

THE NEURAL NETWORKS OF THE TYPE *MLP* AND *RBF* AS CLASSIFYING TOOLS IN PICTURE ANALYSIS

Summary

*The neuronal identification of pictorial data, with special emphasis on both quantitative & qualitative analysis, is more frequently utilized to gain & deepen the empirical data knowledge. Extraction & then classification of selected picture features, enables one to create computer tools in order to identify these objects presented as, for example, digital pictures. In relationship from this, it seems to be purposeful the search of the modern methods helping educational process in the range of construction as well as exploitation of neuronal models in context of their utilization in picture analysis process. The additional aim of the work was the comparison of neural network of the type *MLP* and *RBF* for indication of the optimum classification tool.*

SIECI NEURONOWE TYPU *MLP* ORAZ *RBF* JAKO NARZĘDZIA KLASYFIKACYJNE W ANALIZIE OBRAZU

Streszczenie

*Neuronowa identyfikacja danych obrazowych, ze szczególnym naciskiem na analizę ilościową oraz jakościową, coraz częściej wykorzystywana jest do pozyskiwania oraz zgłębiania wiedzy zawartej w danych empirycznych. Ekstrakcja, a następnie klasyfikacja wybranych cech obrazu, pozwala na wytworzenie informatycznych narzędzi do identyfikacji wybranych obiektów, prezentowanych np. w postaci obrazu cyfrowego. W związku z tym, celowym wydaje się być poszukiwanie nowoczesnych metod wspomagających proces edukacyjny w zakresie konstrukcji oraz eksploatacji modeli neuronowych w kontekście ich wykorzystania w procesie analizy obrazu. Dodatkowym celem pracy było porównanie jakości sieci *MLP* oraz *RBF* mające na względzie wskazanie optymalnego instrumentu klasyfikacyjnego.*

Wprowadzenie

Rozpoznawanie obrazu oparte na technikach klasyfikacyjnych, to jedna z dziedzin informatycznych, w której sieci neuronowe znajdują szerokie zastosowanie. Obserwowany ostatnio szybki rozwój programów typu *OCR (Optical Character Recognition)*, czyli aplikacji do rozpoznawania znaków (również pisanych ręcznie) jest możliwy głównie sztucznych oparciu o zastosowanie sztucznych sieci neuronowych. Dzięki m.in. zdolności do samouczenia się, sztuczne sieci neuronowe stały się w tej dziedzinie narzędziem wiodącym. Rozpoznawanie obrazu przy użyciu modeli neuronowych znalazło zastosowanie w wielu dziedzinach. Również w rolnictwie jest ono coraz częściej wykorzystywane, m.in. do identyfikacji szkodników, uszkodzeń ziaren czy chorób roślin.

W pracy badano możliwości wykorzystania sieci neuronowych typu *RBF (Radial Basis Function)* oraz *MLP (MultiLayer Perceptrons)*, jako narzędzi klasyfikacyjnych do analizy obrazu prezentowanego sieci w postaci dwu wymiarowej mapy binarnej (obraz czarno-biały). Analizie poddano zbiór cyfr arabskich napisanych różnego rodzaju czcionkami przy użyciu wytworzonej uprzednio aplikacji do rozpoznawania cyfr. Dla poprawnego nauczania właściwej topologii sieci neuronowej, konieczne było wykorzystanie odpowiedniej rozdzielczości matrycy, a co za tym idzie wygenerowanie (w postaci wektorów uczących) dużej ilości danych wejściowych. W celu usprawnienia tego procesu użyto jako narzędzia pomocniczego wytworzonego programu „Czcionka”, dzięki któremu możliwa była konwersja obrazów różnych czcionek w formie bitmapy na ciągi cyfrowe oraz zapisanie ich jako plik tekstowy, akceptowalny w procesie uczenia przez symulator sztucznych

sieci neuronowych, zaimplementowany w pakiecie STATISTICA v.7.1.

Celem pracy było wytworzenie oraz testowanie, zgodnie ze standardami inżynierii oprogramowania, interaktywnej aplikacji komputerowej „Rozpoznawanie cyfr”, wspomagającej proces edukacyjny w zakresie konstrukcji oraz eksploatacji podstawowych topologii sztucznych sieci neuronowych w aspekcie użycia ich jako narzędzi identyfikacyjnych. Wytworzony edukacyjny system informatyczny ma przybliżyć wybrane zagadnienia z zakresu sieci typu perceptron i sieci radialnej oraz praktycznie prezentować zasadę działania tych sieci.

Wiadomości podstawowe

Obraz cyfrowy w ujęciu analizy komputerowej to dwuwymiarowa tablica, w której zawierają się wartości poszczególnych elementów obrazu (punkty i piksele). Najprostszym przypadkiem jest obraz binarny, gdzie piksele mogą przyjmować dwie wartości: 0 lub 1. Rozpatrując obraz kodowany w stopniach szarości, wartości te odpowiednio się zwiększają w zakresie narzuconym przez długość słowa maszynowego, np. w przypadku kodowania za pomocą 8 bitów zakres wynosi od 1 do 256. Jeśli chodzi natomiast o obraz kolorowy, sposób reprezentacji punktu (piksela) odnosi się na ogół do trzech składowych: czerwonej, zielonej i niebieskiej (w przypadku modelu koloru RGB).

W analizie obrazu cyfrowego wyróżnić można trzy rodzaje zagadnień: polepszenie jakości, segmentacja oraz rozpoznanie obiektów. Zazwyczaj występują one w trzech etapach, zachowując podaną wyżej chronologię.

Polepszenie jakości obrazu cyfrowego

Do polepszenia jakości obrazu cyfrowego służą różnego rodzaju techniki filtracyjne. Ich Zasada ich działania oparta jest na tzw. operacji konwolucji, będącej w istocie operacją kontekstową, która uwzględnia sąsiedztwo danego punktu obrazu. Polega to na tym, że przyjmując pewną wartość sąsiedztwa, definiuje się jego kształt (najczęściej jest to kwadrat, np. 3• 3, 5• 5 lub uogólniając: n• n pikseli).

Dla przykładowego sąsiedztwa 3• 3 (rys. 1.) operacja konwolucji przebiega w następujący sposób:

- a) okno o rozmiarze 3• 3 jest przemieszczane po obrazie cyfrowym (źródłowym),
- b) aktualny piksel centralny i jego sąsiedztwo jest przekształcane według zadanej reguły filtracji algorytmicznej lub przez sieć neuronową,
- c) nowa wartość piksela wpisywana jest w odpowiednie miejsce obrazu wyjściowego (x).

Segmentacja obrazu

Segmentacja obrazu jest etapem analizy, polegającym na wydzieleniu fragmentów składających się na obraz, które w następnej kolejności będą odrębnie poddane rozpoznaniu. Jeśli rozpatrywane są sceny statyczne, wówczas najpopularniejszą techniką segmentacji jest binaryzacja, ewoluująca od prostego progowania, w stronę złożonych algorytmów wielu i zmiennoprogowych. Inne stosowane metody to: detekcja i łączenie krawędzi obiektów z analizą powstałych linii i konturów. Obrazy w których istotne są tekstury, analizowane są specjalnie do tego celu stworzonymi metodami, bazującymi na analizie statystycznej podobszarów (okien) obrazu [2].

Rozpoznanie obiektów

Obiekty wydzielone w poprzednich dwóch etapach, poddawane są w następnej kolejności rozpoznaniu. Jest to ostatni krok w procesie „komputerowego widzenia”. Operacja ta, może być realizowana w dwojaki sposób. Pierwszym rozwiązaniem jest bezpośrednie porównanie ze wzorcem. Częściej jednak, stosuje się porównanie na podstawie cech, które zostały określone na etapie segmentacji.

Podsumowując należy zauważyć, że w zadaniu rozpoznawania dąży się do określenia przynależności rozmaitego typu obiektów do wyselekcjonowanych wcześniej klas. Rozpoznawanie to realizowane jest w sytuacji braku informacji *a priori* na temat reguł przynależności obiektów do poszczególnych klas. Jedyna informacja, możliwa do wykorzystania przez algorytm lub maszynę rozpoznającą jest zawarta w ciągu uczącym, złożonym z obiektów, dla których znana jest prawidłowa klasyfikacja. Brak apriorycznej wiedzy oznacza to, że maszyna uczy się rozpoznawania na podstawie przedstawionych jej przykładów [4].

Ogólnie analiza obrazu składa się z dwóch podstawowych etapów:

- wytypowanie oraz pomiar cech określających obiekt,
- analiza tych cech.

Jednym z podstawowych problemów jest właściwy wybór cech reprezentatywnych, na podstawie których realizowany będzie proces klasyfikacji. Możliwości doboru tych cech, są ograniczone praktycznie tylko wyobraźnią twórców metod poszukiwania. Przykładem może być zagadnienie rozpoznawania cyfr. Rozmieszczenie punktów informa-

tywnych (skrzyżowań, rozgałęzień i zakończeń linii), topologia zamkniętych i otwartych konturów - to tylko niektóre cechy, jakie można użyć w tym przypadku. Ważne jest przy tym, aby propozycje ich doboru pochodziły od specjalistów, tzn. ludzi będących ekspertami w danej dziedzinie. Przy rozpoznawaniu np. wad serca należy skorzystać z wiedzy kardiologa. Niepowodzenia w analizie obrazu, szeroko opisywane w literaturze, były najczęściej spowodowane brakiem potrzebnego specjalisty.

Techniki klasyfikacyjne

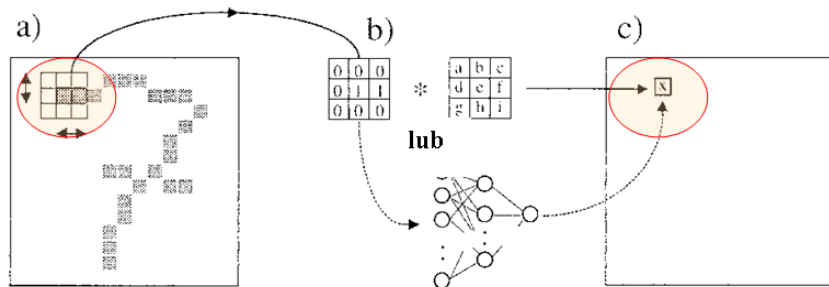
Przez pojęcie klasyfikacji rozumie się dzielenie dowolnego zbioru elementów na grupy, do których zalicza się elementy różniące się, ale podobne tj. mające własności wyróżniające daną grupę. Zbiór elementów należących do jednej grupy nazywany jest klasą, a każdy element klasy - obiektem. Elementy klasy mogą różnić się między sobą, z wyjątkiem tych własności, na których opiera się proces klasyfikacji. W zależności od rodzaju dostępnej informacji w ramach klasyfikacji można wyodrębnić dwa zagadnienia:

1. klasyfikację wzorcową gdy struktura kategorii jest znana, czyli dysponuje się charakterystyką klas, z których pochodzą obiekty. W technikach neuronowych zagadnienie to nazywane jest uczeniem (lub rozpoznawaniem) z nauczycielem lub też pod nadzorem,
2. klasyfikację bezwzorcową, określaną też jako taksonomia albo analiza skupień (klasteryzacja). W technikach neuronowych znaną również jako uczenie (lub rozpoznawanie) bez nauczyciela.

Wykorzystanie sieci neuronowych w procesie analizy obrazu

Badania wykazują, że sieci neuronowe są bardziej wydajnym narzędziem klasyfikacyjnym, niż, powszechnie stosowane, tradycyjne metody dyskryminacyjne. W związku z powyższym, rozpoznawanie obrazów to jedna z dziedzin informatycznych, w której sztuczne sieci neuronowe znajdują coraz szersze zastosowanie. Jednym z potwierdzeń powyższego jest obserwowany dynamiczny rozwój programów typu *OCR (Optical Character Recognition)*, który jest możliwy głównie dzięki zastosowaniu sztucznych sieci neuronowych. Rozpoznawanie obrazów przy użyciu sieci neuronowych znalazło zastosowanie w wielu dziedzinach. W dostępnej literaturze można znaleźć wiele informacji, dotyczących wykorzystania sieci neuronowych do analizy obrazu, począwszy od identyfikacji pojedynczych znaków kończąc na analizie skomplikowanych obrazów kolorowych oraz rozpoznawaniu różnego typu złożonych obiektów. Warto zauważyć, że również w rolnictwie neuronowe techniki identyfikacyjne są coraz częściej wykorzystywane m.in. do identyfikacji szkodników, uszkodzeń ziaren czy chorób roślin [6].

Rozpoznawanie obrazów jest najstarszym z zastosowań sieci neuronowych. W zagadnieniu rozpoznawania obrazu wyróżnia się szczególnie dwa procesy: wyodrębnianie obiektów i ich identyfikację. Wyodrębnianie obiektów polega na ich wykryciu na analizowanym obrazie. Wykonuje się przy tym kilka działań takich jak: wykrywanie krawędzi, wyznaczanie konturów, ocenę powierzchni. W realizacji tych zadań wygrywają sieci neuronowe w szczególności dzięki ich zdolności do funkcjonowania w warunkach zakłóceń. W odróżnieniu od pozostałych (tradycyjnych) algo-



Rys. 1. Zasada realizacji operacji konwolucji (Mirkut Zb., Tadeusiewicz R. 2000)

Fig. 1. Principle of realization of convolution operation (Mirkut Zb., Tadeusiewicz R. 2000)

rytmów analizy obrazów sieci neuronowe nie potrzebują wstępnej filtracji ("odszumiania") obrazów. Identyfikacja obiektów polega na wskazaniu do której klasy obiektów należy obiekt. Klasyfikacja może mieć różny poziom szczegółowości ograniczony jedynie przez właściwości zastosowanej sieci.

W oparciu o dane literaturowe wiadomo, że sieci neuronowe jednowarstwowe można nauczyć bezbłędnego rozpoznawania znaków drukowanych (Bishop C., 1995). Przy ich nauczaniu każda klasa reprezentowana jest przez jeden obraz i dzięki temu klasy są liniowo separowalne, co w sposób zasadniczy ułatwia rozwiązanie problemu. Gdy jednak w czasie nauki sieci klasa jest reprezentowana przez kilka wzorcowych obrazów znaków, co ma miejsce np. przy różnych krojach pisma lub przy prezentowaniu sieci obrazów zakłóconych, klasy mogą nie być liniowo separowalne i konieczne staje się użycie sieci warstwowych. W związku z tym zdecydowano się na badanie reprezentatywnych topologii sieci neuronowych typu **RBF** i **MLP**, jako narzędzi klasyfikacyjnych do analizy obrazu na przykładzie procesu identyfikacji cyfr arabskich występujących w postaci różnych formatów graficznych.

Budowa modelu neuronowego

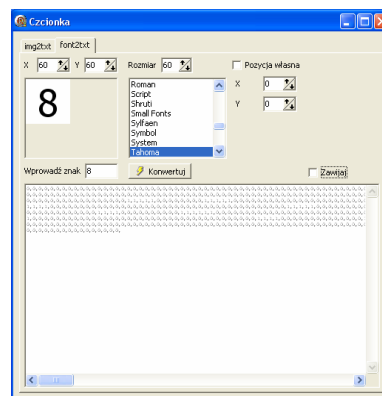
W procesie tworzenia modelu neuronowego do identyfikacji cyfr, kluczowym zadaniem było zbudowanie właściwej struktury zbioru uczącego. Do jego wygenerowania użyto programu wytworzonego „Czcionka”. Dzięki tej aplikacji, można zamienić obraz dowolnej czcionki na postać binarną, która następnie zostaje zwrócona w postaci gotowego ciągu uczącego, akceptowalnego przez symulator SNN, zawarty w pakiecie STATISTICA v.7.1.

Zakładka font2txt (rys. 2.) umożliwia nam przekonwertowanie na postać cyfrową wszystkich rodzajów czcionek dostępnych w pakiecie *Microsoft Office*. Wygenerowano obrazy cyfr od 0 do 9 w 29 różnych czcionkach, co dało 290 przypadków uczących. Do analizy użyto tablicy 30 • 30, co w efekcie generuje ciąg uczący o 900 elementach.

Topologie sieci wygenerowanych za pomocą w.w. symulatora przedstawiają się następująco:

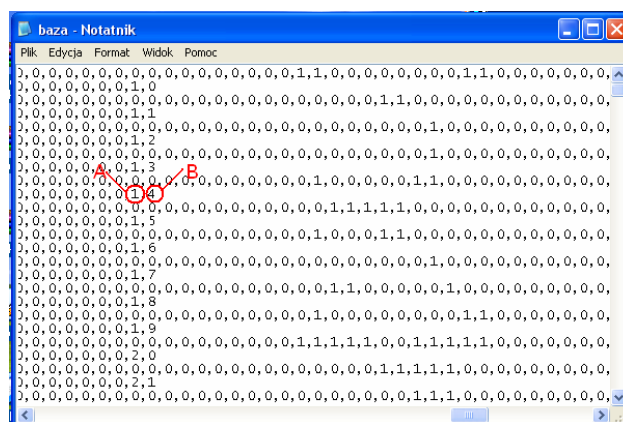
Optymalna sieć **MLP** składała się z 3 warstw:

- warstwa wejściowa – liczba neuronów 929 – funkcja aktywacji liniowa,
- warstwa ukryta – liczba neuronów 102 – funkcja aktywacji hiperboliczna,
- warstwa wyjściowa – liczba neuronów 10 – funkcja aktywacji softmax.



Rys. 2. Konwertowanie bitmapy do postaci binarnej w postaci wektora uczącego za pomocą aplikacji „Czcionka”

Fig. 2. The converting the bitmap to binary map in the form of teaching vector with the help of the application "Type"

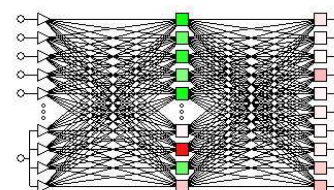


Rys. 3. Fragment wektora uczącego w postaci pliku .txt

Fig. 3. The fragment of teaching vector in the form of .txt file

gdzie:

- A oznacza numer jednej z 29 użytych czcionek,
- B oznacza cyfrę, którą dany ciąg reprezentuje (zmienna wyjściowa).

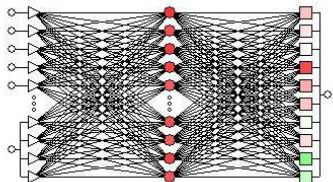


Rys. 4. Topologia otrzymanej sieci typu **MLP**

Fig. 4. Topology of received neural network of the type **MLP**

Optymalna sieć **RBF** składała się z 3 warstw:

- warstwa wejściowa – liczba neuronów 929 – funkcja aktywacji liniowa,
- warstwa ukryta – liczba neuronów 45– funkcja aktywacji wykładnicza,
- warstwa wyjściowa – liczba neuronów 10 – funkcja aktywacji liniowa.

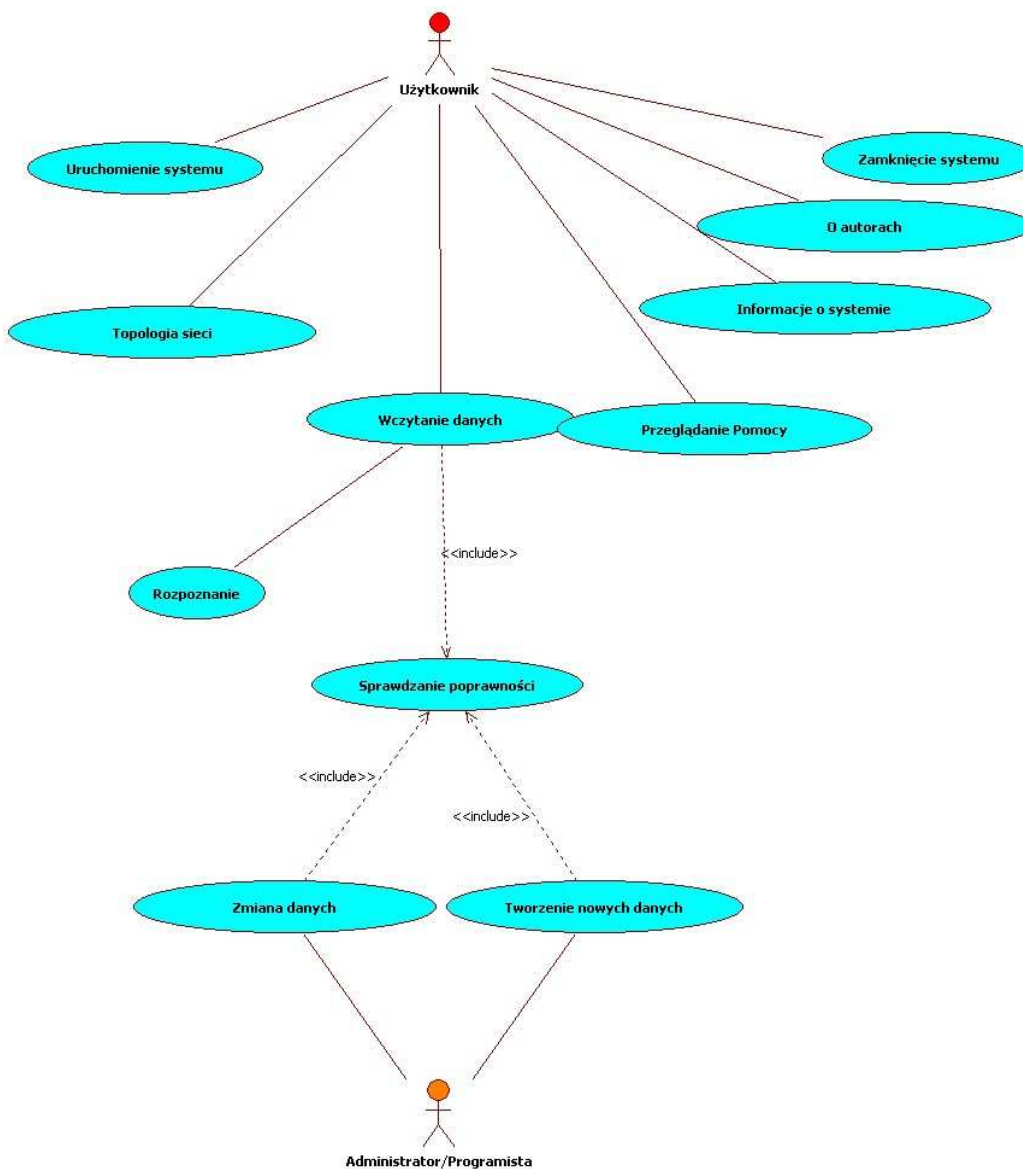


Rys. 5. Topologia otrzymanej sieci typu **RBF**
 Fig. 5. Topology of received neural network of the type **RBF**

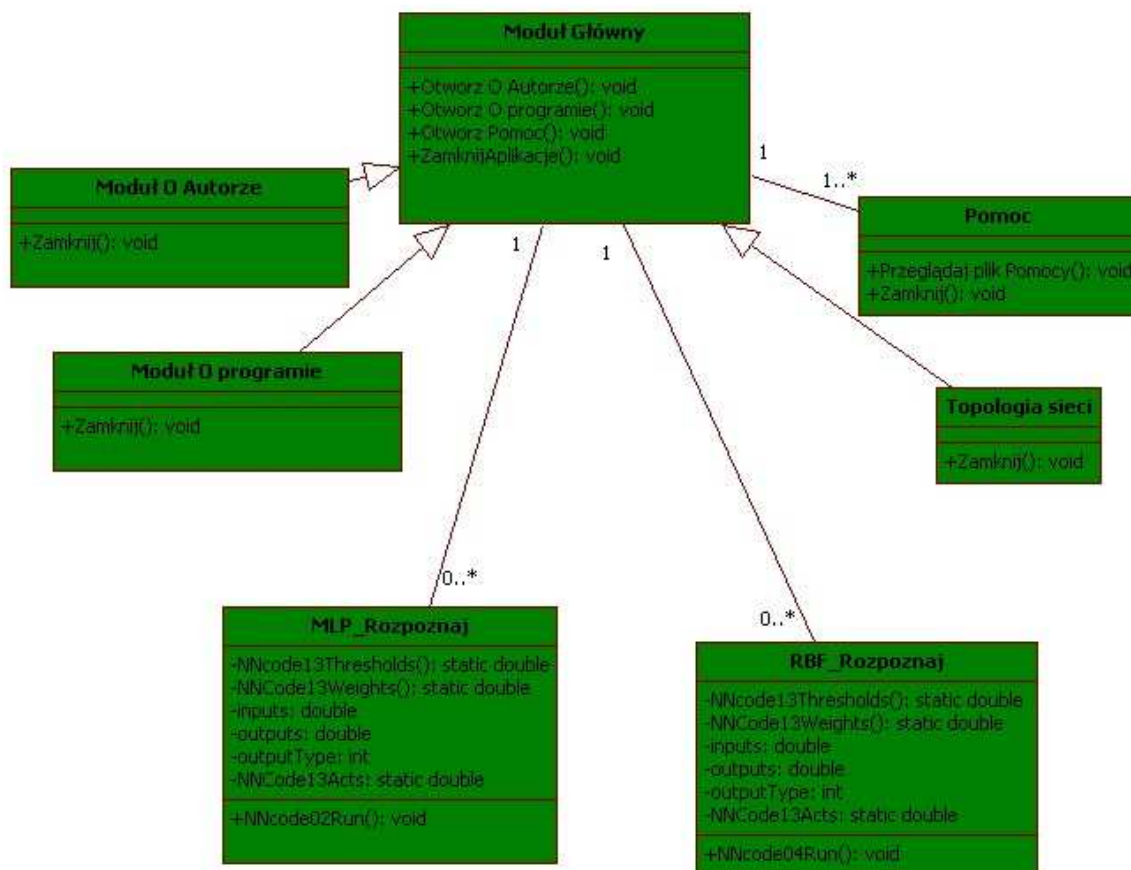
Opis systemu informatycznego „Rozpoznawanie cyfr”

Wstępną fazą budowy systemu informatycznego było określenie wymagań stawianych projektowanej aplikacji. Postawione zostały określone cele, zaproponowany został zakres oraz zaprezentowano ramowy opis systemu informatycznego „Rozpoznawanie cyfr”. W fazie projektowania, korzystając z pomocy programu *Posejdon for UML*, wykonane zostały podstawowe diagramy mające na celu wizualizację implementacji systemu.

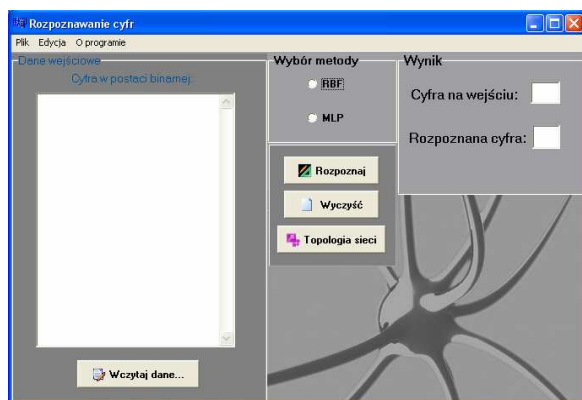
Spośród wielu dostępnych narzędzi programistycznych wybrany został *Borland C++ Builder v.6.0*. Kod symulujący działanie sieci neuronowych został wygenerowany z wykorzystaniem modułu „Generator kodu” zaimplementowanego w komercyjnym pakiecie *STATISTICA v.7.1*. Otrzymany kod w języku *C* został poddany drobnym zmianom adaptacyjnym, w celu lepszego dostosowania go do środowiska programistycznego *Borland C++ Builder v.6*. Wytworzony następnie system informatyczny „Rozpoznawanie cyfr” opiera się na tym kodzie i stanowi jądro wytworzonej aplikacji.



Rys. 6. Diagram przypadków użycia systemu informatycznego „Rozpoznawanie cyfr”
 Fig. 6. The case diagram of using of the computer system „Rozpoznawanie cyfr”



Rys. 7. Diagram klas systemu informatycznego „Rozpoznawanie cyfr”
 Fig. 7. The class diagram of the computer system „Rozpoznawanie cyfr”



Rys. 8. Wygląd okna głównego programu „Rozpoznawanie cyfr”
 Fig. 8. Main window of computer system „Rozpoznawanie cyfr”

Aplikacja została przetestowana zarówno dla wygenerowanego kodu jak również po utworzeniu poszczególnych jej modułów.

Wymagania niefunkcjonalne

System posiada prosty i intuicyjny interfejs umożliwiający użytkownikowi szybkie zapoznanie się z możliwościami aplikacji. Pozwala także na dokonywanie klasyfikacji przy użyciu sieci nauczonych za pomocą różnych algo-

rytmów lub metod. Aby system działał bezproblemowo, potrzebne są:

- komputer klasy PC z systemem operacyjnym Windows®,
- pamięć operacyjna: co najmniej 64MB,
- rozdzielczość ekranu: co najmniej 800• 600 pikseli.

Uwagi końcowe

Pozytywne wyniki uczenia sieci typu *MLP* oraz *RBF*, jakie uzyskano w programie STATISTICA v.7.1. pokazały, że wspomniane typy sieci sprawdzają się jako narzędzia klasyfikacyjne do analizy obrazów binarnych. Zarówno w przypadku sieci *RBF* jak i *MLP* otrzymano niewielkie błędy uczenia, co w efekcie przełożyło się na poprawną identyfikację prezentowanych sieci obrazów cyfr w postaci binarnej. Dlatego też, w wytworzonym systemie informatycznym „Rozpoznawanie cyfr”, wykorzystano w celach porównawczych oba typy sieci. Wyniki działania aplikacji potwierdziły rezultaty uzyskane w procesie uczenia sieci. Prezentowane systemowi informatycznemu dane wejściowe (w postaci ciągu binarnego), zostały poprawnie rozpoznane jako konkretne cyfry, zarówno przy wykorzystaniu sieci perceptronowej typu *MLP* jak i sieci radialnej typu *RBF*. Aplikacja „Rozpoznawanie cyfr” jest przyjaznym użytkownikowi systemem edukacyjnym wspomagającym procesy poznawcze z zakresu budowy oraz eksploatacji najpopularniejszych topologii sztucznych sieci neuronowych.

Literatura

- [1] Tadeusiewicz R., Flasiński M. (1991). Rozpoznawanie obrazów: Wydawnictwo Naukowe PWN, Warszawa.
- [2] Mirkut Zb., Tadeusiewicz R. (2000). Sieci neuronowe tom 6: Akademicka Witryna Wydawnicza EXIT, Warszawa.
- [3] Boniecki P. (2004). Sieci neuronowe typu MLP oraz RGB jako komplementarne modele aproksymacyjne w procesie predykcji plonu pszenżyta: Journal of Research and Applications in Agricultural Engineering, Poznań, (1'2004), Vol. 49(1), str. 28-33.
- [4] Kohorda P., Tadeusiewicz R. (1999). Komputerowa analiza i przetwarzanie obrazów: Drukarnia Narodowa, Kraków.
- [5] Tadeusiewicz R., (1993). Sieci neuronowe: Akademicka Oficyna Wydawnicza, Warszawa.
- [6] Boniecki P., Piekarska-Boniecka H. (2004). The *SOFM* Neural Network in the Process of Identification of Selected Orchard Pests: Journal of Research and Applications in Agricultural Engineering, Vol 49(4), p. 5-10, Poznań.
- [7] Bishop C., (1995). Neural Networks for Pattern Recognition: Oxford University Press.