów jest większa (rys. 5.12b),
3. nałożenie – zasłonięcie części ciała jednego osobnika przez drugiego występujące np. podczas wspinania się (rys.5.12c).



RYS. 5.12. Przykłady złączenia obiektów a)złączenie minimalne, b)złączenie małe, c) nałożenie

Algorytm wyboru stopnia złączenia zilustrowano na rys. 5.13.

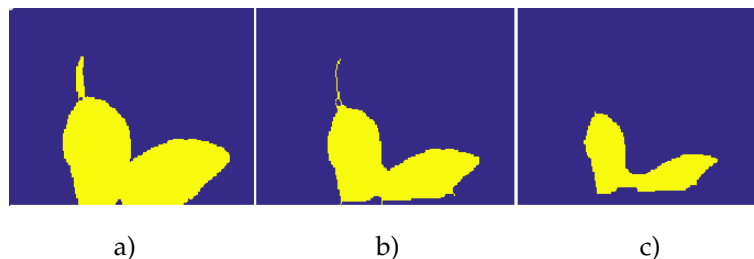


RYS. 5.13. Algorytm wyboru stopnia złączenia

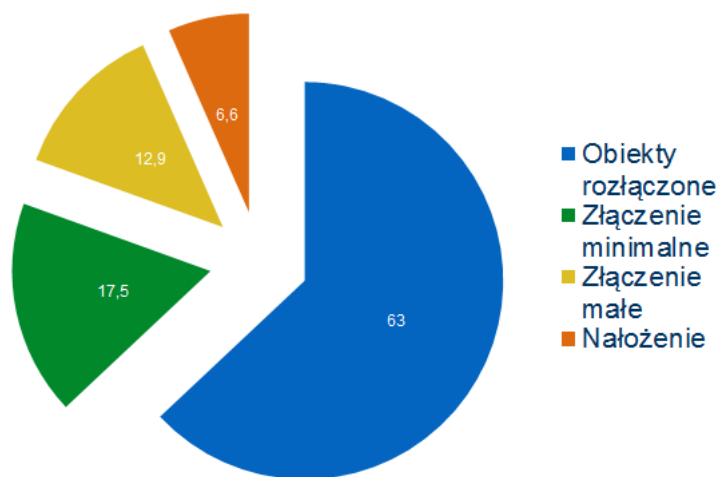
Pierwszym krokiem algorytmu jest n-krotna szkieletyzacja, gdzie n jest określona na podstawie testów stałą. Szkieletyzacja sprowadza obraz wąskich elementów (np. ogon) do szkieletu. Następnie wykonywana jest erozja. Przykłady szkieletyzacji i erozji przedstawione są na rys. 5.14. Jeżeli w wyniku powstaną 2 oddzielne obiekty, klatka zostaje zaklasyfikowana jako złączenie minimalne.

Jeżeli wciąż istnieje tylko jeden obiekt, zostaje wykonane porównanie całkowitego obwodu obiektu w aktualnej klatce do średniego obwodu obiektu obliczonego z klatek poprzednich. W sytuacji, gdy obie wartości są zbliżone (w pewnym zakresie), obraz jest rozpoznany jako złączenie małe. Jeśli obwód jest mniejszy od średniego obwodu, obraz jest sklasyfikowany jako nałożenie.

Średnia liczba wszystkich rodzajów złączeń w trakcie nagrań interakcji socjalnych dwóch dorosłych osobników męskich przetrzymywanych wcześniej osobno,



RYS. 5.14. Zastosowanie na obrazie operacji: a) binaryzacji, b) n-krotnej szkieletyzacji, c) m-krotnej erozji



RYS. 5.15. Wykres średnich procentowych liczby różnych rodzajów wzajemnego ułożenia obiektów w trakcie 20 minutowego nagrania

wpuszczonych do nieznannej klatki, obliczone na podstawie wcześniej opisanego algorytmu wyboru stopnia złączenia, przedstawione są w tabeli 5.2 oraz na rys. 5.15.

TABELA 5.2. Średnia liczba różnych rodzajów zachowań w trakcie nagrań

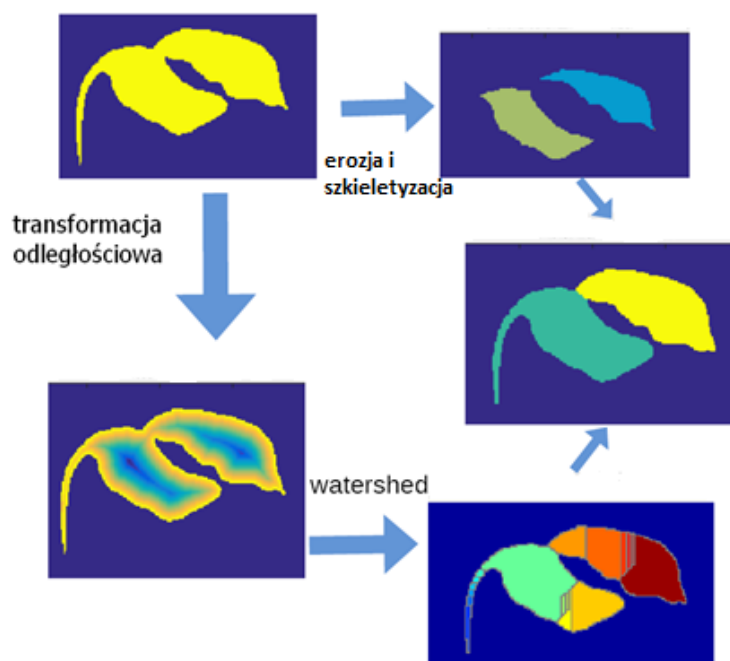
Procentowy udział wszystkich złączeń w całym nagraniu	Procentowy udział poszczególnych złączeń wśród wszystkich złączeń	Rodzaj złączenia
37% ± 2%	47% ± 18%	złączenie minimalne
	34,8% ± 13%	złączenie małe
	18% ± 5%	nałożenie

Średnio 37%±2% klatek nagrania zostało sklasyfikowanych jako złączenie. Złączenie minimalne występowało średnio w 47%±18% wszystkich złączeń, a złączenie małe w 34,8 % ± 13%. Najmniej zanotowano wystąpień nałożenia: 18%±5% ze wszystkich przypadków złączenia.

Algorytmy rozdzielenia obiektów dla poszczególnych stopni złączenia są różne, zostały przedstawione poniżej.

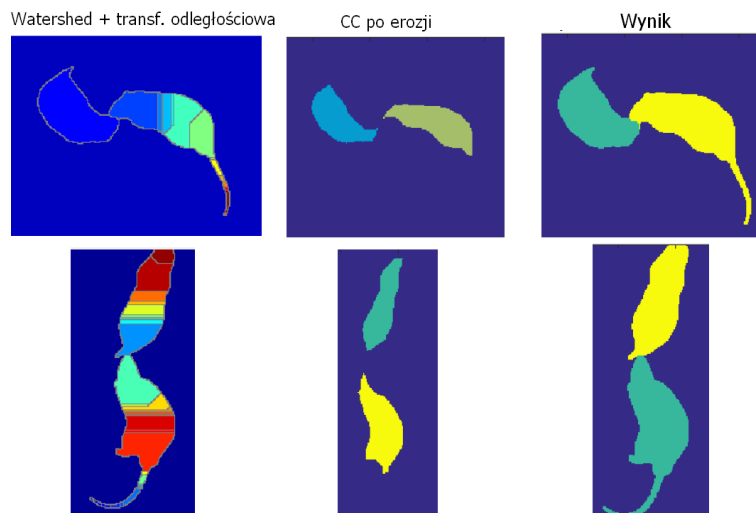
1. Algorytm segmentacji obiektów dla złączenia minimalnego (rys. 5.16) –

obraz oryginalny poddawany jest dwóm rodzajom operacji. Pierwsza z nich to szkieletyzacja wraz z erozją, których wynikiem są 2 oddzielne obiekty o powierzchniach mniejszych niż powierzchnie obiektów oryginalnych. Drugą operacją jest algorytm działów wodnych wykonany na transformacji odległościowej obrazu oryginalnego. Wynikiem tych przekształceń jest podzielenie obiektów na kilka podobszarów. Algorytm działów wodnych poprawnie wyznacza granice obiektów, ale też tworzy inne, liczne i zbędne granice podobszarów wewnątrz obiektów (rys. 5.7, 5.8). Transformacja odległościowa ma na celu ograniczenie liczby podobszarów uzyskanych po zastosowaniu algorytmu działów wodnych, polega na wyznaczeniu dla każdego piksela minimalnej odległości do granicy obiektu. Ostatnim krokiem jest porównanie i korekta etykiet obszarów po zastosowaniu transformacji odległościowej oraz działów wodnych zgodnie z etykietami obrazu po erozji i szkieletyzacji. Wynikiem jest obraz oryginalny z poprawnie rozdzielonymi obiektami. Wyniki działania algorytmu na dwóch różnych obrazach zostały przedstawione na rys. 5.17. Pierwsza kolumna rys. 5.17 przedstawia wynik zastosowania algorytmu działów wodnych na transformacji odległościowej obrazu oryginalnego. Złączone obiekty zostały podzielone na podobszary z poprawnym rozdzielaniem dwóch osobników. Druga kolumna to wynik zastosowania metody połączonych komponentów (CC) na obrazie poddanemu wcześniej szkieletyzacji oraz erozji. Ostateczny wynik algorytmu, przedstawiony w kolumnie trzeciej, powstaje po przepisaniu etykiet obszarów kolumny pierwszej na etykiety kolumny drugiej dla pikseli części wspólnej obrazów oraz stworzeniu macierzy przejść, a następnie uaktualnieniu pozostałych etykiet (w obszarach utraconych po szkieletyzacji i erozji) na podstawie tej macierzy.



RYS. 5.16. Algorytm rozdzielania obiektów dla złączenia minimalnego

2. **Algorytm segmentacji obiektów dla złączenia małego** (rys. 5.18) – na obrazie binarnym po erozji zostają wykonane dwie transformacje. Pierwsza z nich



RYS. 5.17. Wyniki działania algorytmu rozdzielania obiektów dla złączenia minimalnego. Pierwsza kolumna przedstawia wynik zastosowania algorytmu działów wodnych na transformacji odległościowej obrazu oryginalnego. Druga kolumna to wynik zastosowania metody Połączonych komponentów na obrazie poddanemu wcześniej szkieletyzacji oraz erozji. Wynik algorytmu segmentacji dla złączenia minimalnego

to identyczna transformacja jak w algorytmie wcześniejszym polegająca na połączeniu transformacji odległościowej z algorytmem działów wodnych, wynikiem której jest zbiór punktów wyznaczających granice pomiędzy podobszarami obiektu - etap 1. Drugim, równoległym krokiem jest wyznaczenie granicy obiektu (obwodu). Następnie, wyznaczana jest linia separacyjna łącząca dwa punkty granicy, dla której suma dwóch ważonych składowych (równanie 5.3) osiąga najmniejszą wartość- etap 2.:

$$linia_{prop} = \min(w_1 \cdot \text{ranking}(d_{lin}) + w_2 \cdot \text{ranking}\left(\frac{Area_1}{Area_2} - 1\right)) \quad (5.3)$$

gdzie:

w_1, w_2 - wagi składowych sumy,

d_{lin} - długość zaproponowanej linii separacyjnej,

$\frac{Area_1}{Area_2}$ - iloraz powierzchni dwóch obszarów, na które dzieli obiekt dana linia.

Ostatnim etapem jest wybranie linii pozyskanej z algorytmu wododziałów najlepiej dopasowanej do zaproponowanej linii separacyjnej $linia_{prop}$ na podstawie miary wyrażonej równaniem 5.4 – etap 3.:

$$linia_{wybr} = \min\left(w_3 \cdot \frac{\overline{\delta_{lin}}}{\min(\delta_{lin})} + w_4 \cdot \frac{Area_1}{Area_2}\right) \quad (5.4)$$

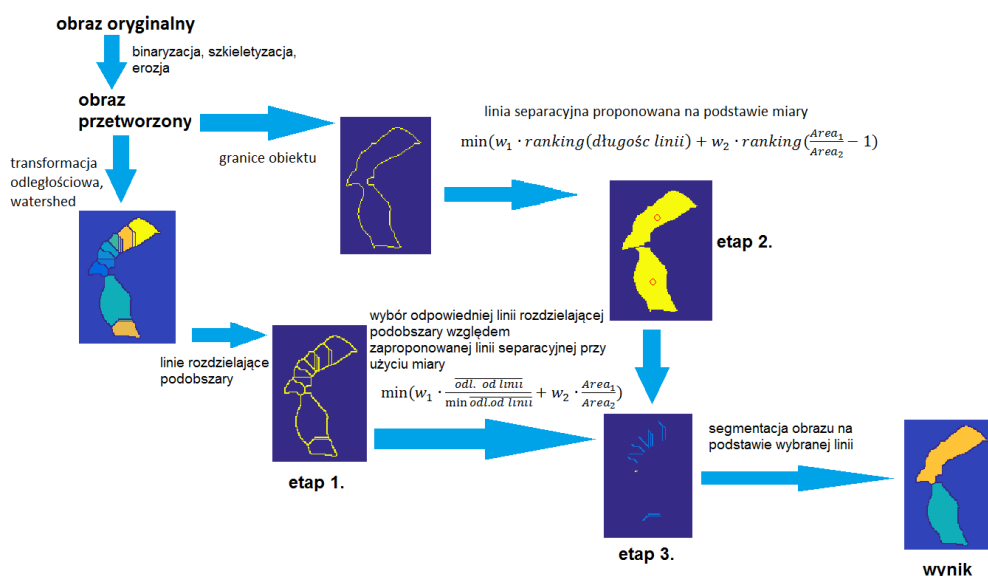
gdzie:

w_3, w_4 - wagi składowych sumy,

$\overline{\delta_{lin}}$ - średnia odległość punktów linii otrzymanych w wyniku algorytmu podziału wodnego od zaproponowanej linii separacyjnej

$\frac{Area_1}{Area_2}$ - iloraz powierzchni dwóch obszarów, na które dzieli obiekt dana linia.

W wyniku algorytm zwraca etykietę linii wyznaczonej poprzez algorytm działów wodnych, która najlepiej rozdziela obrazy dwóch osobników. Założeniem algorytmu jest względna równość powierzchni obu osobników. Warunek ten nie zawsze jest spełniony ze względu na elastyczność ciała gryzoni, które potrafi w pewnym zakresie zmieniać swoje rozmiary np. podczas czyszczenia się czy stania na łapach (zmniejszenie powierzchni ciała) lub eksplorację (wydłużenie ciała). Jednak miary wprowadzone do algorytmu zapewniają poprawny wybór linii separacyjnych nawet w wyżej wymienionych przypadkach.



RYS. 5.18. Schemat algorytmu segmentacji obiektów dla złączenia małego

Wyniki działania algorytmu segmentacji obiektów dla złączenia małego zostały porównane z wynikami obserwatora-człowieka. Dla każdego punktu linii_{wybr} powstałej w wyniku działania algorytmu (etap 3.) obliczono odległość euklidesową od każdego punktu linii zaproponowanej przez obserwatora i wybrano wartość minimalną. Miara odległości dwóch linii M_l to wartość średnia minimalnych odległości dla wszystkich punktów linii, którą przedstawia równanie 5.5.

$$M_l = \frac{\sum_{i=1}^n d(A_i, B_j))}{n} \quad (5.5)$$

gdzie:

$$d(A_i, B_j) = \min(d(A_i, b \in B)),$$

A - n-elementowy zbiór punktów mierzonej linii,

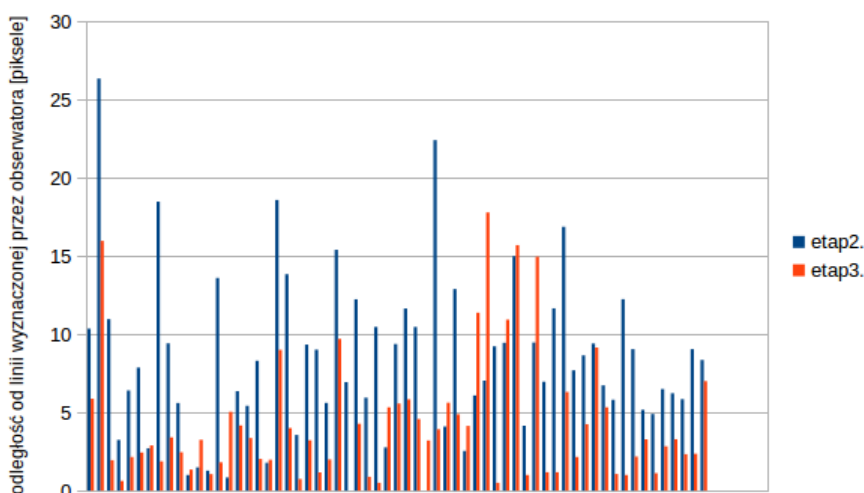
B - m-elementowy zbiór punktów linii obserwatora.

W tabeli 5.3 oraz na rys. 5.19 przedstawione zostały wyniki pomiaru odległości linii_{wybr} otrzymanej w etapie 3. (równanie 5.4) oraz linii obserwatora przy użyciu zaprezentowanej wcześniej miary M_l (równanie 5.5) dla dwóch różnych

par zwierząt. Dla porównania przedstawione zostały również wyniki odległości (przy użyciu tej samej miary) zaproponowanej linii separacyjnej powstałej w etapie 2. ($linia_{prop}$, równanie 5.3) od linii obserwatora.

TABELA 5.3. Porównanie miar M_l odległości linii otrzymanych w wyniku działania algorytmu do linii wybranej przez obserwatora dla dwóch różnych par zwierząt

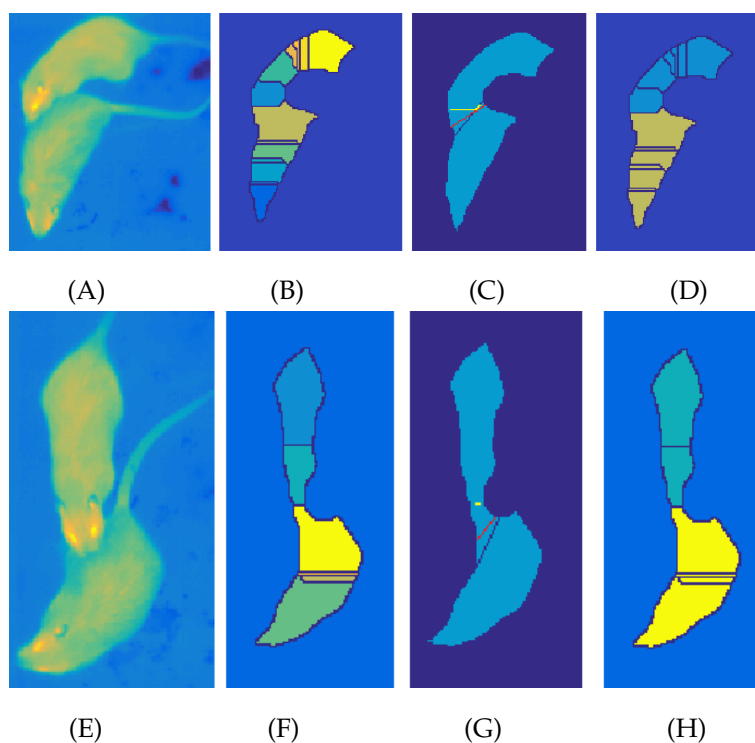
		M_l odległości $linii_{wybr}$ od linii obserwatora	M_l odległości $linii_{prop}$ od linii obserwatora
Druga para zwierząt	wartość średnia	4,29	8,41
	odchylenie standardowe	4,02	5,16
Pierwsza para zwierząt	wartość średnia	5,66	9,18
	odchylenie standardowe	4,95	5,43



rys. 5.19. Wyniki pomiaru odległości linii otrzymanej w etapie 2. algorytmu oraz w etapie 3. – końcowym

Wartości zawarte w tabeli 5.3 jasno wykazują, że $linia_{wybr}$ otrzymana jako ostateczny wynik omawianego algorytmu (etap 3.) ma średnio mniejszą odległość M_l od linii zaproponowanej przez obserwatora (4,29 pikseli dla pierwszej pary i 5,66 pikseli dla drugiej pary) niż $linia_{prop}$ obliczona na podstawie jedynie jej długości oraz stosunku pól powstałych obszarów (etap 2. - 8,41 oraz 9,18 pikseli odpowiednio dla każdej z par).

Jak widać na wykresie rys. 5.19, istnieją takie przypadki, gdzie $linia_{prop}$ daje lepszy wynik niż $linia_{wybr}$ (22% przypadków, z czego 57% z nich to różnica wyniku nie przekraczająca 2 pikseli). Dwa z takich przypadków zostały przedstawione na rys. 5.20. Wyniki pośrednich etapów działania algorytmu przedstawione są na rys. 5.20c oraz 5.20g w postaci linii. Linia czerwona oznacza linię wybraną przez obserwatora, niebieska – $linia_{prop}$, żółta – $linia_{wybr}$. Dla pierwszego przypadku odległość M_l linii niebieskiej od linii czerwonej wyniosła

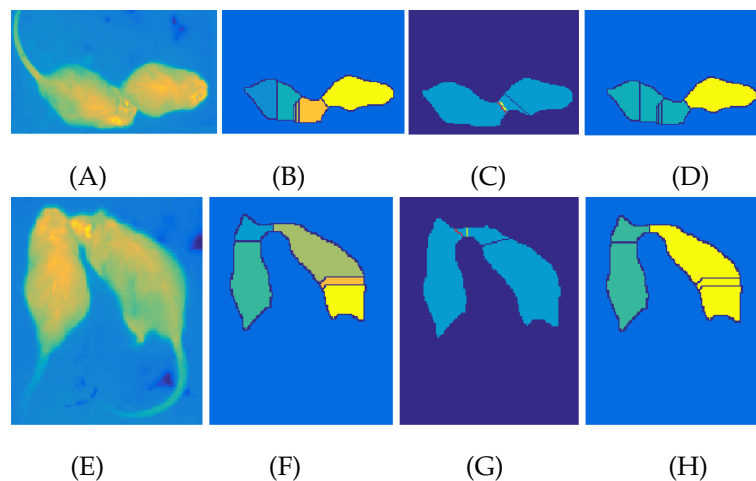


RYS. 5.20. Przykłady wyników wyboru linii rozdzielającej 2 obiekty w przypadku, gdy algorytm działów wodnych prowadzi do błędnego wyniku. Rys. a) i e) przedstawiają obrazy oryginalne; b) i f) wynik algorytmu działów wodnych; c) i g) linie otrzymane w wyniku działania algorytmu (linia niebieska – $linia_{prop}$, żółta – $linia_{wybr}$) oraz wybrane przez obserwatora (linia czerwona); d) i h) ostateczne wyniki segmentacji

2,53 piksele, a odległość M_l linii żółtej od czerwonej 4,13 pikseli. Odległości M_l pomiędzy liniami dla drugiego przypadku wynoszą odpowiednio 6,08 pikseli dla linii niebieskiej i 11,38 pikseli dla żółtej. Wyniki dla tych przykładów wskazują, że wprowadzenie algorytmu działów wodnych pogorszyło wynik. Na rys. 5.20b i 5.20f przedstawiony został podział obiektu na podobszary przy użyciu algorytmu działów wodnych. Wyraźnie widać, że nie ma linii odpowiadającej linii obserwatora, a $linia_{wybr}$ była najbliższą poprawnemu wynikowi linią podziału na podobszary.

Jednak w większości przypadków wprowadzenie algorytmu działów wodnych poprawia ostateczny wynik działania całego algorytmu i koryguje położenie linii separacyjnej. Dla przypadku przedstawionego na rys. 5.21 odległość linii niebieskiej ($linia_{prop}$) od czerwonej (linia obserwatora) wynosi 13,60 dla rys. 5.21c oraz 22,41 dla rys. 5.21g. Odległość linii żółtej ($linia_{wybr}$) od czerwonej to odpowiednio 1,80 dla rys. 5.21c i 3,92 dla rys. 5.21g.

Algorytm rozdzielania obiektów złączenia małego został zastosowany również do rozdzielania obiektów nałożonych. W wyniku 39% klatek z nałożeniem obiektów zostało rozdzielone linią separacyjną, która znajdowała się nie dalej niż 10 pikseli (używając miary M_l) od linii obserwatora. Algorytm ten działa bardzo dobrze dla przypadków, gdy granica pomiędzy obiektami nie jest większa niż szerokości obiektów.



Rys. 5.21. Przykłady wyników wyboru linii rozdzielającej 2 obiekty w przypadku gdy algorytm działów wodnych poprawia wynik ostateczny. Rys. a) i e) przedstawiają obrazy oryginalne; b) i f) wynik algorytmu działów wodnych; c) i g) linie otrzymane w wyniku działania algorytmu (linia niebieska – *linia_{prop}*, żółta – *linia_{wybr}*) oraz wybrane przez obserwatora (linia czerwona); d) i h) ostateczny wynik segmentacji

3. **Algorytm segmentacji obiektów nałożonych** opiera się na algorytmach uczenia głębokiego, przedstawiony został w następnym podrozdziale.

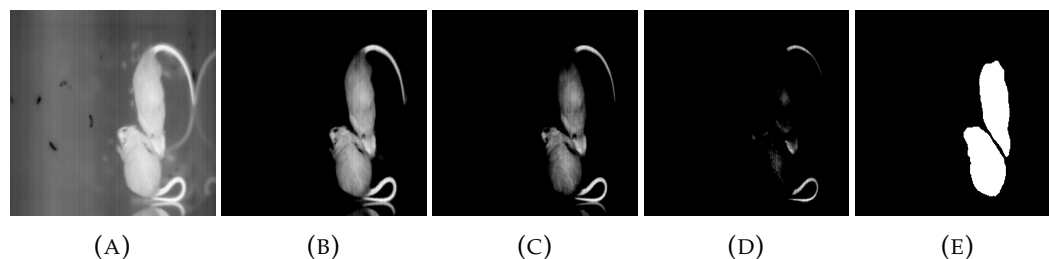
5.5. Segmentacja przy użyciu uczenia głębokiego

W celu rozdzielenia obszarów, które nachodzą na siebie w znaczny sposób, zdecydowano się na segmentację opartą na sieciach neuronowych. W wyniku przeprowadzonych eksperymentów wybrano 4 architektury. Zaadaptowano 2 sieci przeznaczone do segmentacji semantycznej: U-Net i V-Net oraz 2 sieci segmentujące instancje: Mask R-CNN oraz TensorMask. Porównano także wyniki segmentacji dla pięciu różnych typów stworzonych obrazów opisanych w rozdziale 5.3 (*ch0*, *ch1*, *ch2*, *ch3*, *16bit*).

5.5.1. Segmentacja semantyczna

Jednym z ważnych problemów uczenia maszynowego jest wymaganie dużej liczby danych uczących. Model sieci U-Net, stworzony w celu segmentacji elementów biologicznych, dopuszcza generowanie nowych obrazów poprzez różne operacje wykonywane na obrazach zarejestrowanych, jest to tzw. augmentacja. Augmentacja jest sposobem tworzenia bardziej zróżnicowanego oraz większego zbioru danych. Zestawy obrazów z baz danych zachowań socjalnych zwierząt zostały poszerzone o nowe, przekształcone obrazy. Dla sieci U-Net wygenerowane zostały obrazy termograficzne zgodnie z przekształceniami takimi jak: obrót (w zakresie 0,2 stopnia), przesunięcie (0,05 całkowitego rozmiaru), powiększenie (w zakresie od 0,95 do 1,05) i odwrócenie w poziomie. Augmentacją dla V-Net były losowe, nieliniowe transformacje obrazu przy użyciu gęstego pola deformacji i interpolacji B-spline.





RYS. 5.22. Przykładowe obrazy (dane wejściowe): a) *ch0*, b) *ch1*, c) *ch2* i d) *ch3* wraz z e) obrazem maski

Architektury U-Net i V-Net zostały wytrenowane z różnymi parametrami. Wszystkie eksperymenty przeprowadzono przy użyciu stacji NVIDIA DGX. Każda z sieci była trenowana pięć razy - dla:

- 4 rodzajów 8-bitowych obrazów termicznych stworzonych z 4 różnych zakresów danych opisanych w sekcji 5.1.2,
- obrazów 16-bitowych zawierających cały zakres rejestrowanych danych.

Każdy zbiór treningowy zawierał 101 obrazów, które przedstawiały gryzonie podczas fizycznego kontaktu o różnym stopniu. Dla każdego modelu treningowego wykonano trzy niezależne treningi. Testowanie przeprowadzono z kolei na 51 obrazach spoza zestawu uczącego.

Dla każdego obrazu treningowego i testowego stworzono obrazy masek z wyraźnym rozgraniczeniem zwierząt. Rys. 5.22 przedstawia przykłady obrazów odpowiednio dla zakresów *ch0*, *ch1*, *ch2* i *ch3* oraz utworzoną dla nich maskę.

W sumie przeprowadzono eksperymenty dla 15 różnych zaproponowanych modeli. Jako funkcję straty zastosowano binarną entropię krzyżową dla U-Net i współczynnik dice dla architektury V-Net. Najlepsze wyniki trenowania zostały osiągnięte dla wartości parametrów przedstawionych w tabeli 5.4.

TABELA 5.4. Parametry trenowania w eksperymentach segmentacji semantycznej

kombinacje wartości parametrów	A (U-Net i V-Net)	B (U-Net i V-Net)	C (tylko V-Net)	D (tylko U-Net)
rozmiar batch	2	10	16	2
liczba kroków	100	100	400	2000
liczba epok	150	150	300	30

W przypadku obu architektur, U-Net i V-Net, trening składający się ze 150 epok, 100 kroków i rozmiaru batch równego 2 (kombinacja A) lub 10 (kombinacja B) okazały się osiągać najlepsze wyniki. V-Net osiągnął również wysoką wydajność w 300 epokach, 400 krokach i rozmiarze batch wynoszącym 16 (kombinacja C). Znaczący wzrost liczby kroków (2000) przy jednoczesnej redukcji epok (30) wykazał pozytywny wpływ na wyniki dla architektury U-Net (kombinacja D). Wyniki tylko dla tych czterech kombinacji parametrów (po 3 dla każdego modelu) przedstawione są

w dalszej części pracy.

Jako miary jakości wyników zastosowano:

1. Intersection over Union (IoU), nazywany także indeksem Jaccard'a - jest miarą ewaluacji często stosowaną do pomiaru dokładności predykcji (P) w porównaniu do danych referencyjnych (ang. ground-truth - GT). Jest to stosunek obszaru wspólnego dla predykcji i referencji do obszaru ich sumy:

$$IoU(P, GT) = J(P, GT) = \frac{|P \cap GT|}{|P \cup GT|} \quad (5.6)$$

gdzie:

P jest obrazem będącym wynikiem predykcji,

GT jest obrazem referencyjnym.

Wartości IoU zawierają się w zakresie pomiędzy 0 a 1. Wartość 1 oznacza całkowite pokrycie się obszarów predykcji (P) i referencji (GT).

2. Współczynnik dice, nazywany także indeksem Sorensen'a lub współczynnikiem F_1 - jest używany jako miara podobieństwa pomiędzy dwoma obiektami. Zdefiniowany jest jako:

$$F_1 = \frac{2|P \cap GT|}{|P| + |GT|} = \frac{2 \times \text{precyzja} \times \text{czułość}}{\text{precyzja} + \text{czułość}} \quad (5.7)$$

gdzie:

$$\text{precyzja} = \frac{|P \cap GT|}{|P|},$$

$$\text{czułość} = \frac{|P \cap GT|}{|GT|}.$$

F_1 , podobnie jak IoU, przyjmuje wartości od 0 do 1.

Związek pomiędzy współczynnikami IoU a dice (F_1) opisuje równanie:

$$IoU = \frac{F_1}{2 - F_1} \quad (5.8)$$

3. Dokładność segmentacji obiektów rozseparowanych (J_{segm}) - parametr wprowadzony do celów tej pracy, zdefiniowany jako stosunek liczby obrazów poprawnie rozseparowanych (n) do liczby wszystkich obrazów testowych (N_{test}):

$$J_{segm} = \frac{n}{N_{test}} \quad (5.9)$$

Obraz poprawnie rozseparowany jest obrazem, który dla danego przypadku analizy zachowań dwóch gryzoni posiada dwa osobne obiekty. Obiekty uważane są za osobne, jeśli algorytm etykietowania (np. połączonych komponentów) oznacza je różnymi etykietami. Dodatkowo, różnica powierzchni obu obiektów nie powinna być większa niż połowa mniejszej powierzchni. Ma to na celu eliminację przypadków, gdzie zwierzęta będą niepoprawnie rozseparowane i/lub na obrazie pojawiają się artefakty, np. odbicia.

Dokładność segmentacji obiektów rozseparowanych została pomierzona dla obrazów wynikowych po progowaniu, ponieważ obraz otrzymany w wyniku inferencji nie był binarny. Zastosowano także dodatkowe przetwarzanie w postaci erozji o elemencie strukturalnym o wielkości 3x3 piksele.



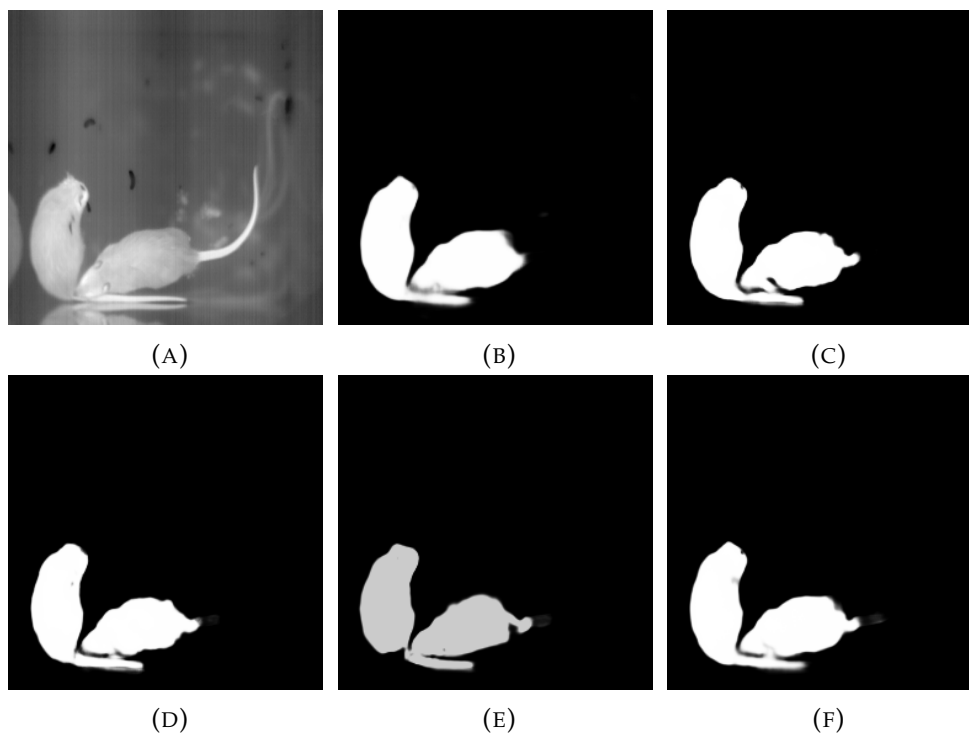
Wyniki treningów przedstawione są w tabeli 5.5 w postaci trzech parametrów: dokładności trenowania zdefiniowanej jako stosunek sumy poprawnych predykcji do całkowitej ich liczby (im wartość bliższa 1, tym lepsza), wartości funkcji kosztu (im wartość bliższa 0, tym lepsza) oraz współczynnika dice (F_1 - im wartość bliższa 1, tym lepsza). Najlepsze wyniki zostały oznaczone w tabeli pogrubieniem.

TABELA 5.5. Wyniki treningu architektury U-Net i V-Net (dokładność, wartość funkcji kosztu oraz współczynnik dice- F_1) dla każdego typu obrazu (*ch0*, *ch1*, *ch2*, *ch3* i *16bit*)

	U-Net	A	B	D	V-Net	A	B	C
dokładność	<i>ch0</i>	0,9893	0,9885	0,9899	<i>ch0</i>	0,9896	0,9898	0,9901
	<i>ch1</i>	0,9897	0,9897	0,99	<i>ch1</i>	0,9896	0,9898	0,9901
	<i>ch2</i>	0,9897	0,9899	0,99	<i>ch2</i>	0,9896	0,9897	0,9901
	<i>ch3</i>	0,9892	0,9897	0,99	<i>ch3</i>	0,9892	0,9897	0,9901
	<i>16bit</i>	0,9894	0,9894	0,9802	<i>16bit</i>	0,9896	0,9897	0,9901
wartość funkcji kosztu	<i>ch0</i>	0,009	0,063	0,007	<i>ch0</i>	0,019	0,017	0,013
	<i>ch1</i>	0,008	0,007	0,01	<i>ch1</i>	0,019	0,017	0,013
	<i>ch2</i>	0,008	0,007	0,006	<i>ch2</i>	0,019	0,017	0,013
	<i>ch3</i>	0,01	0,029	0,007	<i>ch3</i>	0,022	0,019	0,014
	<i>16bit</i>	0,009	0,009	0,082	<i>16bit</i>	0,019	0,017	0,013
F_1	<i>ch0</i>	0,979	0,975	0,985	<i>ch0</i>	0,981	0,983	0,987
	<i>ch1</i>	0,983	0,983	0,986	<i>ch1</i>	0,982	0,983	0,987
	<i>ch2</i>	0,982	0,982	0,986	<i>ch2</i>	0,981	0,984	0,987
	<i>ch3</i>	0,978	0,978	0,985	<i>ch3</i>	0,978	0,981	0,986
	<i>16bit</i>	0,980	0,980	0,905	<i>16bit</i>	0,981	0,983	0,987

Model C dla sieci V-Net uzyskał najlepsze rezultaty wszystkich mierzonych parametrów, prawdopodobnie ze względu na dużą liczbę epok. W przypadku modelu U-Net najlepsze wartości dokładności i współczynnika dice uzyskał model z najmniejszą liczbą epok, ale największą liczbą kroków (model D). Wartość funkcji straty była również najniższa dla modelu D z wyjątkiem obrazu *ch1*. Wyniki dla różnych obrazów są bardzo podobne. Architektura U-Net uzyskała w większości przypadków najlepsze wyniki trenowania dla obrazów *ch1* i *ch2*. Wyniki uzyskane dla obrazu *ch3*, czyli obrazu zawierającego najmniejszą ilość danych (największy histogram), są porównywalne z wynikami uzyskanymi dla obrazu *ch0* - obrazu uzyskanego z całego zakresu surowych danych. Obraz 16-bitowy osiągnął tylko nieznacznie lepsze wyniki niż obraz *ch0*, ale gorsze niż *ch1* i *ch2*. Wyniki dla różnych obrazów wykorzystujących architekturę V-Net są bardzo do siebie zbliżone, jedynie wyniki dla obrazu *ch3* różnią się nieznacznie od pozostałych. Wyraźną różnicę podczas trenowania obu modeli można zauważyć dla wartości funkcji straty, która w przypadku sieci U-Net jest znacznie mniejsza. Jednak funkcja straty jest obliczana na podstawie różnych parametrów, dlatego bezpośrednie porównanie jej wartości nie jest miarodajnym parametrem.

Na rysunku 5.23 przedstawione zostały przykładowe wyniki segmentacji wykonanej przed model A dla U-Net. Najdokładniejsze rozdzielanie obiektów zostało wykonane w tym przypadku dla obrazu *ch2* (rys. 5.23d). Obraz *ch1* (rys. 5.23c) prawie całkowicie rozdzielił obiekty, jednak niepoprawnie stworzył przerwę wewnątrz jednego z obiektów. Obrazy *ch0* (rys. 5.23b) oraz *16-bit* (rys. 5.23f) utworzyły rozmytą granicę, która przy źle dobranym progu binaryzacji może zostać utracona.



RYS. 5.23. Przykładowe wyniki segmentacji dla sieci U-Net wytrenowanej z parametrami A dla a) obrazu oryginalnego: b) obraz *ch0*, c) obraz *ch1*, d) obraz *ch2*, e) obraz *ch3*, f) obraz 16-bit

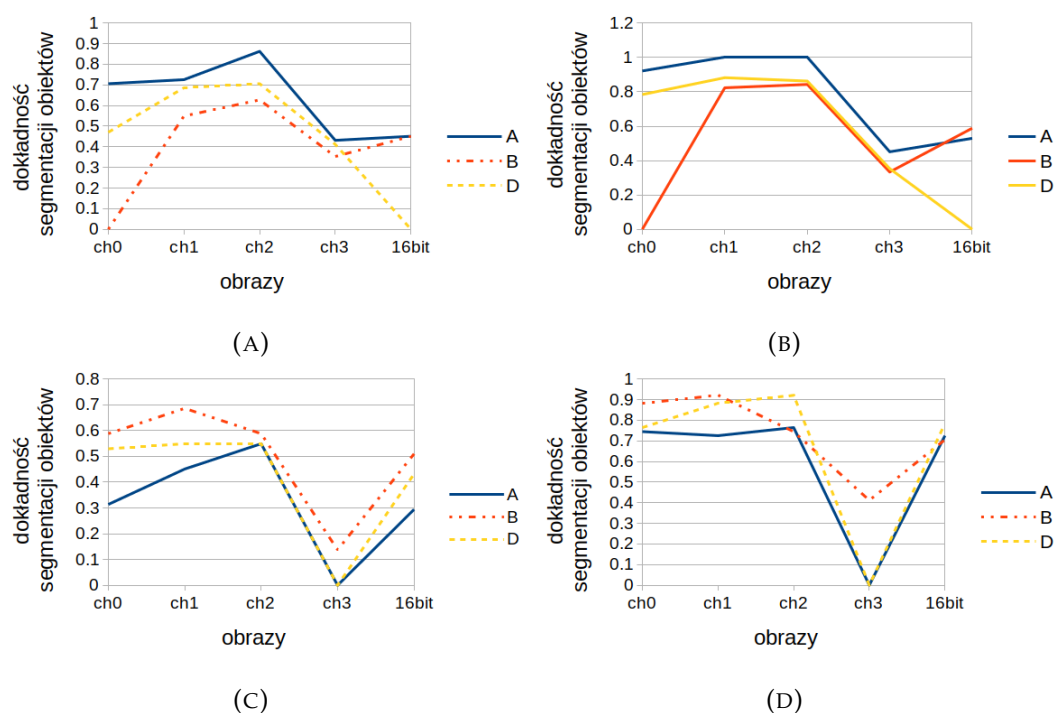
Najmniej dokładnie zostały wyznaczone oraz rozseparowane obiekty na obrazie *ch3* 5.23e.

TABELA 5.6. Wyniki parametrów Intersection over Union (IoU) oraz F_1 dla testowania architektur U-Net i V-Net). Pogrubiono najlepsze wyniki

	ch0	ch1	ch2	ch3	16bit
IoU					
U-Net					
A	0,817±0,023	0,800±0,0213	0,776±0,0189	0,618±0,1542	0,813±0,0153
B	0,855±0,0188	0,794±0,0333	0,78±0,0163	0,600±0,1435	0,816±0,0228
D	0,813±0,0201	0,796±0,026	0,783±0,0179	0,598±0,1684	0,833±0,0185
V-Net					
A	0,793±0,0127	0,802±0,0225	0,783±0,0183	0,141±0,0066	0,787±0,0135
B	0,783±0,0123	0,803±0,0212	0,792±0,0178	0,581±0,2102	0,802±0,0126
C	0,802±0,0114	0,803±0,0198	0,793±0,0179	0,126±0,1374	0,801±0,0131
F_1					
U-Net					
A	0,900±0,0139	0,889±0,0132	0,874±0,0121	0,751±0,1314	0,897±0,0093
B	0,922±0,0109	0,885±0,021	0,876±0,0102	0,739±0,1253	0,898±0,0138
D	0,897±0,0122	0,886±0,0163	0,878±0,0113	0,733±0,1506	0,909±0,011
V-Net					
A	0,885±0,0079	0,890±0,0139	0,878±0,0116	0,247±0,0101	0,880±0,0085
B	0,878±0,0077	0,890±0,0131	0,884±0,0111	0,708±0,2032	0,890±0,0078
C	0,89±0,007	0,891±0,0122	0,885±0,0111	0,200±0,2015	0,892±0,008

Tabela 5.6 przedstawia wyniki eksperymentów weryfikacji (testowania) opracowanych modeli. Zaprezentowano w niej średnią wartość IoU oraz współczynnika F_1 wraz z odchyleniem standardowym dla wszystkich testowanych obrazów odpowiednio przy użyciu algorytmu U-Net i V-Net. Parametry te obliczane były pomiędzy obrazami wynikowymi po inferencji oraz obrazami referencyjnymi. Wartości parametru IoU dla sieci U-Net są najwyższe dla oryginalnych obrazów i maleją wraz ze spadkiem zakresu temperatur histogramu. 16-bitowy obraz osiąga lepsze wyniki niż obraz *ch0* wyłącznie dla modelu D, w przypadku pozostałych dwóch modeli jego wynik jest nieco niższy. Algorytm V-Net uzyskał najlepsze wyniki dla obrazu *ch1*, a następnie dla obrazów *ch0* lub *16-bit*. Obraz *ch3* uzyskał bardzo niską wartość IoU dla modeli A i C. Współczynniki dice (F_1) potwierdzają wcześniejsze wnioski wyciągnięte z wyników wartości IoU, z wyjątkiem najwyższej wartości F_1 dla obrazu 16-bitowego dla modelu C.

Wyniki pomiaru dokładności segmentacji obiektów (J_{segm}) przedstawione są na rys. 5.24. Rys. 5.24a i 5.24b prezentują wyniki dla architektury U-Net bez i z erozją dla elementu strukturalnego 3x3 piksele. Odpowiednie wyniki dla sieci V-Net przedstawiono na rys. 5.24c i 5.24d.



RYS. 5.24. Wyniki dokładności segmentacji obiektów (J_{segm}) dla modeli A, B, C i D oraz algorytmów a) U-Net bez erozji, b) U-Net z erozją dla elementu strukturalnego 3x3 piksele, c) V-Net bez erozji, d) V-Net z erozją dla elementu strukturalnego 3x3 piksele

Wyniki segmentacji po erozji są wyższe niż te bez erozji o 13% dla U-Net i o 25% dla V-Net. Jednak kształt i zależności między wykresami dla tych samych modeli są podobne, co pozwala wnioskować, że erozja jest dobrym sposobem na poprawę segmentacji. Wyniki dla obrazów *ch3* potwierdzają wyniki wcześniejszych pomiarów: obraz o najmniejszym zakresie temperatur uzyskał znacznie gorsze wyniki IoU i współczynnika F_1 a także J_{segm} niż inne obrazy. Jest to szczególnie widoczne w przypadku architektury V-Net, gdzie wyniki segmentacji *ch3* są najgorsze (rys.

5.24c i 5.24d) i jedno z najgorszych dla U-Net (rys. 5.24a i 5.24b). Aż połowa modeli V-Net oraz prawie 30% modeli U-Net wytrenowanych na obrazach *ch3* nie była w stanie wygenerować obrazów z segmentacjami. Obraz *ch0* sprawił także wiele trudności modelowi B architektury U-Net, 69% wszystkich treningów i inferencji zakończyło się brakiem segmentacji na obrazach wynikowych, mimo iż kształty zwierząt zostały wykryte poprawnie (parametr IoU powyżej 0,78; F_1 powyżej 0,87, patrz tabela 5.6). Parametr dokładności segmentacji mierzy jakość separacji dwóch osobników, czyli ich prawidłową identyfikację. Jeśli obiekty zostaną poprawnie posegmentowane, ale bez odseparowania wszystkich instancji od siebie, wartość IoU będzie wysoka, a dokładność segmentacji niska. Tylko wysokie wartości parametru dokładności segmentacji gwarantują poprawną separację obiektów.

W większości przypadków obu architektur segmentacja dała lepsze wyniki dla obrazów *ch1* niż *ch0*. Jednak najlepsze wyniki segmentacji dla modeli U-Net oraz niektórych modeli V-Net zostały osiągnięte przez obraz *ch2*. Może to wynikać z faktu, że przy takiej konwersji, w której pomija się niższe temperatury ciała, pozostały zakres temperatur ma lepszą rozdzielczość, która może wyraźniej wskazywać granice między obiektami. Obraz *16-bit* uzyskał wynik segmentacji lepszy niż *ch0* tylko dla modelu B architektury U-Net. We wszystkich innych przypadkach osiągnął gorsze lub porównywalne wyniki. Zaobserwowano ponadto nieprawidłową segmentację elementów tła w obrazach *16-bit*, czego nie zauważono w obrazach *ch0*.

Szereg wyników badań z adaptacją modeli U-Net i V-Net opublikowano w pracy [95].

5.5.2. Segmentacja instancji

Dla potrzeb pracy zaadaptowano algorytmy segmentacji instancji. W badaniach wykorzystano jeden jedno- i jeden dwu-etapowy algorytm: TensorMask i Mask R-CNN. Użyto implementacji Mask R-CNN oraz TensorMask z Detectron2 [147] a także implementacji Mask R-CNN z biblioteki Matterport [1]. Wszystkie obliczenia zostały wykonane na NVIDIA DGX-1 Station z Ubuntu 18.

Przygotowane zostały dwa zbiory treningowe (zawierające 200 oraz 500 obrazów) oraz jeden 50-elementowy zbiór testowy. Mniejszy zbiór treningowy został użyty do uczenia transferowego, w którym wykorzystana została wstępnie wytrenowana sieć Mask R-CNN oparta na sieci ResNet-FPN złożona z 50 warstw wytrenowanych na zestawie danych COCO oraz podobnie TensorMask z 50 warstwową siecią ResNet-FPN. 500 obrazów zostało wykorzystanych do trenowania całych architektur zaczynając od przypadkowych wartości początkowych wag. Podczas takich szkoleń liczba iteracji musiała zostać zwiększona, aby modele mogły uzyskać zbieżność. Jednak czasem, nawet pomimo bardzo dużej liczby iteracji, architektura TensorMask miała problemy ze zbieżnością: funkcja kosztu rosła w czasie. Dla tych przypadków normalizacja Frozen Batch była zastępowana normalizacją grupową (Group Normalization) zgodnie z [58]. Dokładność normalizacji grupowej jest niewrażliwa na rozmiary wartości batch [146] i umożliwia pomyślne trenowanie całego modelu TensorMask od podstaw.

Dla każdej kombinacji parametrów modeli wykonane zostały 3 niezależne treningi. Maski regionów referencyjnych w postaci zamkniętych wielokątów wykonane zostały ręcznie jako pliki Json w oprogramowaniu VGG Image Annotator (VIA) [39]. Przykładowe obrazy segmentacji referencyjnej oznaczonej żółtą linią przedstawione są

w dalszej części pracy na rys. 5.25b i 5.25d. W celu sprawdzenia jakości trenowania, porównano wyniki trenowania z danymi modelami wytrenowanymi na popularnych bazach: MS COCO Dataset [84], Citiscapes [32] i LVIS [54]. Porównano także wyniki uczenia transferowego z wynikami trenowania od podstaw, zgodnie z pracą [58], w której autorzy udowadniają, że dobranie odpowiednich parametrów trenowania pozwala modelom trenowanym od podstaw osiągnąć wyniki zbliżone do wyników modeli wcześniej przetrenowanych na dużych zbiorach danych.

Łącznie przetrenowano 24 różne konfiguracje modeli dla trenowania wykorzystującego wstępnie przetrenowane wagi oraz 22 konfiguracje dla trenowania od podstaw. Tabela 5.7 przedstawia parametry modeli, dla których uzyskane zostały najlepsze wyniki. Tylko dla tych modeli przedstawiono wyniki w dalszej części rozdziału. Wartości pogrubione w tabeli oznaczają parametry, które zostały wybrane do walidacji krzyżowej (ang. cross-validation). W przeprowadzonych eksperymentach trenowanie sieci Mask R-CNN od podstaw wymagało około 16 razy większej liczby epok (16 000 lub 24 000) niż w przypadku wstępnie przetrenowanego modelu (1000 lub 2000), aby osiągnąć podobne wyniki. Problemy ze zbieżnością architektury TensorMask wymusiły konieczność zastosowania dużo większej liczby epok podczas trenowania od podstaw niż w przypadku tego samego trenowania modelu Mask R-CNN. Aby porównać efektywność obu architektur, wykonano dodatkowe szkolenie Mask R-CNN z optymalnymi parametrami dla TensorMask (batch=2, 100 000 epok).

TABELA 5.7. Parametry trenowania, dla których osiągnięto najlepsze wyniki

Mask R-CNN				TensorMask			
uczenie transferowe		trenowanie od podstaw		uczenie transferowe		trenowanie od podstaw	
batch	liczba epok	batch	liczba epok	batch	liczba epok	batch	liczba epok
2	1 000	4	16 000	2	2 000	2	100 000
4	1 000	4	24 000	4	1 000	4	100 000
4	2 000	8	16 000	4	2 000		
		8	24 000	8	2 000		
		2	100 000				

Do oceny wyników użyto miar oceny jakości segmentacji instancji wprowadzonych w pracy [84] i uznanych za standard ilościowej oceny metod wykrywania obiektów i segmentacji instancji [87, 106]. Wykorzystano trzy różne miary zdefiniowane w dalszej części pracy: uśrednionej średniej precyzji (ang. mean Average Precision) - mAP; średniej precyzji (AP) dla progu IoU wynoszącego 0,5 (AP_{50}) oraz dla IoU = 0,75 (AP_{75}). Wszystkie parametry liczono osobno dla detekcji obiektu oznaczanego przez prostokąt określający położenie obiektu (bbox) i poziomą segmentacji (segm).

Ogólna definicja średniej precyzji polega na znalezieniu obszaru pod krzywą precyzja-czułość zdefiniowanymi w następujący sposób:

$$\text{Precyzja} = \frac{TP}{TP + FP}; \quad \text{Czułość} = \frac{TP}{TP + FN} \quad (5.10)$$

gdzie:

TP = obiekt prawdziwie dodatni (ang. True Positive), tzn. metoda wykryła obiekt obecny w zbiorze referencyjnym,

FP = obiekt fałszywie dodatni (ang. False Positive), tzn. metoda wykryła obiekt

nieobecny w zbiorze referencyjnym,

FN = obiekt fałszywie ujemny (ang. False Negative), tzn. metoda nie wykryła obiektu obecnego w zbiorze referencyjnym.

W celu określenia poprawności zdefiniowania wyników TP, FP, FN i FP standardowo używana jest miara IoU (wzór 5.6). Predykcja jest uważana za prawdziwie dodatnią, jeśli wartość IoU jest większa niż wartość progowa. Jednak dobranie stałej wartości progowej jest rozwiązaniem mało uniwersalnym, dlatego tworzone są krzywe precyzja-czułość, dla których możliwe jest obliczenie powierzchni pod krzywą zdefiniowanej jako AP:

$$AP = \int_0^1 p(r)dr, \quad (5.11)$$

gdzie: $p(r)$ to krzywa precyzja-czułość.

Miara mAP została zdefiniowana jako średnia z pomiarów dla IoU w zakresie od 0,5 do 0,95 z krokiem 0,05. AP_{50} to średnia precyzja dla predykcji, która została uznana za prawdziwie dodatnią przy $IoU > 0,5$. Analogicznie, AP_{75} zaklasyfikuje predykcję tylko, jeśli IoU będzie większe niż 0,75. Im większy próg, tym większe wymaganie na dokładność dopasowania predykcji z referencją.

W tablicach 5.8 i 5.9 przedstawiono porównanie wyników testowych uczenia transferowego oraz trenowania od podstaw dla Mask R-CNN. Analogiczne pomiary dla architektury TensorMask znajdują się w tablicach 5.10 i 5.11.

TABELA 5.8. Wyniki testowania dla uczenia transferowego Mask R-CNN dla 200 obrazów. Pogrubiono najlepsze wyniki

Mask R-CNN uczenie transferowe	mAP	AP ₅₀	AP ₇₅	mAP	AP ₅₀	AP ₇₅
2 batch, 1 000 epok	bbox			segm		
ch0	88,641	100	98,99	89,007	100	98,99
ch1	87,332	100	100	89,289	100	100
ch2	85,252	100	99	87,323	100	100
ch3	63,719	99,114	73,409	52,901	93,649	59,722
16-bit	88,128	100	99,01	89,604	100	100
4 batch, 1 000 epok						
ch0	89,025	100	100	89,444	100	100
ch1	86,942	100	100	89,715	100	100
ch2	85,847	100	98,98	88,256	100	100
ch3	69,028	99,417	80,98	55,181	97,284	60,683
16-bit	89,965	100	100	89,194	100	100
4 batch, 2 000 epok						
ch0	88,781	100	99	89,068	100	99
ch1	90,205	100	100	89,655	100	100
ch2	87,915	100	98,98	88,893	100	100
ch3	73,908	99,952	87,824	65,687	98,902	83,376
16-bit	90,549	100	100	89,417	100	100

TABELA 5.9. Wyniki testowania dla trenowania od podstaw Mask R-CNN dla 500 obrazów. Pogrubiono najlepsze wyniki

Mask R-CNN trenowanie od podstaw	mAP	AP ₅₀	AP ₇₅	mAP	AP ₅₀	AP ₇₅
4 batch, 16 000 epok	bbox			segm		
ch0	81,811	100	98,307	85,586	100	100
ch1	82,085	100	100	83,666	100	100
ch2	80,629	100	98,95	80,896	100	100
ch3	60,947	97,766	73,004	48,183	93,325	47,4
4 batch, 16 000 epok	82,826	100	98,842	84,561	100	100
4 batch, 24 000 epok						
ch0	82,567	100	100	86,864	100	100
ch1	84,368	100	100	85,821	100	100
ch2	82,979	100	100	84,834	100	100
ch3	67,739	98,113	79,816	58,782	96,882	69,92
16-bit	84,29	100	100	86,111	100	100
8 batch, 16 000 epok						
ch0	81,658	100	100	84,728	100	98,584
ch1	81,426	100	100	84,038	100	100
ch2	80,983	100	100	79,931	100	100
ch3	65,355	97,874	74,09	54,855	96,873	64,968
16-bit	83,493	100	97,82	85,313	100	100
8 batch, 24 000 epok						
ch0	84,465	100	100	86,224	100	100
ch1	85,153	100	98,772	86,389	100	100
ch2	83,091	100	100	84,428	100	100
ch3	66,014	98,4	78,156	57,452	95,887	63,699
16-bit	85,596	100	100	86,607	100	100
2 batch, 100 000 epok						
ch0	86,695	100	100	89,569	100	100
ch1	86,564	100	100	87,208	100	100
ch2	82,55	100	97,55	82,397	100	100
ch3	66,369	98,504	84,891	56,116	94,891	67,155
16-bit	84,862	100	100	85,877	100	100

TABELA 5.10. Wyniki testowania dla uczenia transferowego TensorMask dla 200 obrazów. Pogrubiono najlepsze wyniki

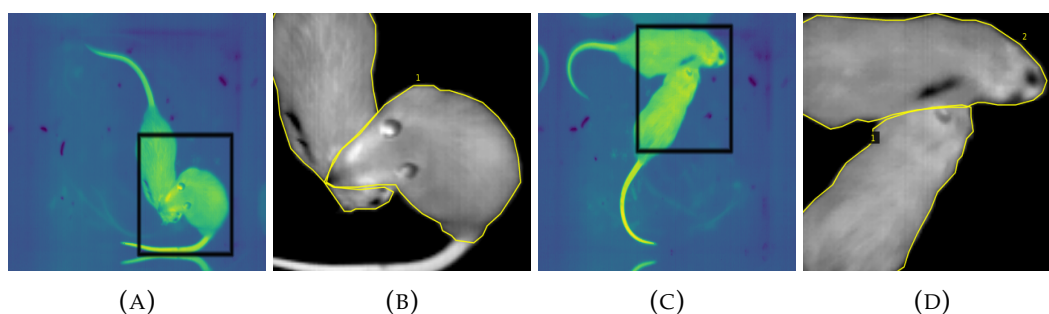
TensorMask uczenie transferowe	mAP	AP ₅₀	AP ₇₅	mAP	AP ₅₀	AP ₇₅
2 batch, 2 000 epok	bbox			segm		
ch0	90,646	100	98,01	89,82	100	100
ch1	89,091	100	100	89,808	100	100
ch2	88,319	100	100	89,496	100	100
ch3	74,37	99,933	88,433	65,062	99,933	80,743
16-bit	91,248	100	100	90,249	100	100
4 batch, 1 000 epok						
ch0	91,264	100	99,01	89,877	100	99,01
ch1	89,624	100	100	89,547	100	99,01
ch2	88,17	100	100	87,949	100	100
ch3	71,325	99,834	84,895	60,251	99,602	69,757
16-bit	90,443	100	98,951	89,857	100	98,951
4 batch, 2 000 epok						
ch0	90,916	100	99,01	90,158	100	100
ch1	90,299	100	100	90,171	100	99,01
ch2	87,912	100	98,931	89,335	100	100
ch3	75,031	99,961	91,558	64,849	99,423	79,397
16-bit	91,609	100	100	90,199	100	100
8 batch, 2 000 epok						
ch0	90,672	100	98,96	89,824	100	100
ch1	90,052	100	100	90,087	100	99,01
ch2	88,828	100	100	89,264	100	100
ch3	76,159	100	91,406	65,307	99	83,105
16-bit	90,939	100	100	90,171	100	98,99

Trenowanie Mask R-CNN od podstaw okazało się nie potrzebować ani bardzo dużej ilości danych ani zasobów obliczeniowych. Tylko nieco większa liczba obrazów i epok pozwala na osiągnięcie wyników treningu wszystkich warstw porównywalnych z wcześniej wytrenowanymi modelami. Zwiększenie liczby epok do 100 000 w przypadku Mask R-CNN poprawiło wyniki wyłącznie dla obrazów *ch0* i *ch1*.

Wartości w tablicach 5.8 i 5.10 wskazują, że najlepsze wyniki segmentacji dla wstępnie wytrenowanych modeli uzyskały obrazy *16-bitowe*, *ch1* oraz *ch0*. Obrazy *ch2* osiągnęły tylko nieznacznie niższe wyniki. Obie wartości mAP (bbox i segm) dla obu architektur były bliskie 90 procent, z niewielką przewagą dla TensorMask. Wyniki dla obrazów *ch3* znacznie odbiegały od pozostałych wyników. Były wyższe dla modelu TensorMask niż Mask R-CNN. Oba algorytmy lepiej także poradziły sobie z detekcją (bbox) niż segmentacją (segm). TensorMask osiągnął lepsze średnie wyniki dla różnych modeli treningowych podczas uczenia transferowego. Jednak podczas trenowania od podstaw to Mask R-CNN wykazał wyższe wartości mAP (powyżej 80 procent dla większości przypadków). Różnice widoczne są szczególnie dla obrazów *ch3*, dla którego najwyższa wartość mAP segmentacji wyniosła 32,752. Obraz *ch3* zawiera bardzo wąski zakres tylko najwyższych temperatur ciała zwierzęcia, w skład którego nie wchodzi obszary położone blisko granic. Taki obiekt może być odnaleziony na obrazie, ale niemożliwa jest jego poprawna segmentacja. Dlatego

TABELA 5.11. Wyniki testowania dla trenowania od podstaw TensorMask dla 500 obrazów. Pogrubiono najlepsze wyniki

TensorMask uczenie od podstaw	mAP	AP ₅₀	AP ₇₅	mAP	AP ₅₀	AP ₇₅
2 batch, 100 000 epok	bbox			segm		
ch0	78,235	100	92,41	79,615	100	95,535
ch1	78,862	100	98,307	77,724	100	90,687
ch2	64,42	99,18	69,516	62,307	96,206	71,867
ch3	44,557	84,367	40,402	13,097	60,133	0,012
16-bit	78,112	99,99	97,29	80,198	99,99	94,435
4 batch, 100 000 epok						
ch0	76,413	99,99	95,435	80,371	99,99	91,678
ch1	77,725	100	94,535	81,259	100	95,188
ch2	75,275	99,833	90,342	74,431	99,833	89,674
ch3	55,982	95,072	57,009	32,752	89,025	16,88
16-bit	76,296	100	90,662	81,267	100	94,887

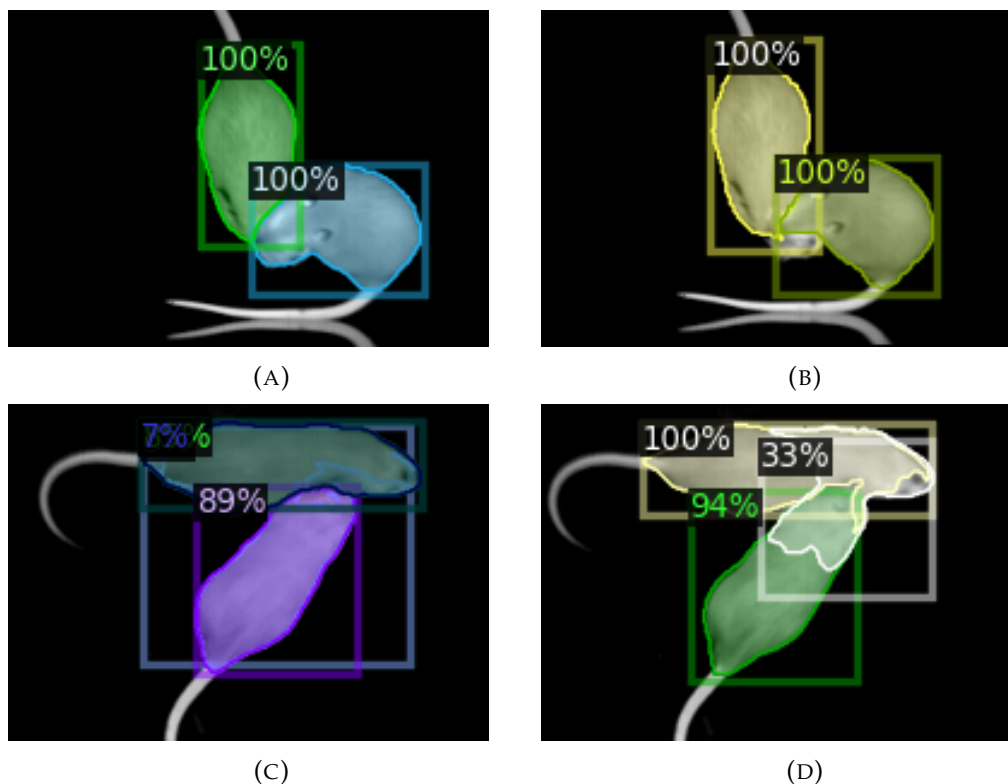


RYS. 5.25. Przykładowe obrazy 2 różnych scen a), c) oryginalne zawierające cały zakres temperatury oraz zaznaczenie powiększenia; b), d) segmentacji referencyjnej w powiększeniu oznaczonej żółtymi liniami

właśnie wartości mAP dla segmentacji obrazu *ch3* osiągają niższe wartości niż dla detekcji, w przeciwieństwie do trenowania od podstaw na pozostałych obrazach. Wyniki w tablicach 5.9 i 5.11 wykazują nieznacznie wyższe wartości dla segmentacji niż detekcji.

Rysunek 5.25 przedstawia przykłady segmentacji referencyjnej (tj. wykonanej przez obserwatora) na obrazach oryginalnych. Instancje tej samej klasy ("rat") oznaczone są żółtym zamkniętym wielokątem. Rys. 5.25b i 5.25d są powiększone w celu dokładniejszego zobrazowania trudnych przypadków segmentacji: nałożenia jednego obiektu na drugi, a także, w przypadku rys. 5.25b, rozłącznych obszarów obiektu będącego pod spodem.

Rysunek 5.26 przedstawia wyniki segmentacji instancji dla różnych typów uczenia. W pierwszej linii znajdują się obrazy po trenowaniu przez architekturę Mask R-CNN. Rys. 5.26a przedstawia wynik uczenia transferowego, a rys. 5.26b trenowania od podstaw. Różnice widoczne są w segmentacji obszaru osobnika znajdującego się pod spodem, który został odłączony od reszty przez pysk drugiego osobnika. Po trenowaniu tylko ostatnich warstw obszar ten został niepoprawnie zaklasyfikowany jako element osobnika znajdującego się na wierzchu. Uczenie całej architektury nieznacznie poprawiło wynik poprzez rozszerzenie obszaru osobnika pod spodem,



RYS. 5.26. Porównanie wyników segmentacji dla różnych typów uczenia. Obrazy przedstawiają wyniki dla: a) uczenia transferowego Mask R-CNN, b) trenowania od podstaw Mask R-CNN, c) uczenia transferowego TensorMask, d) trenowania od podstaw TensorMask

jednak wciąż nie jest to całkowicie poprawna klasyfikacja.

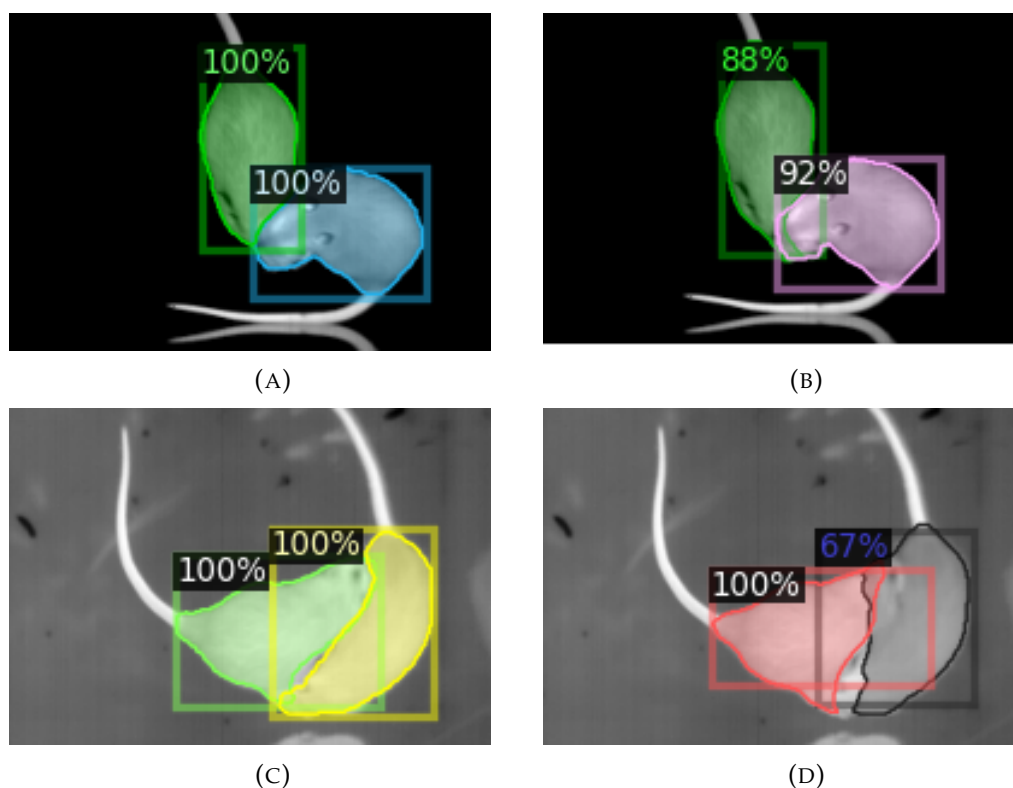
Drugi wiersz rysunku 5.26 przedstawia wyniki dla TensorMask po uczeniu transferowym (5.26c) oraz treningu od podstaw (5.26d). Obszary zwierząt na obu rysunkach wyodrębnione są prawidłowo, jednak w obu przypadkach pojawiają się dodatkowe obiekty uznane także za szczury (choć z mniejszym prawdopodobieństwem). Uczenie transferowe spowodowało detekcję obiektu będącego połączeniem obu instancji, trenowanie od podstaw natomiast wyodrębniło dodatkowy obszar obejmujący pyski wraz z przednimi partiami ciała obu zwierząt.

Rysunek 5.27 porównuje wyniki segmentacji dla różnych architektur. Rys. 5.27a oraz 5.27b prezentują wyniki segmentacji dla uczenia transferowego odpowiednio Mask R-CNN i TensorMask. Obie sieci niepoprawnie dołączyły część jednego osobnika do obszaru osobnika znajdującego się na wierzchu. TensorMask jednak część tego problematycznego obszaru prawidłowo zaklasyfikował do obiektu pod spodem.

Według wcześniejszych wyników przedstawionych w tabelicy 5.11 trenowanie od podstaw architektury TensorMask osiągnęło najgorsze wyniki. Widoczne jest to także na rys. 5.27d. Każdy z obiektów pozbawiony jest części obszaru, który powinien do niego należeć, posiada także obszar drugiego osobnika. Pysk zwierzęcia znajdującego się po prawej stronie nie został przypisany do żadnego z obszarów. Znacznie lepiej wygląda segmentacja wykonana po trenowaniu od podstaw sieci Mask R-CNN. Obszary są dość dokładnie wysegmentowane, obiekty rozpoznane z dużą pewnością.

Choć wyniki segmentacji przedstawionej powyżej nie zawsze są zadowalające, należy

pamiętać, że prezentowane były tutaj trudne przypadki. Dla zdecydowanej większości obrazów predykcja pokrywała się z obrazami referencyjnymi w znacznym stopniu.



RYS. 5.27. Porównanie wyników segmentacji dla różnych architektur. Obrazy przedstawiają wyniki dla: a) uczenia transferowego Mask R-CNN, b) uczenia transferowego TensorMask, c) trenowania od podstaw Mask R-CNN obrazów 16-bit, d) trenowania od podstaw TensorMask obrazów 16-bit

Tabela 5.12 przedstawia wyniki inferencji dla obrazów *ch0* przy użyciu modeli przetrenowanych na bazach MS COCO Dataset, Citiscapes i LVIS oraz dla dwóch różnych implementacji Mask R-CNN przetrenowanej na bazie COCO. Modele nie były dotrenowane na obrazach gryzoni. Osiągnięte wyniki są znacznie gorsze niż prezentowane w tabelach 5.8 - 5.11.

TABELA 5.12. Wyniki segmentacji obrazów *ch0* dla modeli wytrenowanych jedynie na obrazach z popularnych baz danych

	mAP	AP ₅₀	AP ₇₅	mAP	AP ₅₀	AP ₇₅
Mask R-CNN	bbox			segm		
MS COCO - implementacja [147]	4,14	13,39	1,4	5,36	14,44	1,05
MS COCO - implementacja [1]	7,62	13,5	7,62	-	-	-
Citiscapes	0,06	0,3	0	0,01	0,09	0
LVIS	0,71	3,45	0,04	0	0	0
TensorMask						
MS COCO - implementacja [147]	1,69	3,41	1,29	3,06	6,31	2,6

Wyniki 3-krotnej walidacji dla parametrów sieci: 4 batch, 2000 epok dla uczenia transferowego oraz 2 batch, 100000 epok dla trenowania od podstaw przedstawione

są w tabeli 5.13. Każdy z trzech osobnych treningów powtórzony został dwukrotnie. Zaprezentowano najlepsze wyniki.

Podobnie jak we wcześniejszych treningach, tak i tym razem, najwyższe miary detekcji i segmentacji zostały osiągnięte przez wcześniej przetrenowany algorytm TensorMask. Sieć Mask R-CNN po uczeniu transferowym również wykazuje większą dokładność niż po trenowaniu od podstaw, jednak tu różnice są mniejsze niż w przypadku TensorMask. Możliwe, że trenowanie TensorMask od przypadkowych wartości wag wymaga większej ilości danych i/lub epok do osiągnięcia wyników porównywalnych z Mask R-CNN lub trenowaniem od podstaw.

TABELA 5.13. Wyniki mAP dla 3-krotnej cross-walidacji dla uczenia transferowego (4 batch, 2000 epok) oraz trenowania od podstaw (2 batch, 100 000 epok) modeli Mask R-CNN i TensorMask

mAP	uczenie transferowe				trenowanie od podstaw			
	Mask R-CNN		TensorMask		Mask R-CNN		TensorMask	
	bbox	segm	bbox	segm	bbox	segm	bbox	segm
ch0								
1. próba	88,8	89,1	90,9	90,2	83,6	85,9	78,2	79,6
2. próba	92,0	89,5	91,9	90,5	84,7	87,6	73,9	70,6
3. próba	89,7	89,2	90,5	90,2	87,7	88,0	70,6	76,6
średnia	90,2	89,3	91,1	90,3	85,3	87,2	74,2	75,6
ch1								
1. próba	90,2	89,7	90,3	90,2	84,6	87,2	78,9	77,7
2. próba	88,4	89,1	91,3	90,4	88,3	88,9	78,7	77,8
3. próba	89,6	89,4	90,2	89,8	88,3	89,0	75	74,4
średnia	89,4	89,4	90,6	90,1	87,1	88,4	77,5	76,6
ch2								
1. próba	87,9	88,9	87,9	89,3	82,6	82,4	64,4	62,3
2. próba	84,7	86,4	89,1	89,0	86,6	88,2	66	57,2
3. próba	89,5	88,1	88,3	88,9	85,0	87,7	65,7	54,3
średnia	87,4	87,8	88,4	89,1	84,7	86,1	65,4	57,9
ch3								
1. próba	73,9	65,7	75,0	64,8	66,4	56,1	44,6	13,1
2. próba	63,0	51,9	70,9	61,9	65,6	57	40,2	11,2
3. próba	72,6	62,9	71,3	62,5	66,6	62,4	36,6	11,2
średnia	69,8	60,2	72,4	63,1	66,2	58,5	40,5	11,8
16-bit								
1. próba	90,5	89,4	91,6	90,2	84,9	85,9	78,1	80,2
2. próba	89,2	89,4	92,5	90,8	85,8	88,5	76,8	77,6
3. próba	90,0	89,1	90,6	89,9	86,6	87,8	71,5	73,9
średnia	89,9	89,3	91,6	90,3	85,8	87,4	75,5	77,2

Szereg przedstawionych wyników badań związanych z segmentacją instancji opublikowano w pracy [91].

5.5.3. Podsumowanie zastosowania metod uczenia głębokiego

Wyniki segmentacji przy użyciu sieci U-Net i V-Net są lepsze niż przy użyciu innych standardowych algorytmów. Nie są też narażone na zbyt szczegółową segmentację, jak wiele innych powszechnie stosowanych metod. Wyniki wskazują, że odpowiednio dobrany zakres temperatur do konwersji do postaci obrazu standardowego w

niektórych przypadkach jest w stanie poprawić wyniki segmentacji. Dużą zaletą tych architektur jest wymaganie niewielkiego zbioru treningowego.

Algorytmy segmentacji instancji używające uczenia głębokiego są w stanie jeszcze lepiej rozróżnić i poprawnie wysegmentować dwie instancje tej samej klasy pozostające w bliskim kontakcie lub nawet podczas częściowego nałożenia jednego obiektu na drugi. Najlepsze wyniki uzyskano dla wstępnie wytrenowanej sieci jednoetapowej - TensorMask. Architektura Mask R-CNN z kolei lepiej działa podczas trenowania od podstaw. Trenowanie modelu Mask R-CNN od podstaw nie jest, ani czasochłonne, ani nie wymaga dużej ilości danych. Może być stosowane do uczenia niestandardowych obrazów. Koszt trenowania TensorMask jest z kolei duży, a mimo to nie daje rezultatów podobnych do wcześniej wytrenowanej wersji. Wstępnie przetrenowany model TensorMask wykazuje najlepszą wydajność. Badania wskazują, że obrazy termiczne mogą być z powodzeniem analizowane za pomocą architektur wstępnie wyszkolonych na standardowych obrazach.

Konwersja danych termicznych z różnych zakresów w tym przypadku nie poprawia jakości segmentacji. Wyniki dla reprezentacji obrazu 16-bitowego (*16-bit*) i 8-bitowego (*ch0*), a także obrazu pozbawionego tła (*ch1*) są porównywalne. Wyniki dla obrazów z węższego zakresu danych (*ch2*) nie różnią się zbytnio od innych. Jedyna zauważalna różnica występuje w przypadku obrazu *ch3*, na którym widoczne są tylko informacje o najcieplejszych częściach ciała.

Adaptacja algorytmów głębokich sieci neuronowych dla obrazów termicznych jest obiecującą metodą segmentacji oraz detekcji obiektów na obrazach zwierząt laboratoryjnych. Pozwala na zapewnienie ciągłości identyfikacji, szczególnie w przypadkach częściowego zasłonięcia obiektów przez siebie nawzajem. Co w konsekwencji umożliwi automatyzację procesu analizy zachowań podczas trwania całego testu. Przedstawione podejście znajdzie zastosowanie wszędzie tam, gdzie wymagana jest identyfikacja i śledzenie zwierząt doświadczalnych, zwłaszcza w większych grupach.

5.6. Podsumowanie

W rozdziale tym przedstawiono metody wstępnego przetwarzania obrazów i śledzenia gryzoni na nagraniach testów laboratoryjnych. Zaproponowano również 3-etapowy algorytm segmentacji obiektów połączonych. Metoda 1. etapu segmentacji bazuje na operacjach morfologicznych. 2. etap polega na odnalezieniu linii rozdzielającej obiekty na podstawie charakterystycznych cech kształtu ciała, jest ona ściśle dostosowana do nagrań i nagrywanych zwierząt. Dwa pierwsze etapy działają w krótkim czasie, nie wymagają dużych zasobów sprzętowych. 3. etap wykorzystuje architektury sieci neuronowych. Wymaga większych zasobów obliczeniowych, większej ilości danych oraz wcześniejszego przygotowania danych do trenowania. Osiąga jednak dość dobre wyniki segmentacji w przypadkach, kiedy ciała zwierząt są częściowo niewidoczne. Konkretną metodę segmentacji 3. etapu można dostosować do własnych warunków. Architektury segmentacji semantycznej (U-Net i V-Net) osiągają gorsze wyniki niż segmentacji instancji (Mask R-CNN, TensorMask), nie wymagają jednak tak dużych zbiorów uczących.

Opisane badania potwierdzają pierwszą część tezy 2. - stosując metody uczenia

maszynowego, możliwa jest automatyczna segmentacja obiektów złączonych; pozwalają także na pomiar parametrów umożliwiających indywidualizację cech niezbędnych w detekcji zachowań socjalnych, co jest przedmiotem drugiej części tezy 2. W odniesieniu do tezy 1. należy podkreślić, że automatyzacja analizy zachowań socjalnych zwierząt laboratoryjnych wymaga m.in. identyfikacji, detekcji punktów charakterystycznych, określenia położenia ciała oraz wzajemnego położenia obu obiektów podczas całego procesu rejestracji badania, w szczególności podczas kontaktu. Wiele systemów przedstawionych w rozdziale 1.1 pomija w analizie moment, kiedy ciała zwierząt widoczne na obrazie są połączone. W rozdziale tym przedstawiono propozycję rozwiązania tego problemu. Segmentacja obiektów podczas różnego rodzaju kontaktów fizycznych pozwala na identyfikację osobników niezależnie od ich wzajemnego ułożenia. Co więcej, umożliwia pomiar parametrów niezbędnych do analizy zachowań, jak np. kształtu ciała, położenia punktów charakterystycznych czy postawy. Poprawna segmentacja typu obiekt-obiekt daje możliwość stosowania wszystkich metod analizy zachowań dotąd używanych jedynie w przypadkach, gdy ciała zwierząt były wyraźnie oddzielone od siebie, teraz również w momentach kontaktów zwierząt. Zastosowanie tych samych metod dla całego procesu obserwacji zwiększa ilość informacji o dane wcześniej pomijane w analizie oraz pozwala utrzymać ciągłość detekcji parametrów w czasie (co daje kolejne możliwości analizy). Niezbędne jest to w szczególności podczas testów socjalnych, gdzie kontakty są częste i szczególnie istotne dla dalszej analizy.

Podsumowując, prawidłowa segmentacja obiekt-obiekt, której propozycja została przedstawiona w tym rozdziale, jest potwierdzeniem pierwszej części i umożliwia wykonanie drugiej części tezy 2. Wnosi ona także swój wkład w potwierdzenie tezy 1. poprzez umożliwienie identyfikacji osobników oraz detekcji i pomiaru różnego rodzaju parametrów zachowań zwierząt, np. punktów charakterystycznych szerzej opisanych w rozdziale 7, na obrazach termograficznych nawet podczas bliskich kontaktów socjalnych zwierząt.

Rozdział 6

Badania symulacyjne

Uzyskanie odpowiednio licznej i zróżnicowanej bazy danych nagrań agresywnych interakcji zwierząt jest trudne. Związane jest to z koniecznością spełnienia specyficznych warunków sprzyjających wystąpieniu zachowań ustalania hierarchii. Takimi warunkami są m.in. poczucie bezpieczeństwa zwierząt, czyli znajome otoczenie i brak czynników uznawanych za niebezpieczne (jasne oświetlenie, brak obecności człowieka lub innych naturalnych wrogów) jak również odpowiednie warunki przetrzymywania. Duży wpływ na zachowania gryzoni ma także pora dnia, w nocy są one bardziej aktywne. Z tego powodu, zbiory obrazów przedstawiających zachowania agresywne jakimi są ugryzienia, poszerzono o obrazy symulowane.

W celu przetestowania algorytmów detekcji śladów ugryzień zaproponowanych w kolejnym rozdziale, stworzono symulacje nagrań testów socjalnych dwóch gryzoni. Zasympulowano kształty (rozdział 6.1) i tekstury (rozdział 6.2) obiektów, ich ruch łącznie z momentami nachodzenia na siebie (rozdział 6.3) oraz ugryzienia wraz ze zmianami temperatury śladu w czasie (rozdział 6.4).

Zaproponowane w tym rozdziale modele służą wyłącznie celom pracy i nie stanowią zaawansowanych metod kinematyki gryzoni.

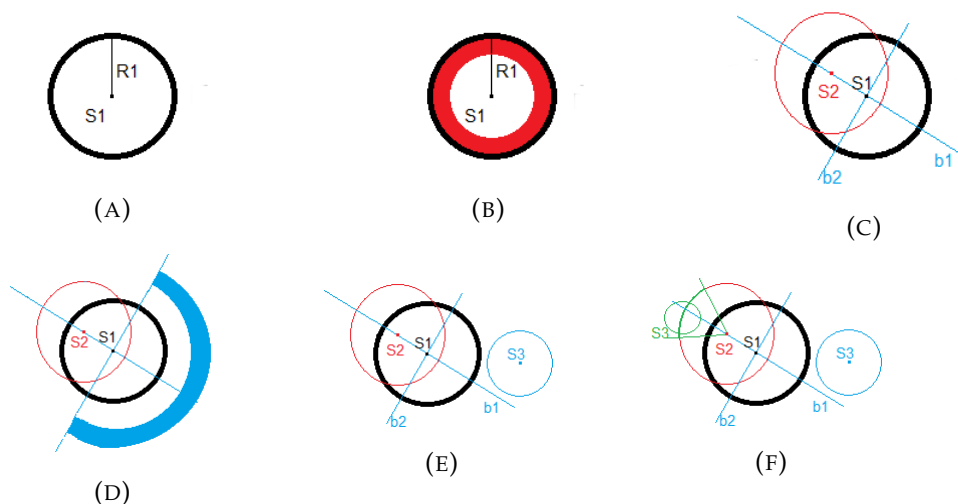
6.1. Symulacja kształtu obiektu

Kształt ciała został uogólniony czterema okręgami, których położenia względem siebie są określone zależnościami przedstawionymi poniżej. Wszystkie stałe wartości wykorzystywane podczas tworzenia kształtu, zostały dobrane na podstawie obserwacji ciała i ruchu zwierząt w laboratorium. Dwa największe wewnętrzne okręgi tworzą ciało zwierzęcia, mniejszy okrąg zewnętrzny to głowa, a najmniejszy zewnętrzny okrąg symuluje ogon. Poszczególne kroki tworzenia kształtu przedstawione są na rys. 6.1.

Symulacja kształtu rozpoczyna się od stworzenia pierwszego okręgu o promieniu R_1 oraz środka w punkcie S_1 (rys. 6.1a). Kształt od razu tworzony jest na obrazie, więc punkt S_1 jest losowo wybieranym punktem obrazu o wymiarach 240×320 pikseli z wyłączeniem marginesów z każdej strony. Marginesy zapobiegają wykroczeniu kształtu poza granice obrazu. Promień R_1 jest wartością losowaną z bardzo małego zakresu liczb.

Krok drugi to losowy wybór środka drugiego okręgu S_2 jedynie wewnątrz czerwonego pola przedstawionego na rys. 6.1b. Dozwolone pole dla środka S_2 jest różnicą pola okręgu pierwszego oraz okręgu o środku w punkcie S_1 i promieniu równym $\frac{2}{3}R_1$. Punkt S_2 nie może wyjść poza obszar okręgu pierwszego. Im szersze jest dozwolone (czerwone) pole, tym bardziej zwarte (ściśnięte) może być ciało zwierzęcia.





RYS. 6.1. Algorytm tworzenia symulowanego kształtu gryzonia w krokach: a) okrąg pierwszy - ciało, b) wybór dozwolonego miejsca losowania środka okręgu drugiego, c) okręgi pierwszy i drugi - ciało, d) wybór dozwolonego miejsca losowania środka okręgu trzeciego, e) okręgi pierwszy i drugi oraz okrąg trzeci - głowa, f) okręgi pierwszy, drugi, trzeci oraz czwarty - ogon

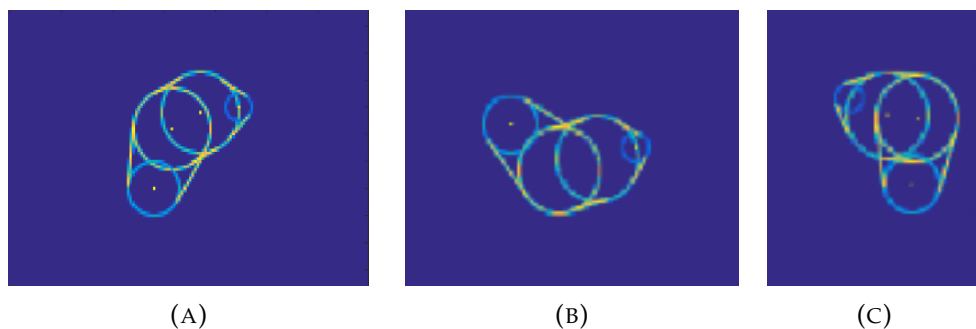
Po wylosowaniu punktu S_2 tworzony jest okrąg drugi o promieniu R_2 wybieranym analogicznie jak promień R_1 . Następnie wyznaczana jest prosta b_1 (rys. 6.1c) przechodząca przez punkty S_1 i S_2 oraz prosta b_2 do niej prostopadła przechodząca przez punkt S_1 . Prosta b_1 wyznacza kierunek ułożenia ciała, a prosta b_2 to prosta pomocnicza do wyznaczenia zakresu ruchu głowy.

Zwrot ułożenia ciała jest zgodny z wektorem o punkcie zaczepienia w S_2 i punkcie końcowym w S_1 . Zgodnie z tym zwrotem wyznaczane jest pole, w obrębie którego wylosowany zostanie punkt S_3 . Pole to, oznaczone kolorem niebieskim na rys. 6.1d, jest ograniczone prostą b_2 , ponieważ trzeci okrąg przedstawia głowę, której ruchy są naturalnie ograniczone. Im dalej punkty pola niebieskiego znajdują się od okręgu pierwszego (czarnego), tym większa jest dozwolona odległość głowy od tułowia. Możliwa jest wtedy symulacja "wydłużenia" ciała, którą gryzonie często wykazują podczas eksploracji nowego terytorium.

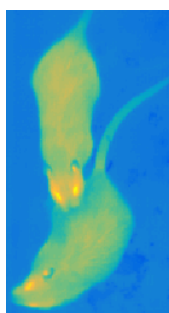
Rys. 6.1e przedstawia dwa okręgi (czarny i czerwony) odpowiadające ciału oraz okrąg niebieski przybliżający głowę.

Ostatnim krokiem jest stworzenie czwartego, najmniejszego okręgu reprezentującego ogon (dokładniej jego nasadę, ponieważ cały, długi ogon, zarówno w większej części analizy, jak i symulacjach został pominięty). Nasada ogona jest nieruchoma, nie ma możliwości przybliżenia lub oddalenia się od reszty ciała. Z tego względu środek czwartego okręgu S_4 zawsze położony jest na obwodzie okręgu drugiego po przeciwnej stronie niż okrąg trzeci. Dokładny punkt obwodu okręgu drugiego jest ograniczony odległością od prostej b_1 . Warunek ten przedstawiony jest na rys. 6.1f za pomocą półprostych oraz wycinka okręgu oznaczonych kolorem zielonym. Ruchomość ogona w obrębie zielonego wycinka symuluje skrętność tułowia.

Po stworzeniu kształtu ogólnego definiowane są styczne dla każdego z sąsiadujących okręgów, które oznaczają granicę obiektu. Rys. 6.2 demonstruje przykłady obliczonych obiektów wraz z granicami.



RYS. 6.2. Przykłady stworzonych kształtów ciała



RYS. 6.3. Przykład rzeczywistej tekstury powierzchni ciała zwierząt na obrazach termograficznych

6.2. Symulacja tekstury

Symulacja tekstury ciała zwierząt ma na celu jak najlepsze odwzorowanie obrazu z kamery. Ze względu na użycie kamer termograficznych, symulowana tekstura powinna odzwierciedlać rozkład powierzchniowy temperatury ciała gryzoni przedstawiony na rys. 6.3. Rozkład ten nie jest jednakowy we wszystkich miejscach ciała. Najwyższą temperaturę można zaobserwować na środku obiektu (wzdłuż grzbietu), najniższą w okolicach granic obiektu. Temperatura zmienia się stopniowo. Odrębnym zagadnieniem są punkty charakterystyczne ciała o znacznej różnicy temperatury, takie jak oczy lub nos. Punkty te zostaną opisane w dalszej części rozdziału.

Bardzo ważną rolę w obserwacji rozkładu powierzchniowego temperatury odgrywa futro - naturalny izolator zwierzęcia. Jego długość oraz ułożenie mają wpływ na detekcję temperatury przez kamerę. Krótkie futro młodych zwierząt układa się regularnie i tworzy jednolity obraz rejestrowanej temperatury. Zaburzenie tej regularności może zostać wywołane poprzez np. ugryzienie lub czyszczenie, kiedy to kępy sierści zostają przesunięte odkrywając lub zasłaniając małe obszary skóry, która ma wyższą temperaturę. W tych miejscach powstają lokalne minimalne cieplejsze lub chłodniejsze miejsca. Cecha ta jest przydatna podczas detekcji zachowań socjalnych.

Symulacja powierzchniowego rozkładu temperatury na obiekcie złożona jest z dwóch etapów:

1. podziału obszaru na podregiony oraz
2. przypisania regionom odpowiedniej temperatury.

Podział obszaru na podregiony stosuje metodę szkieletyzacji opartą na równoległym pocienianiu obiektu opisaną w [53] oraz [78]. Algorytm ten iteracyjnie usuwa zewnętrzne piksele (tzw. punkty proste) w taki sposób, że obiekt kurczy się do środkowej linii (grzbietu).

Działanie pojedynczego kroku iteracji algorytmu jest podzielone na dwie poditeracje, z których każda usuwa punkty w określonym kierunku:

1. Piksel p jest usunięty jeżeli spełnione są warunki:

$$(a) \quad X_H(p) = 1 \quad (6.1)$$

gdzie:

$$X_H(p) = \sum_{i=1}^4 b_i,$$

$$b_i = \begin{cases} 1, & \text{gdzy } x_{2i-1} = 0 \text{ oraz } (x_{2i} = 1 \text{ lub } x_{2i+1} = 1) \\ 0, & \text{w przeciwnym wypadku} \end{cases}$$

x_1, x_2, \dots, x_8 - wartości sąsiednich pikseli piksela p , dla których numeracja rozpoczyna się od piksela leżącego po prawej stronie i kontynuowana przeciwnie do ruchu wskazówek zegara

$$(b) \quad 2 \leq \min\{n_1(p), n_2(p)\} \leq 3 \quad (6.2)$$

gdzie:

$$n_1(p) = \sum_{k=1}^4 x_{2k-1} \vee x_{2k}$$

$$n_2(p) = \sum_{k=1}^4 x_{2k} \vee x_{2k+1}$$

$$(c) \quad (x_2 \vee x_3 \vee \bar{x}_8) \wedge x_1 = 0 \quad (6.3)$$

2. piksel p jest usunięty jeżeli spełnione są warunki:

(a) warunek ze wzoru 6.1,

(b) warunek ze wzoru 6.2,

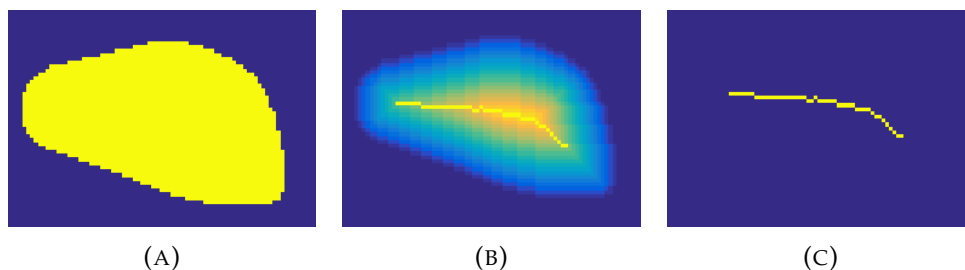
(c) obrót o 180° warunku ze wzoru 6.3:

$$(x_6 \vee x_7 \vee \bar{x}_4) \wedge x_5 = 0 \quad (6.4)$$

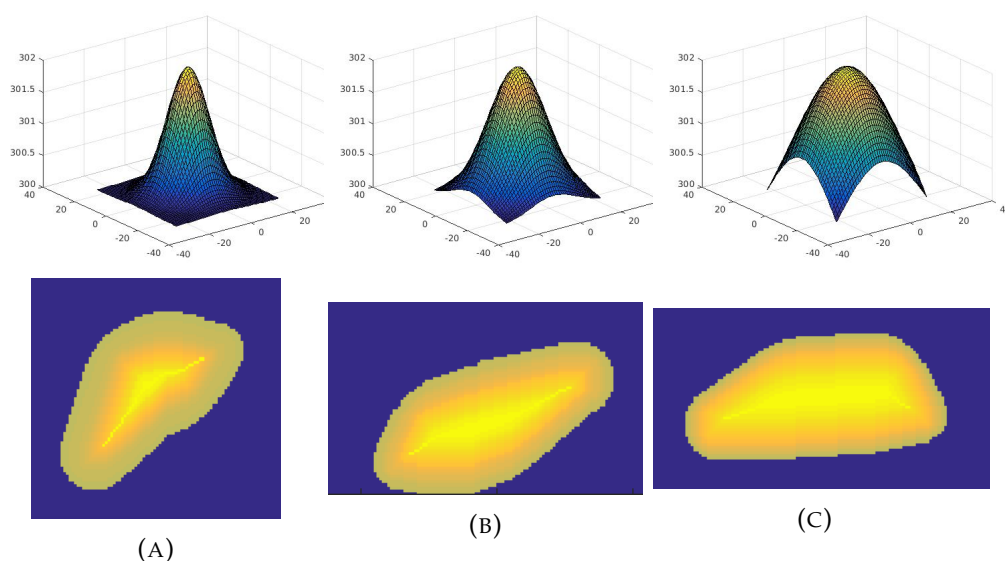
Wynik działania algorytmu został przedstawiony na rys. 6.4. Rys. 6.4a przedstawia kształt początkowy, kolejne iteracje zostały oznaczone na rys. 6.4b przy pomocy różnych kolorów. Rys. 6.4c to obraz ostatniej - 24. iteracji algorytmu.

Przypisanie otrzymanym regionom odpowiedniej temperatury wymaga stworzenia rozkładu Gaussa, ponieważ zaobserwowano, że bardziej naturalnie odwzorowuje on rozkład temperatury na ciele zwierzęcia niż np. rozkład liniowy (przedstawiony na rys. 6.4b). Stworzony rozkład jest spróbkowany zgodnie z liczbą kroków iteracji oraz znormalizowany zgodnie z obserwowanym zakresem temperatury powierzchniowej ciała. Każdemu kolejnemu podobszarowi wyznaczonemu poprzez algorytm szkieletyzacji opisany powyżej nadawana jest kolejna wartość temperatury





RYS. 6.4. Działanie algorytmu szkieletyzacji. A) obraz początkowy, B) kolejne iteracje, każda oznaczona innym kolorem, C) obraz końcowy otrzymany po 24 iteracjach



RYS. 6.5. Wpływ zmian wartości wariancji rozkładu na symulowaną teksturę: A) $Var = (\frac{bin}{3})^2$, B) $Var = (\frac{bin}{2})^2$, C) $Var = (bin)^2$, gdzie: bin - liczba przedziałów próbkowania równa liczbie iteracji algorytmu szkieletyzacji

zgodnie ze stworzonym rozkładem.

Zmiana wartości wariancji rozkładu powoduje zmianę stopnia zmienności symulowanej tekstury (rys. 6.5). Ostatecznie, do badań symulacyjnych został wybrany rozkład przedstawiony na rys. 6.5c.

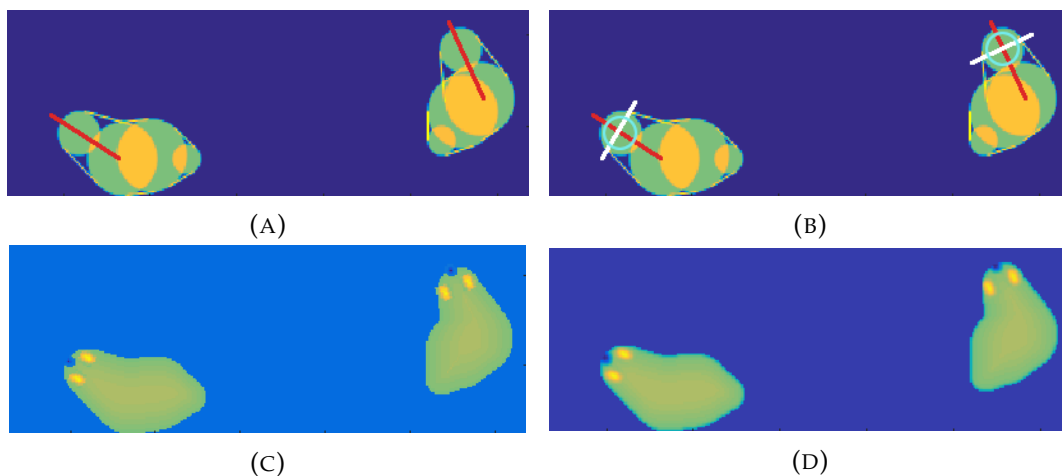
6.2.1. Punkty charakterystyczne ciała

Punktami charakterystycznymi ciała określono stałe miejsca na ciele gryzonia o wyraźnie różnej temperaturze niż reszta obszaru ciała. Takimi naturalnymi punktami są nos i oczy.

Nos został zasymulowany poprzez nałożenie na teksturę okrągłego śladu gradientowego o niskiej temperaturze w punkcie przecięcia okręgu 3. określającego głowę oraz prostej przechodzącej przez środki okręgów głowy i ciała (S_3 i S_1) oznaczonej na czerwono na rys. 6.6a

Oczy oznaczone zostały jako elipsy o gradientowej teksturze usytuowane w miejscu przecięcia się prostej prostopadłej do prostej wyznaczającej nos i przechodzącej przez środek okręgu 3. (oznaczona kolorem białym na rys. 6.6b) oraz okręgu o środku w punkcie S_3 i promieniu 5 pikseli mniejszym niż promień okręgu 3. (okrąg oznaczony

kolorem jasno-niebieskim na rys. 6.6b). Rys. 6.6c przedstawia obraz po nałożeniu punktów charakterystycznych ciała na teksturę. Rys. 6.6d to obraz wynikowy - stworzony po rozmyciu oraz maskowaniu obrazu rys. 6.6c. Rozmycie zostało wykonane przy użyciu 2D filtru gaussowskiego o $\sigma = 1$.



RYS. 6.6. Obliczanie miejsc położenia A) nosa i B) oczu. C) Obraz po nałożeniu okręgu i elips na teksturę, D) Obraz wynikowy, po filtracji i maskowaniu

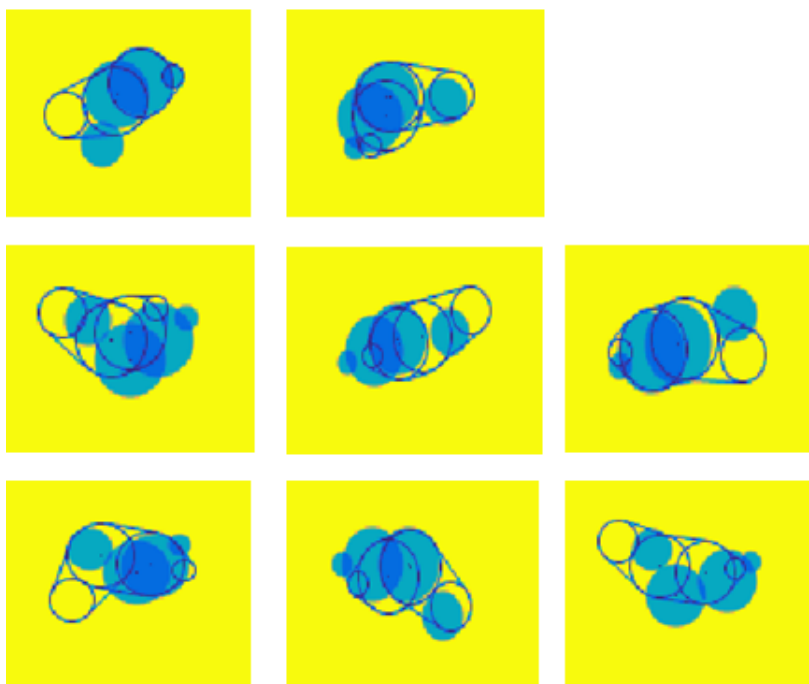
6.3. Symulacja ruchu

Ruch wyznaczany jest przez wektor $\overrightarrow{S1S3}_t$ o punkcie początkowym w $S1$ oraz końcowym w $S3$ w czasie t . Dla czasu $t + 1$ punkt $S1$ przesuwany jest zgodnie z wektorem $\overrightarrow{S1S3}_t$. Środki pozostałych okręgów obliczane są w sposób przedstawiony powyżej, jednak z uwzględnieniem maksymalnej odległości nowego punktu od punktu z poprzedniej klatki. Przykładowe ruchy w ciągu dwóch klatek zostały przedstawione na rys. 6.7. Wypełnione okręgi są okręgami w czasie t , okręgi bez wypełnienia obliczone są dla czasu $t + 1$.

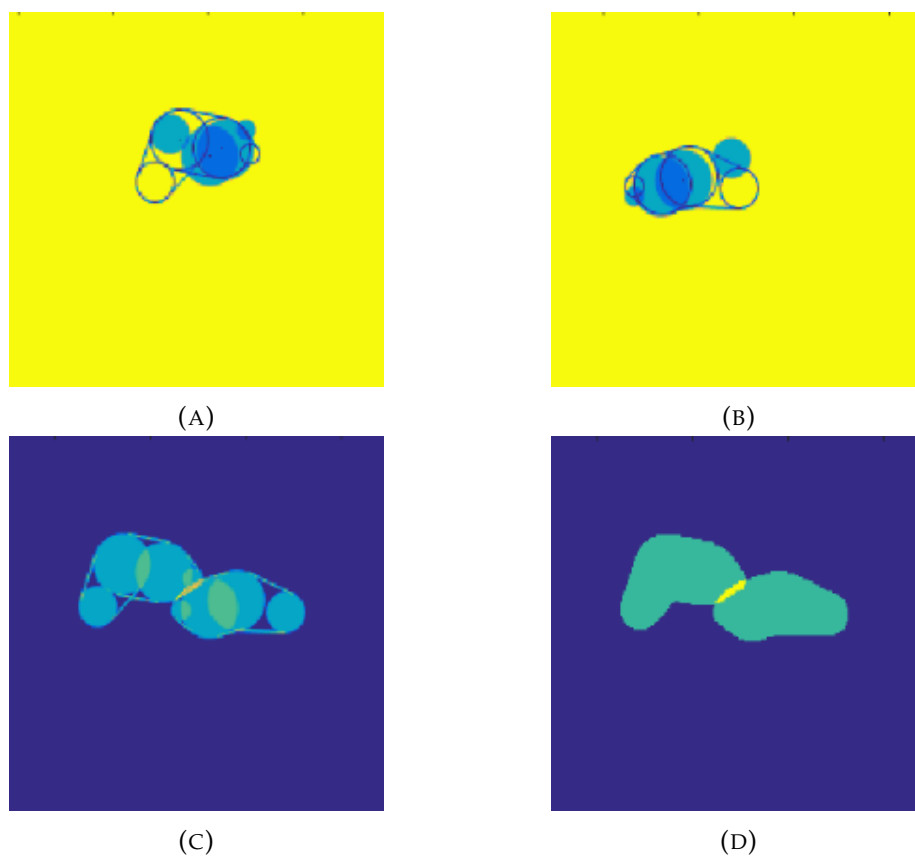
Symulacja obecności dwóch gryzoni w klatce przedstawiona jest na rys. 6.8. Oba kształty obliczane są niezależnie od siebie zgodnie z wcześniej przedstawionymi zależnościami między okręgami i kolejnymi klatkami (rys. 6.8a i 6.8b). Następnie oba wyniki aktualnej klatki przenoszone są na wspólny obraz (rys. 6.8c). Ostatnim etapem jest nadanie pikselom wewnątrz każdego obiektu odpowiedniej etykiety (granica definiowana jest przez styczne okręgów) oraz detekcja ewentualnych części wspólnych obiektu (rys. 6.8d). Ciała zwierząt podczas przeprowadzanych nagrań bardzo często nachodziły na siebie w skutek np. zachowań takich jak wspinanie się czy pozycja ofensywna (tabela 3.1), dlatego w symulacjach możliwe jest wystąpienie kolizji obiektów.

6.3.1. Detekcja granic obrazu

Granice obrazu symulują ściany klatki badań, więc obiekt znajdujący się w ich pobliżu, po pewnym czasie powinien od nich odejść. Jeśli ciało obiektu (pierwszy i drugi okrąg) znajdzie się bardzo blisko granicy obrazu zwrócony do niej czołem, algorytm może nie mieć możliwości znalezienia punktu $S3$ zgodnego ze wszystkimi założeniami i mieszczącego się w granicach obrazu. W takim przypadku, gdy po



RYS. 6.7. Przykłady symulacji ruchu



RYS. 6.8. Proces tworzenia dwóch obiektów na obrazie: a) ,b) Dwa osobne kształty, c) oba kształty na jednym obrazie, d) wypełnione kształty tworzące obiekty z detekcją nałożenia

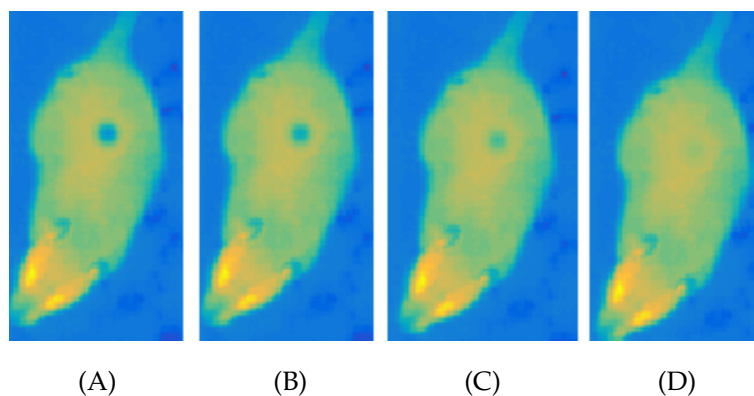
określonej liczbie prób odnalezienia punktu S_3 wciąż nie ma poprawnego wyniku, obiekt wykonuje maksymalny skręt głową. Punkt S_3 zostaje ustawiony na przecięciu prostej b_2 oraz okręgu pierwszego po stronie, dla której kąt utworzony przez płaszczyznę ściany oraz prostą b_1 jest rozwarty. W kolejnych krokach symulacji ciało podąży za wektorem $\overrightarrow{S_1S_3}$ wzdłuż ściany lub oddali się od niej.

6.4. Symulacja śladów śliny

Zgodnie z celami pracy, istotnym elementem oceny jest detekcja śladów śliny będącej efektem zachowań agresywnych.

W celu analizy możliwości detektorów punktów charakterystycznych wykonano symulacje śladów śliny pozostawionych na futrze zwierząt. Symulacje różnią się pod względem:

- temperatury: 297, 299, 301 oraz 303 K (rys. 6.9),
- kształtu: okrągły i podłużny (rys. 6.10 i 6.11),
- wielkości oraz gradientu intensywności pikseli: 4 stopnie (rys. 6.10 i 6.11).



RYŚ. 6.9. Ślad okrągły o największym rozmiarze w temperaturze a) 297K, b) 299K, c) 301K i d) 303K

Symulowane ślady miały charakter rozkładu Gaussa w obu płaszczyznach. Cechy i postać śladu zostały zweryfikowane na podstawie pomiarów laboratoryjnych. Ślady o większej wariancji charakteryzowały się większym rozmiarem oraz mniej stromym zboczem (a więc mniejszym gradientem intensywności pikseli w otoczeniu punktu) niż ślady o mniejszej wariancji. Dla śladów podłużnych ustawione zostały różne wariancje dla każdej płaszczyzny (OX i OY) (rys. 6.11 b, c; rys. 6.11 f, g; rys. 6.11 j, k; rys. 6.11 n, o). Każdy ślad okrągły posiadał taką samą wariancję w osi x oraz y. Rys. 6.10 przedstawia sygnał traktowany jako ślad okrągły:

- jego rozkład intensywności pikseli - rys. 6.10 a, d, g, j,
- przekrój poprzeczny wykresu powierzchniowego ilustrujący stromość zbocza (a tym samym gradient intensywności pikseli w otoczeniu punktu) dla obu płaszczyzn – rys. 6.10 b, e, h, k,
- obraz oryginalny z nałożonym sygnałem oraz powiększenie sygnału i jego otoczenia – rys. 6.10 c, f, i, l.

Stworzono cztery ślady okrągłe o różnych wielkościach:

- wielkość 1 – ślad o wielkości 1 piksela – rys. 6.10 a,
- wielkość 2 – ślad o promieniu 2 pikseli – rys. 6.10 d,
- wielkość 3 – ślad o promieniu 3 pikseli – rys. 6.10 g,
- wielkość 4 – ślad o promieniu 4 pikseli – rys. 6.10 j.

Rys. 6.11 przedstawia sygnał traktowany jako ślad podłużny:

- jego rozkład intensywności pikseli - rys. 6.11 a, e, i, m,
- przekrój poprzeczny wykresu powierzchniowego w płaszczyźnie OX– rys. 6.11 b, f, j, n,
- przekrój poprzeczny wykresu powierzchniowego w płaszczyźnie OY– rys. 6.11 c, g, k, o,
- obraz oryginalny z nałożonym sygnałem oraz powiększenie sygnału i jego otoczenia – rys. 6.11 d, h, l, p.

Stworzono cztery ślady podłużne o różnych wielkościach i różnych wartościach odchylenia standardowego wzdłuż obu osi (x i y):

- wielkość 1 – ślad o długości 3 i szerokości 1 piksela,
- wielkość 2 – ślad o długości 5 i szerokości 1 piksela,
- wielkość 3 – ślad o długości 7 i szerokości 1 piksela,
- wielkość 4 – ślad o długości 9 i szerokości 3 pikseli.

Zasymulowane ślady zostały zastosowane na 2 sposoby, dla:

1. rzeczywistych obrazów termograficznych z zebranej bazy danych (przykład - rys. 6.9) - użyte podczas analizy wpływu parametrów detektorów na detekcję (rozdział 7.3.4);
2. obrazów symulacji zwierząt (przykład - rys. 6.12) - użyte podczas analizy dynamicznych zmian temperatury śladów (rozdział 7.4.2).

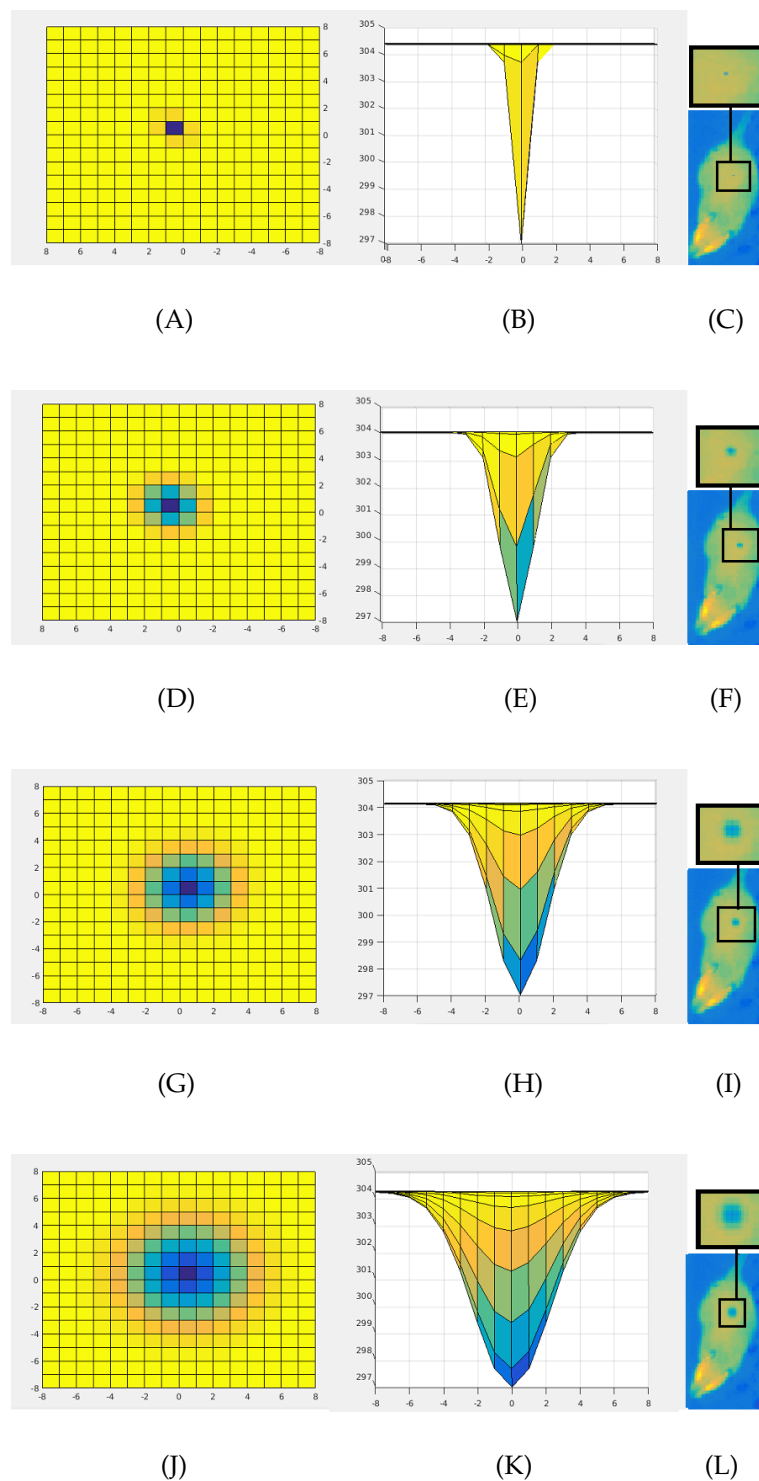
Temperatury brzegów wszystkich śladów, niezależnie od obrazu, na który zostały nałożone, zostały dopasowane do temperatury obszaru tak, aby gradient temperatur na granicach był jak najbardziej zbliżony do warunków naturalnych.

Ślady, które analizowane były w określonym przedziale czasu (naniesione na symulacje obiektów) musiały także poruszać się zgodnie z ruchem zwierząt.

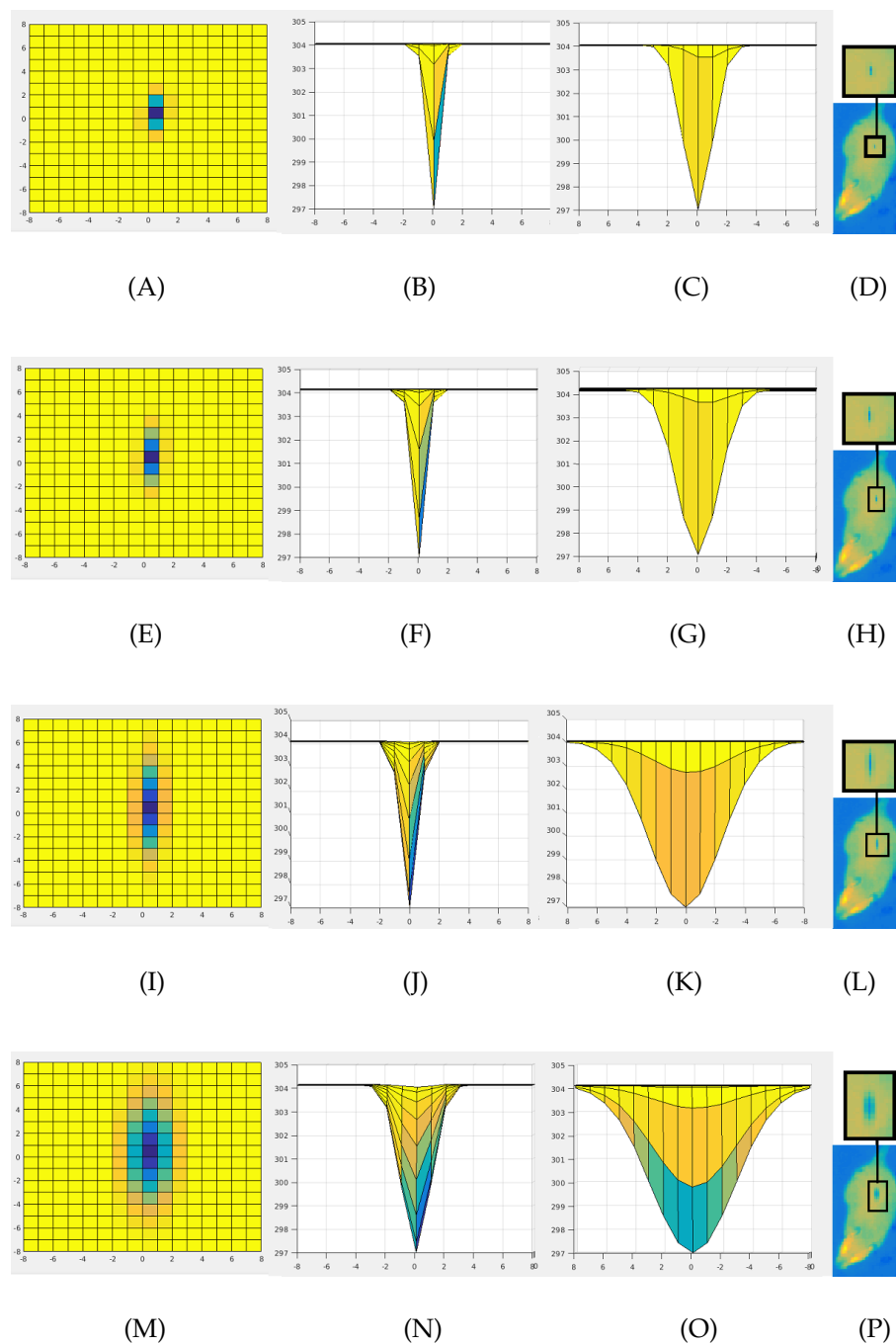
6.4.1. Symulacja ruchu śladu śliny

Ruch śladów śliny związany z ruchem zwierząt został zasymulowany poprzez nałożenie śladu na symulowane kształty obiektu opisane w rozdziale 6.1. Ślad przemieszcza się zgodnie z wektorem przemieszczenia obiektu pomiędzy klatkami, zachowuje jednak stałe położenie względem okręgów składowych obiektu oraz względem kierunku ułożenia. Przedstawione jest to na rys. 6.13.

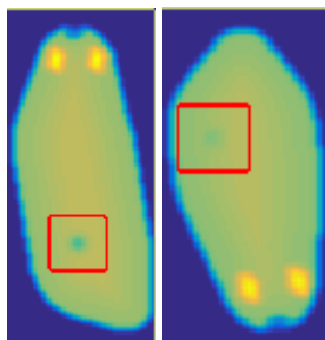
Obiekt ma dwa kierunki ułożenia: kierunek ciała oraz kierunek głowy przedstawione na rys. 6.13 a) odpowiednio kolorami zielonym oraz niebieskim. Obliczany jest następnie kąt, o który obrócona została odpowiednia linia kierunku ułożenia kształtu względem poprzedniej klatki (rys. 6.13 c). Ślad jest przesuwany zgodnie z wektorem przemieszczenia, po czym obracany zgodnie z kątem obliczonym w poprzednim kroku (rys. 6.13 d).



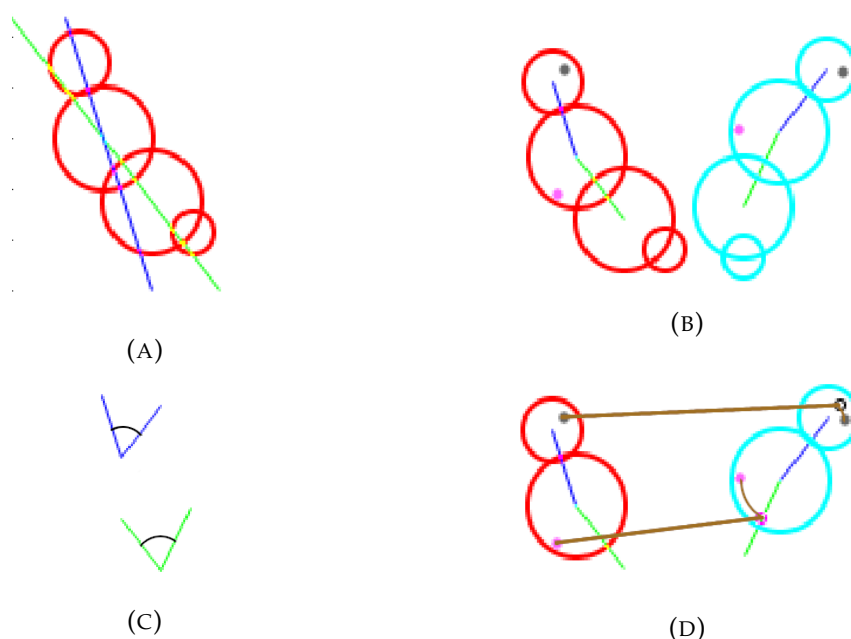
RYS. 6.10. Ślad okrągły o temperaturze 297 K i różnych rozmiarach: a) widok śladu o promieniu 1 piksela (wielkość 1), b) przekrój poprzeczny śladu o wielkości 1, c) widok nałożenia śladu o wielkości 1 na obraz; d) widok śladu o promieniu 2 pikseli (wielkość 2), e) przekrój poprzeczny śladu o wielkości 2, f) widok nałożenia śladu o wielkości 2 na obraz; g) widok śladu o promieniu 3 pikseli (wielkość 3), h) przekrój poprzeczny śladu o wielkości 3, i) widok nałożenia śladu o wielkości 3 na obraz; j) widok śladu o promieniu 4 pikseli (wielkość 4), k) przekrój poprzeczny śladu o wielkości 4, l) widok nałożenia śladu o wielkości 4 na obraz



RYS. 6.11. Ślad podłużny o temperaturze 297 K i różnych rozmiarach: a) widok śladu o długości 3 pikseli i grubości 1 piksela (wielkość 1), b) przekrój poprzeczny śladu o wielkości 1 wzdłuż osi x, c) przekrój poprzeczny śladu o wielkości 1 wzdłuż osi y, d) widok nałożenia śladu o wielkości 1 na obraz; e) widok śladu o długości 4 pikseli i grubości 1 piksela (wielkość 2), f) przekrój poprzeczny śladu o wielkości 2 wzdłuż osi x, g) przekrój poprzeczny śladu o wielkości 2 wzdłuż osi y, h) widok nałożenia śladu o wielkości 2 na obraz; i) widok śladu o długości 6 pikseli i grubości 1 piksela (wielkość 3), j) przekrój poprzeczny śladu o wielkości 3 wzdłuż osi x, k) przekrój poprzeczny śladu o wielkości 3 wzdłuż osi y, l) widok nałożenia śladu o wielkości 3 na obraz; m) widok śladu o promieniu 7 pikseli i grubości 3 pikseli (wielkość 4), n) przekrój poprzeczny śladu o wielkości 4 wzdłuż osi x, o) przekrój poprzeczny śladu o wielkości 4 wzdłuż osi y, p) widok nałożenia śladu o wielkości 4 na obraz



RYS. 6.12. Przykłady nałożenia śladów śliny o różnych temperaturach na symulacje zwierząt. Miejsce śladu oznaczono czerwonym kwadratem



RYS. 6.13. Symulacja przesuwania śladu na obiekcie: a) kształt z zaznaczonymi liniami kierunku ułożenia kształtu, linia zielona - główny kierunek kształtu, linia niebieska - kierunek ułożenia głowy względem ciała, b) przemieszczenie w czasie kształtu oraz śladów - wypełnione czarne i różowe okręgi, c) kąty obrotów kierunków ułożenia ciała w kolejnych klatkach, d) przemieszczenie śladów ugryzienia w kolejnych klatkach: 1. przemieszczenie liniowe zgodne z wektorem przesunięcia kształtu, 2. obrót o wcześniej obliczony kąt

6.4.2. Symulacja zmian temperatury śladu w czasie

Ślady śliny, w przeciwieństwie do innych charakterystycznych punktów na ciele, mają zmienną temperaturę i ostatecznie wysychają. Zmiany temperatury śladów śliny pozostawionych na futrze szczegółowo opisane zostały w rozdziale 7.4.1 oraz pracy [96]. W wyniku przeprowadzonych badań wykazano, że ślady można zamodlować dwoma odrębnymi procesami :

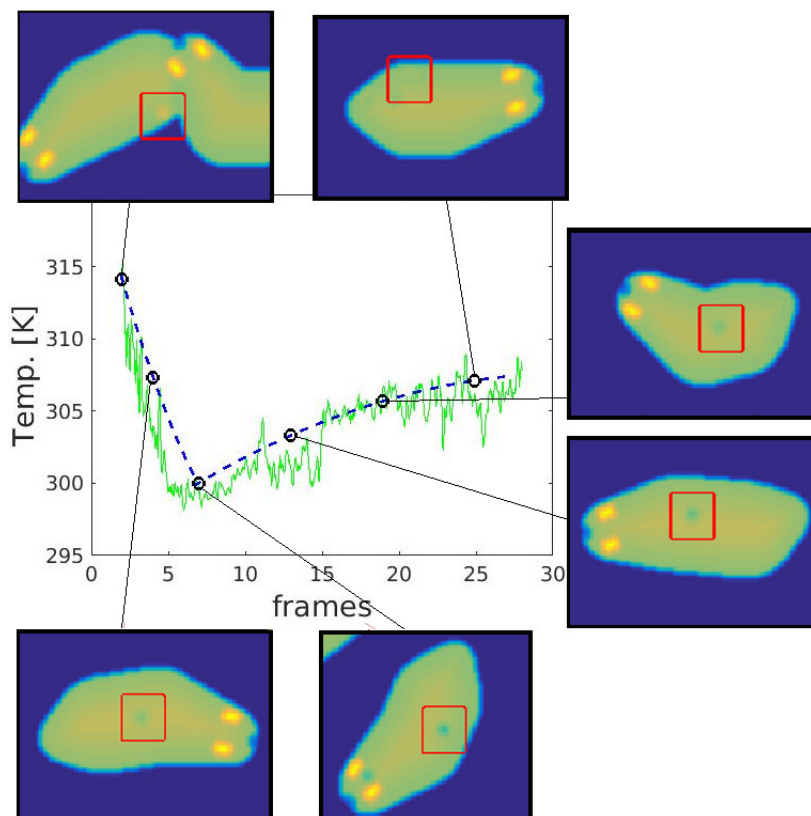
- chłodzeniem - bardzo krótki i ledwo dostrzegalny proces wynikający z szybkiej utraty ciepła śladu;
- ogrzewaniem - wolniejszy wzrost temperatury do wartości otoczenia (w tym wypadku temperatury powierzchni ciała zwierzęcia) spowodowany wysychaniem śladu, jego skutkiem jest powolne zanikanie widoczności śladu w zakresie częstotliwości kamer termowizyjnych.

Przyjęto następujące założenia:

- oba procesy, którymi poddane są ślady śliny są nierozłączne - występują zawsze razem, chociaż proces chłodzenia może być bardzo krótki;
- nieznany jest wpływ parametrów jednego z procesów na drugi, dlatego analiza symulacji traktuje oba procesy osobno;
- różne ślady wykazywały dopasowanie do funkcji kwadratowej o różnych parametrach, jednak przyjęto, że wszystkie symulacje są jednakowe i wzorowane na parametrach śladu *Trace₂*, ponieważ jest to jedyny ślad zaobserwowany w obu procesach w całości;
- procesy chłodzenia i ogrzewania śladów śliny są bardzo złożone i zależą od wielu czynników, m.in. temperatury i wilgotności otoczenia, temperatury obu osobników, długości i grubości futra (które z kolei zależne są od wieku zwierzęcia) czy cech osobniczych zwierząt. Czynniki te nie są uwzględniane w przedstawionych symulacjach. Symulacje mają na celu jedynie zdefiniowanie warunków poprawnej detekcji śladów i ich predykcji w momencie utraty widoczności.

Symulacja dynamiki śladu polegała na naniesieniu śladu na ciało po zasłonięciu obiektu przez pysk drugiego osobnika. Od momentu nałożenia, temperatura śladu w każdej następnej klatce zmieniała się zgodnie z wykresem przedstawionym niebieską linią na rys. 6.14. Podczas symulacji dla wszystkich śladów zastosowano tę samą funkcję zmian temperatury w czasie. Ślady miały również jednakowy rozmiar 2 pikseli oraz temperaturę 299 K.

Zielona linia na rys. 6.14 przedstawia przykładową obserwację zmian temperatury w czasie. Pomiary innych śladów wykazały, że dynamika i czas widoczności śladów są zróżnicowane, jednak wszystkie pomiary obu procesów wykazują najlepsze dopasowanie do wielomianu drugiego rzędu [96]. Sześć wybranych klatek na rys. 6.14 ilustruje widok śladów symulowanej śliny na ciele zwierzęcia w 2, 4, 7, 13, 19 i 25 klatce symulowanej sekwencji. Początkowo temperatura śladu jest wysoka, wyższa niż temperatura powierzchni ciała, ślad jest widoczny w postaci pomarańczowego punktu. Potem temperatura szybko spada poniżej temperatury ciała i zmienia kolor na niebieski - jest to proces chłodzenia. Od 7. klatki następuje wolniejszy proces wysychania, w którym temperatura śladu rośnie do temperatury powierzchni ciała i ostatecznie przestaje być widoczna (klatka 25.).



Rys. 6.14. Wykres zmian temperatury dla przykładowej obserwacji (zielona linia) oraz symulacji (niebieska przerywana linia). Przykładowe obrazy przedstawiają symulacje śliny podczas chłodzenia: klatki 2, 4 i 7 oraz ogrzewania: klatki 13, 19 i 25. Obrazy nie pochodzą z tej samej sekwencji symulacji

6.5. Podsumowanie

W rozdziale 6 został przedstawiony sposób wykonania symulacji obrazów termograficznych szczurów podczas interakcji socjalnych. Zasymulowany został kształt i wygląd każdego osobnika, sposób ich poruszania się oraz kontaktu. Szczególną uwagę zwrócono na symulację śladów śliny pozostawionych po ugryzieniach. Na podstawie obserwacji nagrań rzeczywistych, stworzono model 2-etapowego procesu, któremu poddane są ślady. Badania symulacyjne zostały stworzone w celu generalizacji, uzyskania większej różnorodności oraz liczby nagrań, które są trudne do uzyskania w naturalnych warunkach. Agresywne zachowania socjalne nie występują często pomiędzy zwierzętami laboratoryjnymi, które mają zapewnione dobre warunki życia i nie mają kontaktu ze swoimi naturalnymi wrogami. Nagrania symulacyjne dodatkowo oferują możliwość kontroli warunków. Można je zastosować do wstępnej oceny algorytmów analizy obrazu.

Zaprezentowane w tym rozdziale badania symulacyjne wykorzystane zostały w rozdziałach 7.3.4 i 7.4.2 do analizy możliwości zdefiniowania uniwersalnych detektorów i parametrów umożliwiających indywidualizację cech niezbędnych w detekcji zachowań socjalnych, co jest przedmiotem tezy 2. W rozdziale 7.4.2 zbadano również wpływ zmian temperatury symulowanych śladów w czasie na możliwość rozpoznania określonych klas zachowań socjalnych na obrazach termograficznych

w odniesieniu do tezy 1.

Rozdział 7

Parametry w detekcji zachowań socjalnych

Obrazy termograficzne przedstawiające nagrania testów socjalnych zwierząt laboratoryjnych dostarczają dodatkowe dane, które mogą zostać wykorzystane do bardziej zaawansowanej analizy tych zachowań. Tą dodatkową, użyteczną informacją jest temperatura. Temperatura całego ciała, jak i narządów peryferyjnych oraz ich zmiany mogą być wskaźnikiem fizycznego lub psychicznego stanu zwierzęcia [137]. Z kolei znaczna różnica temperatury punktów charakterystycznych znajdujących się na powierzchni ciała, jak np. oczy, uszy, nos czy nasada ogona, pozwala na szybką i łatwą detekcję tych punktów. Poprawna lokalizacja określonych części ciała jest kluczową informacją o jego ułożeniu, niezbędną podczas klasyfikacji określonych zachowań (patrz. tabela 3.1). Ślady śliny o zmiennej w czasie temperaturze pozostawione na futrze po ataku lub czyszczeniu są czasem jedynym dowodem na wystąpienie agresywnego zdarzenia pomiędzy dwoma osobnikami.

Rozdział ten poświęcony jest:

- przetestowaniu możliwości automatyzacji procesu analizy złożonych zachowań socjalnych, w szczególności agresywnych, na drodze analizy sekwencji obrazów termograficznych w celu oceny prawdziwości tezy 1.
- zaproponowaniu i przetestowaniu parametrów niezbędnych w detekcji zachowań agresywnych umożliwiających automatyczną detekcję nietypowych zdarzeń w celu oceny prawdziwości drugiej części tezy 2.

W punkcie 7.1 przedstawiono wykorzystane detektory punktów charakterystycznych oraz użyte miary. Działanie tych detektorów na obrazach termograficznych gryzoni zostało porównane w rozdziale 7.2. Rozdział 7.3 przedstawia propozycje deskryptorów przydatnych podczas statycznej detekcji śladów śliny. Analiza zmian detekcji śladów śliny w czasie zawarta jest w rozdziale 7.4.

7.1. Detektory punktów charakterystycznych

Jedną z metod analizy obrazu jest ekstrakcja przestrzennych obszarów zainteresowania w formie cech oraz ich późniejsza analiza. Analiza zachowania sprowadza się tu do postrzegania zachowania jako zbioru cech obrazu. Decyzja o podobieństwie dwóch zachowań zostaje tym sposobem zastąpiona odpowiednią detekcją oraz opisem zbioru cech. Przedstawienie obrazu jako parametrów wybranych punktów w niektórych przypadkach okazuje się być skuteczniejsze niż jako parametrów całego obrazu podczas rozpoznania zachowania [126].

Obszary zainteresowania mogą mieć charakter krawędzi, obiektów (potocznie blobów) lub punktów. Ze względu na zaproponowany algorytm, charakter obrazów, który pozwala na niezbyt skomplikowaną detekcję obiektów i ich konturów oraz fakt, iż obrazy analizowane w tej pracy przedstawiają powierzchniowy rozkład temperatury, zastosowane zostały detektory punktów charakterystycznych.

Istnieje wiele metod detekcji przestrzennych punktów zainteresowania opisanych np. w [125]. Jedną z nich jest algorytm Harris'a [57]. Dla każdego punktu obrazu wyznaczana jest macierz kowariancji gradientu intensywności, a następnie jej wartości własne λ_1 i λ_2 . Na podstawie wzoru 7.1 obliczana jest miara R dla każdego piksela.

$$R = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2 \quad (7.1)$$

gdzie: λ_1 i λ_2 - wartości własne macierzy kowariancji gradientu intensywności pikseli. Jeśli wartość R jest większa niż wcześniej ustalony próg, piksel zostaje uznany za narożnik (punkt charakterystyczny).

Shi i Tomasi w pracy [129] przedstawili modyfikację algorytmu Harris'a. Miara R wyznaczana jest ze wzoru 7.2 i oznacza, że mniejsza wartość własna macierzy kowariancji gradientu musi być powyżej zadanego progu aby punkt został uznany za narożnik.

$$R = \min(\lambda_1, \lambda_2) \quad (7.2)$$

Inną metodą detekcji punktów jest algorytm FAST (Features from Accelerated Segment Test) przedstawiony w [122]. Detektor ten sprawdza intensywność szesnastu pikseli wokół potencjalnego narożnika p o intensywności I_p . Jeżeli wśród tych 16-tu pikseli istnieje n narożników, które spełniają jeden z warunków równania 7.3: są jaśniejsze niż $I_p + t$ lub ciemniejsze niż $I_p - t$

$$I_x > I_p + t \quad \text{lub} \quad I_x < I_p - t \quad (7.3)$$

gdzie:

I_x - intensywność jednego z szesnastu pikseli wokół potencjalnego narożnika,

I_p - intensywność potencjalnego narożnika p ,

t - wartość progu,

wtedy punkt p uznawany jest za narożnik.

7.2. Porównanie detektorów punktów charakterystycznych

W celu adaptacji na potrzeby tej pracy najbardziej odpowiedniego detektora przeprowadzono badania porównawcze 3 algorytmów: Harris'a, Shi-Tomasi oraz FAST dla różnych obrazów i parametrów detektorów. Obrazy, na których przeprowadzono testy podzielono na 3 grupy:

- obrazy przedstawiające obiekt ze śladem po ugryzieniu,
- obrazy przedstawiające obiekt ze śladem po czyszczeniu futra,
- obrazy osobników bez śladów śliny na futrze.

Parametry detektorów zostały dobrane na zasadzie eksperymentalnej i wynoszą:

1. Harris:

- minimalna akceptowalna jakość narożników (R) – zakres $[0,1] = 0,0001; 0,0005$ i $0,001$
- rozmiar filtra Gaussa = 3; 5; 7 i 9 pikseli

2. Shi-Tomasi:

- minimalna akceptowalna jakość narożników (R) – zakres $[0,1] = 0,005; 0,01$ i $0,03$
- rozmiar filtra Gaussa = 3; 5; 7 i 9 pikseli

3. FAST

- minimalna akceptowalna jakość narożników (R) – zakres $[0,1] = 0,05; 0,1$ i $0,2$
- minimalna różnica intensywności pomiędzy punktem a jego sąsiedztwem - zakres $[0,1] = 0,01; 0,05; 0,075$ i $0,1$

Badania eksperymentalne miały na celu sprawdzenie, który z detektorów najdokładniej wykryje ślady pozostawione na futrze, jak również inne punkty charakterystyczne osobników takie jak: oczy, uszy, nasadę ogona, łapy. Wszystkie narożniki odnalezione przez detektory zostały przefiltrowane na podstawie położenia; jedynie punkty znajdujące się wewnątrz obiektów były dalej analizowane. Za miarę jakości detekcji punktu uznano jego ranking. Im niższa wartość rankingu, tym łatwiejsza detekcja punktu (wyższa miara R – wzór 7.1 dla algorytmu Harris'a, wzór 7.2 dla algorytmu Shi-Tomasi, liczba punktów spełniających warunek 7.3 algorytmu FAST).

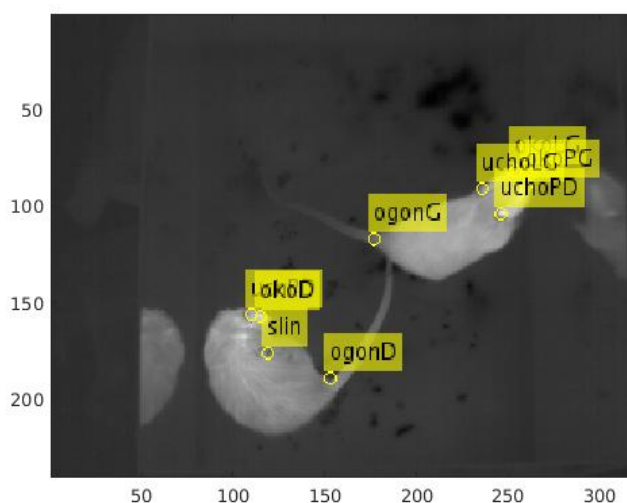
Tabela 7.1 przedstawia rankingi wykrytych punktów dla 3 różnych klatek (nie są to kolejne klatki) oraz 3 algorytmów. Ranking śladów śliny dla algorytmu Harris'a to największa lub jedna z największych wartości, oznacza to, że odpowiednio 27, 15 i 10 innych punktów uważanych jest za bardziej charakterystyczne. Dla klatki pierwszej różnica między największą wartością rankingu śladu śliny (algorytm Harris'a) a najmniejszą (algorytm FAST) wynosi 14 punktów. Dla klatki 3. algorytmy Shi-Tomasi oraz FAST klasyfikują punkt po ugryzieniu wśród pierwszych pięciu najbardziej charakterystycznych punktów, natomiast algorytm Harris'a – wśród pierwszych dziesięciu.

TABELA 7.1. Średnie rankingi dla punktów charakterystycznych - osobno naturalnych pkt. ciała i śladów śliny - dla trzech klatek i trzech algorytmów

	naturalne pkt. ciała			ślady śliny		
	Klatka 1	Klatka 2	Klatka 3	Klatka 1	Klatka 2	Klatka 3
Harris	$7,5 \pm 5,1$	$7,0 \pm 4,5$	$9,8 \pm 10,6$	27	15	10
Shi-Tomasi	$8,0 \pm 7,0$	$8,7 \pm 5,6$	$11,5 \pm 10,9$	21	12	5
FAST	$10,6 \pm 13,5$	$11,9 \pm 12,8$	$36,25 \pm 79$	13	15	5

Wśród punktów o wyższym rankingu znajdują się inne punkty opisujące ułożenie obiektów. Oprócz punktu pozostawionego przez ślinę gryzonia wykryte zostały również oczy, uszy i nasady ogonów. Wartości przedstawione w tabeli 7.1 dla naturalnych punktów ciała to średnie wartości rankingów wszystkich pozostałych punktów charakterystycznych będących naturalnymi punktami ciała przedstawionych na rys. 7.1.





RYS. 7.1. Wynik detekcji punktów charakterystycznych detektorem Harris'a

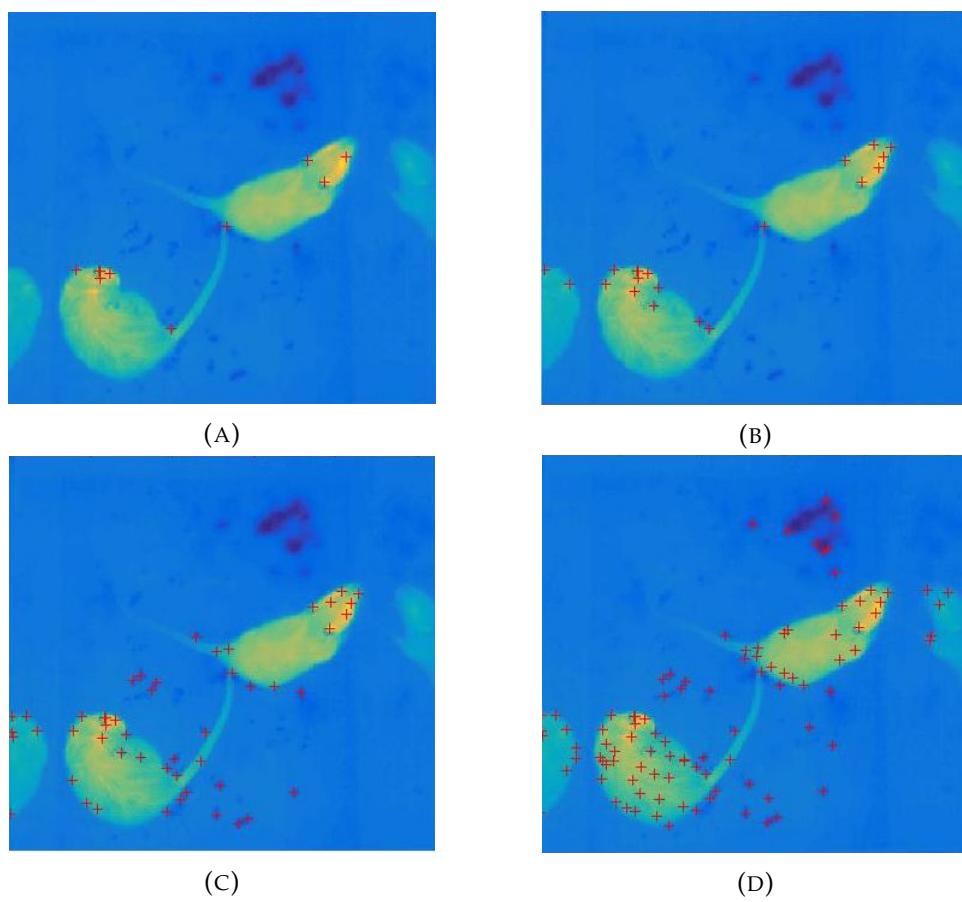
Widoczne jest to, że wyniki dla poszczególnych algorytmów są do siebie zbliżone. Jednak bardziej szczegółowa analiza wykazuje, iż algorytm FAST osiąga wyjątkowo niskie rankingi dla większości punktów. Istnieją jednak takie punkty, które uzyskują bardzo wysokie rankingi tym samym podwyższając średnią (największe odchylenie standardowe jest zauważalne dla algorytmu FAST). Dla klatki przedstawionej na rys. 7.1 takim problematycznym punktem była nasada ogona osobnika znajdującego się wyżej (etykieta „ogonG”). Algorytm FAST określił go jako punkt o średnim numerze równym 245, a algorytmy Harris i Shi-Tomasi odpowiednio jako 37 i 38. Ta cecha algorytmu FAST działa na niekorzyść analizy, ponieważ liczba odszukanych punktów zależna jest od jednego z ustawianych parametrów. W przypadku nieprawidłowo dobranych ustawień punkty o wyższych rankingach zostają pominięte w końcowej analizie. Rozkład narożników w zależności od ich rankingów został przedstawiony na rys 7.2.

Algorytm Shi-Tomasi osiągnął wyniki bardzo zbliżone, tylko nieznacznie gorsze niż algorytm Harris. Specyfika detekcji obu tych algorytmów była bardzo podobna: takie same narożniki odnajdywane w tej samej kolejności. Jediną różnicą były wartości miar wynikające ze wzorów 7.1 i 7.2. Z tego powodu w dalszych badaniach algorytm Shi-Tomasi był pomijany.

Szereg wyników przedstawiających analizę porównawczą detektorów opublikowano w pracy [90].

7.3. Analiza detekcji śladów śliny

Osobno została przeprowadzona analiza śladów śliny, które powstały w wyniku czyszczenia się lub ugryzienia. W rozdziale tym zostaną przedstawione analizy śladów. Na początku wykonano ogólną analizę wszystkich widocznych śladów śliny zarejestrowanych podczas testów nie uwzględniając zmiany temperatury w czasie (rozdział 7.3.1). Następnie porównano jakość detekcji różnych obserwowanych śladów na tych samych obrazach (rozdział 7.3.2). Kolejnym krokiem było zdefiniowanie różnic pomiędzy śladami będącymi skutkiem zachowań agresywnych, a śladami pozostawionymi po czyszczeniu (rozdział 7.3.3). Ostatni punkt analizy



RYS. 7.2. Rozkład punktów odnalezionych przez detektor Harris'a dla a) 10, b) 20, c) 50, d) 100 narożników o najwyższym rankingu

obejmował zbadanie wpływu parametrów detektorów na jakość detekcji śladów na symulacjach (rozdział 7.3.4).

7.3.1. Ogólna analiza śladów śliny

Punkty odnalezione przez detektory nie zawsze pokrywały się z pikselem o najniższej temperaturze w otoczeniu. Zakładając, że najzimniejsze miejsce oznacza miejsce o największej ilości śliny, czyli środek ugryzienia, a temperatura otaczających go pikseli stopniowo rośnie wraz ze wzrostem odległości, wprowadzono dodatkowy punkt $Corner_{LOW}$ umieszczony w miejscu piksela o najniższej temperaturze w otoczeniu 5 pikseli od punktu wyznaczonego przez detektor. Temperatura punktu $Corner_{LOW}$ oznaczana jest przez T_{cLow} , natomiast temperatura odnalezionego narożnika przez T_{cp} .

Statystyczna analiza śladów śliny została przedstawiona w tabeli 7.2. Ranking oraz wartość R zostały obliczone dla algorytmu Harris'a.

TABELA 7.2. Analiza statystyczna śladów śliny

	Wartość średnia	Odchylenie standardowe	Wartość minimalna	Wartość maksymalna
Temperatura punktu narożnika - T_{cp} [K]	300,71	1,24	298,12	304,84
Temperatura punktu $Corner_{LOW}$ - T_{cLow} [K]	299,83	1,02	297,32	301,90
Wartość R	2,33	2,6	0,10	12,31
Ranking	20,65	14,43	1	70

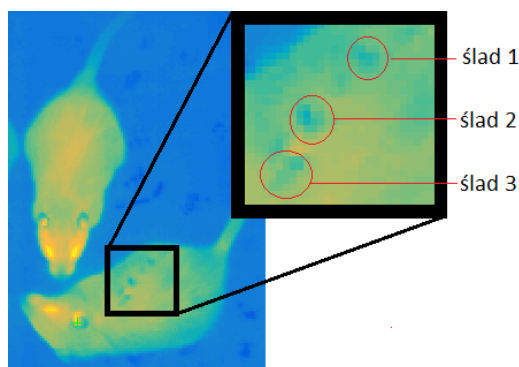
Średnia wartość temperatury wszystkich wykrytych śladów śliny wyniosła 300,71 K, podczas gdy temperatura ciała szczura wynosiła średnio 301,62 K – w zależności od osobnika [90]. Średnia temperatura punktu $Corner_{LOW}$ to 299,83 K. Wyniki wskazują, że temperatura wykrytych punktów śladów śliny jest o prawie 2 K niższa niż średnia temperatura ciała. Ranking oraz wartość R charakteryzują się dużym rozrzutem (odchylenie standardowe wynosi odpowiednio 14,43 i 2,6). Te parametry są silnie zależne od ułożenia ciała obiektu. Niewłaściwe ułożenie może spowodować częściowe lub całkowite zasłonięcie punktu zainteresowania. Ranking zależy również od detekcji pozostałych punktów obrazu. Punkty dające się lepiej odróżnić niż ślad śliny to w większości: oczy, których temperatura jest wyższa niż temperatura ciała, nos i końcówki uszu, które z kolei mają charakterystyczne, względnie stałe ułożenie na ciele. Innym czynnikiem wpływającym na wartość R i ranking jest czas.

Szereg zaprezentowanych wyników badań opublikowano w pracy [94].

7.3.2. Porównanie detekcji różnych punktów na tym samym obrazie

Zmiany temperatury punktu w czasie mają wpływ na jakość detekcji tego samego punktu, jak zostanie to przedstawione w rozdziale 7.4. Jednak różne ślady śliny mogą osiągać różne rankingi w detekcji, od wartości 1 aż do 70 (tabela 7.2) w tym samym czasie. Poniżej przedstawiono analizę obrazu zawierającego 3 ślady po ugryzieniu (rys. 7.3) w ciągu 11 klatek. Kolor niebieski oznacza temperaturę niższą, kolor żółty – wyższą. Ślad 1 znajduje się po prawej stronie kłębca gryzonia, jest

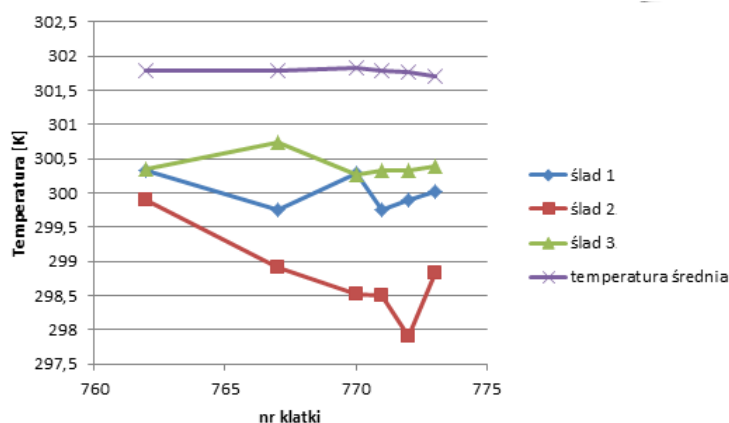
oddalony od pozostałych śladów, ma kształt zbliżony do owalnego. Ślad 2 jest najlepiej widoczny ze wszystkich śladów, znajduje się pomiędzy pozostałymi dwoma śladami w niewielkiej odległości od śladu 3, ma kształt zbliżony do owalnego. Ślad 3 ma kształt podłużny, jest słabo widoczny, jego górny lewy róg praktycznie łączy się z innym, nieanalizowanym śladem. Pomiedzy 2. a 3. śladem znajduje się miejsce o wyższej temperaturze (kolor żółty), które mogło powstać w wyniku ugryzienia i przesunięcia futra przez napastnika odsłaniając tym samym cieplejszą skórę.



RYS. 7.3. Obraz trzech różnych analizowanych śladów

Rys. 7.4 przedstawia wykresy średnich temperatur dla wszystkich śladów oraz dla średniej temperatury ciała obiektu. Temperatury określone zostały dla punktu $Corner_{LOW}$ (piksel o najmniejszej temperaturze znajdujący się w otoczeniu 5 od punktu wyznaczonego przed detektor narożników), ponieważ najwyższe wartości parametru R w okolicach punktu 3 osiągał punkt w górnej granicy śladu ze względu na bliskość pikseli o wysokich temperaturach (duża wartość gradientu).

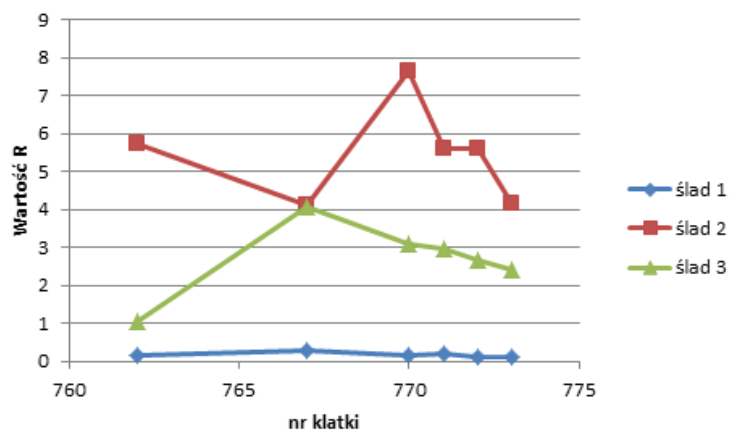
Ślad 2 ma widocznie niższą temperaturę niż pozostałe ślady, co jest przyczyną jego dobrej widoczności na obrazie. Temperatury śladów 1 i 3 są zbliżone.



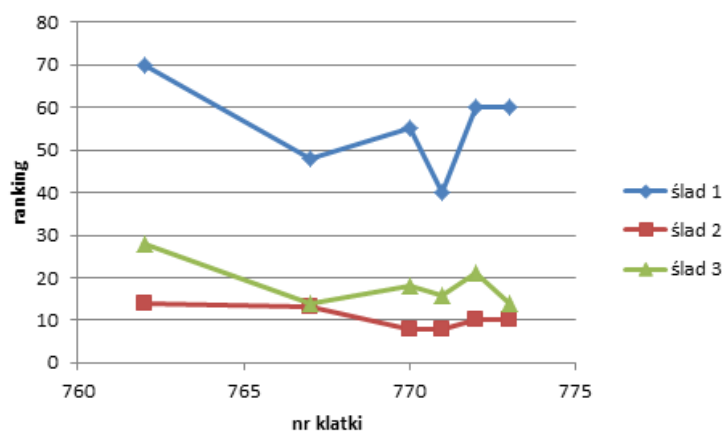
RYS. 7.4. Temperatura T_{cLow} śladów oraz średnia temperatura obiektu

Rys. 7.5 i 7.6 przedstawiają wartości R oraz rankingi uzyskane przez ślady podczas detekcji punktów charakterystycznych detektorem Harris'a.

Wykresy wyraźnie przedstawiają dużą różnicę w wynikach dla śladu 1 w porównaniu do śladów 2 i 3. Ślad 1 uzyskał najniższą wartość parametru R (średnio 0,17) oraz najwyższą wartość rankingu (średnio 55,5). Przyczyną tego nie jest zbyt wysoka wartość temperatury, która dla tego śladu wynosiła mniej niż dla śladu 3.



RYS. 7.5. Wartości R śladów

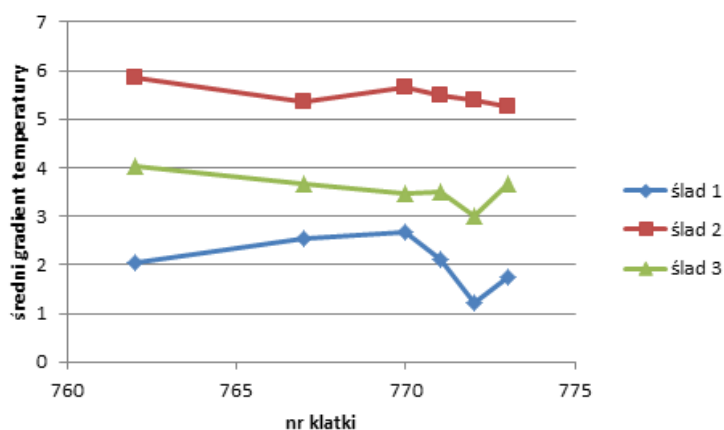


RYS. 7.6. Rankingi śladów

Ślad 3 uzyskał ranking niewiele większy (średnia 18,5) niż ranking śladu 2 (średnio 10,5), jak również zbliżone wartości R (średnia dla śladu 3 = 2,71 oraz dla śladu 2 = 5,49).

Rys. 7.7 przedstawia średni gradient temperatury w obszarze śladu dla sąsiedztwa 5 pikseli. Ślad 1 charakteryzuje się najniższym gradientem temperatury, ponieważ jest dość duży, a jego granice są rozmyte. Ślad 2, którego kształt jest zbliżony do kształtu śladu 1 osiągnął najlepszy ranking, ze względu na wysoki gradient temperatury. Niska temperatura śladu oraz wysoka temperatura otoczenia powodują również jego dobrą widoczność na rys. 7.3. Gradient temperatury śladu 3 plasował się pośrodku pozostałych dwóch gradientów, co odpowiada uzyskanemu (również średniemu) wynikowi parametru R, mimo, że temperatura śladu 3 była najbardziej zbliżona do temperatury średniej zwierzęcia. Na zwiększenie gradientu temperatury śladu 3 prawdopodobnie duży wpływ miało cieplejsze miejsce powyżej śladu.

Przedstawione wyniki wskazują cechy istotne podczas pomiaru jakości detekcji. Większy wpływ na wartości miar narożników będzie miało otoczenie śladu niż jego temperatura. Znaczenie mają także same granice punktu oraz inne obszary znajdujące się w bliskim otoczeniu.

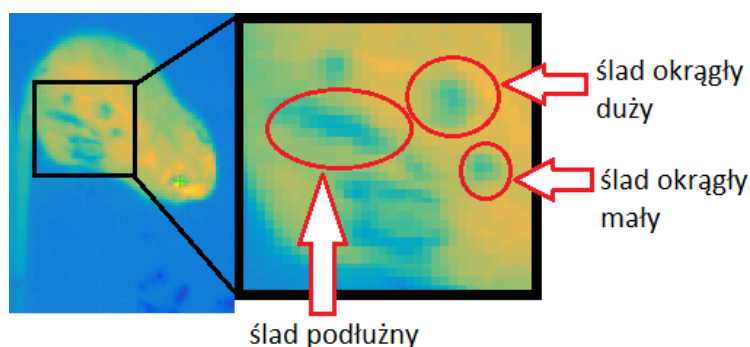


RYS. 7.7. Średni gradient temperatury w obszarze śladów

7.3.3. Porównanie cech śladów pozostawionych przez ugryzienie oraz grooming

Mokre ślady na futrze, poza ugryzieniem, może spowodować także czyszczenie się gryzonia (tzw. grooming). Grooming to zachowanie, które składa się z szeregu czynności pielęgnacyjnych, jest czynnością wrodzoną. Szczegółowy opis groomingu szczurów został zamieszczony w [140] oraz [11]. Modelowy przebieg groomingu polega na myciu pyska oraz okolic uszu przednimi łapami, a następnie pielęgnacji boków ciała m.in. poprzez lizanie. Badania wykazały, że grooming może również odzwierciedlać stan emocjonalny zwierzęcia. Wymuszony groomig może być wynikiem strachu, a jego wykonywanie ma na celu rozładowanie stresu [71], [149].

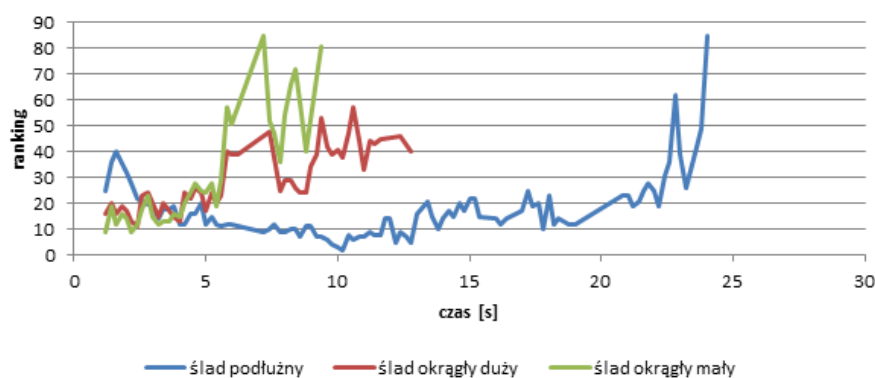
Grooming może być w łatwy sposób automatycznie odróżniony od innych zachowań ze względu na swój indywidualny charakter. Kształt ciała zwierzęcia podczas czyszczenia się jest owalny (posiada współczynnik cyrkularności, zwany także zmodyfikowanym współczynnikiem Malinowskiej, bliski 0), zwierzę porusza się, co powoduje delikatne fluktuacje jego środka ciężkości w przestrzeni i czasie. Grooming prawie nigdy nie zachodzi podczas kontaktu z innym osobnikiem. Również ślady na futrze pozostawione po czyszczeniu są inne od tych pozostawionych przez ugryzienie, są liczne i różnorodne, często podłużne. Rys. 7.8 przedstawia przykładowe ślady pozostawione po czyszczeniu futra z trzema wyodrębnionymi miejscami, dla których wykonana została analiza. Do analizy wybrano dwa ślady okrągłe o różnej wielkości oraz jeden ślad podłużny.



RYS. 7.8. Ślady śliny pozostawione po groomingu

Wyniki zmian rankingów śladów w czasie zostały przedstawione na rys. 7.9, a zmian temperatury na rys. 7.10. Oba wykresy przedstawiają również czas widoczności każdego śladu przez detektor. Ślad podłużny był najdłużej rozpoznawalny (przez 24 sekundy), co wynika z jego wielkości, kształtu oraz temperatury (rys. 7.8). Ślad ten miał najniższą temperaturę oraz najwolniej wysychał, może to świadczyć o największej ilości śliny naniesionej w to miejsce. Wysokie wartości rankingu oraz temperatury w ciągu pierwszych 2 sekund związane były z utrudnioną detekcją ze względu na liczne powierzchniowe ślady śliny w tym miejscu o temperaturze wyższej niż ślad analizowany, które zanikły po krótkim czasie. Średni ranking dla tego śladu wyniósł 17,6. Powierzchnia śladu stopniowo zmniejszała się w ciągu 24 s. (zachodziły duże zmiany w długości, mniejsze w szerokości), jednak nie wpłynęło to znacząco na jakość detekcji aż do momentu, gdy wartość temperatury wykrytego punktu zbliżyła się do średniej temperatury ciała, a cały ślad zaczął zanikać.

Temperatury śladów okrągłych na początku detekcji były tylko nieznacznie wyższe od temperatury śladu podłużnego, ale szybko rosły, aby w 7 s. detekcji osiągnąć temperaturę równą średniej temperaturze ciała. Szybki wzrost temperatury świadczy o tym, że w tych miejscach warstwa śliny była dużo mniejsza niż w przypadku śladu podłużnego. Detekcja małego śladu okrągłego osiągnęła najwyższy ranking średni (32,8) i zakończyła się po 9 s. Ślad okrągły większy przestał być widoczny po 13 s. ze średnim rankingiem 29,7. Możliwość detekcji śladów okrągłych o temperaturze większej od średniej temperatury ciała wynika z nierównomiernego rozkładu temperatury gryzonia co jest widoczne na rys. 7.8. Prawa strona ciała zwierzęcia (ta, którą zwierzę czyściło) ma wyższą temperaturę (jaśniejszy kolor) niż strona lewa. Przyczyny tego zjawiska mogą być dwie. Pierwsza z nich to fakt, iż mokre futro skleja się w kępki odsłaniając tym samym skórę, która ma wyższą temperaturę niż powierzchnia futra. Drugim wytłumaczeniem może być lepsze ukrwienie czyszczonej części ciała.

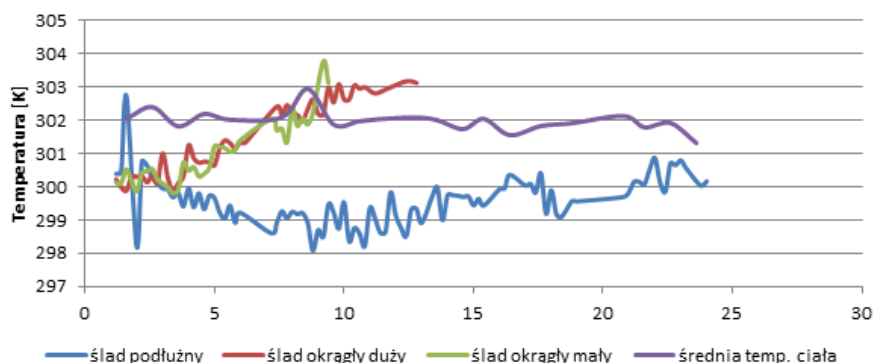


RYS. 7.9. Zmiany rankingu śladów w czasie

Duże oscylacje wartości rankingu i temperatury śladów na rys. 7.9 oraz 7.10 są wynikiem poruszania się obiektu oraz zmian ułożenia jego ciała.

Warto zauważyć, że przez kilka ostatnich klatek detekcji ślady były jeszcze rozpoznawane przez detektor (z dużymi rankingami), a już całkowicie niemożliwe do dostrzeżenia przez oko człowieka (niezależnie od zastosowanej mapy kolorów).

Rys. 7.9 oraz 7.10 wskazują, że istnieją dwa rodzaje śladów pozostawionych po czyszczeniu się. Pierwsze to ślady podłużne z dużą ilością śliny pozostawione po czyszczeniu (lizaniu) futra językiem. Drugie to punktowe ślady okrągłe o znacznie mniejszej ilości śliny oraz krótszym czasie wysychania.



RYS. 7.10. Zmiany temperatury wykrytych punktów śladów oraz średniej temperatury ciała w czasie

Rozróżnienie pochodzenia śladów na podstawie jedynie ich obserwacji oraz zmian temperatury jest możliwe. Tabela 7.3 przedstawia porównanie cech w zależności od przyczyny powstania. Ślady po groomingu są liczne i utrzymują się na futrze przez dłuższy okres czasu niż te po ugryzieniu, niezależnie czy są to ślady podłużne, które mają najdłuższy czas wysychania, czy okrągłe. Ślady pozostawione po ugryzieniu są krótkotrwałe. Na czas ich detekcji wpływa także to, że powstają one zaraz po lub jeszcze w trakcie walki zwierząt (a co najmniej ich kontaktu) podczas ich szybkich ruchów i często są zasłaniane przez drugiego osobnika, co jeszcze bardziej skraca możliwy czas detekcji. Różnica czasów wysychania wynika z samego celu i charakteru obu zdarzeń. Czyszczenie się ma za zadanie nanieść ślinę na futro, aby je zwilżyć, ślina pozostawiona po ugryzieniu jest tam pozostawiona przypadkiem, w małej ilości, po złapaniu fałdów skóry zębami, nie ma na celu czyszczenia futra.

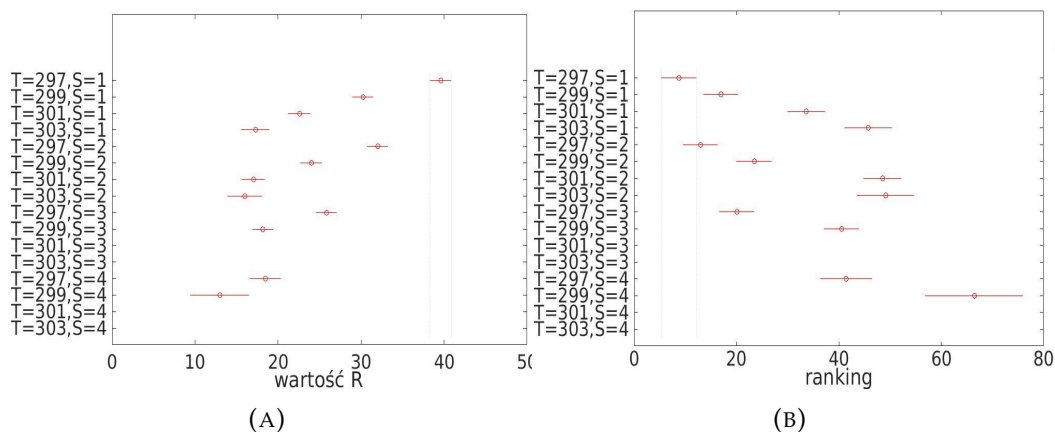
TABELA 7.3. Tabela porównania cech śladów po groomingu i ugryzieniu

	Ślady po ugryzieniu	Ślady po groomingu
Liczność	pojedyncze	liczne
Wielkość	punktowe (kilka pikseli)	duże
Kształt	okrągłe	okrągłe i podłużne
Zmiany temperatury	szybkie	powolne

7.3.4. Analiza wpływu detektora na detekcję - badania symulacyjne

W celu zbadania wpływu detektora na detekcję śladów śliny przeprowadzono analizę wyników detekcji algorytmów Harris i FAST na zasymulowanych śladach opisanych w rozdziale 6.4. Wygenerowane zostały ślady o 4 różnych temperaturach (297, 299, 301 i 303 K) oraz 4 wielkościach (1, 2, 3 i 4 piksele). Wyniki dla kształtów podłużnych okazały się być jednakowe z wynikami dla kształtów okrągłych o średnicach odpowiadających szerokościom kształtów podłużnych, dlatego ślady o kształtach podłużnych zostały pominięte z dalszej analizy.

Dla detektora Harris zastosowano 3 różne wielkości filtra: 3x3, 5x5 i 7x7 pikseli. Jako parametry ilościowe opisujące detekcję przyjęto ranking oraz wartości R obliczane na podstawie wzoru 7.1 dla algorytmu Harris i 7.3 dla FAST.



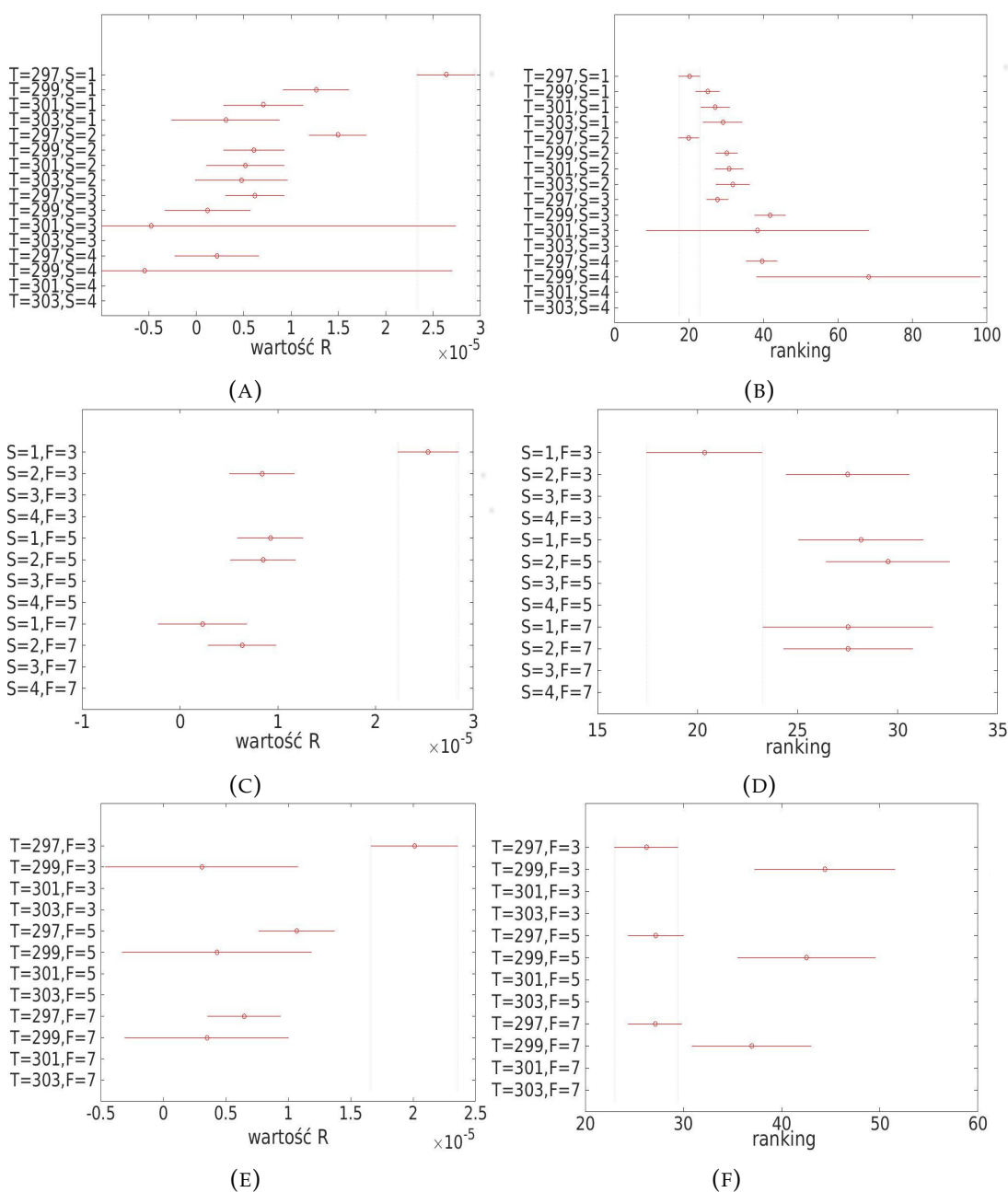
RYS. 7.11. Wyniki analizy detektorem FAST w zależności od temperatury (T) i wielkości śladu (S); a) wartość parametru R, b) ranking

Symulacje wykonano dla 400 różnych obrazów w 16 konfiguracjach kształtu i temperatury śladów (4 temperatury \times 4 wielkości śladu).

Rys. 7.11 i 7.12 przedstawiają testy porównań wielokrotnych. Każda średnia grupowa jest reprezentowana przez okrąg, a linia reprezentuje przedziały ufności dla poziomu istotności $\alpha = 0,05$. Dwie średnie grupowe są statystycznie różne, jeśli ich przedziały są rozłączne. Wyniki detekcji w postaci wartości R i rankingu w zależności od temperatury (T) i wielkości śladu (S) dla detektora FAST przedstawiono na rys. 7.11. Widoczna jest zależność wartości R od temperatury oraz wielkości śladu (rys. 7.11a). Dla śladów o jednakowej wielkości temperatura jest czynnikiem decydującym o wyniku detekcji: im niższa temperatura, tym większa wartość R. Analogicznie jest w przypadku śladów o jednakowej temperaturze: im mniejszy ślad, tym większa wartość R. Ślady o dużym rozmiarze i wysokiej temperaturze nie zostały odnalezione. Wyniki tych zależności widoczne są również dla rankingu (rys. 7.11b): im mniejszy i chłodniejszy ślad, tym niższy ranking. W tym przypadku jednak, średnie grupowe nie są tak często statystycznie od siebie różne jak w przypadku wartości R.

Wykresy na rys. 7.12 dodatkowo uwzględniają różne wielkości filtrów (F) podczas zastosowania detektora Harris. Podobnie jak algorytm FAST, Harris również nie wykrył śladów o wysokiej temperaturze i dużym rozmiarze. Niewielka liczba pomiarów dla niektórych grup (T=301 i S=3 lub T=299 i S=4) powodują duże zakresy przedziałów ufności. Parametr R osiąga najwyższe wartości dla najmniejszego śladu o najniższej temperaturze podczas detekcji filtrem o najmniejszych wymiarach (rys. 7.12a, 7.12c i 7.12e). Średnia dla tej grupy jest statystycznie różna od innych grup. Średnie pozostałych grup nie różnią się tak znacznie od siebie, widoczna jednak jest zależność detekcji od rozpatrywanych parametrów dla śladów, które zostały odnalezione. Wartości rankingów na rys. 7.12b i 7.12d prezentują prawie dokładnie takie same zależności jak odpowiadające im wartości R. Wykres 7.12f prezentuje wyraźne zależności wyników rankingów od temperatury, które nie były widoczne dla wartości R (rys. 7.12e).

Wszystkie wykresy przedstawiają zależności pomiędzy dwoma czynnikami. Wyniki trójczynnikowej analizy ANOVA dla poziomu istotności $<0,095$ zawarte są w tabeli 7.4.



RYS. 7.12. Wyniki analizy detektorem Harris w zależności od temperatury (T), wielkości śladu (S) oraz wielkości filtra (F); a) wartość parametru R w zależności od temperatury i wielkości śladu, b) ranking w zależności od temperatury i wielkości śladu, c) wartość parametru R w zależności od wielkości śladu i filtra, d) ranking w zależności od wielkości śladu i filtra, e) wartość parametru R w zależności od temperatury i wielkości filtra, f) ranking w zależności od temperatury i wielkości filtra

Wynik analizy wartości R pozwala odrzucić hipotezę o braku wpływu temperatury oraz wielkości śladu na wartość R zarówno w przypadku zastosowania detektora Harris (poziomy istotności odpowiednio 0,0018 i 0,0001) jak i FAST (poziomy istotności = 0). Również wyniki dla rankingów sugerują odrzucenie powyższej hipotezy. Rozmiar zastosowanego filtra nie ma wpływu na wartość R ($p=0,7081$), ma natomiast wpływ na ranking ($p=0,0134$). Nie stwierdzono również współdziałania wszystkich czynników na ranking ($p=0,9852$). Wysoce istotne jest natomiast współdziałanie czynników: temperatury z wielkością śladu, temperatury z rozmiarem filtra oraz wielkości śladu z rozmiarem filtra.

TABELA 7.4. Wyniki trójczynnikowej analizy ANOVA dla wartości R oraz rankingów w postaci poziomu prawdopodobieństwa p

	Harris		FAST	
	wartość R	ranking	wartość R	ranking
temperatura	0,0018	0	0	0
wielkość śladu	0,0001	0,0087	0	0
rozmiar filtra	0,7081	0,0025	-	-
temperatura * wielkość śladu	0,0004	0,0134	0	0
temperatura * rozmiar filtra	0	0,0066	-	-
wielkość śladu * rozmiar filtra	0	0	-	-
temperatura * wielkość śladu * rozmiar filtra	0	0,9852	-	-

Oba testowane detektory nie poradziły sobie z detekcją dużych śladów ($S=3$ i $S=4$) o wysokiej temperaturze ($T=301$ i $T=303$). Takie temperatury były zbliżone do temperatury ciała zwierząt, co zmniejszało widoczność śladów. Zbyt duży rozmiar śladu działał w tym momencie na niekorzyść detekcji, ponieważ dodatkowo zmniejszył gradient temperatury w okolicy środka śladu. Wyniki dla różnych rozmiarów filtra detektora Harris były zbliżone. W większości przypadków najmniejszy rozmiar filtra okazał się być najlepszym rozwiązaniem.

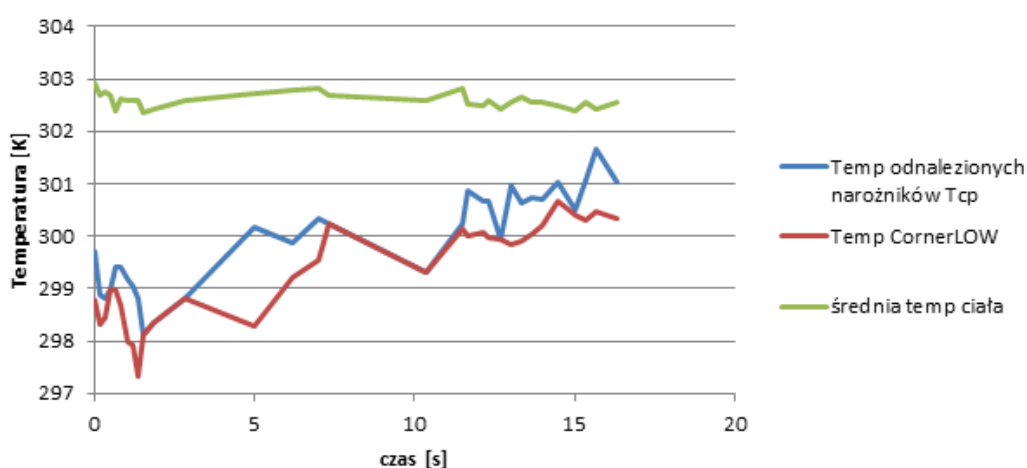
7.4. Dynamiczna analiza detekcji śladów śliny

Punkty charakterystyczne zwierząt o temperaturze różnej od powierzchni ciała, takie jak np. oczy czy końcówki uszu, mają względnie stałą temperaturę. Ewentualne zmiany wywoływane są przez procesy fizjologiczne (np. termoregulacja) lub czynniki zewnętrzne (np. strach czy podanie leków) i nie są one przedmiotem badań tej pracy. Ślady śliny, tak samo jak i inne przypadkowe ślady naniesione na futro i widoczne jedynie w zakresie częstotliwości kamery termowizyjnej (np. pozostawione po dotknięciu poidelka), z czasem wysychają. Zmiany te zaobserwowano na nagraniach rzeczywistych, a następnie zbadano ich charakter na przykładzie kilku przypadków. Wyniki zostały przedstawione w rozdziale 7.4.1. Z uwagi na to, że przykładów było niewiele, dodatkowo nie wszystkie ślady były widoczne podczas całego procesu wysychania, przygotowano symulację procesów zachodzących podczas obserwacji śladów śliny opisaną w rozdziale 6.4.2. Następnie, w rozdziale

7.4.2 dokonano analizy zasymulowanych śladów i określono optymalne warunki detekcji.

7.4.1. Analiza zmian temperatury na obrazach rzeczywistych

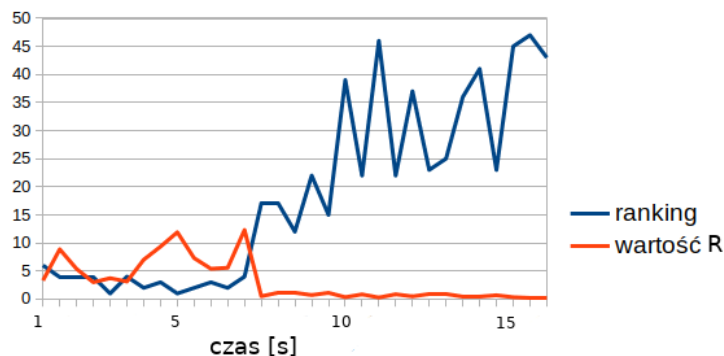
Niewątpliwym i bardzo ważnym czynnikiem w analizie ugryzień jest czas. Ślina pozostawiona na futrze wysycha i zmienia swoją temperaturę, co wstępnie zostało przedstawione w rozdziale 7.3.3 na rys. 7.10. Cały proces wysychania najlepiej widoczny jest dla śladu podłużnego (linia niebieska), który był największy (największa była także ilość śliny) i zmiany jego temperatury zachodziły najwolniej. Wyraźnie widać także, że temperatura najpierw obniżała się, aby w około 9 sekundzie zacząć powoli rosnać. Rys. 7.13 przedstawia dokładniejszą zależność temperatury śladu od czasu - linia niebieska. Obliczono również temperaturę dla najzimniejszego punktu w okolicy odnalezionego narożnika - linia czerwona. Dla porównania, zieloną linią zaznaczono średnią temperaturę ciała zwierzęcia. Temperatury T_{cp} oraz T_{cLow} rosną w czasie, podczas gdy średnia pozostaje względnie stała na poziomie 302,59 K z odchyleniem standardowym wynoszącym 0,14 K. Widoczny jest także nagły spadek temperatury w pierwszych sekundach rejestracji, dużo mniej dostrzegalny niż dla śladu podłużnego przedstawionego w rozdziale 7.3.3.



RYS. 7.13. Zmiany temperatury T_{cp} , T_{cLow} i temperatury średniej w czasie

Rys. 7.14 przedstawia zmiany rankingu oraz wartości R w czasie dla przykładowego śladu śliny dla algorytmu Harrisa. Ranking ma stabilne wartości przez pierwsze 5 s., które następnie zaczynają rosnać, aby ostatecznie osiągnąć wartości powyżej 40 po 15 s. Wartości rankingu silnie zależą od innych wykrytych punktów. Nagłe zmiany rankingu mogą być spowodowane przez np. tymczasowe zasłonięcie lub pojawienie się innych punktów. Bardziej obiektywnym parametrem jest wartość R (równanie 7.1, 7.2 lub 7.3), który wskazuje na spadek wartości R po 6 s., co jest zgodne ze wzrostem rankingu w tym czasie. Dane dla rys. 7.13 i rys. 7.14 zostały otrzymane z analizy jednego śladu śliny.

Podczas analizy śladów śliny popularnymi detektorami narożników należy zwrócić uwagę na zanikanie śladów w czasie. Względnie niski wzrost temperatury skutkuje znacznym wzrostem rankingu i spadkiem wartości R. Stopniowa zmiana temperatury jest cechą przydatną podczas śledzenia śladu śliny, np. po zasłonięciu go. Ta dynamiczna właściwość może być użyta do automatycznego rozróżniania punktów śladów śliny od innych punktów charakterystycznych otrzymanych jako wynik



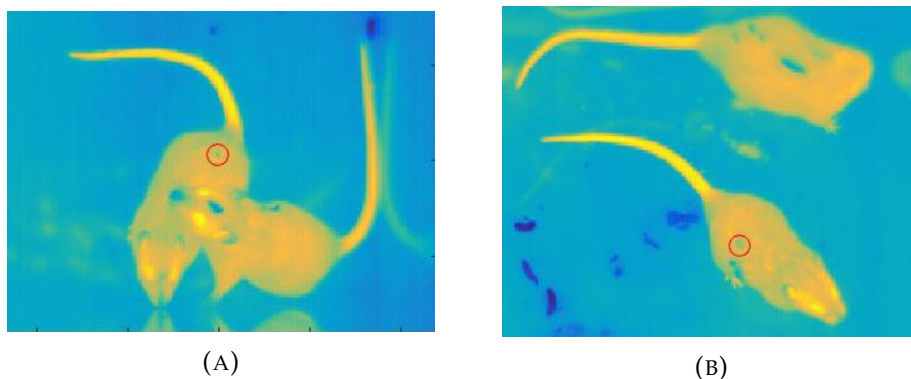
RYS. 7.14. Zmiany rankingu i wartości R w czasie

detekcji narożników.

Szereg zaprezentowanych wyników badań opublikowano w pracy [94].

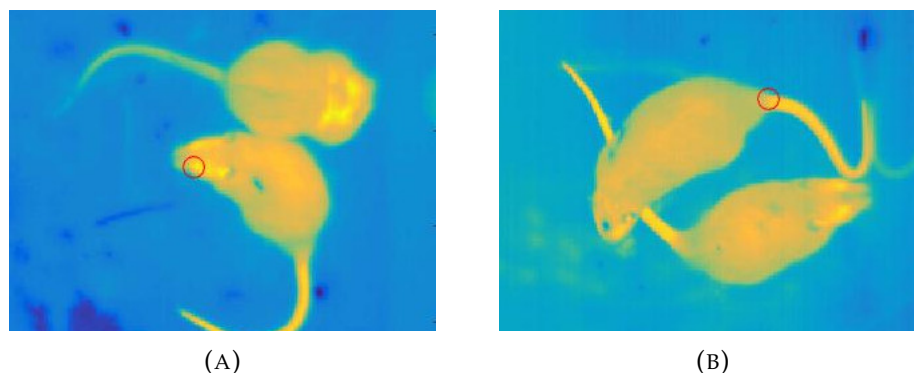
W celu analizy zmian temperatury śladów wyodrębniono kilka interesujących zdarzeń interakcji między zwierzętami na nagraniach wraz z odpowiednimi punktami zainteresowania znajdującymi się na ciele osobników. Niektóre ślady zostały przedstawione na rys. 7.15 i 7.16. Rozpatrywanymi śladami były:

1. Ślad 1. ($Trace_1$) - mały okrągły ślad śliny o niskiej temperaturze i długim czasie powrotu temperatury do stanu początkowego, czas trwania detekcji = 141,4 s. (rys. 7.15a);
2. Ślad 2. ($Trace_2$) - ślad śliny o czasie trwania detekcji = 26 s. (rys. 7.15b) ;
3. Ślad 3. ($Trace_3$) - krótki ślad śliny całkowicie zasłonięty podczas procesu ochładzania, czas trwania detekcji = 14,5 s.;
4. Ślad 4. ($Trace_4$) - mały i krótki ślad śliny zasłonięty podczas procesu wzrostu temperatury, czas trwania detekcji = 6,3 s.;
5. Ślad 5. ($Trace_5$) - ślad śliny częściowo zasłonięty podczas procesu wzrostu temperatury, czas trwania detekcji = 3 s.;



RYS. 7.15. Obrazy przykładowych punktów śladów śliny wybranych do analizy (czerwone okręgi) przedstawiające a) ślad $Trace_1$, b) ślad $Trace_2$

6. Punkt referencyjny 1. ($RefP_1$) - punkt referencyjny - oko, czas trwania detekcji = 31,1 s. (rys. 7.16a);



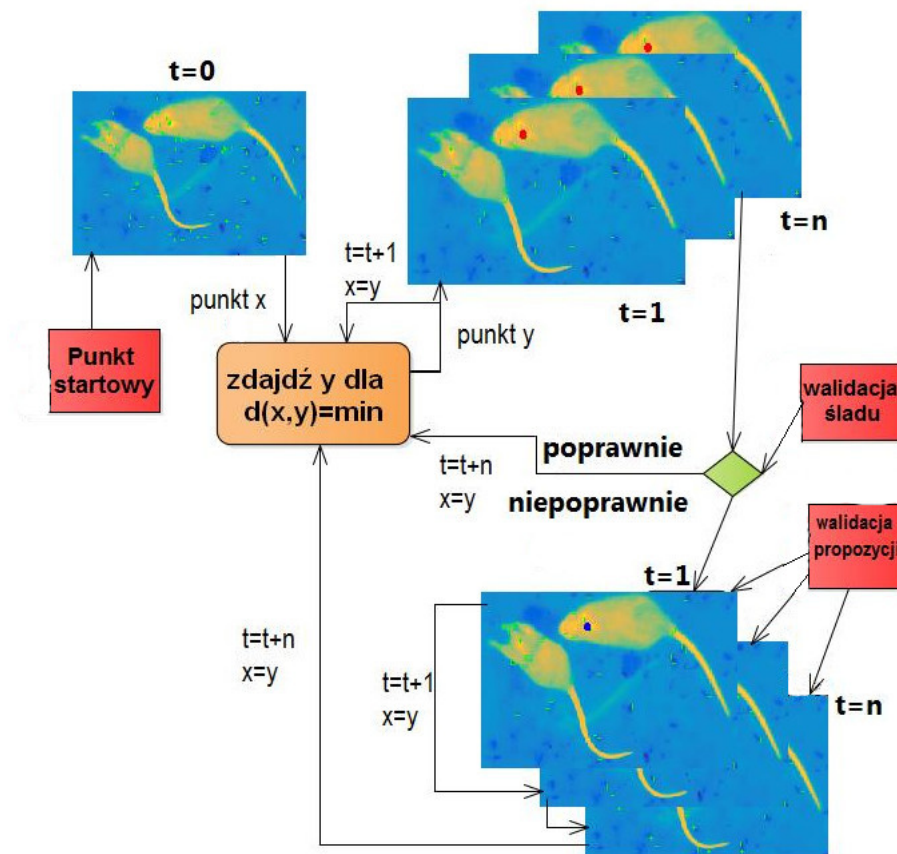
RYS. 7.16. Obrazy przykładowych punktów referencyjnych wybranych do analizy (czerwone okręgi) przedstawiające a) punkt referencyjny - oko $RefP_1$, b) punkt referencyjny - nasada ogona $RefP_3$

7. Punkt referencyjny 2. ($RefP_2$) - punkt referencyjny - nasada ogona, dane zebrane z tego samego nagrania co Ślad 1. dla drugiego osobnika, czas trwania detekcji = 115,5 s.;
8. Punkt referencyjny 3. ($RefP_3$) - punkt referencyjny - nasada ogona, dane zebrane z tego samego nagrania co Ślad 1. dla tego samego osobnika, na którym zarejestrowano Ślad 1., czas trwania detekcji = 141,4 s. (rys. 7.16b).

W celu rejestracji danych dla wybranych śladów w ciągu całego czasu ich trwania, stworzono algorytm półautomatycznego śledzenia wybranych punktów. Został on schematycznie przedstawiony na rys. 7.17. Wszystkie konieczne interwencje użytkownika zaznaczono sygnałami w kolorze czerwonym. W początkowej fazie śledzenia użytkownik ręcznie zaznacza punkt zainteresowania (punkt x) wśród wszystkich punktów wykrytych przez detektor narożników. Algorytm wykrywa w następnej klatce punkt o najmniejszej odległości wg. miary euklidesowej od punktu startowego (punkt y zaznaczony kolorem czerwonym na serii klatek), uaktualnia pozycję wyszukanego punktu ($x=y$), a następnie powtarza te kroki dla $n-1$ następnych klatek. Po detekcji punktu na wszystkich n klatkach użytkownik powinien sprawdzić poprawność śledzenia (walidacja śladu). Jeżeli punkt został poprawnie wykryty, algorytm przystępuje do analizy kolejnych klatek w dokładnie taki sam sposób traktując ostatni znaleziony punkt jako punkt początkowy. W przypadku stwierdzenia "zagubienia" punktu, algorytm powraca od pierwszej z n klatek i wykonuje detekcję jeszcze raz, ale tym razem każda klatka wymaga ręcznego zaakceptowania przez użytkownika zaproponowanego punktu (punkt niebieski- walidacja propozycji). Manualna walidacja jest zakończona po n klatkach, po której cała procedura zaczyna się od początku dla klatki $n+1$.

Parametrami, które można dostosowywać w celu sprawniejszego działania algorytmu są:

1. liczba automatycznie analizowanych klatek - n - w zależności od obrazu, można je zwiększać, gdy śledzenie automatyczne daje lepsze wyniki, lub zmniejszać, gdy śledzenie trzeba często korygować - dostosowanie tego parametru pozwala usprawnić procedurę i zmniejszyć czas poświęcony na analizie.
2. wielkość obszaru poszukiwania śledzonego punktu - N - wielkość obszaru sąsiedztwa, w którym algorytm wyszukuje nowe położenie punktu startowego



Rys. 7.17. Algorytm półautomatycznego śledzenia wybranego narożnika

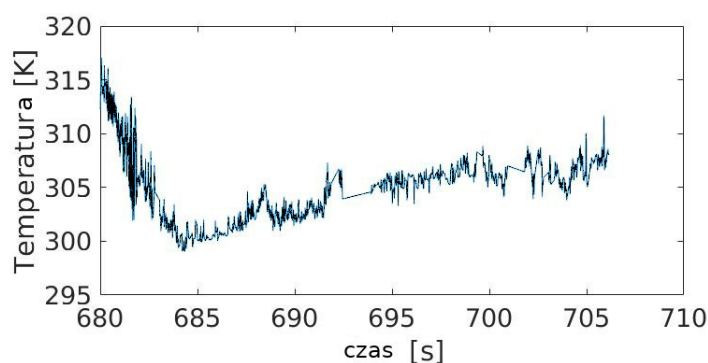
z poprzedniej klatki, zależna od przestrzennego położenia śladu. Parametr ten może zostać wprowadzony w przypadku, gdy ślad umiejscowiony jest niedaleko innych wyraźnych punktów (oczu, uszu, innych mokrych śladów na futrze lub krawędzi obiektu). Wielkość obszaru poszukiwania należy wtedy zmniejszyć, aby zapobiec błędnej detekcji i zmiany punktu śledzenia. Parametr ten wpływa na poprawność detekcji, optymalnie dobrana wielkość obszaru poszukiwania pozwala uniknąć korekcji manualnej.

3. maksymalna liczba punktów wykrytych przez detektor - k - optymalizacja tego parametru pozwala zaoszczędzić czas i zwiększyć dokładność detekcji. Dla każdego punktu wykrytego przez detektor obliczana jest miara R , a następnie punkty są ustawiane w rankingu według tej miary. Maksymalna liczba punktów wykrytych przez detektor jest to liczba k pierwszych punktów z listy rankingowej (o największych wartościach R). Jeżeli punkt jest wyraźny (występuje duży gradient temperatury w jego sąsiedztwie) - ma wysoką wartość R (niski ranking) w porównaniu z innymi wykrytymi narożnikami - wtedy maksymalna liczba punktów wykrytych przez detektor może być zmniejszona.

Analiza poszczególnych punktów była przeprowadzana po wcześniejszej optymalizacji parametrów. Parametry były dostosowywane, na podstawie obserwacji, zarówno dla każdego z punktów osobno (ze względu np. na miejsce ułożenia

ślądu), jak również do ogólnych warunków całego nagrania (np. do średniej temperatury obiektów i otoczenia).

Rys. 7.18 przedstawia wynik pomiaru śladu $Trace_2$. Temperatura śledzonego punktu została uśredniona w sąsiedztwie N4. Na początku można zaobserwować chłodzenie śladu, średnia temperatura zmniejsza swoją wartość z 315 do 300 K. Po krótkim czasie następuje ogrzanie śladu związane z wysychaniem śliny, temperatura stopniowo zwiększa swoją wartość do 305 K. Proces chłodzenia jest kilkukrotnie krótszy niż proces ogrzewania śladu. 15-stopniowy spadek temperatury dla śladu $Trace_2$ trwał około 4 sekund, podczas gdy ogrzanie śliny o 8 stopni zajęło 20 sekund.



RYS. 7.18. Średnia temperatura śladu $Trace_2$ w sąsiedztwie N4

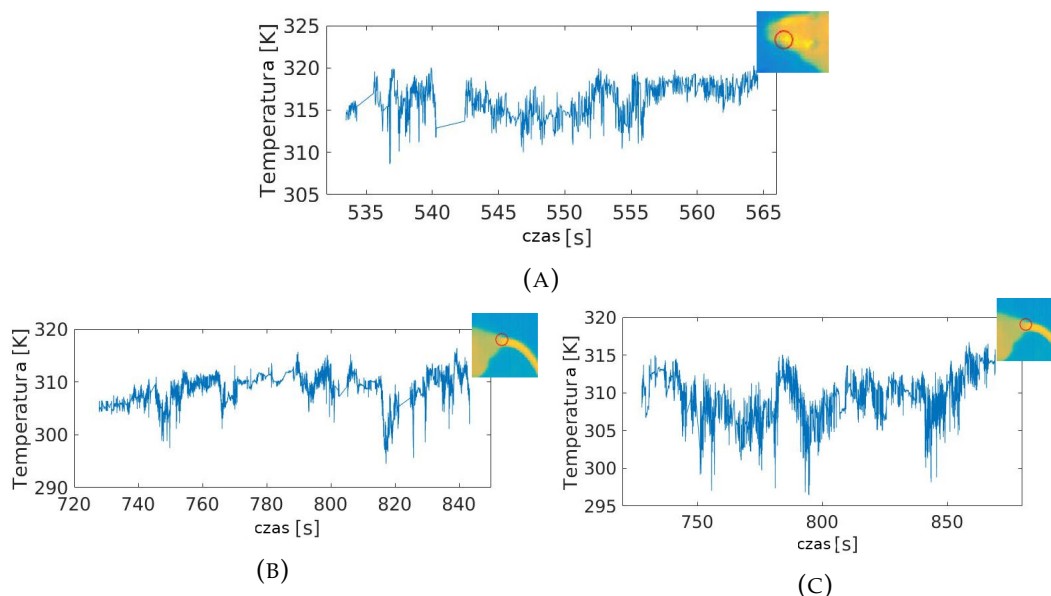
W celu porównania, wykonano również detekcję punktów referencyjnych: oka ($RefP_1$) oraz nasad ogonów dla różnych osobników na tym samym nagraniu ($RefP_2$ oraz $RefP_3$). Wykresy średnich temperatur w sąsiedztwie N4 dla punktów referencyjnych przedstawione są na rys. 7.19.

Nagłe zmiany temperatury na rys. 7.18 i 7.19 mogą wynikać z niedokładności detekcji lub zmian sąsiedztwa spowodowanych ruchem obserwowanego obiektu. Zmiany te są szczególnie widoczne podczas śledzenia punktów referencyjnych. Wszystkie punkty znajdujące się blisko granicy obiektu są narażone na duże zmiany średniej temperatury otoczenia w związku z ruchem ciała i faktem, że podczas przesunięcia się punktu w stronę granicy obiektu, jego sąsiedztwem N4 może stać się tło o znacznie niższej temperaturze. Punkty umiejscowione na nasadzie ogona ($RefP_2$ i $RefP_3$) znajdują się blisko granicy obiektu (rys. 7.16b).

Średnia temperatura punktu referencyjnego oka również wykazuje dużą zmienność (rys. 7.19a). Wynika to także z umiejscowienia punktu. W pobliżu oka oprócz granicy obiektu, znajdują się również inne punkty o ekstremalnych temperaturach (rys. 7.16a): punkty chłodne - nos, końcówki uszu i punkty ciepłe - wewnątrz ucha. Uśrednienie temperatury śledzonego punktu w określonym sąsiedztwie skutkuje, w wypadku analizy oka, nie tylko błędem spowodowanym niskimi wartościami tła, lecz także innymi, naturalnymi obszarami ciała. Punkt oka był też często gubiony przez algorytm, ponieważ pysk zwierzęcia nie zawsze był dobrze widoczny na obrazie.

Stopniowe zmiany temperatury, takie jak te widoczne np. pomiędzy 780. a

790. sekundą na rys. 7.19c, są prawdopodobnie wynikiem termoregulacji zwierzęcia. Odpowiadające zmiany w tym samym czasie dla tego obiektu można również zaobserwować na wykresie średniej temperatury całego obiektu oraz punktu śliny $Trace_1$.



RYS. 7.19. Średnia temperatura punktów referencyjnych w sąsiedztwie N4 dla: a) oka - punkt $RefP_1$, b) nasady ogona dla pierwszego osobnika - punkt $RefP_2$, c) nasady ogona dla drugiego osobnika - punkt $RefP_3$

Wszystkie obserwacje chłodzenia i ogrzewania śladów śliny zostały dopasowane do krzywych. Wizualna obserwacja profili zmian temperatury pozwoliła na ocenę trendu zmian. Sprawdzone model eksponencjalny oraz wielomian pierwszego i drugiego rzędu używając współczynnika determinacji R^2 (równanie 7.4) jako kryterium dopasowania modelu.

$$R^2 = \frac{\sum_{t=1}^n (f_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (7.4)$$

gdzie:

y_t - rzeczywista wartość zmiennej w momencie t ,

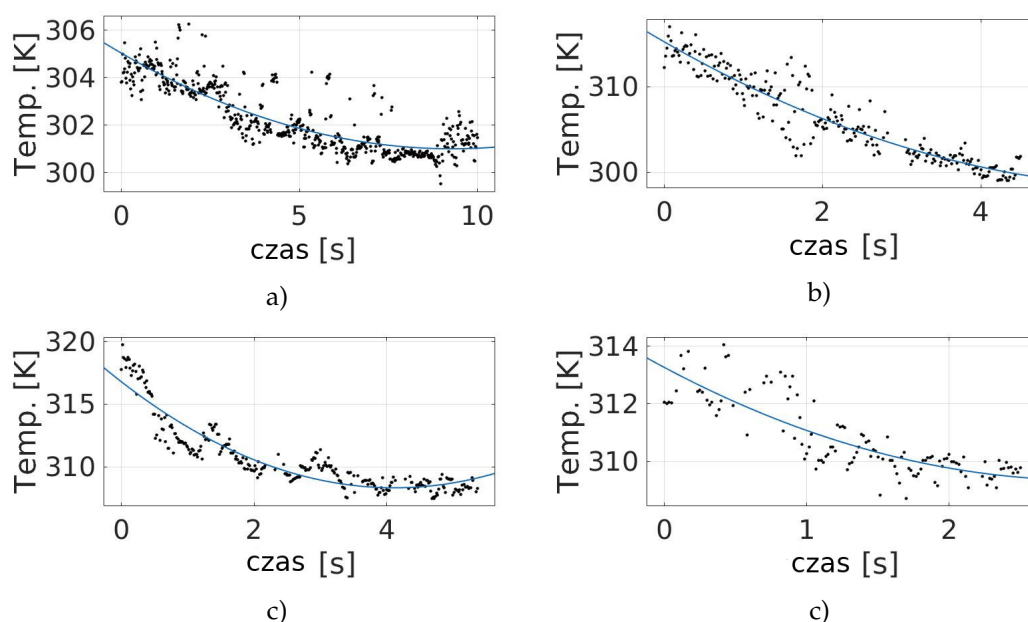
f_t - estymowana wartość zmiennej (na podstawie modelu),

\bar{y} - średnia arytmetyczna wartość rzeczywistej zmiennej (pomiaru).

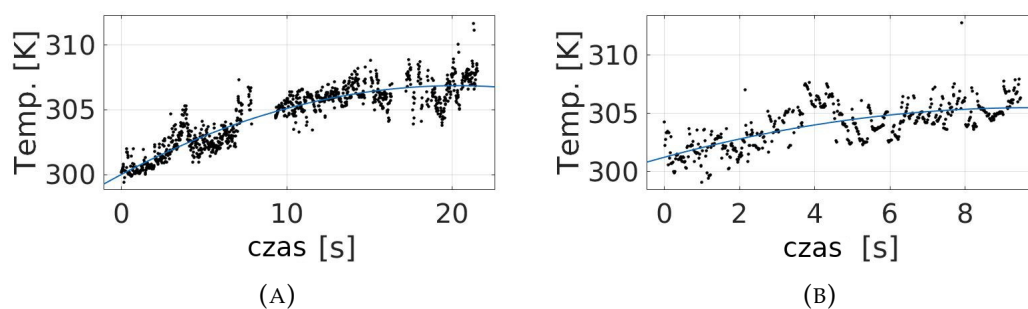
Wybraną metodą dopasowania były liniowa i nieliniowa metoda najmniejszych kwadratów.

Wyniki dopasowania danych do modelu przedstawione są na rys. 7.20 dla procesu chłodzenia oraz rys. 7.21 dla procesu ogrzewania.

Najlepiej dopasowanym modelem dla wszystkich danych okazał się być wielomian drugiego stopnia, który osiągnął najwyższe wartości kryterium dopasowania. Tabela 7.5 przedstawia wyniki dopasowania danych do funkcji $f(x) = a * x^2 + b * x + c$, gdzie a , b i c oznaczają obliczone współczynniki równania, a R^2 miarę dopasowania.



RYS. 7.20. Dopasowanie krzywych dla procesu chłodzenia śladów:
a) *Trace*₁, b) *Trace*₂, c) *Trace*₄, d) *Trace*₅



RYS. 7.21. Dopasowanie krzywych dla procesu ogrzewania śladów:
a) *Trace*₂, b) *Trace*₃

Współczynnik a przyjmuje wartości dodatnie dla procesów chłodzenia i ujemne dla ogrzewania. Najszybszy spadek temperatury obserwowany jest dla śladu *Trace*₄, dla którego współczynnik a przyjmuje wartość równą 0,5009. Najmniejsza wartość współczynnika a wyniosła 0,0467 - dla śladu *Trace*₁ - prezentowała najwolniejszy spadek temperatury wśród obserwowanych punktów.

Procesy ogrzewania trwały znacznie dłużej niż procesy chłodzenia. Temperatura rosła zgodnie z narastającym zboczem paraboli o negatywnym współczynniku a . Współczynnik a dla śladu *Trace*₂ wyniósł -0,0167, a dla śladu *Trace*₃ -0,0448. Ślad *Trace*₃ uzyskał również najniższy współczynnik determinacji wynoszący 0,49, podczas gdy miara dopasowania dla wszystkich innych punktów zawierała się pomiędzy wartościami od 0,71 do 0,89.

Kształt rozkładu temperatury jest jednakowy dla wszystkich analizowanych punktów, jednak prędkość chłodzenia i ogrzewania prawdopodobnie zależą od warunków takich jak: ilość śliny pozostawionej na futrze, lokalizacja śladu na ciele (mniejsze owłosienie np. na głowie spowoduje szybsze ogrzanie śladu dzięki słabszej izolacji cieplej skóry) lub średnia temperatura ciała badanego osobnika. Termoregulacja

TABELA 7.5. Wyniki obliczonych współczynników dopasowanej funkcji oraz miara dopasowania R^2

proces		a	b	c	R^2
chłodzenie	śladu $Trace_1$	0,047	-0,87	305	0,75
	śladu $Trace_2$	0,416	-5,32	315,3	0,89
	śladu $Trace_4$	0,501	-4,12	316,8	0,86
	śladu $Trace_5$	0,442	-2,63	313,3	0,71
ogrzewanie	śladu $Trace_2$	-0,0167	0,6758	300	0,84
	śladu $Trace_3$	-0,0448	0,8726	301,2	0,49

obserwowana w procesie ogrzewania dla śladu $Trace_1$ oraz punktu referencyjnego $RefP_3$, także wpływa na wyniki pomiarów. Dane dotyczące procesu termoregulacji mogą być przydatne w analizie stresu i lęku w badaniach behawioralnych [137].

Szereg zaprezentowanych wyników badań zmian temperatury w czasie opublikowano w pracy [96].

Na podstawie powyższych obserwacji stworzono symulację procesu chłodzenia i ogrzewania śliny. Opis tych symulacji zawarty jest w rozdziale 6.4.2.

7.4.2. Analiza zmian temperatury na symulacjach

Na podstawie obrazów symulacji opisanych w rozdziale 6.4.2 wykonano analizę detekcji zmian temperatury śladu w czasie przy użyciu detektora Harris o rozmiarze filtra równym 3 i minimalnej akceptowanej jakości narożników ustawionej na 0,000001. Konieczność tak niskiej wartości minimalnej akceptowanej jakości spowodowana była bardzo dobrym dopasowaniem temperatury granic śladu do temperatury ciała, co skutkowało brakiem wyraźnej granicy.

Ślady śliny odnajdywane były wśród wszystkich wyników detektora wykorzystując fakt, że znane było dokładne położenie symulacji. Narożnik uznawany był za ślad, gdy znajdował się w pewnej odległości od znanego miejsca symulacji. Ustawienie obszaru poszukiwań zamiast dokładnej lokalizacji śladu pozwoliło na detekcję zbliżoną do rzeczywistej, gdy inny punkt o dużym gradiencie (np. granica obiektu lub części ciała o ekstremalnej temperaturze) znajdujący się w bliskiej odległości od śladu uzyska wynik wyższy niż szukany ślad (patrz rys 7.23e).

Tutaj przyjęto uproszczony algorytm detekcji, ponieważ badania miały na celu sprawdzenie możliwości znalezienia trendu zmian temperatury dla niepełnych obserwacji, a nie samej detekcji. Docelowy algorytm identyfikacji śladów śliny wśród wszystkich wykrytych narożników powinien obejmować m.in. następujące kroki:

1. obliczanie znaczenia narożników za pomocą określonej miary, np. wartości R lub rankingu - co umożliwi wstępne odrzucenie nieistotnych punktów;
2. eliminacja narożników niespełniających określonych warunków: np. wartość temperatury w określonym zakresie, położenie na ciele zwierzęcia;
3. przewidywanie lokalizacji śladów na podstawie wektorów przemieszczenia zwierzęcia z poprzednich klatek.

Wygenerowane ślady śliny zostały nałożone na symulowane kształty podczas symulowanego ruchu dwóch obiektów, dlatego czasami ślad był niekompletny w

wyniku zasłonięcia go przez ciało drugiego osobnika. Mając to na uwadze, wprowadzono parametr długości śladu LR (ang. length ratio) będący stosunkiem przedziału czasu, w którym ślad śliny został wykryty do przedziału czasu trwania śladu:

$$LR = \frac{\text{przedział czasu, w którym ślad śliny został wykryty}}{\text{przedział czasu trwania śladu}} \quad (7.5)$$

Jednostką miary obu przedziałów czasu jest liczba klatek. Przedział czasu, w którym ślad był widoczny określa łączną liczbę klatek w których ślad został odnaleziony przez detektor. Przedział czasu dla symulowanych danych wynosi 27 ramek (rys.6.14).

Na 53 zasymulowane sekwencje śladów (każda po 2 procesy), 7,6 % miało tylko jeden pomiar na proces (ślad był widoczny tylko przez jedną klatkę podczas całego procesu) i dlatego zostało wykluczone z dalszej analizy. 18,9 % procesów chłodzenia i 9,4 % ogrzewania osiągnęło ujemny współczynnik determinacji, co wskazuje na słabe dopasowanie danych do modelu. W rzeczywistości negatywny wynik był często spowodowany jednym błędnym pomiarem i/lub małą liczbą wszystkich pomiarów (patrz rys. 7.23d chłodzenie). Średni stosunek długości LR dla pozostałych pomiarów wyniósł $0,44 \pm 0,12$ dla procesów chłodzenia i $0,47 \pm 0,25$ dla ogrzewania.

Atak i ugryzienie ma miejsce zawsze podczas bliskiego kontaktu i często ślad można dostrzec dopiero wtedy, gdy zwierzęta oddalają się od siebie. To sprawia, że dokładny moment ugryzienia jest bardzo trudny do zauważenia podczas obserwacji interakcji zwierząt w zakresie częstotliwości promieniowania widzialnego. Ślad śliny pozostawiony na futrze jest dowodem ugryzienia lub groomingu socjalnego. Dokładny czas tych zdarzeń możliwy jest do oszacowania na podstawie krzywej dopasowanej do pomiarów temperatury narożnika.

Dla każdego odnalezionego narożnika zaklasyfikowanego jako ślad śliny zastosowano procedurę dopasowania krzywej. Następnie, dla każdego procesu obliczono jakość dopasowania krzywej do symulowanej funkcji wielomianu przy użyciu współczynnika determinacji R^2 obliczonego ze wzoru:

$$R^2 = 1 - \frac{\sum_{t=1}^n (y_t - f_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (7.6)$$

gdzie:

y_t - rzeczywista wartość zmiennej w momencie t ,

f_t - estymowana wartość zmiennej (na podstawie modelu),

\bar{y} - średnia arytmetyczna wartość rzeczywistej zmiennej (pomiaru).

W celu zbadania liniowej korelacji pomiędzy długością śladu a jakością dopasowania zastosowano współczynnik korelacji Pearsona, zdefiniowany jako:

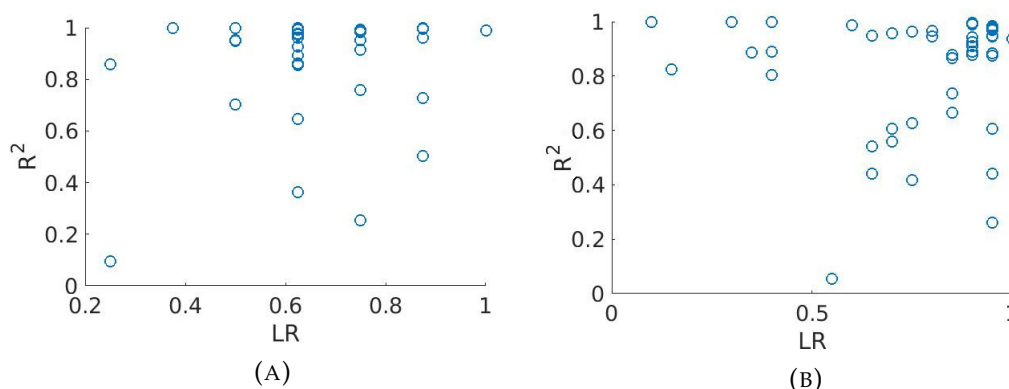
$$\rho = \frac{1}{n-1} \sum_{t=1}^n \left(\frac{y_t - \bar{y}}{\sigma_y} \right) \left(\frac{f_t - \bar{f}}{\sigma_f} \right) \quad (7.7)$$

gdzie:

n to liczba obserwacji,

\bar{y} i σ_y to średnia i odchylenie standardowe pomiarów,

\bar{f} i σ_f to średnia i odchylenie standardowe wartości modelowanych.



Rys. 7.22. Wykresy rozrzutu dla wpływu długości śladu (LR) na współczynnik dopasowania R^2 dla procesu a) chłodzenia i b) ogrzewania

Średnie wartości współczynnika determinacji (R^2) między obserwacjami a symulowanymi wartościami dla śladów wyniosły $0,87 \pm 0,22$ dla chłodzenia i $0,81 \pm 0,23$ dla procesów ogrzewania. $R^2 = 1$ to idealne dopasowanie funkcji, wynik powyżej 0,8 oznacza dość dobre dopasowanie.

Współczynnik korelacji Pearsona przedstawiający wpływ długości śladu na wartość R^2 wyniósł $\rho = 0,15$ dla chłodzenia i $\rho = 0,08$ dla ogrzewania. Wykresy rozrzutu dla obu procesów przedstawiono na rys. 7.22.

Jak widać na rys. 7.22 długość śladu (LR) nie wpływa na jakość dopasowania detekcji do symulacji. Współczynniki korelacji ρ (0,15 i 0,08) wskazują, że nie ma liniowej korelacji pomiędzy długością obserwowanego śladu a jakością dopasowania. Nawet ślady zawierające mniej niż 30 % obserwacji mogą zostać poprawnie dopasowane do rzeczywistego procesu. Warunkiem koniecznym jest jednak widoczność w obu procesach, czyli osobno podczas chłodzenia i ogrzewania. Było to założenie symulacji śladów o zmiennej w czasie temperaturze ze względu na zbyt małą liczbę rzeczywistych śladów, które umożliwiłyby określenie zależności pomiędzy procesami.

Lokalizacja śladu ma większy wpływ na jakość dopasowania modelu do obserwacji niż współczynnik LR. Jeśli w sąsiedztwie śladu znajdują się inne regiony o dużym gradencie temperatury, wtedy w momentach gorszej widoczności śladu te miejsca mogą zostać uznane za obserwację (patrz rys. 7.23e). Podczas poruszania się zwierzęcia, nieruchomy ślad umieszczony na futrze zmienia swoje położenie względem innych charakterystycznych punktów i granic ciała. W praktyce okazało się, że najczęstszą przyczyną błędów detekcji była nieduża odległość śladu od granicy obiektu. Błędną detekcję zaobserwowano również podczas analizy śladów śliny zlokalizowanych na głowie, gdzie znajduje się wiele punktów o skrajnych temperaturach. Z kolei ślady o temperaturze zbliżonej do temperatury otoczenia i niewidoczne na obrazie termograficznym dla ludzi bez odpowiedniej zmiany kontrastu (patrz klatka 25 na rys. 6.14), często były poprawnie identyfikowane.

Rys. 7.23 przedstawia kilka przykładów dopasowania obserwacji (temperatury narożników odnalezionych przez algorytm Harris) do krzywych. Czarne okręgi oznaczają obserwacje, czyli odnalezione narożniki, niebieska przerywana linia to krzywa najlepiej dopasowana do obserwacji, czerwona ciągła linia przedstawia symulowaną (modelowaną) krzywą. Obserwacje na rys. 7.23a wykazują wysokie dopasowanie do modelu, parametr R^2 dla chłodzenia wynosi 0,96, a dla ogrzewania

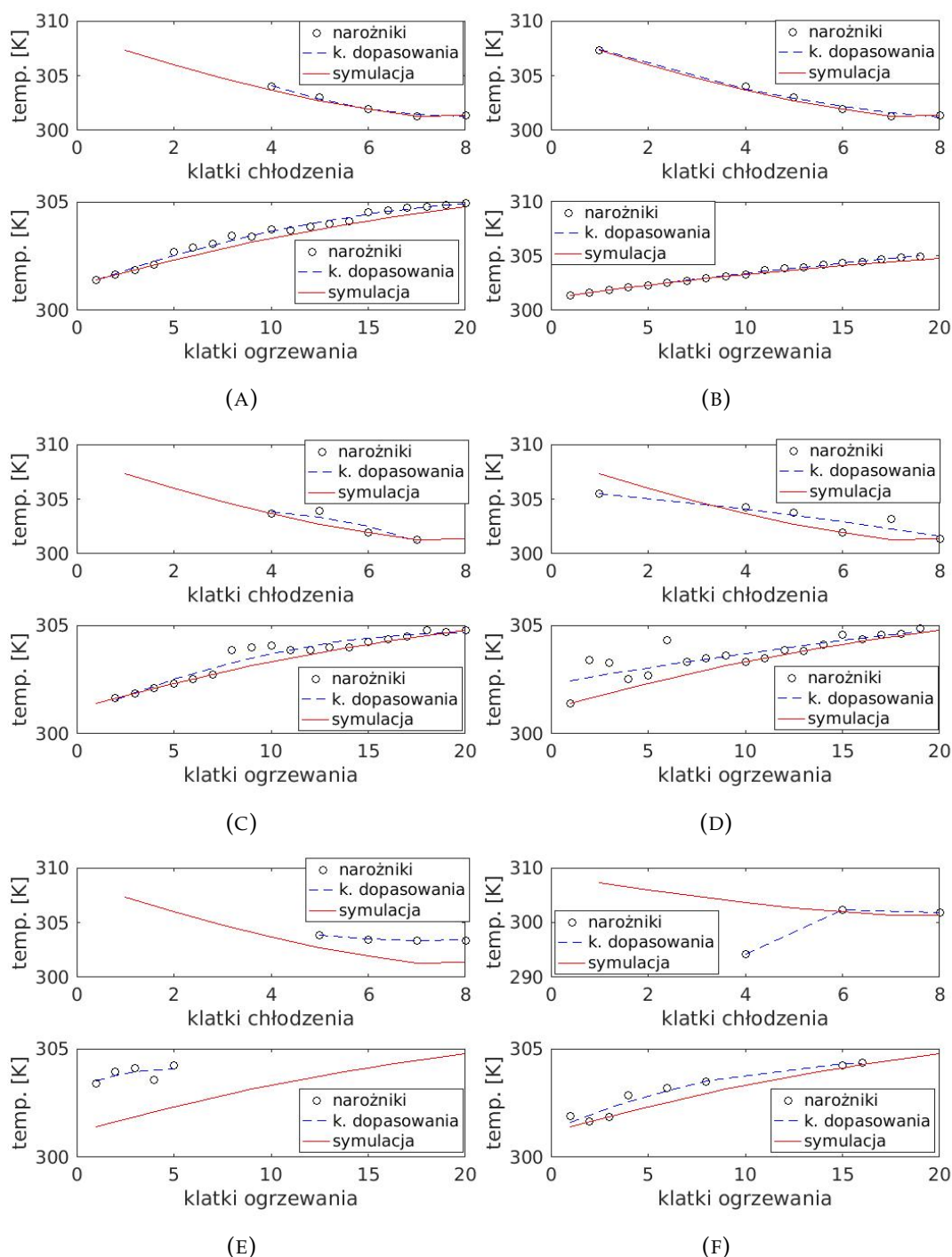
0,94. Dopasowanie przykładu 7.23b jest jeszcze wyższe, R^2 w obu procesach osiągnął wartości powyżej 0,98. Gorsze wyniki dopasowania przedstawione zostały na rys. 7.23c, gdzie kilka pomiarów w kolejnych klatkach ogrzewania i jeden pomiar chłodzenia odbiegają od symulacji, oraz na rys. 7.23d, na którym większość obserwacji chłodzenia i znaczna ilość ogrzewania nie zgadzają się z symulacją, powodując tym samym spadek wartości współczynnika determinacji poniżej 0,3. Takie chwilowe oscylacje pomiarów wokół zasymulowanych wartości temperatury często powodowane były przez ruch zwierzęcia i/lub położenie śladu blisko granicy ciała. Dwa ostatnie przykłady prezentują ujemne wyniki R^2 . Pomiary zarówno dla chłodzenia ($R^2=-68,43$) jak i ogrzewania ($R^2=-38,07$) przedstawione na rys. 7.23e wskazują, że za ślad po ślinie prawdopodobnie został uznany inny narożnik o odmiennej, wyższej temperaturze. Takie przypadki zdarzały się, kiedy ślad został naniesiony na obszar głowy, gdzie w niedalekiej odległości od siebie umiejscowione są liczne punkty o różnych temperaturach (oczy, uszy, nos). Proces chłodzenia na rys. 7.23f także osiągnął ujemną wartość parametru $R^2=-1,14$. Spowodowane jest to jednak wyłącznie jednym nieprawidłowym pomiarem, który w znacznym stopniu wpłynął na wynik ze względu na małą liczbę obserwacji.

Wizualizacja rozkładu cech statystycznych dla wartości R oraz rankingu pogrupowanych według kolejnych klatek symulacji została przedstawiona na rys. 7.24 w postaci wykresów skrzynkowych. Dolny i górny bok każdego pudełka oznaczają odpowiednio 25. i 75. percentyl próbki (pierwszy i trzeci kwartył). Czerwona linia wewnątrz to mediana próbki. Wąsy rozciągają się do wartości, które odpowiadają około $\pm 2.7\sigma$. Obserwacje poza granicami wąsów są wartościami odstającymi oznaczonymi jako czerwony znak + wyłącznie na rys. 7.24b. Obserwacje odstające dla wartości R zostały pominięte na rys. 7.24a, w celu lepszej widoczności pozostałych danych.

Zmiany wartości R w czasie (rys. 7.24a) wyraźnie odwzorowują zmiany temperatury śladu. Podczas chłodzenia wartość R rośnie wraz ze spadkiem temperatury, aż do osiągnięcia maksymalnej wartości dla 7. klatki, w której ślad osiąga najniższą temperaturę. Z kolei wzrost temperatury w trakcie wysychania powoduje powolne zanikanie śladu na futrze i spadek wartości R. Tendencje zmian są zbliżone do symulowanej dynamiki temperatury, z wyjątkiem wyniku dla klatki nr 1, dla której wartość R jest równa 0,7. Na ten odmienny wynik wpływ miała prawdopodobnie niewielka ilość danych dla pierwszej klatki (tylko 8,8 % wszystkich detekcji) spowodowana częstymi zasłonięciami śladu podczas interakcji zwierząt. Wykres skrzynkowy dla rankingu nie wykazuje oczekiwanych zależności, jak pokazano na rys. 7.24b. Jedynie wzrost średniego rankingu w kilku ostatnich klatkach może sugerować trudniejszą detekcję spowodowaną wzrostem temperatury śladu. Inne narożniki obrazu również wpływają na wynik rankingu. W zależności od ich wyników, rankingi śladów o podobnych parametrach mogą się znacznie różnić w kolejnych klatkach sekwencji. Wartości średnie rankingów dla różnych klatek nie są statystycznie różne, w przeciwieństwie do większości wartości R.

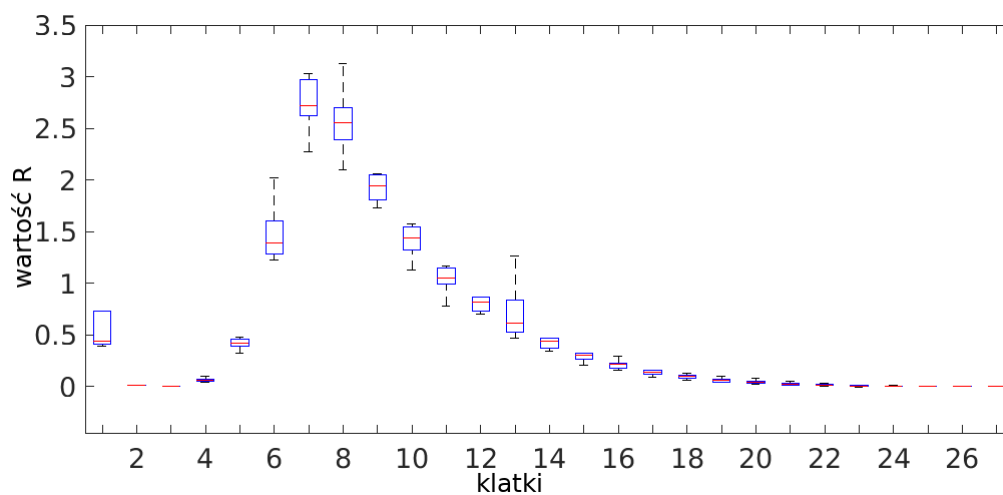
7.5. Podsumowanie

W powyższym rozdziale wykazano, że możliwa jest detekcja śladów śliny w postaci punktów charakterystycznych i rozróżnienie ich od innych elementów obrazu oraz

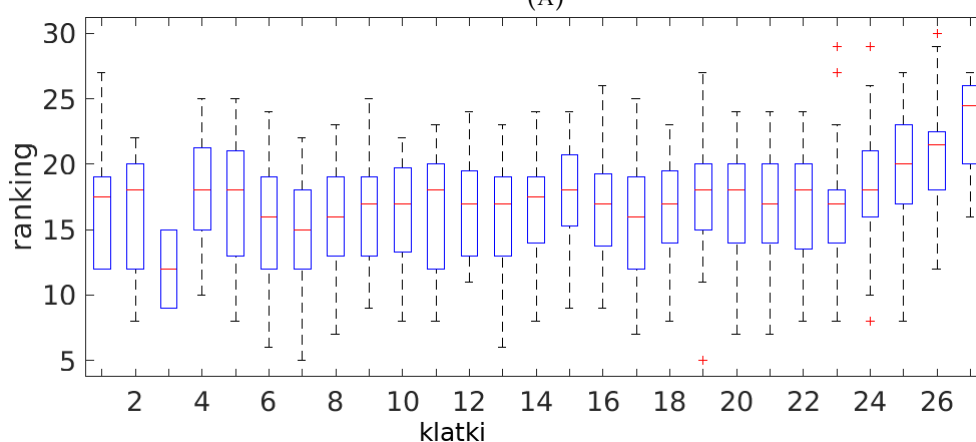


RYS. 7.23. Przykłady dopasowania obserwacji (czarne okręgi) do modelu (linia czerwona) dla procesu chłodzenia (wykres górny) i ogrzewania (wykres dolny) dla 8 różnych śladów o parametrach R^2 odpowiedni dla procesu chłodzenia i ogrzewania równych: a) 0,96 i 0,94; b) 0,99 i 0,98; c) 0,7 i 0,87; d) 0,25 i 0,26; e) -68,43 i -38,07; f) -1,14 i 0,81. Niebieska przerywana linia oznacza krzywą wielomianu drugiego stopnia najlepiej dopasowaną do obserwacji

dopasowanie modelu chłodzenie-ogrzewanie do większości symulowanych sekwencji ze śladami śliny, nawet takich, które zawierają dużą liczbę klatek ze śladami niewidocznymi w wyniku interakcji zwierząt. Proponowana metoda detekcji ugryzień



(A)



(B)

RYS. 7.24. Wykresy średnich a) wartości R i b) rankingów dla wszystkich 27 klatek symulacji.

pozwala nie tylko na wykrycie tego zjawiska, zwykle niewidocznego dla obserwatora, ale także na oszacowanie momentu jego wystąpienia. Model umożliwia identyfikację czasu ugryzienia nawet wtedy, gdy zdarzenie ugryzienia nie jest bezpośrednio rejestrowane. Pozwala także oszacować temperaturę śladu w przypadku, gdy jest on niewidoczny. Wykazano, że wpływ temperatury i wielkości śladu na wyniki detekcji pozwala na automatyczne rozpoznanie interesujących punktów. Jednak czasami może pojawić się wpływ innych niestałych punktów charakterystycznych obrazu na wartość rankingową wykazaną podczas analizy czasowej. Z tego powodu wartość R, miara niezależna od otoczenia, jest bardziej użytecznym parametrem niż ranking.

Automatyzacja procesu analizy złożonych zachowań socjalnych zwierząt doświadczalnych, będąca tematem tezy 1., możliwa jest dzięki detekcji określonych zdarzeń nietypowych, jakim jest np. ugryzienie lub agresywne czyszczenie socjalne. W przypadku powyższych zdarzeń zagadnienie detekcji może zostać zastąpione detekcją punktów charakterystycznych przedstawiających ślady śliny pozostawione po ugryzieniu. W rozdziale tym przedstawiono i scharakteryzowano detektory



umożliwiający odnalezienie śladów śliny na obrazie termograficznym. Przeprowadzone analizy z wykorzystaniem danych pomiarowych i danych symulacyjnych pozwoliły wykazać, że parametry te wspomagają detekcję złożonych zachowań socjalnych, jakimi są ugryzienia lub agresywne czyszczenie. W ten sposób potwierdzono tezę 1.

W odniesieniu do tezy 2:

W pracy zaproponowano parametry umożliwiające indywidualizację cech niezbędnych w detekcji zachowań socjalnych, w celu wykorzystania w analizie zachowań.

W szczególności zaproponowano następujące parametry:

- parametry położenia śladu- w obszarze wysegmentowanego obiektu (segmentacja zwierząt podczas kontaktu opisana w rozdziale 5),
- wartość R - opisana w rozdziałach 7.3.1, 7.3.2, 7.3.3, 7.3.4 oraz 7.4.1,
- ranking - opisany w rozdziałach 7.3.1, 7.3.2, 7.3.3, 7.3.4 oraz 7.4.1,
- wartość temperatury narożnika odpowiadającemu śladowi śliny T_{cp} - opisana w rozdziałach 7.3 i 7.4,
- wartość temperatury punktu $Corner_{LOW} - T_{cLow}$ - opisana w rozdziałach 7.3 i 7.4.1,
- parametry zmian temperatury w czasie - zgodne z oczekiwanymi parametrami modelu procesu wysychania opisanego w rozdziałach 6.4.2 i 7.4.

Parametry te umożliwiają indywidualizację cech niezbędnych w detekcji zachowań socjalnych, co jest potwierdzeniem drugiej części tezy 2.

Rozdział 8

Podsumowanie

Automatyzacja badań laboratoryjnych staje się dziś standardem. Zarówno zrobotyzowane urządzenia jak i odpowiednie oprogramowanie pozwalają nie tylko na optymalizację przepływu zadań i poprawę produktywności, ale, co ważniejsze, uzyskują lepszą dokładność, powtarzalność i precyzję niż badania wykonywane manualnie.

Metody wspomaganie automatycznej analizy zachowań zwierząt laboratoryjnych mają istotne znaczenie w inżynierii biomedycznej. Umożliwiają przyspieszenie prac, obiektywizację oraz parametryzację obserwacji związanych z pracami nad chorobami psychicznymi i neurologicznymi, stresem, strachem czy skutkami stosowania produktów leczniczych. Zachowania socjalne oraz ich zmiany niosą wiele informacji nie tylko o poważnych schorzeniach takich jak np. depresja, schizofrenia czy uszkodzenia mózgu, ale także o stanie zdrowia, czy samopoczuciu osobnika.

Badania laboratoryjne przeprowadzane na zwierzętach bardzo często polegają na ocenie ich zachowań. Obserwator zwraca uwagę na czynności wykonywane przez zwierzę, a następnie klasyfikuje je jako określone zachowanie. Niektóre typy zachowań mogą w prosty sposób zostać zautomatyzowane za pomocą analizy obrazu przedstawiającego scenę testu. Są to np. bieg, odpoczynek, chodzenie, które to charakteryzowane są przez prędkość przemieszczenia obiektu [69], [23]. Na podstawie parametrów przemieszczenia, położenia oraz wzajemnych odległości możliwa jest także detekcja prostych zachowań socjalnych takich jak np. zbliżanie się, śledzenie, pogoń (patrz tabela 3.2). Parametry kształtu ciała oraz położenie i zmiany punktów charakterystycznych ciała pomocne są z kolei w detekcji np. czyszczenia się (groomingu), eksploracji czy stania na tylnych łapach [23], [35], [69], [70]. Największym jednak problemem są wciąż złożone zachowania socjalne typu: gryzienie, atak, grożenie czy ofensywna pozycja boczna charakteryzujące poziom agresji lub lęku zwierzęcia. Dla takich zachowań brak jest zdefiniowanych parametrów umożliwiających detekcję automatyczną. Dodatkowo, znaczna ich część zachodzi podczas bliskiego kontaktu osobników, co powoduje konieczność rozdzielania obiektów na obrazie przed wykonaniem analizy.

Zaprezentowana rozprawa doktorska przedstawia propozycje miar oraz algorytmów pomocnych w automatycznej analizie złożonych zachowań socjalnych zwierząt laboratoryjnych. Skupia się na dokładnej segmentacji zwierząt podczas kontaktu fizycznego, podczas którego często występują znaczne zasłonięcia obiektów. Możliwość identyfikacji osobników oraz wykonywania pomiarów także w trakcie złączenia dostarczają dane istotne z punktu widzenia analizy zachowań socjalnych. Praca proponuje także metodę detekcji zachowań agresywnych na podstawie obrazu termograficznego śladu śliny pozostawionej na futrze gryzonia przez drugiego osobnika. Jest to prawdopodobnie jedyny dostrzegalny dowód wystąpienia agresywnej interakcji między zwierzętami.

Na potrzeby weryfikacji skuteczności zaproponowanych algorytmów stworzono bazę danych składającą się z ponad 500 minut nagrań sekwencji termicznych oraz filmów w zakresie częstotliwości widzialnych zawierających testy socjalne. Baza danych została opisana w rozdziale 4.5, natomiast testy socjalne w rozdziale 3. Dodatkowo wykonano także symulacje nagrań szczegółowo przedstawione w rozdziale 6.

W pracy przedstawiono dwie tezy. Teza 1. zakłada, iż:

1. Możliwe jest zautomatyzowanie procesu analizy złożonych zachowań socjalnych zwierząt doświadczalnych, na drodze budowy scenariuszy testowych i cyfrowej rejestracji obrazu sceny eksperymentu oraz zastosowanie procedur cyfrowej analizy sekwencji obrazów termograficznych do rozpoznawania określonych klas zachowań socjalnych.

Scenariusze testowe eksperymentu w postaci cyfrowej rejestracji obrazu zostały omówione w rozdziałach 3 i 4. Scenariusze symulacyjne opisano w rozdziale 6.

Badania zawarte w rozdziale 3.3 wykazały, że możliwa jest automatyzacja procesu analizy zachowań prostych przy użyciu łatwych w obliczeniu parametrów ruchu czy kształtu ciała oraz klasycznych klasyfikatorów wykorzystujących metody uczenia maszynowego. System opisany w rozdziale 3.4 potwierdził z kolei możliwość detekcji zachowań nietypowych na podstawie pomiaru jedynie parametrów motorycznych zwierząt.

Badania opisane w rozdziale 5 wykazały, że precyzyjna segmentacja pozwala na pomiar podstawowych parametrów przedstawionych w rozdziale 3.1, istotnych z punktu widzenia dalszej analizy typowych zachowań socjalnych zwierząt, osobno dla każdego zwierzęcia. Dzięki temu, możliwa jest automatyzacja analizy nawet podczas złączenia się obiektów na obrazie, które to wcześniej było często powodem zaniechania detekcji i pomiaru tych danych. Jest to szczególnie ważne w przypadku analizy testów socjalnych, gdzie zachowania podczas kontaktu są istotą testu.

Badania przeprowadzone w rozdziale 7 wykazały, że detekcja nietypowych zachowań socjalnych, na których skupiono się w pracy - zachowań agresywnych, może zostać wykonana pośrednio poprzez detekcję śladów śliny pozostawionych po zachowaniach zawierających ugryzienia. W rozdziale 7.2 zaproponowano odpowiednie parametry oraz przeanalizowano ich działanie w rozdziale 7.3.4.

Potwierdzenie tezy 2. brzmiące:

2. Stosując metody uczenia maszynowego, możliwa jest automatyczna segmentacja obiektów złączonych oraz detekcja i pomiar parametrów umożliwiających indywidualizację cech niezbędnych w detekcji zachowań socjalnych

przedstawione zostało głównie w rozdziale 5 oraz 7. Prace zawarte w rozdziale 5 wykazały, że możliwa jest dokładna segmentacja zwierząt na obrazach termograficznych nawet podczas nałożenia i częściowego zasłonięcia obiektów. Zostało to zaprezentowane w rozdziałach 5.4 oraz 5.5. Zaproponowano 3-etapowy algorytm segmentacji obiekt-obiekt na podstawie stopnia złączenia zwierząt. Przetestowano także najpopularniejsze algorytmy zarówno metod segmentacji semantycznej (rozdział 5.5.1), które są w stanie wyznaczyć granicę pomiędzy gryzoniami, jak i segmentacji instancji (rozdział 5.5.2), które z kolei dodatkowo wykonują detekcję każdej instancji osobno. Do przeprowadzenia eksperymentów dotyczących segmentacji wykorzystano metody uczenia maszynowego, w szczególności architektury uczenia głębokiego.

Wstępne prace dotyczące podstawowych parametrów, które jednak nie są wystarczające do detekcji złożonych zachowań, mogą ją jedynie wspomagać, zaprezentowano w rozdziale 3.

Za zdarzenia złożone w zakresie tej pracy uznano wystąpienie śladów śliny na futrze pozostawione przez agresywnego osobnika, widoczne na obrazach termograficznych. W rozdziale 7.3 przeanalizowano charakterystykę śladów rzeczywistych (rozdziały 7.3.1, 7.3.2 i 7.3.3) oraz zbadano wpływ parametrów detektorów i śladów na wyniki detekcji na obrazach symulacyjnych (rozdział 7.3.4). Zaproponowano także parametry odpowiednie do opisu i identyfikacji analizowanych śladów: parametry położenia śladów (rozdział 7.3), temperatura śladu (T_{cp}) oraz punktu $Corner_{LOW}$ (T_{cLow}) (rozdział 7.3), wartość R i ranking (rozdziały 7.3.7.4), zmiany temperatury i wartości R w czasie (rozdział 7.4) oraz parametry modeli dopasowanych (rozdziały 6.4.2 i 7.4). Na podstawie badań opisanych w rozdziale 7.4.1 opracowano model zmian temperatury śladu (rozdział 6.4.2), a następnie wykonano analizę tych śladów w rozdziale 7.4.2. Wykazano również, że możliwa jest predykcja temperatury śladu śliny w przypadku braku jego widoczności (rozdział 7.4).

Powyższe badania przedstawiają dodatkową zaletę zastosowania technologii termografii w analizie zwierząt laboratoryjnych.

8.1. Osiągnięcia i znaczenie pracy

Wyniki przedstawione w pracy mające na celu potwierdzenie słuszności tez prowadzą do następujących stwierdzeń:

1. Możliwa jest segmentacja zwierząt laboratoryjnych podczas interakcji społecznych skutkujących bliskim kontaktem fizycznym z ponad 90% zgodnością wysegmentowanych obszarów z obrazami referencyjnymi dla przypadków nałożenia się obiektów (rozdziały 5.4 i 5.5), co z kolei umożliwia pomiar parametrów położenia, przemieszczenia oraz kształtu zwierząt;
2. Do segmentacji obrazów termograficznych można zastosować algorytmy stosowane podczas segmentacji obrazów z zakresu promieniowania widzialnego wstępnie przetrenowane na takich obrazach uzyskując jakość porównywalną z jakością modelu przetrenowanego od podstaw na obrazach termograficznych (rozdział 5.5);
3. Odwzorowanie oryginalnych danych termicznych w standardowe zakresy danych obrazowych w niektórych przypadkach poprawia dokładność segmentacji semantycznej nawet o 25% (rozdział 5.5.1);
4. Na ślady śliny pozostawione po interakcji pomiędzy osobnikami oddziałują dwa procesy: krótki proces chłodzenia oraz dłuższy ogrzewania, co może być wykorzystane w detekcji śladów śliny (rozdział 7.4.1);
5. Zmiany temperatury dla obu tych procesów mogą być opisane funkcjami kwadratowymi drugiego stopnia uzyskując stopień dopasowania R^2 średnio ponad 0,9 (rozdział 7.4.1);
6. Jakość dopasowania obserwacji temperatury śladu do trendu rzeczywistego w nieznanym stopniu ($\rho < 0,2$) zależna jest od czasu widoczności śladu (rozdział 7.4.2);
7. Zmiany wartości R są odwrotnie proporcjonalne do zmian temperatury (rozdział 7.4.2).

8. Zaproponowane parametry mogą wspomagać automatyczną detekcję zachowań (rozdział 7).

W ramach niniejszej pracy doktorskiej przeprowadzono szereg badań proponując, adaptując i weryfikując algorytmy analizy obrazów w badaniach zwierząt laboratoryjnych. W szczególności:

- przeanalizowano przydatność parametrów przemieszczenia i kształtu ciała w klasyfikacji zachowań typowych,
- stworzono system dedykowany analizie zachowań motorycznych (prosty) myszy,
- przygotowano stanowiska badawcze do obserwacji i nagrań zwierząt,
- zebrano i opracowano bazę danych zawierającą ponad 500 minut nagrań szczurów i 192 h nagrań myszy,
- zaproponowano metody przetwarzania wstępnego (wraz z segmentacją obiektów) oraz przekształcenia danych termicznych na obrazy,
- przeprowadzono analizę porównawczą klasycznych metod śledzenia obiektów,
- zaproponowano 3-etapową metodę segmentacji obiekt-obiekt w zależności od stopnia złączenia zwierząt,
- zaadaptowano i zweryfikowano modele segmentacji semantycznej, U-Net i V-Net, do segmentacji obiektów w obrazach termograficznych,
- opracowano etap algorytmu separacji zwierząt,
- przeanalizowano wpływ sposobu tworzenia obrazu z danych termicznych na wyniki segmentacji przez powyższe algorytmy,
- zaadaptowano i zweryfikowano modele segmentacji instancji, Mask R-CNN i TensorMask, do segmentacji i detekcji obiektów (gryzoni) w obrazach termograficznych,
- opracowano modele symulacji obrazów termograficznych przedstawiających dynamikę interakcji zwierząt z możliwością regulacji: kształtu, rozmiaru i tekstury ciała; temperatury ciała oraz jego punktów charakterystycznych; prędkości i sposobu przemieszczania; częstości wystąpień ugryzień; wielkości, kształtu i temperatury śladów po ugryzieniach; charakteru zmian temperatury śladu w czasie,
- przeanalizowano wyniki detekcji śladów śliny przez popularne detektory punktów charakterystycznych na obrazach termograficznych,
- przeanalizowano wpływ parametrów śladu oraz detektora na wyniki detekcji na obrazach symulacyjnych,
- przeanalizowano zmiany temperatury śladu w czasie oraz stworzono 2-procesowy model wysychania śliny,
- przeanalizowano możliwości detekcji śladu zmiennego w czasie oraz możliwości jego predykcji w momentach zasłonięcia.



Podsumowując, cel pracy został osiągnięty. Zaproponowano metody poprawy segmentacji zwierząt w czasie interakcji socjalnych, które pozwalają na dalszą automatyzację analiz. Przedstawiono także parametry pomocne w rozpoznaniu określonych zdarzeń agresywnych nie zastosowane dotąd w żadnym istniejącym rozwiązaniu.

Złożoność problemów przedstawionych w rozprawie wskazuje, iż w kontynuacji prac należałoby przeprowadzić większą liczbę nagrań rzeczywistych zachowań agresywnych, w szczególności nagrań ugryzień. Rzadki charakter tych zdarzeń nie pozwolił na zbadanie zależności pomiędzy dwoma procesami wysychania śladu śliny. Przedstawiony model procesu wysychania zaproponowany został na podstawie obserwacji niewielkiej liczby śladów rzeczywistych. Powinien zostać zweryfikowany na większej liczbie obserwacji oraz nagraniach w innych warunkach otoczenia (inna temperatura otoczenia, zwierzęta w innym wieku - co wpływa na długość futra). Przydatnym byłoby też zaproponowanie deskryptora rozróżniającego pochodzenie śladu wyłącznie na podstawie jego parametrów (np. temperatury lub czasu wysychania) bez uwzględniania kontekstu całej sekwencji zachowania. W pracy nie zbadano także możliwości rozpoznania pozy ciała podczas częściowego zasłonięcia. Zaproponowane algorytmy segmentacji poprawnie segmentują obiekty, szczególnie, jeżeli te nie zasłaniają się nawzajem. Jeśli, zwierzęta są częściowo zasłonięte (np. część ciała jednego osobnika znajduje się pod drugim), algorytmy wykonują poprawną detekcję, jednak segmentacja będzie uwzględniała tylko widoczne regiony. Detekcja innych parametrów, jak położenie, widoczne punkty charakterystyczne itp. powinny zostać wykonane prawidłowo. Natomiast wartości deskryptorów pozy ciała mogą różnić się od wartości, dla których zwierzę widoczne jest w całości. Tematyka badań jest zatem wciąż aktualna.

Bibliografia

- [1] W. Abdulla. "Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow". In: *GitHub repository* (2017).
- [2] S. Akita, S. Tsuchihiro, and H. Takemura. "Detection of Rapid Mouse's Scratching Behavior Based on Shape and Motion Features". In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2019, pp. 925–928.
- [3] M. E. Albonetti and F. Farabollini. "Effects of single and repeated restraint on the social behaviour of rats." In: *Physiol. Behav.* 53 (1993), pp. 937–942.
- [4] M. E. Albonetti and F. Farabollini. "Social stress by repeated defeat: effects on social behavior and emotionality." In: *Behav Brain Res* 62 (1994), pp. 187–193.
- [5] D. Anopas et al. "Automatic Inference of Rat's Hindlimb Trajectory to Synchronize with Forelimb Gait Through Phase". In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2019, pp. 4615–4618.
- [6] M. S. Arulampalam et al. "A tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking." In: *IEEE Transactions on Signal Processing*. Vol. 50. 2. 2002, pp. 174–187.
- [7] S. Aslani et al. "Day and night: diurnal phase influences the response to chronic mild stress". In: *Front Behav Neurosci.* (2014), pp. 8–82.
- [8] E. Z. Barcelos et al. "A Combined Method for Segmentation and Registration for an Advanced and Progressive Evaluation of Thermal Images". In: *Sensors* 14.11 (2014), pp. 21950–21967.
- [9] C. A. Barnes. "Memory deficits associated with senescence: a neurophysiological and behavioral study in the rat." In: *J Comp Physiol Psychol.* 93.1 (1979), pp. 74–104.
- [10] Y. Benjamini et al. "Ten ways to improve the quality of descriptions of whole-animal movement". In: *Neuroscience and Biobehavioral Reviews* (2010), pp. 1351–1365.
- [11] K. C. Berridge and I. Q. Whishaw. "Cortex, striatum and cerebellum: control of serial order in a grooming sequence." In: *Exp. Brain Res.* (1992), pp. 275–290.
- [12] O. Berton et al. "Behavioral reactivity to social and nonsocial stimulations: A multivariate analysis of six inbred rat strains." In: *Behavior Genetics* 27.2 (1997), pp. 155–166.
- [13] M.M. Blanchard, D. Mendelsohn, and J.A. Stamp. "The HR/LR model: Further evidence as an animal model of sensation seeking". In: *Neurosci Biobehav Rev.* 3.7 (2009), pp. 1145–54.
- [14] R. J. Blanchard and D. C. Blanchard. "Aggressive behavior in the rat." In: *Behavioural Biology* (1977), pp. 197–224.



- [15] R. J. Blanchard et al. "Attack and defense behavior in the albino rat." In: *Animal Behaviour* 25 (1977), pp. 622–634.
- [16] R. J. Blanchard et al. "Conspecific aggression in the laboratory rat." In: *Journal of Comparative and Physiological Psychology* 89 (1975), pp. 1204–1209.
- [17] G. Bradski and A. Kaehler. *Learning OpenCV*. O'Reilly Media. Inc., 2008.
- [18] G. R. Bradski. *Computer Vision Face Tracking for Use in a Perceptual User Interface*. 1998.
- [19] P. F. Brain, K. H. McAllister, and S. Walmsley. *Psychopharmacology*. Ed. by W. A. A. Boulton and G. B. Baker and A. J. Greenshaw. Vol. 13. Neuromethods. Clifton: Humana Press., 1989.
- [20] K. Branson. "Tracking multiple mice through severe occlusions." PhD thesis. University of California at San Diego, 2007.
- [21] K. Branson and S. Belongie. "Tracking multiple mouse contours (without too many samples)". In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. Vol. 1. 2005, pp. 1039–1046.
- [22] K. Branson, V. Rabaud, and S. Belongie. "Three brown mice: See how they run". In: *VS-PETS Workshop at ICCV*. 2003.
- [23] P. Burgos-Arizzu X. nad Dollar et al. "Social behavior recognition in continuous video." In: *25th Conference on Computer Vision and Pattern Recognition*. 2012, pp. 1322–1329.
- [24] M. Casarrubea et al. "Temporal Patterns of Rodent Behavior in the Elevated Plus Maze test". In: *Proceedings of Measuring behavior*. 2012, pp. 350–353.
- [25] M. Casarrubea et al. "Temporal structure of rat behavior in the social interaction test". In: *Proceedings of Measuring Behavior*. 2014.
- [26] A. Chaudhary and D. Chaturvedi. "Efficient Thermal Image Segmentation for Heat Visualization in Solar Panels and Batteries using Watershed Transform". In: *International Journal of Image, Graphics and Signal Processing* 9 (Nov. 2017), pp. 10–17.
- [27] F. de Chaumont, S. Granon, and J.C. Olivio-Marin. "Video analysis of social interactions in mice: an integrated solution." In: *Proceeding of Measuring Behavior*. 2014.
- [28] F. de Chaumont et al. "Computerized video analysis of social interactions in mice." In: *Nature Methods* 9 (2012), pp. 410–417.
- [29] K. Chen et al. "Hybrid Task Cascade for Instance Segmentation". In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019), pp. 4969–4978.
- [30] X. Chen et al. "TensorMask: A Foundation for Dense Object Segmentation". In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)* (2019), pp. 2061–2069.
- [31] D. Comainciu and P. Meer. "Mean Shift Analysis and Applications." In: *IEEE Int'l Conf. Comp. Vis.* Vol. 2. 1999, pp. 1197–1203.
- [32] M. Cordts et al. "The Cityscapes Dataset for Semantic Urban Scene Understanding". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), pp. 3213–3223.
- [33] I. Dagher and K. E. Tom. "WaterBalloons: A hybrid watershed Balloon Snake segmentation." In: *Image and Vision Computing* 26 (2008), pp. 905–912.

- [34] A. Dalmia, S.T. Kakileti, and Manjunath G. "Exploring Deep Learning Networks for Tumour Segmentation in Infrared Images". In: *14th Quantitative InfraRed Thermography Conference*. 2018, pp. 521–530.
- [35] E. van Dam et al. "An automated system for the recognition of various specific rat behaviours." In: *Journal of Neuroscience Methods* (2013), pp. 214–224.
- [36] D. Dayakshini et al. "Segmentation of Breast Thermogram Images for the Detection of Breast Cancer: A Projection Profile Approach." In: *Third International Conference on Intelligent Mechatronics and Automation*. Vol. 3. 2015.
- [37] M. B. Dillencourt, H. Samet, and M. Tamminen. "A general approach to connected-component labeling for arbitrary image representations." In: *Journal of the ACM (JACM)*. Vol. 39. 1992, pp. 253–280.
- [38] P. Dollar, V. Rabaud, and G. Cottrel i S. Belongie. "Behavior Recognition via Sparse Spatio-Temporal Features". In: *Proceedings 2nd Joint IEEE International Workshop on VS-PETS*. 2005, pp. 65–72.
- [39] A. Dutta and A. Zisserman. "The VIA Annotation Software for Images, Audio and Video". In: *Proceedings of the 27th ACM International Conference on Multimedia*. MM '19. 2019.
- [40] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. London: Chapman & Hall. London, 1993.
- [41] S. E. File. "The use of social interaction as a method for detecting anxiolytic activity of chlordiazepoxide-like drugs." In: *J. Neurosci. Methods* 2 (1980), pp. 219–238.
- [42] S. E. File and J. R. Hyde. "Can social interaction be used to measure anxiety?" In: *Br. J. Pharmacol.* 62 (1978), pp. 19–24.
- [43] S. E. File and P. Seth. "A review of 25 years of the social interaction test." In: *European Journal of Pharmacology* 483 (2003), pp. 35–53.
- [44] J. Fleuret et al. "A Real Time Animal Detection And Segmentation Algorithm For IRT Images In Indoor Environments". In: *Quantitative InfraRed Thermography* (2016).
- [45] N. Franco et al. "ThermoLabAnimal – A high-throughput analysis software for non-invasive thermal assessment of laboratory mice". In: *Physiology & Behavior* 207 (2019), pp. 113–121.
- [46] N. H. Franco et al. "ThermoLabAnimal – A high-throughput analysis software for non-invasive thermal assessment of laboratory mice". In: *Physiology and Behavior* 207 (2019), pp. 113–121.
- [47] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Boston: Academic Press., 1990.
- [48] B. Q. Geuther et al. "Robust mouse tracking in complex environments using neural networks". In: *Communications Biology* 2.1 (2019), p. 124.
- [49] G. Gheusi et al. "Social and individual recognition in rodents: methodological aspects and neurobiological bases." In: *Behavioural Processes* 33 (1994), pp. 59–87.
- [50] L. Giancardo et al. "Automatic Visual Tracking and Social Behaviour Analysis with Multiple Mice." In: *PLoS ONE* (2013).
- [51] R. B. Girshick. "Fast R-CNN". In: *2015 IEEE International Conference on Computer Vision (ICCV)* (2015), pp. 1440–1448.

- [52] E. C. Grant and J. H. MacKintosh. "A comparison of the social postures of some common laboratory rodents." In: *Behaviour* 21 (1963), pp. 246–259.
- [53] Z. Guo and R. W. Hall. "Parallel thinning with two-subiteration algorithms". In: *Comm. ACM* 32.3 (1989), pp. 359–373.
- [54] A. Gupta, P. Dollár, and R. Girshick. "LVIS: A Dataset for Large Vocabulary Instance Segmentation". In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019), pp. 5351–5359.
- [55] C. Hammels et al. "Defeat stress in rodents: From behavior to molecules". In: *Neurosci Biobehav Rev* 59 (2015), pp. 111–40.
- [56] A. F. Hanson. *Rat behavior*. dostęp 2017.09.20. URL: http://www.ratbehavior.org/norway_rat_ethogram.htm#Social.
- [57] C. Harris and M. Stephens. "A combined corner and edge detector." In: *Alvey Vision Conferen.* 1988, pp. 147–152.
- [58] K. He, R. Girshick, and P. Dollar. "Rethinking ImageNet Pre-Training". In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. 2019, pp. 4917–4926.
- [59] K. He et al. "Mask R-CNN". In: *IEEE International Conference on Computer Vision (ICCV)* (2017), pp. 2980–2988.
- [60] D. Heeren and A. Cools. "Classifying postures of freely moving rodents with the help of Fourier descriptors and a neural network". In: *Behavior Research Methods, Instruments and Computers* 32 (2000), pp. 56–62.
- [61] B. K. Horn and B. G. Schunck. "Detrmining Optical Flow." In: *Artificial Intelligence* 17 (1981), pp. 185–203.
- [62] Z. Huang et al. "Mask Scoring R-CNN". In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019), pp. 6402–6411.
- [63] Clever Sys inc. *Clever Sys*. dostęp 2020.09. URL: <http://cleversysinc.com/about-our-software>.
- [64] Med Associates Inc. *Arena Video Tracking*. dostęp 2020.09. URL: <http://www.med-associates.com/product-category/open-field-activity-accessories/>.
- [65] Med Associates Inc. *Wspomaganie uczenia*. dostęp 2020.09. URL: <https://www.med-associates.com/product-category/active-passive-avoidance-learned-helplessness/>.
- [66] San Diego Instruments. *Any-Maze*. dostęp 2020.09. URL: <http://www.sandiegoinstruments.com/any-maze-video-tracking/>.
- [67] M. Isard and A. Blake. "CONDENSATION - Conditional Density Propagation for Visual Tracking." In: *Int. J. Computer Vision* 29.1 (1998), pp. 5–28.
- [68] E.H. Jang et al. "Analysis of physiological signals for recognition of boredom, pain, and surprise emotions". In: *J Physiol Anthropol.* 34 (2015), p. 25.
- [69] H. Jhuang et al. "Automated home-cage behavioural phenotyping of mice." In: *Nature Communications* 1.6 (2010).
- [70] Z. Jiang. et al. "Behavior Recognition in Mouse Videos using Contextual Features Encoded by Spatial-temporal Stacked Fisher Vectors". In: *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods - Volume 1: ICPRAM, 2017*, pp. 259–269.

- [71] J. Jolles, J. Rompa-Barendregt, and W. H. Gispen. "Novelty and grooming behavior in the rat." In: *Behav. Neur. Biol.* 25 (1979), pp. 563–572.
- [72] M. Kabra et al. "Jaaba: interactive machine learning for automatic annotation of animal behavior." In: *Nat Methods* 10 (2013), pp. 64–67.
- [73] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. "Snakes: Active contour models". In: *INTERNATIONAL JOURNAL OF COMPUTER VISION* 1.4 (1988), pp. 321–331.
- [74] Z. Khan, T. Balch, and F. Dellaert. "Mcmc-based particle filtering for tracking a variable number of interacting targets." In: *IEEE Trans. on Pattern Analysis and Machine Intelligence* 27 (2005), pp. 1805–1819.
- [75] W. Kim, Y. B. Cho, and S. Lee. "Thermal Sensor-Based Multiple Object Tracking for Intelligent Livestock Breeding". In: *IEEE Access* 5 (2017), pp. 27453–27463.
- [76] J.M. Koolhaas et al. "The resident-intruder paradigm: a standardized test for aggression, violence and social stress". In: *J Vis Exp* 77 (2013).
- [77] G. Kramida et al. "Automated Mouse Behavior Recognition using VGG Features and LSTM Networks". In: *Proc. Vis. Observ. Anal. Vertebrate Insect Behav. Workshop (VAIB)*. 2016, pp. 1–3.
- [78] L. Lam, S. Lee, and C. Y. Suen. "Thinning methodologies-a comprehensive survey". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14.9 (1992), pp. 869–885.
- [79] B. Latane. "Gregariousness and fear in laboratory rats." In: *J. exp. Social Psychol.* 5 (1969), pp. 61–69.
- [80] V. A. Le and K. Murari. "Recurrent 3D Convolutional Network for Rodent Behavior Recognition". In: *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2019, pp. 1174–1178.
- [81] C. Lee, W. Gao, and P. Lui. "Rat Grooming Behavior Detection with Two-stream Convolutional Networks". In: *2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA)*. 2019, pp. 1–5.
- [82] S. Lee et al. "Comparison of Deep Learning and Image Processing for Tracking the Cognitive Motion of a Laboratory Mouse". In: *2019 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. 2019, pp. 1–4.
- [83] K.R. Lezak, G. Missig, and W.A. Carlezon Jr. "Behavioral methods to study anxiety in rodents". In: *Dialogues Clin Neurosci.* 19.2 (2017), pp. 181–191.
- [84] T. Lin et al. "Microsoft COCO: Common Objects in Context". In: (2014), pp. 740–755.
- [85] S. Liu et al. "Path Aggregation Network for Instance Segmentation". In: (2018), pp. 8759–8768.
- [86] M. Lorbach et al. "Learning to recognize rat social behavior: Novel dataset and cross-dataset application". In: *Journal of Neuroscience Methods* 300 (2018), pp. 166–172.
- [87] Y. Lu, C. Lu, and C. Tang. "Online Video Object Detection Using Association LSTM". In: *2017 IEEE International Conference on Computer Vision (ICCV)*. 2017, pp. 2363–2371.

- [88] B. D. Lucas and T. Kanade. "An iterative image registration technique with an application to stereo vision." In: *Proceedings of the 1981 DARPA Imaging Understanding Workshop*. 1981, pp. 121–130.
- [89] M. W. Mathis and A. Mathis. "Deep learning tools for the measurement of animal behavior in neuroscience". In: *Current Opinion in Neurobiology* 60 (2020). *Neurobiology of Behavior*, pp. 1–11.
- [90] M. Mazur-Milecka. "Thermal imaging in automatic rodent's social behaviour analysis". In: *QIRT*. Gdańsk, 2016, pp. 563–569.
- [91] M. Mazur-Milecka, Tomasz Kocejko, and J. Ruminski. "Deep Instance Segmentation of Laboratory Animals in Thermal Images". In: *Applied Sciences* 10 (2020), p. 5979.
- [92] M. Mazur-Milecka and A. Nowakowski. "An automated system for analysis of mouse movement activity." In: *Biocybernetics and Biomedical Engineering* 30.2 (2010), pp. 55–65.
- [93] M. Mazur-Milecka and A. Nowakowski. "Comparison of tracking methods in respect of automation of an animal behavioral test." In: *Metrology and Measurement Systems* 18.1 (2011), pp. 91–104.
- [94] M. Mazur-Milecka and J. Ruminski. "Automatic analysis of the aggressive behavior of laboratory animals using thermal video processing". In: *IEEE Conference of the Engineering in Medicine and Biology Society, EMBC*. 2017, pp. 3827–3830.
- [95] M. Mazur-Milecka and J. Ruminski. "Deep learning based thermal image segmentation for laboratory animals tracking". In: *Quantitative InfraRed Thermography Journal* (2020), pp. 1–18.
- [96] M. Mazur-Milecka and J. Ruminski. "The analysis of temperature changes of the saliva traces left on the fur during laboratory rats social contacts". In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2018), pp. 2607–2610.
- [97] F. Meyer. "Topographic distance and watershed lines." In: *Signal Processing* 38 (1994), pp. 113–125.
- [98] F. Milletari, N. Navab, and S.-A. Ahmadi. "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation". In: (2016), pp. 565–571.
- [99] N. Mir et al. "Thermal Image Segmentation of Facial Thermograms Using K-Means Algorithm in Evaluation of Orofacial Pain". In: *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAL-CVB)*. 2019, pp. 565–572.
- [100] R. Morris. "Developments of a water-maze procedure for studying spatial learning in the rat." In: *Journal of neuroscience methods* 11.1 (1984), pp. 47–60.
- [101] T. Nath et al. "Using DeepLabCut for 3D markerless pose estimation across species and behaviors". In: *Nature Protocols* 14 (June 2019), 2152–2176.
- [102] N. Nguyen et al. "Applying Deep Learning Models to Mouse Behavior Recognition". In: *Journal of Biomedical Science and Engineering* 12 (2019), pp. 183–196.
- [103] H. Noh, S. Hong, and B. Han. "Learning Deconvolution Network for Semantic Segmentation". In: *abs/1505.04366* (2015), 1520–1528.



- [104] Noldus. *EthoVisionXT*. dostęp 2020.09. URL: <http://www.noldus.com/animal-behavior-research/products/ethovision-xt>.
- [105] S. Ohayon et al. "Automated multi-day tracking of marked mice for the analysis of social behaviour". In: *Journal of Neuroscience Methods* (2013), pp. 10–19.
- [106] K. Oksuz et al. "Localization Recall Precision (LRP): A New Performance Metric for Object Detection". In: *Computer Vision – ECCV 2018*. 2018, pp. 521–537.
- [107] N. Otsu. "A Threshold Selection Method from Gray-Level Histograms". In: *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (1979), pp. 62–66.
- [108] S. M. Pellis and V. C. Pellis. "Play-fighting differs from serious fighting in both target of attack and tactics of fighting in the laboratory rat *Rattus norvegicus*." In: *Aggressive behavior* 13 (1987), pp. 227–242.
- [109] S.M. Peters, H.J. Pothuizen, and B.M Spruijt. "Ethological concepts enhance the translational value of animal models". In: *European Journal of Pharmacology* 759 (2015), pp. 163–170.
- [110] R. Petryniak. "Algorytmy komputerowej detekcji dużych obiektóww obrazach o wysokim poziomie szumui niejednorodności." PhD thesis. Kraków: Akademia Górniczo-Hutnicza, 2011.
- [111] G. Pierelli et al. "Uncoupling Protein 2: A Key Player and a Potential Therapeutic Target in Vascular Diseases". In: *Oxidative Medicine and Cellular Longevity* 2017 (2017).
- [112] H. Pistori et al. "Mice and larvae tracking using a particle filtering with an auto-adjustable observation model." In: *Pattern Recognition Letters* 31 (2010), pp. 337–346.
- [113] Y. Qiao, Z. Wei, and Y. Zhao. "Thermal Infrared Pedestrian Image Segmentation Using Level Set Method". In: *Sensors* 17 (Aug. 2017), p. 1811.
- [114] S. Ren et al. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks". In: 28 (2015), pp. 91–99.
- [115] S. Retana-Marquez, E. D. Salazar, and J Velazquez-Moctezuma. "Effect of acute and chronic stress on masculine sexual behavior in the rat." In: *Psychoneuroendocrinology* 21 (1996), pp. 39–50.
- [116] C. A. Richardson. "The power of automated behavioural homecage technologies in characterizing disease progression in laboratory mice: A review". In: *Applied Animal Behaviour Science* 163 (2015), pp. 19–27.
- [117] A. Rodriguez et al. "ToxTrac: A fast and robust software for tracking organisms". In: *Methods in Ecology and Evolution* 9.3 (2018), pp. 460–464.
- [118] A. Roedel et al. "Effects of light or dark phase testing on behavioural and cognitive performance in DBA mice". In: *Lab Anim.* 40.4 (2006), pp. 371–81.
- [119] F. Romero-Ferrero et al. "idtracker.ai: Tracking all individuals in large collectives of unmarked animals". In: *Nature Methods* 16 (Feb. 2019), pp. 179–182.
- [120] O. Ronneberger, P. Fischer, and T. Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (2015), pp. 234–241.
- [121] A. Rosenfeld and J. L. Pfaltz. "Sequential operations in digital picture processing." In: *ACM* 4.13 (1966), pp. 471–494.

- [122] E. Rosten and T. Drummond. "Fusing Points and Lines for High Performance Tracking." In: *Proceedings of the IEEE International Conference on Computer Vision*. 2005, pp. 1508–1511.
- [123] J. Rousseau, P. vanLochem, and W. Gispen i B. Spruijt. "Classification of rat behavior with an image-processing method and a neural network". In: *Behavior Research Methods, Instruments and Computers* 1 (2000), pp. 63–71.
- [124] A. Schaefer and A. Claridge-Chang. "The surveillance state of behavioral automation." In: *Current Opinion in Neurobiology* 22 (2012), pp. 170–176.
- [125] C. Schmidt, R. Mohr, and C. Bauckhage. "Evaluation of interest point detectors." In: *IJCV* 37.2 (2000), pp. 151–172.
- [126] C. Schuldt, I. Laptev, and B. Caputo. "Recognizing human actions: a local SVM approach." In: *ICPR* (2004), pp. 32–36.
- [127] J. Serra. *Image Analysis and Mathematical Morphology*. Londyn: Academic Press, 1982.
- [128] E. Shelhamer, J. Long, and T. Darrell. "Fully Convolutional Networks for Semantic Segmentation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.4 (2017), pp. 640–651.
- [129] J. Shi and C. Tomasi. "Good features to track." In: *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Seattle, 1994, pp. 593–600.
- [130] BP. Sokolov, CW. Schindler, and JL. Cadet. "Chronic methamphetamine increases fighting in mice". In: *Pharmacol Biochem Behav* 77.2 (2004), pp. 319–26.
- [131] D. Sona et al. "Automated Discovery of Behavioural Patterns in Rodents." In: *Proceedings of Measuring Behavior*. Wageningen, 2014.
- [132] V. H. Sridhar, D. G. Roche, and S. Gingins. "Tracktor: Image-based automated tracking of animal movement and behaviour". In: *Methods in Ecology and Evolution* 10.6 (2019), pp. 815–820.
- [133] P. Steensland et al. "Amphetamine-induced aggression is enhanced in rats pre-treated with the anabolic androgenic steroid nandrolone decanoate". In: *Steroids* 70.3 (2005), pp. 199–204.
- [134] TSE Systems. *Rotametr*. dostęp 2020.09. URL: <http://www.tse-systems.com/products/behavior/motor-function/rotameter.htm>.
- [135] TSE Systems. *Siła uchwytu*. dostęp 2020.09. URL: <http://www.tse-systems.com/products/behavior/motor-function/grip-strength.htm>.
- [136] TSE Systems. *Video Mot2*. dostęp 2020.09. URL: <http://www.tse-systems.com/products/behavior/videotracking/videomot2.htm>.
- [137] Mufford J. T. et al. "The development of a non-invasive behavioral model of thermal heat stress in laboratory mice (*Mus musculus*)". In: *J. Neurosci. Methods* (2016), pp. 189–195.
- [138] C.L. Tan and ZA Knight. "Regulation of Body Temperature by the Nervous System". In: *Neuron* 98.1 (2018), pp. 31–48.
- [139] G. T. Taylor, J. Weiss, and R. Rupich. "Male rat behaviour, endocrinology, and reproductive physiology in a mixed-sex, socially stressful colony." In: *Physiol. Behav.* 39 (1987), pp. 429–433.

- [140] D. Thiessen, M. Pendergrass, and R. K. Young. "Development and expression of autogrooming in the Mongolian gerbil, *Meriones unguiculatus*." In: *J. Comp. Psychol.* 97 (1983), pp. 187–190.
- [141] Z. Tian et al. "FCOS: Fully Convolutional One-Stage Object Detection". In: *Proc. Int. Conf. Computer Vision (ICCV)*. 2019, pp. 9626–9635.
- [142] S. Umapathy, S. Vasu, and N. Gupta. "Computer Aided Diagnosis Based Hand Thermal Image Analysis: A Potential Tool for the Evaluation of Rheumatoid Arthritis". In: *Journal of Medical and Biological Engineering* 38.4 (2018), pp. 666–677.
- [143] E. A. van Dam, L. P.J.J. Noldus, and M. A.J. van Gerven. "Deep learning improves automated rodent behavior recognition within a specific experimental setup". In: *Journal of Neuroscience Methods* 332 (2020), p. 108536.
- [144] L. Vincent and P. Soille. "Watershed in digital space: an efficient algorithm based on immersion simulation." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13.6 (1991), pp. 583–598.
- [145] D. Wahlsten. "Different data from different labs: lessons from studies of gene-environment interaction". In: *Journal of Neurobiology* (2002), pp. 283–311.
- [146] Y. Wu and K. He. "Group Normalization". In: *International Journal of Computer Vision* 128 (July 2019), 742–755.
- [147] Y. Wu et al. *Detectron2*. <https://github.com/facebookresearch/detectron2>. 2019.
- [148] H. Wuerbel. "Behavioral phenotyping enhanced - beyond (environmental) standardization". In: *Genes, brain, and behavior* (2001), pp. 3–8.
- [149] M. Węsierska and K. Turlejski. "Spontaneous behavior of the gray short-tailed opossum (*Monodelphis domestica*) in the elevated plus-maze: comparison with Long-Evans rats." In: *Acta Neurobiol. Exp.* 60 (2000), pp. 479–487.
- [150] X. Xie et al. "Rodent Behavioral Assessment in the Home Cage Using the SmartCage™ System". In: *Animal Models of Acute Neurological Injuries* 2 (Apr. 2012), pp. 205–222.
- [151] Q. Yang and W. Kang. "General Research on Image Segmentation Algorithms". In: *Graphics and Signal Processing* 1 (Oct. 2009), pp. 1–8.
- [152] H. Ying et al. "EmbedMask: Embedding Coupling for One-stage Instance Segmentation". In: (2019). arXiv: [1912.01954](https://arxiv.org/abs/1912.01954).
- [153] C. G. Zhao and T. G. Zhuang. "A hybrid boundary detection algorithm based on watershed and snake." In: *Pattern Recognition Letters* 26 (2005), pp. 1256–1265.

Dodatek A

A. Publikacje związane z tematem pracy

- [1] M. Mazur-Milecka, Tomasz Kocejko, and J. Rumiński. "Deep Instance Segmentation of Laboratory Animals in Thermal Images". *Applied Sciences* 10 (2020), s. 5979. [70 punktów, dla czasopisma według listy MNiSW]
- [2] M. Mazur-Milecka and J. Rumiński. "Deep learning based thermal image segmentation for laboratory animals tracking". *Quantitative InfraRed Thermography Journal* (2020), s. 1–18. [140 punktów, dla czasopisma według listy MNiSW]
- [3] M. Mazur-Milecka and J. Rumiński. "The analysis of temperature changes of the saliva traces left on the fur during laboratory rats social contacts". *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2018), s. 2607–2610. [20 punktów, dla materiałów konferencyjnych według listy MNiSW]
- [4] M. Mazur-Milecka and J. Rumiński. "Automatic analysis of the aggressive behavior of laboratory animals using thermal video processing". *IEEE Conference of the Engineering in Medicine and Biology Society, EMBC. 2017*, s. 3827–3830. [20 punktów, dla materiałów konferencyjnych według listy MNiSW]
- [5] M. Mazur-Milecka. "Thermal imaging in automatic rodent's social behaviour analysis". *QIRT Conference. Gdańsk, 2016*, s. 563–569.
- [6] M. Mazur-Milecka and A. Nowakowski. "Comparison of tracking methods in respect of automation of an animal behavioral test." *Metrology and Measurement Systems* 18.1 (2011), s. 91–104. [100 punktów, dla czasopisma według listy MNiSW]
- [7] M. Mazur-Milecka and A. Nowakowski. "An automated system for analysis of mouse movement activity." *Biocybernetics and Biomedical Engineering* 30.2 (2010), s. 55–65. [100 punktów, dla czasopisma według listy MNiSW]

W recenzji

- [8] M. Mazur-Milecka, J. Rumiński, W. Glac, N. Głowacka. "Detection and model of thermal traces left after aggressive behavior of laboratory rodents".