

mgr inż. Dariusz Grabowski, prof. dr hab. inż. Andrzej Czyżewski, Politechnika Gdańska

Badanie stanu nawierzchni drogowej z wykorzystaniem uczenia maszynowego

W artykule opisano budowę systemu informowania o stanie nawierzchni drogowej z wykorzystaniem metod cyfrowego przetwarzania obrazów oraz uczenia maszynowego. Efektem wykonanych prac badawczych jest eksperymentalna platforma, pozwalająca na rejestrację uszkodzeń na drogach, system do analizy, przetwarzania i klasyfikacji danych oraz webowa aplikacja użytkownika do przeglądu stanu nawierzchni w wybranej lokalizacji.



Fot. 1. Pojazd wyposażony w czujnik światła strukturalnego i kamerę. Żółtą linią zaznaczony emitowany wzorzec [2]

Każdego roku samorządy polskich miast wydają od 4 do 8% [1] swojego budżetu na utrzymanie dróg publicznych, co przekłada się na miliony złotych nakładów na ten cel. Główną przyczyną takiego stanu rzeczy jest stosunkowo szybka degradacja stanu jezdni. Ważnymi czynnikami wpływającymi na uszkodzenia są warunki atmosferyczne, wciąż zwiększające

Determining pavement condition using machine learning

Damage to the road surface is caused by many factors: from atmospheric conditions to high traffic to erosion. Poor road conditions cause damage to vehicles, high fuel consumption and accidents. Investigations of this condition, due to their high costs, are often performed manually. The steps of designing and implementing a system for the automatic identification of road pavement and creating a web application for informing the user about the road condition are presented in the paper. A set of algorithms for processing RGB and depth images was created. A neural network model has been trained and used to classify road defects. The obtained research results show 83% efficiency of using digital image processing in discarding images without any damages. In the case of pavement defects classification, the achieved average efficiency approximated 70%.

Keywords:

image processing, neural networks, pothole detection, Kinect



Fot. 2. Wykorzystywany przez GDDKiA profilograf laserowy [6]



Fot. 3. Sposób montażu systemu stereowizyjnego [7]



Fot. 4. Samochód z zamontowanym sensorem Kinect. Źródło: opracowanie własne

się natężenie ruchu, erozja i inne. Ciągły niedostatek funduszy potrzebnych na remonty dróg pogłębia ten stan. Słaba jakość dróg powoduje uszkodzenia pojazdów (w tym taboru miejskiego), zwiększenie zużycia paliwa oraz kolizje.

Poza koniecznością inwestowania środków istotnym problemem jest też wybór dróg, które powinny być wyremontowane w pierwszej kolejności oraz ocena obecnego stanu poszczególnych odcinków. Generalna Dyrekcja Dróg Krajowych i Autostrad (GDDKiA) co roku dokonuje pomiarów zarządzanych przez siebie jezdni, wykorzystując w tym celu kosztowną aparaturę. Samorządy natomiast, dla których ów sprzęt jest poza zasięgiem, ocenę stanu dróg muszą wykonywać manualnie. Ta powszechnie stosowana metoda jest czasochłonna, kosztowna w realizacji, i obciążona błędem wynikającym z subiektywnego podejścia inspektorów. Zebrane w ten sposób dane należy dodatkowo ujednolicić i przetworzyć, aby wydobyć potrzebne informacje. Pojawia się więc potrzeba automatyzacji opisanego procesu, z uwzględnieniem niskich kosztów wytworzenia i utrzymania systemu.

Rozwiązania oparte o rekonstrukcję 3D

Problem badania stanu nawierzchni nie jest nowym zagadnieniem. Opracowano wiele rozwiązań z wykorzystaniem różnej aparatury. Pierwszym z systemów wartych wspomnienia oparty jest o technikę światła strukturalnego. Na uniwersytecie w Carnegie opracowano platformę do detekcji uszkodzeń nawierzchni. W tym celu wykorzystano laserowy sensor emitujący wzorzec, kamerę rejestrującą obraz w częstotliwości 15 Hz oraz GPS do określenia lokalizacji. Czujniki te zostały zamontowane na pojeździe (fot. 1), który przeskanował wybrany odcinek drogi, poruszając się z prędkością 14,3 m/s. Dzięki temu w procesie analizy udało się wygenerować mapę 3D skanowanej drogi pozwalającą na detekcję uszkodzeń [2].

Innym rozwiązaniem stosowanym m.in. przez GDK-DiA [3] jest profilograf laserowy. Jednym z takich urządzeń jest produkt firmy Dynatest, model 5051 Mk-II Test System. Sprzęt w postaci belki zamontowany na przedzie pojazdu (fot. 2) wyposażony jest w gęsto rozmieszczone

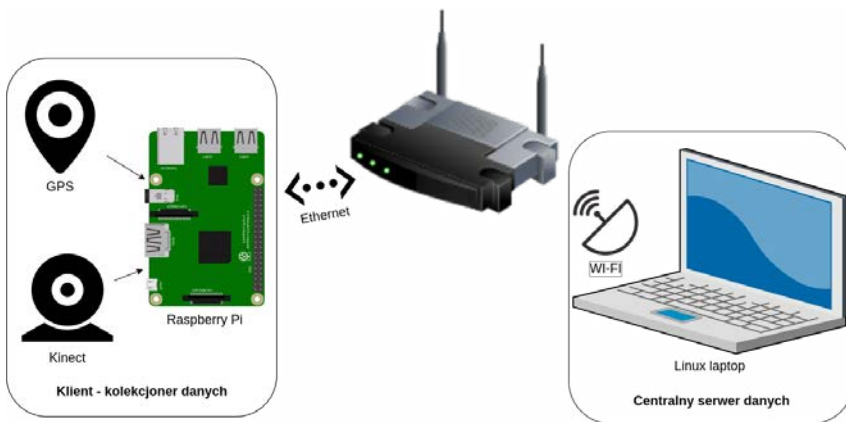
czujniki laserowe. Skanują one powierzchnię pasa jezdni co ok. 5 mm, a dane zapisywane są w pamięci stałej. Test polega na rejestracji maksymalnej wielkości prześwietu pomiędzy uszkodzoną nawierzchnią a poziomo umieszczoną belką [4]. Zwykle pomiary wykonuje się przy prędkości 50 km/h, ale możliwe jest też przeprowadzanie badań w zakresie prędkości od 20 km/h do 110 km/h, bez znacznego pogorszenia wyników [5]. Pozwala to na prowadzenie badań również na trasach szybkiego ruchu czy autostradach bez powodowania utrudnień.

Poza czujnikami laserowymi urządzenie ma na wyposażeniu akcelerometrię, żyroskopy, odbiornik GPS, a także dystansomierz. Pozwala to na badanie podłużnych, jak i poprzecznych profili jezdni, krzywizny łuków itp. [6].

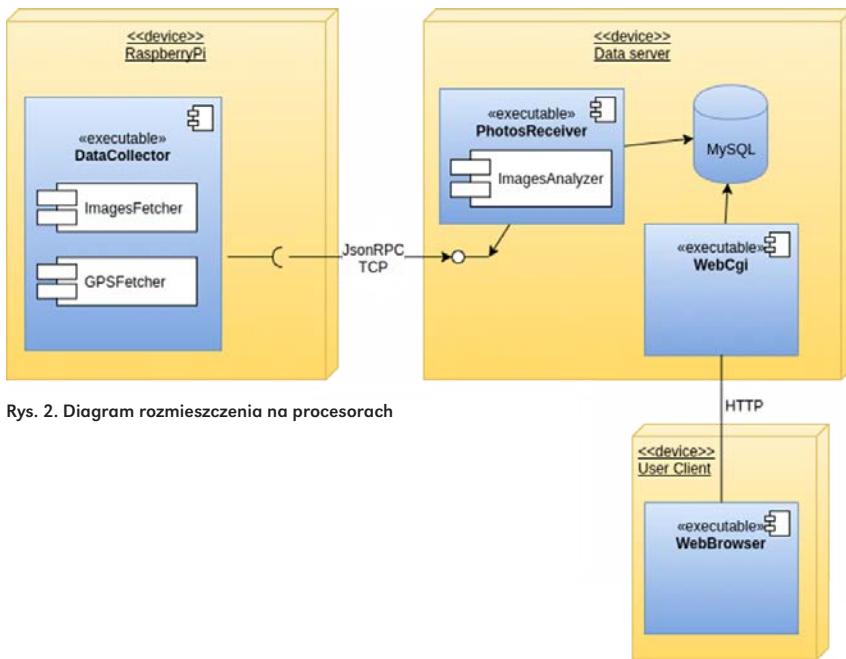
Kolejnym rozwiązaniem, o którym warto wspomnieć, jest system stereowizyjny skonstruowany na Politechnice Śląskiej [7]. Platforma realizuje ocenę stanu nawierzchni na podstawie stereoskopowych obrazów – sekwencyjnie wykonywanych zdjęć z dwóch kamer CCD, podczas przemieszczania się pojazdu (przedstawiono na fot. 3). Dodatkowo wyposażona jest w sensory do identyfikacji przemieszczenia i zmiany kąta obrotu koła samochodu. W procesie analizy z zespołu czujników uzyskiwany jest obraz wraz z informacją dotyczącą głębi. Rozwiązanie to zestawiono i porównano z pomiarem tej samej nawierzchni dalmierzem laserowym dla 160 różnych przekrojów pomiarowych. Wykazano, że różnice pomiędzy metodami mają charakter jedynie przypadkowy, co dowodzi skuteczności działania systemu. Warto wspomnieć, że opracowana metoda może być stosowana niezależnie od warunków pogodowych – odwrotnie niż w przypadku badania laserem, gdy nawierzchnia musi być sucha [7].

System przetwarzania danych o stanie nawierzchni

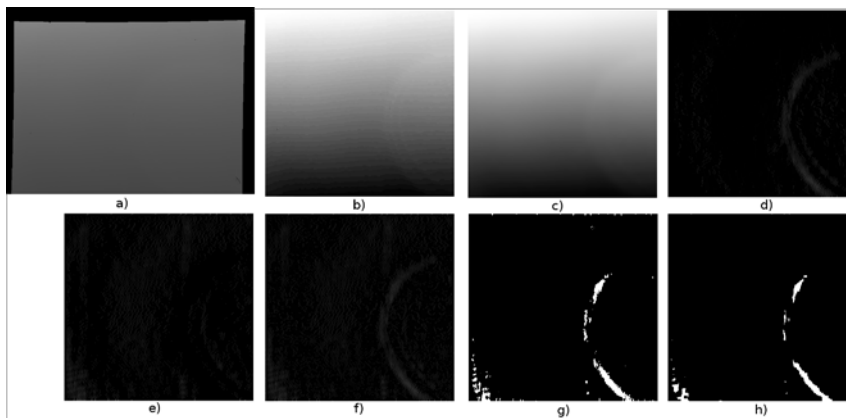
Na podstawie analizy dostępnych rozwiązań oraz występujących problemów zaproponowano rozwiązanie systemu agregującego i przetwarzającego dane o stanie nawierzchni drogowej. Składowe platformy sprzętowej i architektury oprogramowania dobrano w taki sposób, aby uzyskać



Rys. 1. Schemat platformy sprzętowej



Rys. 2. Diagram rozmieszczenia na procesorach



Rys. 3. Kolejne etapy przetwarzania zdjęć głębi



Rys. 4. Przepływ zdjęć głębi przez system: na zielono – elementy znajdujące się w aplikacji klienckiej, na żółto – elementy aplikacji serwerowej

jak najlepszy stosunek ceny i łatwości adaptacji danego rozwiązania do skuteczności w realizacji zamierzonego celu. Głównym założeniem projektowym dla platformy sprzętowej była mobilność rozwiązania. Jako system jezdny wybrano samochód osobowy, stąd konieczne było wyselekcjonowanie takich podzespołów, które mogą być zasilane z jego akumulatora (za pomocą ładowarki lub dołączonej przetwornicy). System przedstawiony na rys. 1 składa się z dwóch głównych elementów:

- Klient – generator danych – umieszczony bezpośrednio na pojeździe moduł odpowiedzialny za zbieranie i wstępne przetwarzanie danych graficznych i pozycji GPS.
- Serwer – strumień danych odpowiedzialny za zbieranie i klasyfikowanie informacji przesyłanych od Klientów.

Ważnym elementem składowym systemu klienckiego do zbierania danych jest sensor RGB-D Kinect v1 wyprodukowany w roku 2010 przez firmę Microsoft. Jest to urządzenie dedykowane do konsoli Xbox 360 tego samego producenta. Jego głównym zadaniem jest generowanie obrazów RGB oraz obrazów odwzorowujących głębie. Czujnik składa się z 3 głównych elementów: kamery RGB, projektora i kamery IR. Ze względu na różną odległość rejestrowanych obiektów od czujnika rzuca wiązka światła IR ulega deformacji. Zawarty w urządzeniu moduł analizuje różnice pomiędzy wyemitowanym a zarejestrowanym światłem IR i na tej podstawie tworzy zdjęcie głębi. Czujnik Kinect zamontowano na masce samochodu osobowego. Częstotliwość pracy sensora pozwalała na jazdę z prędkością ~ 30 km/h. Przygotowany do jazdy testowej pojazd przedstawiono na fot. 4.

Oprogramowanie systemu

System został zaprojektowany w taki sposób, aby mógł pracować w środowisku scentralizowanym. Ideą było

umożliwienie jednoczesnej pracy dla wielu jednostek klienckich zamontowanych na pojazdach. Powinny one mieć możliwość przesyłania w trybie ciągłym zebranych danych do centralnego serwera, w celu ich dalszego przetwarzania. Z drugiej strony konieczne było też udostępnienie interfejsu, dzięki któremu użytkownicy/zarządcy dróg mogliby bez większych opóźnień podglądać aktualny stan jezdni. Architekturę oprogramowania przedstawiono na rys. 2.

Aplikacja ze względu na ograniczenia sprzętowe działa w dwóch trybach – dziennym i nocnym. W pierwszym przypadku używane są jedynie zdjęcia RGB, a w drugim zdjęcia głębi. W ciągu dnia promienie słoneczne w znacznym stopniu zakłócają pracę kamery IR, natomiast kamera RGB po zmroku generuje zdjęcia o bardzo słabej jakości.

Analiza zdjęć głębi

Aplikacja, działając w trybie nocnym, z określoną częstotliwością generuje zdjęcia głębi. Jednym etapów procesu przetwarzania jest analiza tych grafik. Głównym celem realizacji tego zadania jest wyodrębnienie tych zdjęć, które zawierają element zniekształceń drogowych. Takie podejście pozwala ograniczyć ilość danych przesyłanych do serwera. Na podstawie 500 pobranych próbek obrazu głębi opracowano zespół algorytmów, które pozwalają wyekstrahować wartościowe dane z zaszumionego obrazu.

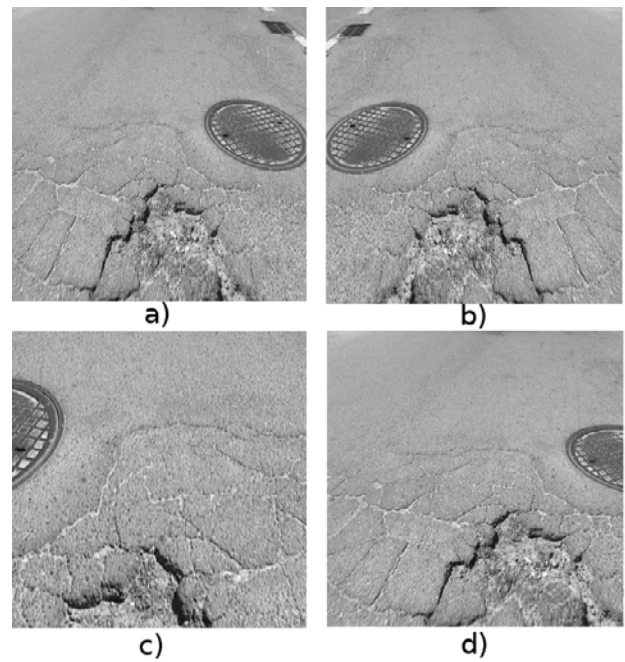
Kolejnymi etapami procesu przedstawionego na rys. 3 są:

- pozyskanie oryginalnego obrazu głębi,
- odcięcie „ramki” i rozciągnięcie histogramu,
- zastosowanie liniowego filtra dolnoprzepustowego typu box z oknem o wielkości 19 x 19 pikseli,
- użycie operatora Sobela z maską zachodnio-wschodnią o wielkości $k = 5$,
- użycie operatora Sobela z maską wschodnio-zachodnią o wielkości $k = 5$,
- połączenie obrazów wynikowych z pkt d) i e),
- binaryzacja obrazu z użyciem wartości stałej $t = 30$,
- zastosowanie dylatacji z maską $k = 3$ oraz dwukrotnej erozji z taką samą maską.

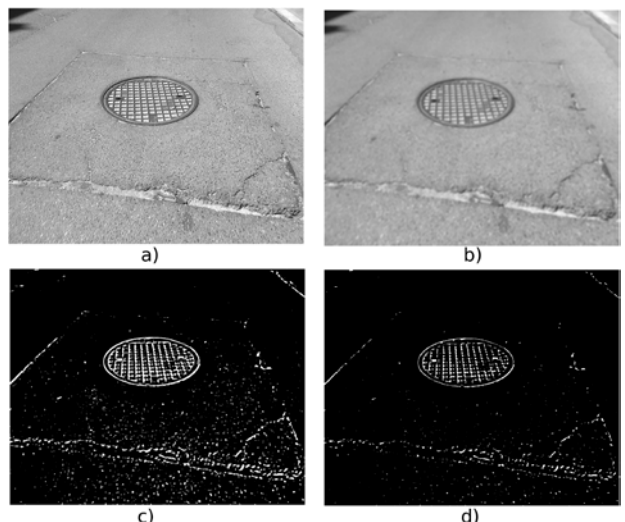
Największym wyzwaniem w powyższej analizie było znalezienie odpowiedniego algorytmu wykrywania krawędzi. Ze względu na kąt natarcia kamery na obrazie można zaobserwować przejście tonalne od ciemnego (na dole) do jasnego (u góry). Większość metod, które dobrze sprawdzały się w przypadku zdjęć RGB, tutaj zawodziły, wykrywając krawędzie pomiędzy różnymi pasami jasności. Jedynie zastosowanie wspomnianych operatorów kierunkowych Sobela umożliwiło rozwiązanie problemu. Przepływ zdjęć głębi przez system przedstawiono na rys. 4.

Wstępna analiza zdjęć RGB

Kolejnym rodzajem zdjęć, dla których konieczna była analiza, są grafiki RGB. W tym przypadku ze względu



Fot. 5. Wybrane transformacje obrazu źródłowego a) obraz oryginalny, b) odbicie lustrzane, c) przybliżenie, d) przesunięcie w prawo



Rys. 5. Kolejne etapy wstępnego przetwarzania obrazu RGB

na moc obliczeniową urządzenia klienckiego zdecydowano się na dokonanie jedynie wstępnej weryfikacji, czy dane zdjęcie może przedstawiać uszkodzenie drogi. Dokładna analiza i klasyfikacja miała, według założenia, odbywać się na serwerze.

Jednym z głównych założeń dla opisywanej analizy był jak najkrótszy czas jej wykonywania, stąd zrezygnowano ze złożonych algorytmów przetwarzania. Na rys. 5 przedstawiono kolejne kroki tego procesu:

- konwersja oryginalnego obrazu do skali szarości,
- zastosowanie liniowego filtra dolnoprzepustowego typu box z oknem o wielkości 11 x 11 pikseli,
- binaryzacja obrazu metodą adaptacyjną Gaussa, z rozmiarem bloku $b = 25$,
- zastosowanie dylatacji z maską $k = 3$ oraz dwukrotnej erozji z taką samą maską.

W drugim kroku przetworzony obraz podlegał ocenie, z jakim prawdopodobieństwem znajdują się na nim istotne informacje. Etap ten zrealizowano poprzez zliczanie białych pikseli i porównywanie z ogólnie przyjętą wartością. Wartość tę wyznaczono metodą doświadczalną, analizując 10 000 zdjęć podzielonych na kategorie: zawierające uszkodzenia i te przedstawiające jedynie jezdnię. Dobry współczynnik pozwala na odrzucenie zdjęć nieistotnych z jednoczesnym zachowaniem zdjęć właściwych. Elementy tego procesu oznaczono kolorem zielonym na rys. 6.

Klasyfikacja obrazów

Kolejnym celem opisywanego systemu jest klasyfikacja zidentyfikowanych na drodze przeszkód (proces oznaczony kolorem żółtym na rys. 6). Ważną informacją dla użytkownika/zarządcy drogi jest to, z jakiego rodzaju uszkodzeniem ma do czynienia. Korzyścią płynącą z klasyfikacji tych obiektów jest możliwość filtrowania w poszukiwaniu interesujących elementów, np. studzienek kanalizacyjnych. Popularnym sposobem na realizację tego typu zadania jest użycie sieci neuronowej. Na podstawie liczego zbioru podzielonych na kategorie danych wejściowych przeprowadza się proces uczenia. Efektem jest identyfikacja przez sieć cech szczególnych obrazów w każdej kategorii. Wykształconą w ten sposób sieć można użyć do klasyfikacji zdjęć spoza zbioru wejściowego. Do realizacji procesu uczenia zastosowano 18-warstwową sieć CNN o nazwie Resnet-18, opracowaną przez firmę Microsoft. Podstawowym parametrem przy wyborze tego rozwiązania był stosunek czasu uczenia i dokonywania predykcji do średniego poziomu błędów.

Uczenie sieci neuronowej

Aby przystąpić do procesu uczenia sieci neuronowej, należy przygotować zbiór zdjęć podzielonych na kategorie. Im liczniejszy i bardziej różnorodny jest ów zbiór, tym lepsze wyniki uczenia można uzyskać. W związku z tym, w trakcie opracowywania materiału zebrano 10 000 zdjęć jezdni, które wyselekcjonowano, a następnie bazując na klasyfikacji GDDKiA [3], podzielono je w sposób manualny na kategorie:

- wyboje – miejscowe ubytki w warstwach bitumicznych,
- studzienka kanalizacyjna,

Zdjęcia słusznie odrzucone	Zdjęcia niesłusznie odrzucone
83,80%	21,67%

Tab. 1. Wyniki testów wstępnej selekcji zdjęć

Kategoria uszkodzenia	Odsetek TP	Odsetek FN
Spękania siatkowe i pojedyncze	93,02%	2,79%
Studzienka kanalizacyjna	83,78%	0,36%
Łaty wbudowane i nałożone	73,79%	2,18%
Wyboje	29,66%	0,13%

Tab. 2. Wyniki klasyfikacji zdjęć z uszkodzeniami

- spękania siatkowe i pęknięcia pojedyncze,
- łaty wbudowane i nałożone – lokalne obszary naprawione przy pomocy dodatkowej warstwy bitumicznej,
- droga bez uszkodzeń.

W przypadku mniej liczego zbioru (wyboje), w ramach uzupełnienia pobrano zdjęcia z Google Images. Aby uzyskać jeszcze lepsze wyniki, postanowiono dokonać kilku prostych przekształceń na zebranych grafikach, powiększając w ten sposób materiał wejściowy i zwiększając wydajność procesu uczenia. Przykłady zamieszczono na fot. 5, natomiast w praktyce wykorzystano różne kombinacje podanych technik.

Na potrzeby uczenia sieci konieczne było przygotowanie dwóch zestawów zdjęć – pierwszy zawierający zbiór danych uczących w wysokości 90% wszystkich zdjęć. Pozostałe 10% stanowiło zawartość drugiego ze zbiorów weryfikacyjnym.

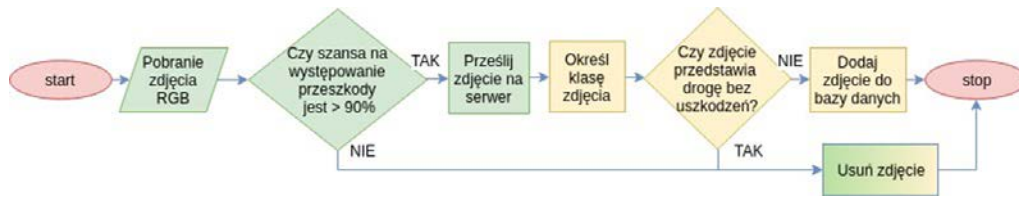
Ostateczny model wygenerowano przy pomocy chmury obliczeniowej Amazon Web Services. Do wykonania zadania wybrano maszynę o nazwie p2.xlarge wyposażoną w procesor GPU NVIDIA K80 z 12 GB pamięci, wspierający technologię CUDA. Do procesu uczenia wykorzystano 4500 zebranych wcześniej zdjęć. Całkowity czas przetwarzania wyniósł 54 min 13 sek. Postęp procesu dla danych testowych i uczących przedstawiono na rys. 7. Końcowy rezultat uczenia dla danych testowych wyniósł 94%.

Aplikacja użytkownika

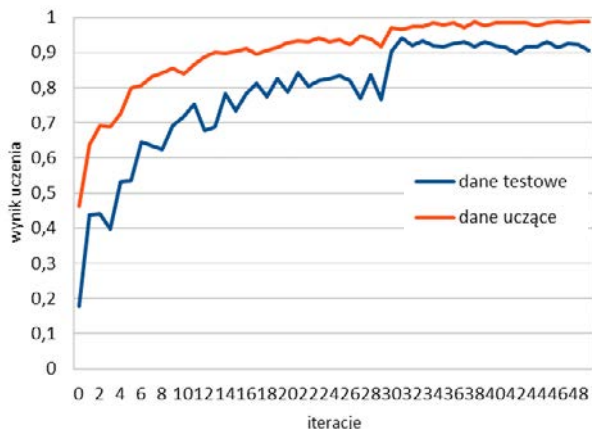
Interfejs użytkownika (rys. 8) podzielony jest na dwie główne części. Po prawej stronie znajduje się mapa z zaznaczonymi w postaci markerów uszkodzeniami dróg. Zielone ikony oznaczają zdjęcia głębi, czerwone zaś zdjęcia RGB. Kategorie uszkodzeń zarejestrowanych za pomocą zdjęć kolorowych oznaczone są kolejnymi literami alfabetu od ‘A’ do ‘D’. Po kliknięciu na wybrany marker pojawia się okno prezentujące wykonane w tym miejscu zdjęcie. W lewym górnym rogu wyświetlana jest data jego wykonania. Interfejs umożliwia także ustawienie zakresu dat, z jakich mają się wyświetlać dane oraz ukrywanie/pokazywanie poszczególnych kategorii zdjęć.

Testy analizy zdjęć

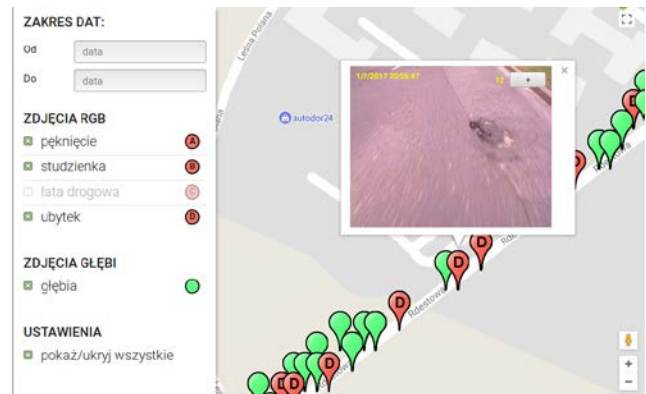
Pierwszy z przeprowadzonych testów dotyczył wstępnego przetwarzania zdjęć RGB na urządzeniu klienckim. Określony w aplikacji współczynnik określał, z jakim prawdopodobieństwem na obrazie znajduje się przeszkoda. Test polegał na uruchomieniu analizy dla reprezentatywnej grupy 1000 zdjęć zawierających i niezawierających uszkodzeń. Jak przedstawiono w tab. 1, oprogramowanie skutecznie eliminuje zdjęcia bez defektów, zmniejszając tym samym ilość koniecznych do przesłania na serwer danych. Warto zaznaczyć, że wartość wspomnianego współczynnika można zdefiniować w pliku konfiguracyjnym, dostosowując czułość algorytmu do potrzeb.



Rys. 6. Przepływ zdjęć RGB przez system: na zielono – elementy znajdujące się w aplikacji klienckiej, na żółto – elementy aplikacji serwerowej. Gradientem oznaczono część występującą w każdej z aplikacji.



Rys. 7. Wyniki uczenia w kolejnych iteracjach dla danych testowych i uczących



Rys. 8. Interfejs webowej aplikacji użytkownika po kliknięciu na znacznik

Kolejny przeprowadzony test dotyczył skuteczności klasyfikacji przesłanych do serwera zdjęć RGB. Ideą było sprawdzenie, jaki odsetek zdjęć został prawidłowo zakwalifikowany do swojej kategorii i jaki jest poziom błędów. Sprawdzeniu podlegały klasy zawierające defekty. Kategorię „droga bez uszkodzeń” potraktowano jako zdjęcia pozytywnie niezakwalifikowane. Do przeprowadzenia testów użyto 2400 zdjęć, które nie brały udziału w procesie uczenia sieci neuronowej. Badania wykazały, że średnio w 70% przypadków zdjęcie zostaje przypisane do odpowiedniej kategorii. Szczegółowe wyniki przedstawiono w tab. 2. Wyjaśnienia wymaga stosunkowo niski wynik dla kategorii „Wyboje”. Głównym powodem takiego stanu rzeczy jest niewielki zbiór danych uczących dla tej kategorii. Próba uzupełnienia zdjęciami z Google Images nie przyniosła oczekiwanego rezultatu. Rozwiązaniem byłoby zwiększenie zbioru o dane pobrane z tego samego czujnika. Wykluczając tę kategorię, globalny średni wynik skuteczności wynosi 83,5%.

Podsumowanie

W ramach opisywanego systemu opracowano eksperymentalny model sprzętowy rejestratora danych. W kolejnym kroku stworzono zespół algorytmów do przetwarzania zdjęć głębi oraz zdjęć RGB. Następnie na podstawie zebranego materiału wejściowego wyuczono model sieci neuronowej służący do klasyfikacji defektów drogowych. Finalnie stworzono oprogramowanie i przeprowadzono testy systemu pod kątem skuteczności i niezawodności działania. Przeprowadzone eksperymenty badawcze wykazały 83% skuteczność użycia cyfrowego przetwarzania obrazów

w odrzucaniu zdjęć niezawierających uszkodzeń. W przypadku rozpoznawania uszkodzeń średnią skuteczność wyniosła 70%. Testy systemu potwierdziły zasadność wyboru zastosowanych mechanizmów. Niewątpliwie zastosowanie uczenia maszynowego w badaniu stanu nawierzchni jest obiecującym kierunkiem rozwoju metod i technologii. Przy wsparciu tanich sensorów lub kamer możliwe jest uzyskanie satysfakcjonujących rezultatów. W procesie uczenia sieci neuronowej uzasadnione jest zastosowanie wydajnych procesorów graficznych. Należy przy tym zwracać uwagę na odpowiednią liczebność i różnorodność zbiorów uczących, aby uzyskać zadowalające efekty. □

Badania dofinansowane w ramach programu „Rozwój Innowacji Drogowych”, współfinansowanego przez NCBR i GDDKiA, nr projektu: OT4-4B/AGH-PG-WSTKT.

Piśmiennictwo

1. Bank Danych Lokalnych: <https://bd.l.stat.gov.pl>, (data dostępu 17.09.2017 r.).
2. Mertz C.: *Continuous Road Damage Detection Using Regular Service Vehicles*. „Proceedings of the ITS World Congress”, 2011, s. 1, 5-8.
3. IBDiM: *Tom 1 Organizacja i standardy kampanii diagnostyki stanu technicznego nawierzchni. Część 4: Wymagane techniki pomiarowe*. Generalna Dyrekcja Dróg Krajowych i Autostrad, 2012, s. 8,15-24
4. Staniek M.: *Metody oceny stanu nawierzchni sieci drogowej*. „Zeszyty Naukowe Politechniki Śląskiej”, 2011.
5. Szpinek S.: *Statystyczna metoda rozpoznawania odcinków jednorodnych. Część II. Przykłady zastosowań w badaniach równości nawierzchni*. „Drogownictwo”, 2012, s. 359.
6. Godlewski D., Szpinek S.: *Diagnostyka nawierzchni jezdni w miejscach zdarzeń drogowych*. „Drogownictwo”, 2015, s. 363.
7. Staniek M.: *Ocena dokładności stereowizyjnej metody inspekcji dróg*. „Budownictwo i Architektura”, 2014, s. 259, 261, 265.