

doi:10.15199/48.2021.11.03

Porównanie skuteczności rozpoznawania obiektów morskich na podstawie obrazów FLIR za pomocą klasyfikatorów typu sieć neuronowa i PCA

Streszczenie. W pracy przedstawiono wyniki badań skuteczności metod rozpoznawania obiektów morskich na podstawie obrazów FLIR za pomocą metod opartych na różnych wariantach transformaty PCA oraz konwolucyjnej sieci neuronowej. Jako bazę wzorów wykorzystano zdjęcia dziewięciu typów obiektów zarejestrowanych na Morzu Bałtyckim. Obie metody potwierdziły swoją skuteczność, przy czym sieci neuronowe wykazały wyższą skuteczność na poziomie 96%, natomiast PCA - 80% - 87%..

Abstract. The paper presents results of the effectiveness research on the of maritime object recognition methods from FLIR images using methods based on different variants of the PCA classifier and a convolutional neural network classifier. Images of nine types of objects recorded in the Baltic Sea have been used as a pattern base. Both methods confirmed their effectiveness, with neural networks showing 96% effectiveness, while PCA 80% - 87%. (**Comparison of maritime objects recognition efficiency upon FLIR images using neural network and PCA classifiers**)

Słowa kluczowe: rozpoznanie obrazów w podczerwieni, klasyfikacja za pomocą sieci neuronowej, klasyfikator PCA.

Keywords: FLIR images recognition, neural network classifier, PCA classifier.

Wstęp

Pasywne sensory podczerwieni FLIR (forward looking infrared) na powietrznych platformach rozpoznania i okrętach używane są przede wszystkim do rozpoznania krótkiego i średniego zasięgu. Pozwalają one na opracowanie systemów identyfikacji obiektów nawodnych (statków i okrętów). Sensor FLIR rejestruje cyfrowe obrazy achromatyczne, w których każdy piksel scharakteryzowany jest jedną liczbą – luminancją piksela, proporcjonalną do temperatury punktu.

Obrazy te z natury rzeczy są achromatyczne, ale ze względu na to, że obserwuje je operator, przekształcane są w obrazy barwne. Barwom z określonej palety przypisane są wartości temperatur. Barwowa skala temperatur umieszczana jest na obrazie. Obrazy te uzupełniane są ponadto przez kamerę o dodatkowe informacje, takie jak symbol środka obrazu oraz informacje dotyczące warunków otoczenia w czasie rejestracji obrazu.

Obrazy otrzymywane z kamer FLIR są najczęściej zaszumione oraz mogą być zniekształcone na skutek rejestracji ich z różnej odległości i pod różnym kątem obserwacji (aspektu). Z tych względów proces rozpoznawania obiektów zawartych w obrazach FLIR jest procesem wieloetapowym.

Wykryte obiekty należące do tej samej klasy mają cechy wspólne, które zwane są cechami dystynktywnymi. Wartości tych cech decydują o identyfikacji obiektów. Wybór konkretnej klasy obiektów zależy od wartości funkcji odległości rozpoznawanego obiektu od obiektów wzorcowych przechowywanych w tzw. bazie wzorców. Cechy dystynktywne obiektów tworzą przestrzeń wartości cech dystynktywnych. Wybór cech dystynktywnych uzależniony jest od przyjętej metody rozpoznania.

Przedmiotem pracy jest rozpoznawanie obiektów morskich na podstawie obrazów FLIR. W zastosowanych metodach rozpoznawania (klasyfikatorach) do zbioru cech dystynktywnych zaliczamy przede wszystkim sylwetkę obiektu oraz rozkład jasności poszczególnych punktów sylwetki. Przykłady wyboru różnych cech dystynktywnych sylwetek zależnych od metody klasyfikacji pokazano w [1-3,18]. Należy podkreślić, że w literaturze przedmiotu można znaleźć definicje kilku klasyfikatorów, które mogą

być użyte do rozpoznawania obiektów za pośrednictwem sensorów FLIR.

Pierwszą metodą rozpoznawania obiektów morskich na podstawie obrazów FLIR zastosowaną w pracy jest klasyfikacja wykonywana przez sieć neuronową [3,4,5,7,8,9,13,14,15,16,17]. Zastosowana sieć neuronowa jest siecią konwolucyjną pozwalającą na redukcję objętości informacji i zmniejszenie złożoności obliczeniowej w kolejnych warstwach sieci. W pierwszym etapie badań tego klasyfikatora należało zrealizować czynność uczenia sieci. W tym celu zbiór dostępnych obrazów podzielono na dwie części: 60% obrazów zakwalifikowano do zbioru uczącego (treningowego), natomiast pozostałe 40% zakwalifikowano do zbioru testowego. Szczegółowy opis sieci i sposobu jej wykorzystania jako klasyfikatora obrazów obiektów morskich został przedstawiony w rozdziale 3.

Drugą metodą rozpoznawania zastosowaną w pracy jest klasyfikacja metodą analizy głównych składowych PCA (Principal Component Analysis), zwaną dalej klasyfikatorem PCA [10, 11]. Metoda PCA pozwala na uszeregowanie cech obrazów (wartości jasności pikseli) w ciąg nierosnący od pikseli o największej rozróżnialności (największej wariancji cechy jasności w zbiorze obrazów wzorcowych) począwszy do pikseli o najmniejszej rozróżnialności (najmniejszej wariancji cechy jasności w zbiorze obrazów wzorcowych). W naturalny sposób można z obrazu wyeliminować piksele, dla których wariancja jasności w zbiorze obrazów wzorcowych wynosi zero lub jest stosunkowo mała. O eliminacji kolejnych pikseli o wariancji jasności większej niż zero należy zdecydować arbitralnie. Proces eliminacji pikseli jest odpowiednikiem stratnej kompresji obrazów i prowadzi do utraty informacji. Zbiór pikseli, które pozostały w obrazie, nazywa się, w przypadku rozpoznawania sylwetek, sylwetką własną. Metoda PCA bazuje na liniowym przekształceniu Karhunen-Loeve'a [12].

Oba klasyfikatory nie rozwiązują problemu rozpoznawania sylwetek okrętów płynących w różnych kierunkach. Problem ten można rozwiązać poprzez porównanie analizowanej sylwetki z sylwetką wzorcową i jej lustrzanym odbiciem (symetria względem prostej).

Oba klasyfikatory wymagają, aby obrazy FLIR poddawane procesowi klasyfikacji zostały wcześniej

poddane procesowi wstępnego przetwarzania, którego celem jest wyeliminowanie zbędnych informacji dotyczących tła obiektu oraz zakłóceń, a także normalizacja sylwetki obiektu. Proces wstępnego przetwarzania obrazu może obejmować segmentację, normalizację jasności, przeskalowanie sylwetki, centrowanie sylwetki oraz poziomowanie sylwetki.

W końcowej części pracy przedstawiono wyniki badań dwóch klasyfikatorów: SN (sieć neuronowa) i trzy warianty klasyfikatora PCA.

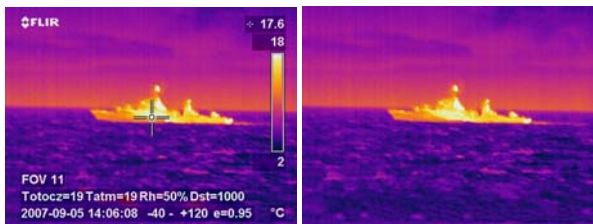
Proces rozpoznawania obrazowego obiektów

Proces klasyfikacji dzieli się zwykle na cztery podstawowe etapy:

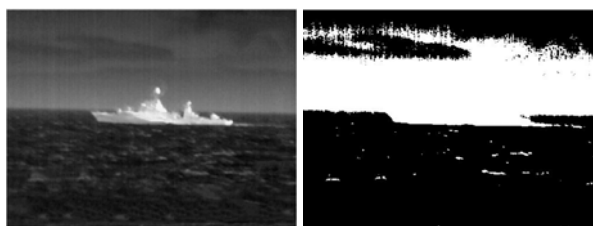
- wstępnego przetwarzania obrazu,
- segmentacji,
- ekstrakcji cech (wyznaczania wartości cech dystynktywnych),
- klasyfikacji właściwej.

W ramach czynności wykonywanych podczas przetwarzania wstępnego obrazów FLIR dokonywane są między innymi następujące czynności: przekształcenie obrazu do postaci obrazu achromatycznego w skali szarości, usunięcie elementów stałych obrazów FLIR (skala temperatury, dane opisujące stan otoczenia, logo producenta, znak środka obrazu), przeskalowanie obrazu, usunięcie szumów poprzez zastosowanie na przykład filtra medianowego, poprawa dynamiki jasności obrazu.

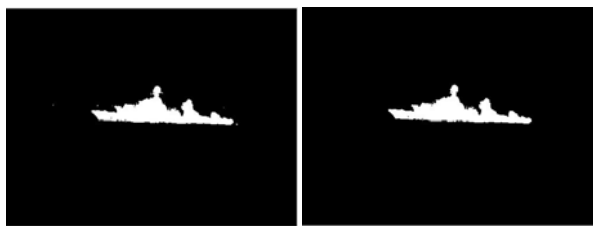
W ramach przetwarzania wtórnego realizowane są operacje poprawiające kształt sylwetki obiektu morskiego (m.in. operacje morfologiczne), poziomowanie sylwetki obiektu, centrowanie sylwetki obiektu, normalizację jasności sylwetki oraz ewentualnie jej powtórne przeskalowanie.



Rys.1. Pierwotny obraz z kamery FLIR oraz z usuniętymi zbytecznymi artefaktami



Rys.2. Obraz w skali szarości oraz obraz po wykonaniu progowania binarnego z progiem wyznaczonym przez algorytm Otsu



Rys.3. Obraz po wykonaniu progowania binarnego z progiem określonym ręcznie oraz obraz po wykonaniu operacji usunięcia artefaktów typu „pieprz i sól” oraz zamknięcia

Obrazy rozpoznawane w pracy zostały poddane przetwarzaniu wstępnemu i segmentacji, którego efektem

była sylwetka obiektu morskiego w skali szarości. Proces segmentacji był wspomagany komputerowo za pomocą modułu segmentacji metodą Otsu. Wynikiem tego działania jest globalny próg binarny, który pozwalał na wykonanie segmentacji binarnej. Niestety w większości przypadków (przede wszystkim w przypadkach, gdy temperatura obiektu była zbliżona do temperatury pewnych obszarów tła) wynik działania algorytmu z automatycznie określonym progiem globalnym był niezadawalający, pozostawiono więc możliwość ręcznej korekty wartości progów. Na rys. 1 – 5 zaprezentowano wyniki realizacji kolejnych etapów przetwarzania obrazu pierwotnego z kamery FLIR.



Rys.4. Iloczyn obrazu binarnego i obrazu w skali szarości oraz obraz po wykonaniu operacji normalizacji wysokości sylwetki i centrowania



Rys.5. Obraz po wykonaniu operacji normalizacji jasności i obcięciu górnego oraz dolnego obszaru o jasności zero

Charakterystyka sieci neuronowej jako klasyfikatora obrazów obiektów morskich

Sieci neuronowe stają się coraz bardziej popularnym narzędziem służącym do klasyfikacji danych. Odnaczają się dużą skutecznością działania w różnych dziedzinach. Przykładem takiej dziedziny jest klasyfikacja obrazów bądź elementów na obrazach. Potwierdzają to wyniki konkursów np. ImageNet LSVR (ang. Large Scale Visual Recognition) Challenge [6], gdzie rozwiązania oparte o sieci neuronowe zajmują czołowe miejsca rankingu.

Jedną z metod uczenia sieci neuronowych jest uczenie nadzorowane, w którym sieć uczy się odpowiednich wzorców na podstawie zaprezentowanych jej przykładów wraz z sugerowanym wynikiem klasyfikacji. W takim trybie uczenia wymagane jest posiadanie relatywnie dużego zbioru danych uczących, który dodatkowo jest już sklasyfikowany. Powoduje to, że sieć neuronowa nie nadaje się do wszystkich problemów, np. takich, w których trudno jest uzyskać ręcznie sklasyfikowaną dużą ilość danych.

Siłą sieci neuronowych w klasyfikacji jest możliwość uczenia się struktur hierarchicznych. Wielowarstwowa sieć neuronowa najpierw rozpoznaje podstawowe wzorce za pomocą swoich początkowych warstw, a następnie na ich podstawie buduje kolejne klasyfikatory. Przykładowo konwolucyjna warstwa sieci neuronowej może pozwalać na przygotowanie obrazu przez jego filtrację, przy czym sama sieć neuronowa wybiera najbardziej odpowiedni filtr.

Do klasyfikowania obrazów 2D najczęściej wykorzystywanym typem sztucznych sieci neuronowych są tzw. sieci konwolucyjne, zwane również splotowymi (ang. convolutional neural networks, CNN). Pozwalają one na ograniczenie wymaganej pamięci poprzez stosowanie filtrów. Główną składową tego rodzaju sieci są warstwy konwolucyjne, których zadaniem jest wykrywanie zależności pomiędzy pikselami znajdującymi się w bliskim sąsiedztwie. Warstwy głębsze, a więc te, które znajdują się bezpośrednio za obrazem wejściowym, mają na celu

wykrycie cech takich jak krawędzie czy gradienty kolorów, natomiast te wyższe łączą je ze sobą, a następnie generują klasyfikację.

Różnica pomiędzy sieciami spłotowymi, a tradycyjnymi jest taka, że poszczególne neurony nie są łączone ze wszystkimi pikselami obrazu wejściowego, a jedynie z ustalonym tzw. polem odbiorczym. Pozwala to na zredukowanie ilości wag i zmniejszenie w trakcie wykonywania uczenia sieci złożoności obliczeniowej. Przy obrazach z kamery termowizyjnej o wymiarach 640x480 pikseli i dużej liczby wzorców, zwiększona byłaby możliwość przekroczenia dostępnych zasobów pamięci. Na podstawie pól odbiorczych tworzona jest mapa aktywacji i określana wartość wagi poszczególnych neuronów.

Pomiędzy kolejnymi warstwami spłotowymi wykorzystuje się warstwę łączącą (ang. pooling layer). Może ona działać na zasadzie uśredniania bądź maksymalizacji wartości w danych obszarach, co pozwala na zmniejszenie rozmiarów i ograniczenie nadmiarowości danych.

Ostatnim ogniwem jest warstwa połączeniowa „każdy z każdym”, która składa się z tylu neuronów, ile klas przewidywanych jest w ramach sieci.

Na potrzeby klasyfikacji obrazów obiektów morskich z kamery termowizyjnej przygotowano w środowisku MATLAB sieć konwolucyjną, której architekturę przedstawia następująca tablica warstw (rys. 6).

```
inputSize = [640 480 1];
numClasses = 11;

layers = [
    imageInputLayer(inputSize)

    convolution2dLayer(8, 8, 'Stride', 2, 'Padding', 'same')
    batchNormalizationLayer
    reluLayer

    maxPooling2dLayer(3, 'Stride', 3)

    convolution2dLayer(8, 8, 'Stride', 2, 'Padding', 'same')
    batchNormalizationLayer
    reluLayer

    maxPooling2dLayer(3, 'Stride', 3)

    convolution2dLayer(8, 16, 'Stride', 2, 'Padding', 'same')
    batchNormalizationLayer
    reluLayer

    fullyConnectedLayer(11)
    softmaxLayer
    classificationLayer];
```

Rys. 6. Architektura sieci konwolucyjnej

Sieć konwolucyjna zbudowana została z trzech warstw konwolucyjnych (convolution2dLayer). W każdej z nich wielkość filtrów ustalono na 8x8 pikseli. Pierwsza z warstw składała się z czterech tego typu filtrów. Pozostałe dwie zbudowane były z kolejno 8 i 16 filtrów. Krok dla przejścia filtru w pionie i poziomie ustalono na wartość 2.

Funkcję warstw łączeniowych pełni warstwa maxPooling2dLayer, która dokonuje próbkowania w dół (ang. down-sampling) poprzez dzielenie obszaru wejściowego na kwadratowe pola (3x3 piksele), a następnie w każdym z nich wylicza maksymalną wartość.

Dla tak ustawionych parametrów poszczególnych warstw sieci i przygotowaniu bazy treningowej wzorców opierającej się na 99 obrazach, podzielonych na 11 klas (po 9 w każdej) wykonano proces uczenia sieci. Zbiór wszystkich obrazów podzielono na część treningową oraz testową w proporcjach 60:40. Po wyznaczeniu maksymalnej liczby epok, które oznaczają przejście algorytmu przez cały zbiór treningowy, równej piętnastu

takim powtórzeniem, uzyskano skuteczność testowania przekraczającą 93%.

Charakterystyka klasyfikatora PCA

Niech X będzie zbiorem obrazów treningowych x^1, x^2, \dots, x^M . Każdy obraz jest wektorem kolumnowym o długości $N=m \cdot n$, stąd

$$(1) \quad X = \{x^i \in \{\mathbf{N} \cup \{0\}\}^N, i = 1, \dots, M\}.$$

gdzie M jest liczbą obrazów wzorcowych.

Kolejnym etapem przekształcenia obrazów w transformacie PCA jest usunięcie wartości średniej. Niech \bar{x} będzie wektorem wartości średnich wektorów obrazów x^i . Wektor średni wyznaczany jest zgodnie ze wzorem

$$(2) \quad \bar{x} = \frac{1}{M} \sum_{i=1}^M x^i.$$

Każdy obraz x^i różni się od obrazu średniego o wektor \bar{x} .

$$(3) \quad r^i = x^i - \bar{x}.$$

Zbiór tych wektorów jest przedmiotem analizy głównych składowych PCA.

Macierz kowariancji zbioru wektorów bazowych (treningowych) wyznaczana jest zgodnie z następującym wzorem

$$(4) \quad C = \frac{1}{M} \sum_{i=1}^M r^i [r^i]^T = RR^T,$$

gdzie macierz R jest zdefiniowana w sposób następujący

$$(5) \quad R = [r^1, r^2, \dots, r^M].$$

Ponieważ macierz C jest macierzą symetryczną o wartościach rzeczywistych, zatem wszystkie jej wartości własne są rzeczywiste i istnieją odpowiadające im wektory własne: rzeczywiste, liniowo niezależne i ortonormalne. Można pokazać, że macierz kowariancji określona wzorem (4) jest nieujemnie określona, co oznacza, że wszystkie jej wartości własne są nieujemne.

Jeżeli obrazy wzorcowe w X są skorelowane, to macierz kowariancji C nie jest macierzą diagonalną.

W celu dekokorelacji danych należy dokonać takiej transformacji danych (macierzy X), aby przekształcone dane miały diagonalną macierz kowariancji. Taka transformacja może być znaleziona poprzez rozwiązanie problemu wyznaczenia wektorów i wartości własnych macierzy C . Dokonuje się tego poprzez rozwiązanie równości

$$(6) \quad CP = P\Lambda.$$

gdzie $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$ jest macierzą diagonalną posiadającą na głównej przekątnej wartości własne macierzy kowariancji C , natomiast macierz P jest macierzą wektorów własnych tej macierzy. Macierze P i Λ mogą być wyznaczone za pomocą jednej z bibliotek obliczeń numerycznych [11].

Wartości i wektory własne powinny być uporządkowane (przeindeksowane) według malejących wartości własnych. W szczególnych przypadkach pewne wartości własne mogą być równe zero. Kolumny macierzy P są wektorami własnymi, które są wektorami ortonormalnymi. Macierz P jest macierzą przekształcenia PCA.

Stosując przekształcenie P możemy otrzymać dane (obrazy) zdekorelowane

$$(7) \quad y^i = P^T r^i,$$

$$(8) \quad \mathbf{Y} = \mathbf{P}^T \mathbf{R}.$$

Jeżeli przyjmiemy zasadę, że wartości i wektory własne powinny być uporządkowane (przeindeksowane) według malejących wartości własnych, to powoduje ona, iż współrzędne y_k^i transformat \mathbf{y}^i ($i = 1, 2, \dots, M$) różnią się między sobą coraz mniej wraz ze wzrostem wartości indeksu k . W szczególnych przypadkach od pewnej wartości indeksu $k = L \leq N$ wartości własne $\lambda_k = 0$, co oznacza, że wartości współrzędnych y_k^i dla tych samych wartości indeksu k są jednakowe dla wszystkich obiektów treningowych. Macierz kowariancji posiadająca zerowe wektory własne jest macierzą osobliwą, co powoduje niemożliwość zastosowania odległości Mahalanobisa w ocenie podobieństwa rozpoznawanego obiektu od tego zbioru obiektów treningowych.

W przypadku analizowanych obrazów sylwetek obiektów morskich oznacza to, iż piksel o indeksie k , dla którego $\lambda_k = 0$, ma taką samą jasność we wszystkich obiektach treningowych. Taki piksel może zostać wyeliminowany z obrazu, ponieważ nie jest cechą zapewniającą możliwość odróżnienia jednego obrazu od drugiego w zbiorze obiektów. Pozostawienie w diagonalnej macierzy wartości własnych wierszy i kolumn o indeksie k , dla którego $\lambda_k = 0$, będzie miało również takie konsekwencje, że nie będzie można wykonać operacji odwrócenia macierzy kowariancji \mathbf{C} .

Zakładając przeindeksowanie współrzędnych wektorów treningowych w sposób zgodny z malejącymi wartościami własnymi, eliminacja współrzędnych o zerowych wartościach własnych oznacza usunięcie z wektorów \mathbf{y}^i współrzędnych o indeksach $k \geq L$.

Oznaczmy zredukowaną macierz wektorów własnych przez $\tilde{\mathbf{P}}$ a zredukowaną diagonalną macierz wartości własnych przez $\tilde{\mathbf{\Lambda}}$. Macierz $\tilde{\mathbf{P}}$ można otrzymać z macierzy \mathbf{P} przez wykreślenie ostatnich $(N-L+1)$ kolumn tej macierzy. Macierz $\tilde{\mathbf{\Lambda}}$ można otrzymać z macierzy $\mathbf{\Lambda}$ przez usunięcie ostatnich $(N-L+1)$ wierszy i kolumn macierzy $\mathbf{\Lambda}$. Zmienioną macierz kowariancji $\tilde{\mathbf{C}}$ zwaną macierzą kowariancji transformaty PCA otrzymujemy ze wzoru:

$$(9) \quad \tilde{\mathbf{C}} = \tilde{\mathbf{P}} \tilde{\mathbf{\Lambda}} \tilde{\mathbf{P}}^T.$$

Wykorzystując zredukowaną macierz $\tilde{\mathbf{P}}$ w przekształceniu PCA otrzymamy nową postać wzorów (7) i (8) określających przekształcone pojedyncze obrazy i ich zbiór

$$(10) \quad \mathbf{y}^i = \tilde{\mathbf{P}}^T \mathbf{r}^i,$$

$$(11) \quad \mathbf{Y} = \tilde{\mathbf{P}}^T \mathbf{R}.$$

Ocena podobieństwa sylwetek obiektów morskich w obrazach FLIR poddanych przekształceniu PCA

W rozpoznawaniu obrazów obiektów morskich posługujemy się wybranym kryterium podobieństwa sylwetek rozpoznawanych obiektów do sylwetek obiektów wzorcowych, zwanych dalej obiektami treningowymi. Kryterium to jest związane z miarą odległości między rozpoznawanym obrazem a pewnym obrazem wzorcowym (lub ich transformatami), który może być pojedynczym obrazem obiektu wzorcowego, bądź tzw. obrazem średnim obiektu wzorcowego, uzyskany z wielu obrazów tego obiektu w drodze operacji uśredniania obrazów.

Kryterium oceny podobieństwa obiektu rozpoznawanego do poszczególnych klas obiektów wzorcowych (treningowych) zależy od dwóch czynników:

- metody oceny podobieństwa obiektu rozpoznawanego do poszczególnych klas obiektów wzorcowych,
- miary odległości w przestrzeni obiektów.

W niniejszej pracy wykorzystano trzy metody oceny podobieństwa obiektu rozpoznawanego do poszczególnych klas obiektów wzorcowych:

- metodę najbliższego sąsiada,
- metodę k najbliższych sąsiadów,
- metodę centroidalną.

Metoda najbliższego sąsiada oceny podobieństwa pojedynczego obiektu testowego do klasy obiektów treningowych określa odległość rozpoznawanego obiektu od najbliższych obiektów w każdej klasie i wybiera tę klasę, dla której odległość najbliższego obiektu jest najmniejsza.

Metoda k najbliższych sąsiadów oznacza wybór tej klasy zbioru treningowego, która dominuje w zbiorze k najbliższych położonych obiektów całego zbioru treningowego w stosunku do rozpoznawanego obiektu.

Metoda centroidalna oceny podobieństwa pojedynczego obiektu testowego do klasy obiektów treningowych określa odległość rozpoznawanego obiektu od środka każdej klasy i wybiera tę klasę, dla której odległość rozpoznawanego obiektu od środka klasy jest najmniejsza.

W niniejszej pracy wykorzystano miarę Mahalanobisa jako miarę odległości w przestrzeni obiektów. Badając przynależność nieznanego obrazu \mathbf{x}^{test} (obraz jest tutaj utożsamiany z pewnym wektorem losowym) do danej klasy, mierzy się jego podobieństwo do wektora $\bar{\mathbf{x}}_i$, uwzględniając przy tym informację o wariancjach poszczególnych składowych oraz korelacjach między nimi. Miarą takiego podobieństwa jest odległość Mahalanobisa, przy czym macierzą wag jest odwrotność macierzy kowariancji:

$$(12) \quad d(\mathbf{x}^{test}, \bar{\mathbf{x}}_i) = \sqrt{(\mathbf{x}^{test} - \bar{\mathbf{x}}_i)^T \mathbf{C}^{-1} (\mathbf{x}^{test} - \bar{\mathbf{x}}_i)}.$$

Macierz kowariancji \mathbf{C} musi być macierzą nieosobliwą, tzn. dodatnio określoną.

Należy zauważyć, że jeżeli macierz kowariancji jest macierzą tożsamościową ($\mathbf{C} = \mathbf{I}$ i $\mathbf{C}^{-1} = \mathbf{I}$), to odległość Mahalanobisa sprowadza się do odległości euklidesowej.

Wyniki rozpoznawania obiektów morskich za pomocą klasyfikatorów typu sieć neuronowa i PCA

W bazie treningowej znajdowało się 99 obrazów FLIR jedenastu typów obiektów morskich. Dla wszystkich obrazów wykonano czynności wstępnego przetwarzania obrazów, segmentacji oraz przetwarzania. Fragmenty bazy treningowej przedstawiono na rys. 7 (po jednym obrazie dla każdej z 11 klas obiektów).



Rys. 7. Wybrane pojedyncze obrazy treningowe dla każdego typu obiektów morskich

W bazie testowej znajdowało się 77 obrazów FLIR tych samych typów obiektów morskich, po 7 obrazów testowych każdej klasy.

Wyniki badań skuteczności rozpoznawania obiektów morskich na podstawie obrazów FLIR dla klasyfikatora typu sieć neuronowa i trzech wariantów klasyfikatora PCA przedstawia tabela 1. Warianty klasyfikatora PCA związane są z trzema metodami oceny podobieństwa obiektu testowego do klasy obiektów treningowych:

- metodą najbliższego sąsiada,
- metodą 6 najbliższych sąsiadów,
- metodą centroidalną.

Tabela 1. Porównanie wyników klasyfikacji obrazów za pomocą różnych klasyfikatorów

Nr klasy	Liczba obrazów testowych	Klasyfikator typu sieć neuronowa	Klasyfikator PCA		
			metoda najbliższego obrazu	metoda 6 najbliższych obrazów	metoda centroidalna
			%	%	%
1	7	100,00	85,71	71,43	42,86
2	7	100,00	71,43	71,43	71,43
3	7	85,71	85,71	100,00	100,00
4	7	100,00	71,43	42,86	28,57
5	7	100,00	85,71	100,00	100,00
6	7	85,71	100,00	85,71	100,00
7	7	100,00	85,71	71,43	0,00
8	7	100,00	100,00	100,00	100,00
9	7	85,71	85,71	85,71	85,71
10	7	100,00	100,00	100,00	100,00
11	7	100,00	85,71	71,43	28,57
Łącznie	77	96,10	87,01	81,82	68,83

Analizując powyższe wyniki można zaobserwować, że sieć neuronowa uzyskała najwyższy wynik (96,1 %) w trakcie przeprowadzanych badań. Przy odpowiednio przetworzonych obrazach, w których optymalnie dobrany został próg binaryzacji, źle rozpoznane zostały jedynie trzy obrazy.

Śśród trzech klasyfikatorów PCA najlepszy wynik (87,01 %) osiągnął klasyfikator wykorzystujący metodę najbliższego obrazu. Nieznacznie gorszy wynik (81,82 %) osiągnął klasyfikator wykorzystujący metodę 6 najbliższych obrazów.

Klasyfikator PCA wykorzystujący metodę centroidalną odbiega wynikami od pozostałych dwóch klasyfikatorów PCA (68,83 %). Wpływ na to może mieć występowanie w bazie treningowej danej klasy takich obiektów, których obrazy zostały wykonane pod innym kątem obserwacji. Skutkiem tego jest przesunięcie środka tej klasy (uśrednionej transformaty PCA obrazów treningowych klasy).

Metoda 6 najbliższych sąsiadów w trakcie badań osiągnęła skuteczność o kilka procent niższą od tej polegającej na wyborze najbliższego obrazu. Należy jednak zaznaczyć, że w przypadku bardziej rozbudowanych baz treningowych, o większej liczbie klas i wzorców, które je reprezentują, algorytm ten zabezpiecza przed wystąpieniem przypadkowego wyniku. Nie jest on bowiem określany na podstawie tylko jednej, najbliższej wartości, ale również kilku kolejnych, które zwiększają prawdopodobieństwo poprawnej klasyfikacji.

Biorąc pod uwagę czas wykonywania operacji, najszybszy jest algorytm wykorzystujący metodę centroidalną. W przypadku klasyfikatorów wykorzystujących metodę najbliższego obrazu i metodę 6 najbliższych obrazów wymagane jest wyliczenie odległości obrazu testowego od wszystkich obrazów treningowych wszystkich rozpoznawanych klas.

Praca została sfinansowana przez Wojskową Akademię Techniczną w ramach projektu UGB 856.

Autorzy: dr inż. Tadeusz Pietkiewicz, Wojskowa Akademia Techniczna, Wydział Elektroniki, Instytut Radioelektroniki, ul. Gen. S. Kaliskiego 2, 00-908 Warszawa,
E-mail: tadeusz.pietkiewicz@wat.edu.pl;
mgr inż. Patryk Dudek, Wojskowa Akademia Techniczna, Wydział Elektroniki, Instytut Radioelektroniki, ul. Gen. S. Kaliskiego 2, 00-908 Warszawa,
E-mail: patryk.dudek@student.wat.edu.pl.

LITERATURA

- [1] Park Y., Sklansky J., Automated Design of Linear Tree Classifiers, *Pattern Recognition*, Vol. 23, No. 12, 1990, 1393-1412
- [2] Valin P., Methods for the Fusion of Multiple FLIR Classifiers, [w] *Proceedings of the Workshop on Signal Processing, Communication, Chaos and Systems: a Tribute to Rabinder N. Madan*, June 20, 2002, Newport, RI, 2002, 117-122
- [3] Valin P., Bossé É., Jouan A., Airborne application of information fusion algorithms to classification, *Defence R&D Canada - Valcartier, Technical Report DRDC Valcartier*, TR 2004-282, May 2006
- [4] Valin P., Reasoning Frameworks, [w] A.K. Hyde, E. Shahbazian, E. Waltz (eds) *Multisensor Fusion. NATO Science Series (Series II: Mathematics, Physics and Chemistry)*, vol 70, pp. 222-245. Springer, Dordrecht, 2002, doi.org/10.1007/978-94-010-0556-2_9.
- [5] Schmidhuber J., Deep learning in neural networks: An overview, *Neural Networks, Elsevier*, No. 61, 2015, 85-117
- [6] ImageNet LSVR Challenge, URL <https://imagenet.org/challenges/LSVRC/2017/index.php> dostęp 19.06.2021
- [7] Osowski S., Sieci neuronowe do przetwarzania informacji, *Oficina Wydawnicza Politechniki Warszawskiej*, Warszawa 2020
- [8] Rutkowski L., Metody i techniki sztucznej inteligencji, *Wydawnictwo Naukowe PWN*, Warszawa, 2006
- [9] Żurada J., Barski M., W. Jędruch, Sztuczne sieci neuronowe. Podstawy teorii i zastosowania, *Wydawnictwo Naukowe PWN*, Warszawa, 1996
- [10] Gouaillier V., Gagnon L., Ship Silhouette Recognition Using Principal Components Analysis, *SPIE Proc. #3164, conference "Applications of Digital Image Processing XX"*, San Diego, 1997
- [11] Osowski S., Metody i narzędzia eksploracji danych, *Wydawnictwo BTC*, Legionowo 2013
- [12] Kwiatkowski W., Metody automatycznego rozpoznawania wzorców, *Bel Studio*, Warszawa, 2007
- [13] Chollet F., Deep Learning. Praca z językiem Pythoni biblioteką Keras. Helion, Gliwice 2019.
- [14] Bonnacorso G., Algorytmy uczenia maszynowego. Zaawansowane techniki implementacji. Helion, 2019
- [15] Deep Learning in MATLAB, URL https://www.mathworks.com/help/deeplearning/ug/deeplearning-in-matlab.html#responsive_offcanvas dostęp 19.06.2021
- [16] Pietkiewicz T., Sikorska-Łukasiewicz K., Comparison of two classifiers based on neural networks and the DTW method of comparing time series to recognize maritime objects upon FLIR images. *Proc. SPIE 11055, XII Conference on Reconnaissance and Electronic Warfare Systems, 110550V* (27 March 2019); doi: 10.1117/12.2524918
- [17] Matuszewski J., Pietrow D., Recognition of electromagnetic sources with the use of deep neural networks, *XII Conference on Reconnaissance and Electronic Warfare Systems, 2019 in Proceedings of SPIE - The International Society for Optical Engineering*, Volume 11055, Article Number: 110550D, DOI: 10.1117/12.2524536
- [18] Matuszewski J., The methods of class pattern forming in order to emitters recognition [Metody tworzenia wzorców klasy dla celów rozpoznawania źródeł emisji], *Przegląd Elektrotechniczny*, 84(5), 2008, 104-108