

ANALIZA ISTOTNOŚCI CECH ZNAMION SKÓRNYCH DLA CELÓW DIAGNOSTYKI CZERNIAKA ZŁOŚLIWEGO

Agnieszka MIKOŁAJCZYK¹, Michał GROCHOWSKI²

1. Politechnika Gdańska, Wydział Elektrotechniki i Automatyki
tel.: 58 347 17 42 e-mail: agnieszka.mikolajczyk@pg.edu.pl
2. Politechnika Gdańska, Wydział Elektrotechniki i Automatyki
tel.: 58 347 29 04 e-mail: michal.grochowski@pg.edu.pl

Streszczenie: Pomimo dynamicznego rozwoju metod uczenia maszynowego i ich wdrażania do praktyki lekarskiej, automatyczna analiza znamion skórnych wciąż jest nierozwiązanym problemem. Poniższy artykuł proponuje zastosowanie algorytmu ewolucyjnego do zaprojektowania, wytrenowania i przetestowania całych populacji klasyfikatorów (sztucznych sieci neuronowych) oraz ich iteracyjnego udoskonalania w każdej kolejnej populacji, w celu osiągnięcia jak najlepszej dokładności klasyfikacji znamion skórnych. Algorytm zwraca optymalny zestaw cech opisujących obraz dermatoskopowy wraz z proponowaną architekturą sieci neuronowej. Uzyskano dokładność równą 85,83%, swoistość równą 79,07% oraz czułość równą 92,60%.

Słowa kluczowe: algorytm ewolucyjny, uczenie maszynowe, sieci neuronowe, systemy wspomaganie decyzji

1. WPROWADZENIE

Rocznie, na terenie samych Stanów Zjednoczonych, czerniak złośliwy pojawia się u około 92 000 ludzi, z czego około 9500 osób umiera [1]. Duża umieralność jest często spowodowana zbyt późnym wykryciem choroby - wcześniej postawiona diagnoza w większości przypadków ratuje ludzkie życie. Różnego rodzaju zmiany skórne są często spotykane zarówno u młodych, jak i u starszych osób, a do postawienia diagnozy wciąż jest potrzebny doświadczony lekarz zaznajomiony z metodami dermatoskopowymi.

Metody dermatoskopowe stanowią opracowane przez dermatologów i onkologów algorytmy pozwalające na sprawne odróżnianie znamion skórnych o charakterze złośliwym od łagodnych zmian skórnych. Skuteczność działania tych metod jest szacowana na około 70% [2], jednak podejmowane przez lekarzy decyzje nie są wyłącznie poparte wiedzą i znajomością kryteriów ale także intuicją.

Analiza znamion skórnych pod kątem występowania cech opisanych w ww. metodach jest bardzo żmudnym i czasochłonnym procesem. Ostatnie lata przyniosły gwałtowny rozwój systemów służących wspomaganie pracy lekarzy. Szczególne osiągnięcia w tej dziedzinie należą do sztucznych sieci neuronowych wspieranych przez zaawansowane metody przetwarzania i analizy obrazów.

Od 2016 roku, Międzynarodowe Stowarzyszenie w Obrazowaniu Skóry (ang. *the International Skin Imaging Collaboration*) corocznie przeprowadza wyzwanie, polegające na stworzeniu systemu wspomagającego automatyczną analizę znamion skórnych pod kątem wykrycia czerniaka złośliwego [3]. Pierwsza edycja w roku

2016 zakończyła się otrzymanie 79 zgłoszeń, z czego najlepszy system osiągnął czułość równą 50,5% oraz swoistość równą 92,0% [4]. W roku 2017 wpłynęło 46 zgłoszeń, a uzyskana przez najlepszy system czułość pobiła rekord z poprzedniego roku i wynosiła 54,7%, natomiast swoistość 95,0% [5].

2. CZERNIAK ZŁOŚLIWY

2.1. Podstawy dermatoskopii

Najpopularniejszymi metodami różnicowania znamion o charakterze łagodnym od czerniaka złośliwego są: kryterium ABCD Stoltza, lista siedmiopunktowa, metoda Menziesa [6]. W kryterium oceny dermatoskopowej ABCD oceniane są następujące cechy zmiany skórnej: symetria znamienia (A - *Assymetry*), krawędź (B - *Border*), kolor (C - *Color*), struktura (D - *Differential Structures*). Cecha A oceniana jest przez przyzmat symetrii w jednej oraz dwóch osiach; oceny krawędzi B polega na podziale znamienia na 8 równych części i ocenę ciągłości jej krawędzi w każdym fragmencie; ocena cechy C wymaga zliczenia występujących na znamieniu kolorów (jasnobrązowy, ciemnobrązowy, czarny, niebieski, czerwony, biały); ostatnia cecha D wymaga odnalezienia i zliczenia struktur występujących na powierzchni znamienia (kropek, smug gałzkowatych, pseudopodium, plam bezstrukturalnych) [6-8]. Po szczegółowej analizie każdej z cech, przydzielane są punkty, które są następnie sumowane z odpowiadającymi im wagami i zliczane. Jeśli suma punktów przekroczy podany w metodzie próg, oznacza to że znamie jest złośliwe. Pomimo przydzielanych poszczególnym cechom wag, które sugerują ich ważność w procesie klasyfikacji, różne badania sugerują skupienie się na innych parametrach. Przykładowo, w [9] autorzy tłumaczą, że jedynymi z najważniejszych cech są: różnorodność kolorystyczna na powierzchni znamienia lub ich nierównomierny rozkład oraz zmieniająca się struktura znamienia. Z kolei w metodzie ABCD punktacja sugeruje, że asymetria, struktura i kolor mają podobny stopień ważności, jednak zdobycie dużej liczby punktów we wskaźniku A jest statystycznie najłatwiejsze. W niniejszym artykule przeanalizowano znaczenie poszczególnych cech znamion skórnych na jakość późniejszej klasyfikacji.

2.2. Baza danych

Zaproponowany w artykule algorytm selekcji cech znamion skórnych został przetestowany na ogólnodostępnej

bazie danych zdjęć dermatoskopowych (ang. *International Skin Imaging Collaboration*) zawierającej ponad 13 000 rekordów [7]. Obrazy zawarte w bazie danych zostały przygotowane i opracowane przez lekarzy specjalistów.

2.3. Wstępne przetwarzanie obrazu

Przed ekstrakcją cech charakterystyczną dla złośliwych zmian skórnych przeprowadzono wstępne przetwarzanie obrazu. Ten etap miał na celu całkowite lub częściowe usunięcie wybranych elementów z obrazu: usunięcie czarnych ramek, szumu, eliminację owłosienia, a następnie wydzielenie znamienia skórnoego z tła którym jest skóra.

Metoda redukcji owłosienia zastosowana przez autorów została opisana w [10] oraz przez innych badaczy m.in. w [9,10], metody segmentacji znamienia zostały opisane w [13], inne metody zostały zaproponowane przez szereg badaczy [14–16].

2.4. Metoda ekstrakcji cech

Opis obrazu przy pomocy wskazania charakterystycznych cech jest bardzo istotnym i złożonym procesem, którego celem jest ułatwienie podjęcia decyzji. Skutecznie przeprowadzona zmiana reprezentacji danych z przestrzeni pikseli opisujących obraz do przestrzeni cech umożliwia znaczne zmniejszenie liczności danych, które muszą zostać przekazane na wejście klasyfikatora, a co za tym idzie, zwiększenie przejrzystości procesu decyzyjnego. To podejście wykorzystano także w niniejszej publikacji: fotografie dermatoskopowe zostały przeniesione do przestrzeni cech, opisującej sfotografowane znamię. Cechy wybrano na podstawie metod i kryteriów stosowanych do oceny dermatoskopowej. Każde znamię zostało opisane przy pomocy 58 cech, w tym: dwie cechy reprezentujące kształt, 9 cech opisujących kolor oraz 47 cech opisujących strukturę znamienia (tabela 1).

Tablica 1. Metody ekstrakcji cech

Rodzaj cechy/metoda ekstrakcji cech	ID cechy
Kształt	
Symetria	1
Krawędź	2
Kolor	
Maksymalna zmiana koloru	3
Powierzchnia znamienia z dominującym kolorem czerwonym/niebieskim/zielonym	4 – 6
Maksymalna zmiana jasności	7 – 9
Struktura	
Detekcja narożników filtrem konwolucyjnym	10 – 12
Metoda minimalnych wartości własnych (ang. <i>Minimum eigenvalue algorithm</i>) [17]	13 – 15
Metoda maksymalnie stabilnych regionów (MSER) [18]	16 – 18
Algorytm FAST [19]	19– 21
Algorytm BRISK [20]	22 – 24
Algorytm Harrisa–Stephensa [21]	25 – 27
Metoda SURF [22]	28 – 30
Wykrywanie cech HOG [23]	31 – 33
Wykrywanie kół, płam i kropli	34 – 58

2.5. Normalizacja danych

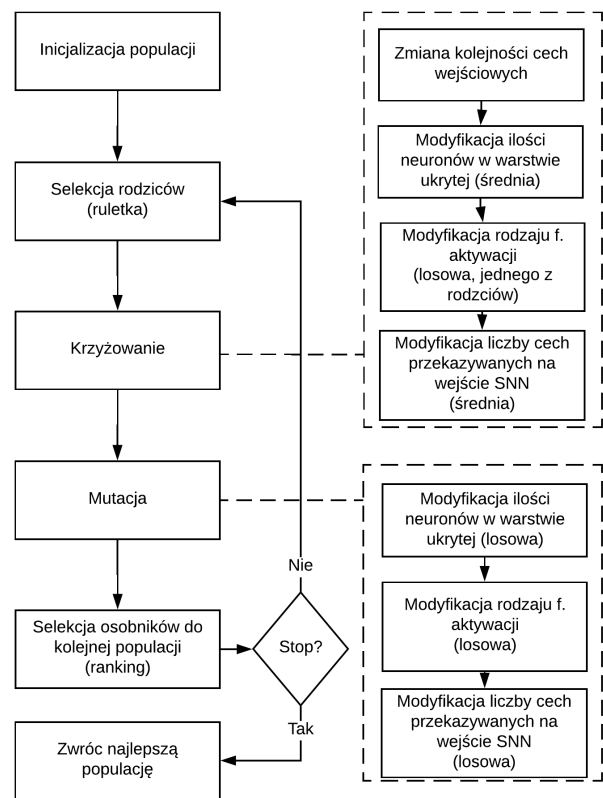
Każda z wyznaczonych cech została znormalizowana celem uzyskania jednakowego rzędu wielkości cech. Cechy opisujące kolor znamienia zostały podzielone przez maksymalną wartość koloru w przestrzeni RGB (255). Każda z cech opisujących strukturę znamienia została podzielona przez powierzchnię znamienia. Dodatkowo, w

celu usunięcia szumu, od każdej cechy odjęto medianę danej cechy (mediana została obliczona na zbiorze treningowym).

3. ALGORYTM EWOLUCYJNY

3.1. Idea systemu

Liczba, rodzaj a także jakość cech podawanych na wejście sieci neuronowej ma duży wpływ na możliwą do osiągnięcia przez sieć dokładność. Zwiększanie liczby podawanych na sieć cech nie zawsze idzie w parze ze wzrostem dokładności - część danych może być ze sobą skorelowana, a korelacja nie zawsze oznacza związek przyczynowy. Powyższe problemy, a także wymagana ekspercka wiedza w zakresie konkretnego aplikacyjnego problemu sprawia, że dobór odpowiednich cech, które powinny być przekazane sieci neuronowej jest skomplikowanym zadaniem. W niniejszym artykule autorzy proponują automatyczny dobór cech, przekazywanych na wejście sieci neuronowej, a także częściowy dobór architektury sieci: dobór funkcji aktywacji, liczby neuronów w warstwie ukrytej, liczby wejść przy pomocy algorytmu neuro-ewolucyjnego. Zdecydowano się na podejście z kodowaniem pośrednim (ang. *indirect encoding*) - w chromosomie zakodowana jest jedynie część informacji, a sam główny szkielet SSN pozostaje taki sam [24].



Rys. 1. Schemat ideowy działania algorytmu ewolucyjnego

W pierwszym kroku przeprowadzane jest wstępne przetwarzanie obrazu oraz segmentacja znamienia z tła, którym jest skóra. Następnie, wyekstrahowano 58 cech opisujących obraz przy pomocy metod przetwarzania obrazu. Wszystkie cechy są normalizowane zgodnie z przyjętą metodologią opisaną w sekcji 2.5. Przygotowane w ten sposób dane będą stanowić zbiór treningowy, walidacyjny i testowy dla projektowanych sieci neuronowych. Zadaniem zastosowanego algorytmu ewolucyjnego jest zaprojektowanie, trenowanie i testowanie

całych populacji SNN oraz ich iteracyjne poprawianie w każdej kolejnej populacji, w celu osiągnięcia jak najlepszej dokładności działania. Taki sposób doboru cech oraz architektury sieci eliminuje czynnik ludzki, ułatwia proces projektowania SNN, nie wymaga wiedzy eksperckiej oraz umożliwia analizę istotności poszczególnych cech.

3.2. Kodowanie chromosomu

Każdy z chromosomów można podzielić na dwie części: część opisującą architekturę sieci oraz opisującą kolejność cech. Pierwsze trzy pola chromosomu definiują: liczbę wejść SNN (od 10 do 58), liczbę neuronów w warstwie ukrytej (od 5 do 30), rodzaj funkcji aktywacji (1-*lin*, 2-*elliot2sig*, 3-*radbas*, 4-*tansig*, 5-*logsig*, 6-*satlin*, 7-*satlins*, 8-*softmax*).

Od czwartego pola chromosomu rozpoczyna się druga część chromosomu opisująca kolejność cech, które wejdą na SNN. Kolejność cech w chromosomie może być dowolna, jednak żadna z cech nie może się powtórzyć oraz żadna z cech nie może zostać pominięta. Każda z cech posiada swój unikatowy numer identyfikacyjny (numer ID – od 1 do 58), który został nadany w celu możliwości identyfikacji poszczególnych cech w chromosomie. Pierwsze n cech, gdzie n to liczba wejść sieci neuronowej podana w polu pierwszym, zostaną przekazane na wejście SSN.

3.3. Operator krzyżowania i mutacji

Operatory krzyżowania i mutacji różnią się od siebie w zależności od części chromosom, na którym pracują. W jego pierwszej części, w której zdefiniowane są: liczba wejść SNN oraz liczba neuronów w warstwie ukrytej przeprowadzane jest uśrednianie poszczególnych parametrów pomiędzy rodzicami z zaokrągleniem w górę. Trzecie pole, rodzaj funkcji aktywacji, jest losowo dziedziczone po jednym z rodziców.

W drugiej części chromosomu przedstawiony jest ciąg numerów ID cech. W pierwszych 20 iteracjach algorytmu zamienianych jest 10 par numerów ID w celu przeszukania większej przestrzeni rozwiązań. W kolejnych krokach zamieniana jest kolejnością jedynie jedna para numerów ID.

Mutacja polega na zmianie losowo wybranego parametru lub parametrów z chromosomu na losową wartość z dozwolonego zakresu. Prawdopodobieństwo krzyżowania wyniosło 70%, mutacji 10%.

3.4. Selekcja

Proces selekcji rodziców jest procesem wyboru osobników z obecnej populacji, którzy zostaną rodzicami dla populacji przyszłej. Zastosowano popularny operator ruletki jako operator selekcji, natomiast jako operator selekcji przetrwania wybrano metodę rankingową z dodatkowym zachowaniem najlepszych osobników z populacji przed wyborem rodziców.

3.5. Ocena funkcji przystosowania

Do każdego osobnika jest przypisany chromosom, w którym jest zakodowana informacja potrzebna do budowy sieci neuronowej. Jakość zaprojektowanej sieci neuronowej m jest przeprowadzana przy pomocy oceny funkcji przystosowania. Dla jednego chromosomu jest tworzonych 5 kopii takiej samej sieci z różnymi losowo dobranymi wagami początkowymi. W celu oceny działania zaprojektowanej sieci, każda z nich jest trenowana na zbiorze treningowym i walidacyjnym, a potem testowana na zbiorze testowym. Dla każdej z 5 kopii sieci przypisanych do aktualnie badanego chromosomu jest obliczania średnia

ważona ze średniej dokładności (waga $w_1=0,3$) oraz najlepszy wynik dokładności dla wszystkich kopii (waga $w_2=0,7$). Uzyskany wynik procentowy jest zapisywany jako wartość funkcji przystosowania. Dodatkowo wprowadzono funkcję kary, której zadaniem jest zmniejszenie prawdopodobieństwa wystąpienia wysoko niezbilansowanych różnic w wartościach swoistości i czułości: jeśli czułość jest niższa niż 75%, wartość funkcji przystosowania jest obniżana o 20 punktów; jeżeli czułość jest niższa niż 55 wartość funkcji jest obniżana o 50 punktów. Kary dla niskiej swoistości są analogiczne.

4. WYNIKI

Przy pomocy opisanej metody wygenerowano 10 różnych propozycji topologii sieci neuronowych wraz z listą wybranych cech. Ustalono maksymalną liczbę generacji równą 300. Każda populacja składała się z 50 chromosomów. Każdy chromosom był inicjalizowany losowymi wartościami. Prawdopodobieństwo krzyżowania wyniosło 70%, a mutacji 10%.

Tablica 2. Wyniki

Nr sprawdzianu	Dokładność	Swoistość	Czułość
1	86,00%	79,00%	93,00%
2	86,00%	79,00%	93,00%
3	86,00%	79,00%	93,00%
4	86,00%	78,00%	94,00%
5	86,00%	80,00%	92,00%
Średnia	85,90%	79,40%	92,40%
MAKS.	86,00%	80,00%	94,00%

Wszystkie topologie porównano ze sobą, badając je podczas 5-krotnego sprawdzianu krzyżowego. Wyniki uzyskane dla najlepszej topologii przedstawiono w tabeli 2.

Najlepsza topologia posiadała 29 wejść. Wejścia najlepszej sieci neuronowej stanowiły cechy o numerach identyfikacyjnych: 1, 2, 4, 7, 9, 10, 11, 12, 14–17, 19, 23, 24, 26, 28, 31–34, 38, 42–44, 47, 48, 55, 58. Dobrano 9 neuronów w warstwie ukrytej oraz funkcję aktywacji *tansig*.

Wśród wszystkich topologii najczęściej pojawiały się takie cechy jak: symetria, kontur, maksymalna zmiana jasności oraz struktury wykryte algorytmami: koewolucyjna detekcja narożników, metoda minimalnych wartości własnych, metoda maksymalnie stabilnych regionów, algorytm BRISK, algorytm Harrisa–Stephensa, wykrywanie cech HOG, wykrywanie kół, plam i kropli. Ponadto, podobieństwo pomiędzy 3 najlepszymi topologiami jest szacowane na około 70%.

5. WNIOSKI KOŃCOWE

Rozwój dermatoskopii, metod przetwarzania i analizy obrazu oraz metod uczenia maszynowego spowodowały jednoczesny wzrost zainteresowania rozszerzaniem metod automatycznej analizy i diagnostyki znamion skórnych przez naukowców. W celu dobrania optymalnego zestawu cech opisującego obraz autorzy zaproponowali wykorzystanie algorytmu ewolucyjnego. Ponadto, algorytm jednocześnie dobierał część parametrów sieci neuronowej. Uzyskano wysokie wyniki końcowe: dokładność równą 85,83%, swoistość równą 79,07%, czułość równą 92,60% [25].

6. BIBLIOGRAFIA

1. Key Statistics for Melanoma Skin Cancer n.d. <https://www.cancer.org/cancer/melanoma-skin-cancer/about/key-statistics.html> (accessed March 16, 2018).
2. Annessi G, Bono R, Sampogna F, Faraggiana T, Abeni D. Sensitivity, specificity, and diagnostic accuracy of three dermoscopic algorithmic methods in the diagnosis of doubtful melanocytic lesions: the importance of light brown structureless areas in differentiating atypical melanocytic nevi from thin melanomas. *Journal of the American Academy of Dermatology* 2007;56:759–767.
3. ISIC Archive n.d. <https://isic-archive.com/> (accessed January 24, 2018).
4. Bi L, Kim J, Ahn E, Feng D, Fulham M. Automatic melanoma detection via multi-scale lesion-biased representation and joint reverse classification. *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on, IEEE; 2016, p. 1055–1058.*
5. Codella NC, Gutman D, Celebi ME, Helba B, Marchetti MA, Dusza SW, et al. Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). *ArXiv Preprint ArXiv:171005006* 2017.
6. Jorh RH. Dermoscopy: alternative melanocytic algorithms—the ABCD rule of dermoscopy, menzies scoring method, and 7-point checklist. *Clinics in Dermatology* 2002;20:240–247.
7. Nachbar F, Stolz W, Merkle T, Cagnetta AB, Vogt T, Landthaler M, et al. The ABCD rule of dermoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. *Journal of the American Academy of Dermatology* 1994;30:551–559.
8. Henning JS, Dusza SW, Wang SQ, Marghoob AA, Rabinovitz HS, Polsky D, et al. The CASH (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy. *Journal of the American Academy of Dermatology* 2007;56:45–52.
9. Jamil U, Khalid S, Akram MU. Dermoscopic feature analysis for melanoma recognition and prevention. *Innovative Computing Technology (INTECH), 2016 Sixth International Conference on, IEEE; 2016, p. 290–295.*
10. Mikołajczyk A, Wykorzystanie metod inteligencji obliczeniowej dla celów diagnostyki i klasyfikacji stanów anormalnych wybrane zastosowania w medycynie. Praca dyplomowa 115272, Politechnika Gdańska, 2017
11. Kang D, Kim S, Park S. Flow-guided hair removal for automated skin lesion identification. *Multimedia Tools and Applications* 2018;77:9897–9908.
12. Abbas Q, Garcia IF, Emre Celebi M, Ahmad W. A Feature-Preserving Hair Removal Algorithm for Dermoscopy Images. *Skin Research and Technology* 2013;19.
13. Mikołajczyk A, Kwasigroch A, Grochowski M. Intelligent system supporting diagnosis of malignant melanoma. *Polish Control Conference, Springer; 2017, p. 828–837.*
14. Attia M, Hossny M, Nahavandi S, Yazdabadi A. Skin melanoma segmentation using recurrent and convolutional neural networks. *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on, IEEE; 2017, p. 292–296.*
15. Zeng G, Zheng G. Multi-scale Fully Convolutional DenseNets for Automated Skin Lesion Segmentation in Dermoscopy Images. *International Conference Image Analysis and Recognition, Springer; 2018, p. 513–521.*
16. Yuan Y, Chao M, Lo Y-C. Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance. *IEEE Transactions on Medical Imaging* 2017;36:1876–1886.
17. Shi J. Good features to track. *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on, IEEE; 1994, p. 593–600.*
18. Matas J, Chum O, Urban M, Pajdla T. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing* 2004;22:761–767.
19. Rosten E, Drummond T. Fusing points and lines for high performance tracking. *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on, vol. 2, IEEE; 2005, p. 1508–1515.*
20. Leutenegger S, Chli M, Siegwart RY. BRISK: Binary robust invariant scalable keypoints. *Computer Vision (ICCV), 2011 IEEE International Conference on, IEEE; 2011, p. 2548–2555.*
21. Harris C, Stephens M. A combined corner and edge detector. *Alvey vision conference, vol. 15, Citeseer; 1988, p. 10–5244.*
22. Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-up robust features (SURF). *Computer Vision and Image Understanding* 2008;110:346–359.
23. Dalal N, Triggs B. Histograms of oriented gradients for human detection. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1, IEEE; 2005, p. 886–893.*
24. Clune J, Stanley KO, Pennock RT, Ofria C. On the performance of indirect encoding across the continuum of regularity. *IEEE Transactions on Evolutionary Computation* 2011;15:346–367.
25. Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. *BMJ: British Medical Journal* 1994;308:1552.

SKIN LESION FEATURES ANALYSIS FOR MALIGNANT MELANOMA CLASSIFICATION

Despite the dynamic development of machine learning methods, automatic analysis of skin lesions is still open issue. The following article proposes the use of an evolutionary algorithm to design, train, and to test a whole population of classifiers (artificial neural networks) and to iteratively improve them in each subsequent population, in order to achieve the best possible accuracy in the classification of skin lesions task. The algorithm returns an optimal set of features describing the dermoscopic image together with the proposed architecture of the neural network. High classification results were obtained, in particular: accuracy equal to 85.83%, specificity 79.07% and sensitivity 92.60%.

Keywords: evolutionary algorithm, neural networks, decision support system, machine learning.