

NEURAL IDENTIFICATION OF SELECTED KINDS OF INSECTS BASED ON COMPUTER TECHNOLOGY FOR THE IMAGES ANALYSIS

Summary

There has been noticed growing explorers' interest in drawing conclusions based on information of data coded in a graphic form. The neuronal identification of pictorial data, with special emphasis on both quantitative and qualitative analysis, is more frequently utilized to gain and deepen the empirical data knowledge. Extraction and then classification of selected picture features, such as color or surface structure, enables one to create computer tools in order to identify these objects presented as, for example, digital pictures. The work presents original computer system designed to digitalize pictures on the basis of color criterion. The system has been applied to generate a reference "learning" file for the neural system to identify selected kinds of insects.

NEURONOWA IDENTYFIKACJA WYBRANYCH OWADÓW Z WYKORZYSTANIEM KOMPUTEROWYCH TECHNIK ANALIZY OBRAZU

Streszczenie

Celem zrealizowanych badań była analiza możliwości wykorzystania sztucznych sieci neuronowych jako instrumentu przeznaczonego do identyfikacji motyli. Rozpoznawane owady reprezentowały gatunki, które są objęte ochroną prawną na terenie Polski. Neuronowej identyfikacji dokonano na podstawie (uprzednio pozyskanych) dwuwymiarowych obrazów, przedstawiających owady z rodziny Papilionidae.

1. Wprowadzenie

Rozwój technik multimedialnych w informatyce oraz liczne ich aplikacje w naukach przyrodniczych, powoduje położenie większego nacisku na rozwiązywanie problemów związanych z przetwarzaniem oraz analizą obrazów. Wzrasta zainteresowanie badaczy zagadnieniami związanymi z szeroko rozumianym wnioskowaniem prowadzonym w oparciu o informację zakodowaną w postaci danych obrazowych, ze szczególnym wskazaniem na metody analizy ilościowej oraz jakościowej. Ważne stają się metody i techniki ekstrakcji cech, sposoby opisu obrazu oraz statystyczne, syntaktyczne i niekonwencjonalne metody rozpoznawania obrazów, w tym wykorzystanie specjalizowanych sieci neuronowych. Metody sztucznej inteligencji, a w szczególności modelowanie neuronowe, znajdują już zastosowanie praktyczne w badaniach naukowych realizowanych w dyscyplinie inżynieria rolnicza [1]. Coraz częściej wykorzystuje się numeryczne symulatory sztucznych sieci neuronowych do informatycznego wsparcia procesu rozpoznawania obrazów, realizowanego w czasie rzeczywistym. W trakcie pozyskiwania, identyfikacji oraz ekstrakcji wiedzy zawartej w danych empirycznych występujących w postaci graficznej, ważną rolę odgrywają sieci neuronowe uczone techniką „z nauczycielem”, takie jak np. sieć typu MLP (ang. *MultiLayer Perceptron* - perceptron wielowarstwowy). Aby tworzony model neuronowy można „nauczyć” właściwej reakcji na zadane wzorce, niezbędne jest posiadanie zbioru danych występujących w formie akceptowalnej przez symulator sieci neuronowych (na ogół w postaci zaetykietowanych wektorów uczących).

Problematyka analizy oraz rozpoznawania obrazów z wykorzystaniem topologii neuronowych jest stosunkowo nowym obszarem zainteresowań informatyków i ściśle wiąże się z badaniami prowadzonymi nad sztuczną inteligencją. Modele neuronowe stanowią uzupełnienie dla klasycznych metod identyfikacyjnych, posiadając przy tym pewne cechy (jak np. szybkość działania), dające istotną przewagę nad technikami tradycyjnymi.

Identyfikacja informacji zakodowanej w postaci graficznej jest często prowadzona w sytuacji braku pełnej wiedzy, pozwalającej na zakwalifikowanie przynależności obiektów, znajdujących się na obrazie, do określonych klas. W procesie generowania topologii neuronowej pełna informacja, możliwa do wykorzystania przez algorytm uczący, jest zawarta w zbiorze uczącym złożonym z obiektów, dla których zdefiniowana jest prawidłowa klasyfikacja [5]. W związku z tym, sztuczna sieć neuronowa jako powszechnie uznane i wykorzystywane narzędzie klasyfikacyjne, okazuje się adekwatnym oraz efektywnym instrumentem, również w procesie rozpoznawania obrazów.

Celem pracy była analiza możliwości wykorzystania sztucznych sieci neuronowych do identyfikacji motyli objętych ustawą o ochronie gatunkowej na terenie Polski, należących do rodziny *Papilionidae*, dokonywanej na podstawie ich dwuwymiarowych obrazów.

2. Motyle z rodziny *Papilionidae*

Z obserwacji prowadzonych od kilkudziesięciu lat wynika jednoznacznie, że ilość motyli występujących w naturze ulega ciągłej oraz znaczącej redukcji. Przejawia się to przede wszystkim w obniżaniu się różnorodności

gatunkowej przy jednoczesnym wzroście liczebności kilku najpospolitszych gatunków. Przyczynę obserwowanych zmian w faunie motyli Polski stanowi wiele różnorodnych czynników. Do najważniejszych należy zaliczyć m. in. zanieczyszczenie powietrza, stosowanie insektycydów, wypalanie traw oraz zmiany stosunków wodnych. W przypadku motyli dziennych szczególne znaczenie ma ewolucja strukturalnego środowiska lęgowego, w którym występuje coraz mniej roślin mogących stanowić pokarm dla gąsienic. Brakuje też przestrzeni niezbędnej dla realizacji czynności życiowych, właściwych dla form dorosłych. W związku z tym w wielu krajach, również w Polsce, prowadzi się działania mające na celu ochronę gatunkową motyli. Obecnie na liście zwierząt chronionych znajdują się wszystkie gatunki krajowe *Papilionidae*, z wyjątkiem *Zerynthia polyxena* (Den. & Schaff.), który to gatunek już od ponad stu lat nie był obserwowany na terenie naszego kraju. Przedstawiciele rodziny *Papilionidae* pojawiają się lokalnie i należą do występujących rzadko. Jedynie *Papilio machano* jest gatunkiem pospolitym, często spotykanym na terenie Polski. Jest to rodzina dużych motyli dziennych, obejmująca na świecie ok. sześćset gatunków. Dwanaście z nich występuje w Europie, natomiast cztery można zaobserwować w Polsce. Omawiane motyle na ogół spotykane są na terenach otwartych. Najczęściej można je spotkać na nasłonecznionych polanach oraz na stokach górskich. Do krajowych fauny *Papilionidae* należą następujące gatunki:

- *Papilio machano* (L.) – paż królowej,
- *Iphiclides podalirius* (L.) – paż żeglarz,
- *Parnassius apollo* (L.) – niepylak apollo,
- *Parnassius mnemosyne* (L.) – niepylak mnemosyna.

3. Identyfikacja gatunków

3.1. *Papilio machano*

Cechy wyróżniające:

- rozpiętość skrzydeł 65-80 mm,
- nasadowa część przedniego skrzydła czarna,
- wzdłuż zewnętrznego brzegu obu par skrzydeł rząd półksiężycowatych, żółtych plamek,
- na skrzydle tylnym, przy tylnym kącie czerwona, okrągła plamka.

Środowisko: nasłonecznione stoki, łąki, polany, pola, ogrody.

Rośliny pokarmowe gąsienicy: różne gatunki z rodziny *Umbelliferae*, najczęściej *Peucedanum oreoselinum* L., *Daucus carota* L., *Anethum graveolens* L. i *Pimpinella saxifraga* L..

Rozsiedlenie w Polsce: pospolity, spotykany na obszarze całego kraju z wyjątkiem wysokich gór.



Rys. 1. *Papilio machano* (L.) – paż królowej
Fig. 1. *Papilio machano* (L.)

3.2. *Iphiclides podalirius*

Cechy wyróżniające:

- rozpiętość skrzydeł 62-75 mm,
- na przednim skrzydle czarne przepaski i klinowate plamy,
- na tylnym skrzydle przy zewnętrznym brzegu niebieskie, półksiężycowate plamy,
- przy tylnym kącie podłużna, pomarańczowa plamka.

Środowisko: skałki, nasłonecznione pagórki, polany, zadrzewienia śródpolne, sady.

Rośliny pokarmowe gąsienicy: różne gatunki drzew i krzewów z rodziny *Rosaceae*, przeważnie *Prunus spinosa* L., a także inne gatunki z rodzaju *Prunus* L., w tym drzewa owocowe.

Rozsiedlenie w Polsce: spotykany lokalnie w południowej części kraju, sporadycznie notowany w okolicach Warszawy i na Pojezierzu Mazurskim.



Rys. 2. *Iphiclides podalirius* (L.) – paż żeglarz
Fig. 2. *Iphiclides podalirius* (L.)

3.3. *Parnassius apollo*

Cechy wyróżniające:

- rozpiętość skrzydeł 70-80 mm,
- na przednim skrzydle pięć czarnych plam,
- na tylnym skrzydle dwie czerwone plamy w czarnej obwódce.

Środowisko: piargi, stoki górskie o wystawie południowej, preferuje miejsca o podłożu wapiennym.

Rośliny pokarmowe gąsienicy: różne gatunki *Serum* L., najczęściej *Serum telephium* L., *S. album* L., *S. fabaria* Koch..

Rozsiedlenie w Polsce: tylko w Tatrach i Pieninach.



Rys. 3. *Parnassius apollo* (L.) – niepylak apollo

Fig. 3. *Parnassius apollo* (L.)

3.4. *Parnassius mnemosyne*

Cechy wyróżniające:

- rozpiętość skrzydeł 50-58 mm,
- mniejszy od *Parnassius apollo*,
- na przednim skrzydle dwie czarne plamy.

Środowisko: leśne łąki, nasłonecznione skraje lasów.

Rośliny pokarmowe gąsienicy: różne gatunki *Corydalis* L.

Rozsiedlenie w Polsce: głównie spotykany we wschodniej części kraju, zwłaszcza w Bieszczadach i w Beskidzie Niskim, w południowo-wschodniej części kraju oraz w Sudetach na rozproszonych stanowiskach.



Rys. 4. *Parnassius mnemosyne* (L.) – niepylak mnemosyna

Fig. 4. *Parnassius mnemosyne* (L.)

W pracy podjęto działania zmierzające do opracowania modelu neuronowego, wspomagającego proces identyfikacji gatunków motyli chronionych na terenie Polski, które są zagrożone wyginięciem.

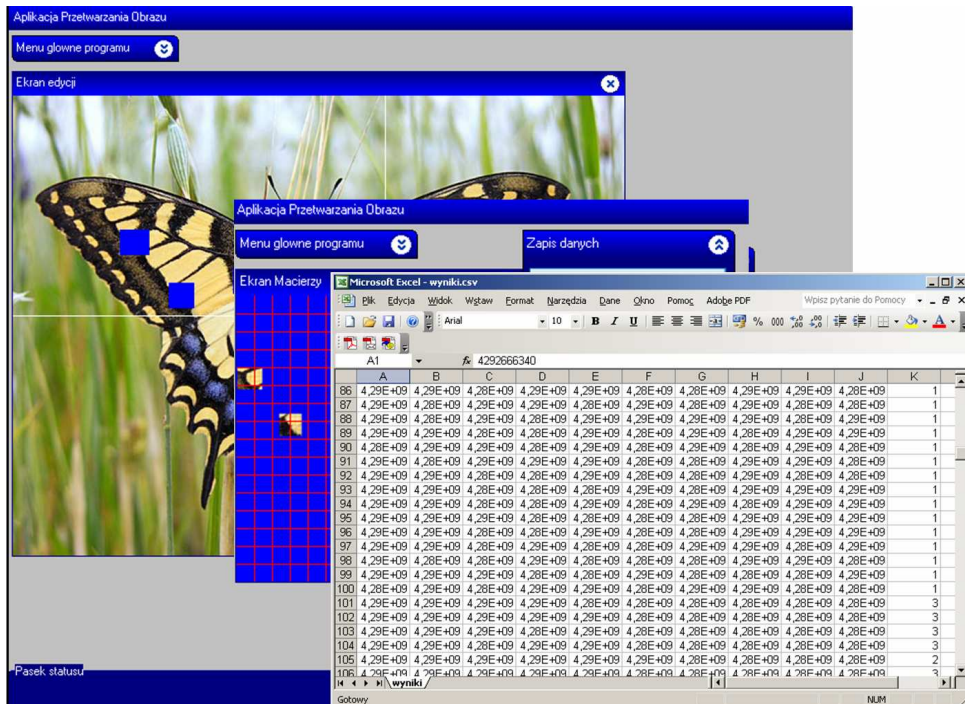
4. Metodyka badawcza

Podstawowymi czynnościami mającymi na celu rozpoznanie obiektów występujących w postaci graficznej jest akwizycja zbioru obrazów a następnie ekstrakcja i identyfi-

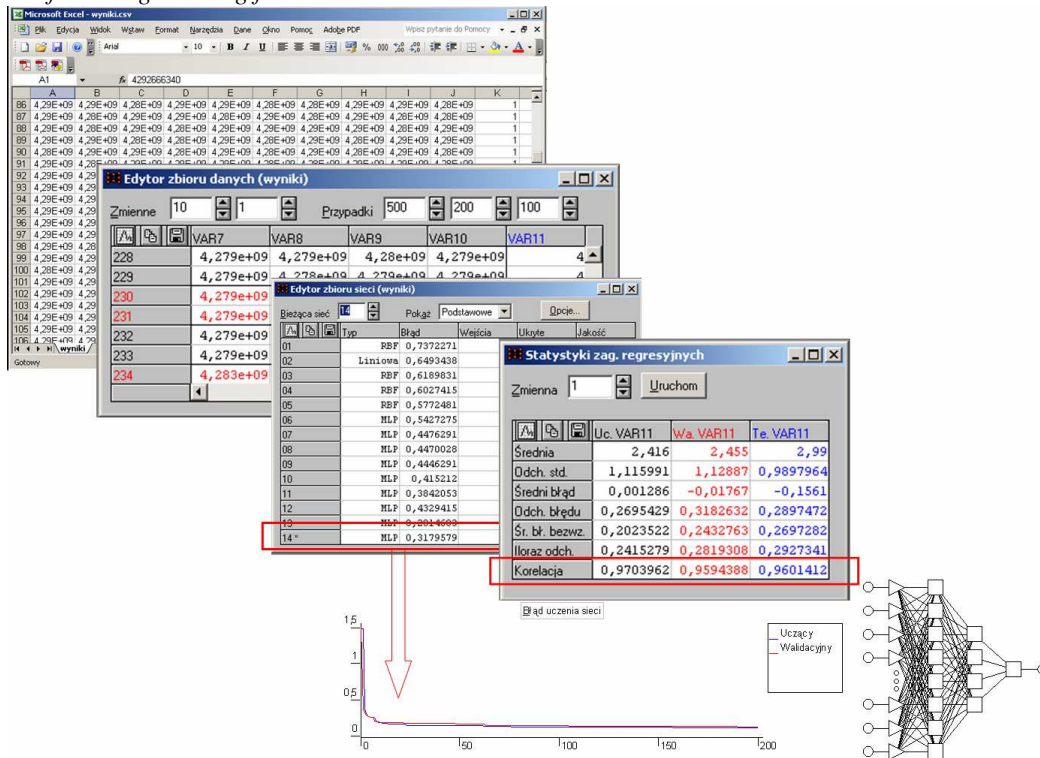
kacja reprezentatywnych cech badanych obiektów. W dalszej kolejności dokonuje się ich analizy, zgodnie ze zdefiniowaną wcześniej funkcją celu [6]. Działania te stanowią podstawę do budowy adekwatnych zbiorów uczących, niezbędnych w procesie tworzenia modeli neuronowych. W celu wygenerowania zbiorów uczących wykorzystano (specjalnie do tego celu wytworzony) oryginalny system informatyczny: „Przetwarzanie Obrazu v.1.0” [3]. Aplikacja ta pozwala m.in. na konwersję zdjęć do postaci zbioru uczącego, akceptowalnego przez cyfrowy symulator sztucznych sieci neuronowych, zaimplementowany w komercyjnym pakiecie statystycznym „STATISTICA v. 4.0”. W celu budowy zbioru uczącego posłużono się narzędziem „Próbnik” zaimplementowanym w wykorzystanej aplikacji. Pozwala on na wygenerowanie zbioru wektorów uczących zapisanych w pliku .csv, który można skonwertować do pliku .sta stanowiącego format właściwy do dalszej obróbki w pakiecie „STATISTICA v. 4.0” (rys. 6).

Jako cechy reprezentatywne uznano dziesięć zmiennych, stanowiących reprezentacje wybranych kolorów oraz artefaktów, charakterystycznych dla opisanych motyli. Wytworzony zbiór uczący zawierał czterysta niezależnych wektorów, standardowo podzielonych losowo na zbiory: uczący, walidacyjny oraz testowy, w proporcji 200:100:100 przypadków.

Po zaimportowaniu wygenerowanych plików do edytora danych symulatora sieci neuronowych pakietu „STATISTICA v. 4.0” wygenerowano sieć neuronowa typu MLP (ang. *MultiLayer Perceptron*) składającą się z czterech warstw i posiadającą strukturę: 10-8-4-1 (dwie warstwy ukryte, rys.7). Uczenie przeprowadzono z wykorzystaniem algorytmu wstecznej propagacji błędów BP (ang. *Back Propagation*) w 1000 epokach oraz algorytmu gradientów sprzężonych CG (ang. *Conjugate Gradients*) w 5 sesjach po 800 epok. Dalsze uczenie sieci zostało zrealizowane za pomocą algorytmu LM (ang. *Lovenberga-Marqardta*), efektywnego dla małych sieci z jednym wyjściem, w trakcie 10 sesji po 1500 epok.



Rys. 6. Proces przetwarzania obrazów do postaci zbiorów uczących
 Fig. 6. Process of creating learning files



Rys. 7. Proces tworzenie modelu neuronowego typu MLP
 Fig. 7. Process of creating neural network model type MLP

5. Wnioski

1. Została wykazana przydatność sieci neuronowych typu MLP do rozpoznawania obrazów wybranych motyli, których nadrzędnym kryterium klasyfikacyjnym jest barwa oraz tekstura.
2. Przeprowadzone badania wykazały, że istotny wpływ na jakość identyfikacji motyli ma struktura danych wejściowych, a w szczególności właściwy dobór cech

reprezentatywnych, charakteryzujących klasyfikowane motyle.

3. Optymalny dobór zmiennych wejściowych oraz liczby przypadków uczących ma istotny wpływ na jakość procesu uczenia sieci neuronowej typu MLP.

6. Literatura

- [1] Boniecki P., Piekarska-Boniecka H.: Neuronowa identyfikacja wybranych szkodników drzew

- owocowych w oparciu o cyfrowe techniki analizy obrazu [Neural identification of selected fruit trees pests based on digital techniques for the analysis of images], *Journal of Research and Applications in Agricultural Engineering* (2) 2004, pp.19-23, Poznań.
- [2] Boniecki P.: Elementy modelowania neuronowego w rolnictwie. WUP Poznań 2008.
- [3] Nowakowski K., Boniecki P., Weres J., Mueller W.: Przetwarzanie graficznych danych empirycznych dla potrzeb edukacji sztucznych sieci neuronowych, modelujących wybrane zagadnienia inżynierii rolniczej [Image processing empirical data for the purpose of education artificial neural network, modeling selected issues in agricultural engineering]. *Inżynieria Rolnicza* 2007, 2(90) str. 243-248 Kraków.
- [4] Ossowski S.: Sieci neuronowe do przetwarzania informacji [Neural networks for processing information], Warsaw 2000.
- [5] Tadeusiewicz R.: Rozpoznawanie obrazów [Recognition of images], Warsaw 1998.
- [6] Tadeusiewicz R., Flasiński M.: Rozpoznawanie obrazów [Recognition of images], Warsaw 1991.