Jolanta CHMIELIŃSKA, Jacek JAKUBOWSKI

BIOMETRYCZNA WERYFIKACJA TOŻSAMOŚCI KIEROWCÓW NA PODSTAWIE OBRAZU TWARZY

W artykule przedstawiono problematykę weryfikacji tożsamości na podstawie obrazu twarzy w kontekście systemu monitorowania kierowców na potrzeby bezpieczeństwa w ruchu drogowym. Zaproponowane zostały dwie metody weryfikacji tożsamości oparte na konwolucyjnej sieci neuronowej, opracowane z wykorzystaniem techniki „transfer learningu". W artykule przedstawione zostały wyniki porównawcze efektywności działania przedstawionych metod a także ich wady oraz zalety.

WSTĘP

Stan przeciążenia psychofizycznego kierowców stanowi jedną z głównych przyczyn wypadków drogowych w Polsce i na świecie. Brak dostatecznej ilości snu oraz wzmożone skupianie uwagi na prowadzeniu pojazdu w dłuższym okresie czasu przy dużym natężeniu ruchu i często w trudnych warunkach atmosferycznych w sposób naturalny sprzyjają obniżeniu koncentracji, spowolnieniu reakcji i pojawieniu się senności [1]. Podobne efekty obserwuje się również nawet w czasie spokojnej ale długotrwałej jazdy, której monotonia może prowadzić do stanu tzw. hipnozy autostradowej [2].

Stąd, niezależnie od zwiększających bezpieczeństwo nowoczesnych rozwiązań technicznych instalowanych już nawet seryjnie w pojazdach jak np. asystent pasa ruchu, aktywny tempomat czy system rozpoznawania znaków drogowych TSR (ang. *Traffic Sign Recognition*), w wielu krajach funkcjonują zrozumiałe i powszechnie akceptowane uregulowania prawne narzucające ograniczanie czasu nieprzerwanej jazdy kierowców. Dotykają one zwłaszcza kierowców zawodowych spędzających w pracy (na trasie) dużo czasu, często nawet powyżej 24h. W egzekwowaniu tych uregulowań pomagają tachografy, czyli rozwiązania techniczne instalowane w pojazdach. Z inżynierskiego punktu widzenia są one urządzeniami pomiarowymi wyznaczającymi i rejestrującymi m.in. prędkość pojazdu, przejechaną odległość oraz czas aktywności kierowców. Jednakże jako urządzenia techniczne stosunkowo często stają się obiektem różnorodnych nielegalnych działań, których celem jest manipulacja zapisów ukrywająca jazdę bez wymaganych prawem przerw [3], [4]. Na podstawie ustaleń Komisji Europejskiej z 2014 szacuje się, że w co trzecim pojeździe dochodzi do oszustw w rejestrowaniu czasu pracy i odpoczynku kierowców [5]. Niezależnie od pogorszenia bezpieczeństwa na drogach, działania takie sprzyjają powstawaniu znaczących nierówności pomiędzy ofertami firm transportowych co prowadzi do nieuczciwej konkurencji. Z tego powodu najnowsze dyrektywy zawarte w Rozporządzeniu Parlamentu Europejskiego [6] wprowadzają od 15 czerwca 2019 r. obowiązek montowania we wszystkich nowych pojazdach tzw. inteligentnych tachografów. Ich najważniejszą cechą ma być integracja z globalnym systemem nawigacji satelitarnej, dzięki której możliwa będzie automatyczna rejestracja położenia pojazdu podczas dziennego okresu pracy umożliwiająca wspieranie działań organów kontrolnych. Zostaną one ponadto wyposażone w urządzenia mikrofalowe do bezprzewodowego transferu danych z tachografów bez konieczności zatrzymywania pojazdu co sprzyjać będzie niepostrzeżonemu wykrywaniu nieprawidłowości.

Wspomniane dyrektywy [6] nie odnoszą się jednak do idei wykorzystania najnowszych rozwiązań z dziedziny biometrii, które oferują potencjalną możliwość ukrócenia procederu posługiwania się cudzymi kartami kierowcy będącym jednym z łatwiejszych i popularnych sposobów manipulacji. Manipulacja taka byłaby niemożliwa w przypadku zapewnienia ciągłej, biometrycznej weryfikacji tożsamości osoby siedzącej za kierownicą. W praktyce oznaczałoby to podanie pokładowemu systemowi kontrolnemu danych identyfikacyjnych kierowcy celem odnalezienia w lokalnej bazie rekordu biometrycznego, który na bieżąco byłby porównywany z rekordem biometrycznym rejestrowanym w czasie jazdy. Innym rozwiązaniem mogłoby być wprost zapisanie rekordu biometrycznego na karcie kierowcy. Wykrycie rozbieżności pomiędzy rekordami zapisanymi i rejestrowanymi skutkowałoby komunikatami świadczącymi o próbie użycia cudzej tożsamości do kontynuowania jazdy.

Rozwiązanie tego typu, z przeznaczeniem również do systemów antykradzieżowych, prezentuje np. [7], w którym jako materiał biometryczny wykorzystany został odcisk palca rejestrowany za pomocą czujnika pojemnościowego. Weryfikacja odciskiem palca charakteryzuje się wysokim poziomem wiarygodności, ale jest mało praktyczna w zadaniu ciągłego monitoringu tożsamości kierowcy. Spośród dostępnych biometryk najlepszym rozwiązaniem jest wykorzystanie obrazu twarzy rejestrowanego za pomocą umieszczonej w aucie kamery. Kamera nie absorbuje uwagi kierowcy a stosowane współcześnie algorytmy widzenia maszynowego realizują zadanie rozpoznawania w czasie rzeczywistym. Jednym z najnowszych podejść stosowanych w tym zakresie jest nienadzorowane poszukiwanie dystynktywnej informacji bezpośrednio w obrazie z zastosowaniem konwolucyjnej sieci neuronowej uczonej z wykorzystaniem odpowiednio dużego zbioru danych. Sieci tego typu mają hierarchiczną strukturę wielowarstwową i wykorzystują mechanizmy tzw. głębokiego uczenia (ang. *Deep Learning*) w procesie rozkładu danych wejściowych na cechy. Podstawą jest operacja splotu (konwolucji), której efektem są maski filtrów wypracowane w procesie uczenia i traktowane jako nauczone cechy. Pierwsze warstwy wydobywają cechy proste, wspólne dla danych o bliskim sąsiedztwie. Następne warstwy wykorzystują je do wydobycia kolejnych, bardziej ogólnych własności danych służących w konsekwencji do klasyfikacji. Uproszczoną postać sieci konwolucyjnej przedstawia rys. 1. Sieci takie można tworzyć samodzielnie dysponując dużym zbiorem danych obrazowych oraz wydajną techniką obliczeniową, ale można też dokonać adaptacji nauczonych już sieci do rozwiązania zupełnie

innego problemu w ramach swoistego transferu wiedzy (ang. *Transfer Learning*). Przykładem konstrukcji sieci o bardzo dobrych własnościach uogólniania jest opracowana przez zespół badawczy Facebooka sieć o nazwie DeepFace [8]. Sieć ta nauczona na bazie 4 milionów fotografii profilowych użytkowników portalu oferuje błąd rozpoznawania w zadaniu weryfikacji na poziomie 3% a więc niższym od błędu człowieka.

Sieć DeepFace jest siecią opracowaną na użytek komercyjny. W niniejszym artykule zaprezentowano wyniki autorskich prac nad możliwością wykorzystania jednej z udostępnionych do badań naukowych sieci konwolucyjnych o nazwie AlexNet [9] do rozwiązania zadania weryfikacji tożsamości kierowcy. Podstawą działań był zbiór danych obrazowych zebranych w warunkach laboratoryjnych za pomocą kamery bliskiej podczerwieni wyposażonej w filtr optyczny dostosowany do długości fali użytego oświetlacza, a więc za pomocą aparatury predestynowanej do wykorzystania w czasie jazdy, niezależnie od warunków oświetleniowych.

1. SIEĆ ALEXNET

AlexNet jest konwolucyjną siecią neuronową, dzięki której jej autorzy w 2012 roku wygrali organizowane corocznie zawody w rozpoznawaniu obrazów – ImageNet Large Scale Visual Recognition Challenge. Sieć ta uzyskała wynik o około 10 % lepszy niż rozwiązanie, które zajęło 2 miejsce w konkursie. Sieć AlexNet rozpoznaje obiekty należące do 1000 kategorii a wytrenowana została z wykorzystaniem ponad miliona obrazów, będących podzbiorem bazy danych ImageNet [9]. W prezentowanych w niniejszej pracy badaniach wykorzystany został wstępnie wytrenowany model sieci AlexNet dostępny w środowisku Matlab firmy MathWorks.

Standardowa struktura sieci konwolucyjnej, zwanej również w literaturze polskojęzycznej siecią splotową [10], składa się z następujących warstw:

- warstwy wejściowej, której wymiary determinują rozmiar przetwarzanych obrazów;
- jednej lub kilku warstw konwolucyjnych (ang. *convolutional layers*), zwanych również splotowymi, które składają się z określonej liczby neuronów-filtrów o zadanych rozmiarach pól recepcyjnych wytwarzających tzw. mapy cech;
- jednej lub kilku warstw ReLu (ang. *Rectified Linear Unit*) utożsamianych z nieliniową funkcją aktywacji, eliminującą wartości ujemne zwracane przez warstwę poprzedzającą poprzez zastąpienie ich zerami;
- jednej lub kilku warstw redukujących (ang. *pooling layers*), które realizują filtrację statystyczną w obrębie maski o zadanych rozmiarach wyznaczając wybraną statystykę;
- jednej lub kilku warstw pełnego połączenia (ang. *fully-connected layer*), w których każdy neuron jest połączony ze wszystkimi wyjściami warstwy poprzedzającej, przy czym ostatnia warstwa pełnego połączenia posiada tyle neuronów ile klas ma rozpoznawać dana sieć;
- warstwy softmax, która wyznacza wartości prawdopodobieństwa przynależności obrazu wejściowego do poszczególnych klas i jest utożsamiana z funkcją aktywacji ostatniej warstwy pełnego połączenia;
- warstwy wyjściowej, podejmującej decyzję o wyniku klasyfikacji na podstawie wartości zwracanych przez warstwę softmaxu.

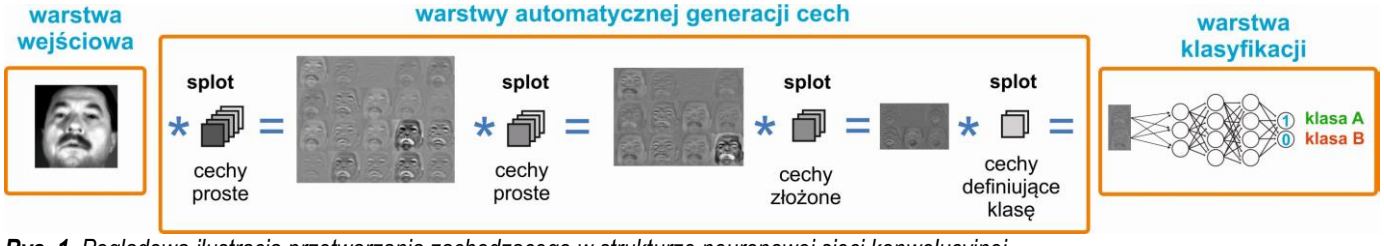
Sieć AlexNet składa się z pięciu warstw konwolucyjnych oraz trzech warstw pełnego połączenia, których parametry podlegają optymalizacji w procesie uczenia. Łącznie sieć ta posiada 60 milionów parametrów, co czyni proces optymalizacji bardzo czasochłonnym i złożonym obliczeniowo. Dlatego procedura uczenia sieci przeprowadzana jest zazwyczaj z wykorzystaniem możliwości obliczeniowych kart graficznych, co znacząco przyspiesza ten proces [9]. Za każdą warstwą konwolucyjną oraz dwoma pierwszymi warstwami pełnego połączenia występuje warstwa ReLU. Zostało udowodnione, że wykorzystanie funkcji ReLU zamiast jej standardowych odpowiedników jako funkcji aktywacji znacząco skraca czas uczenia sieci [11]. W strukturze sieci AlexNet występują również trzy warstwy redukujące, wyznaczające wartość maksymalną w obrębie maski o zadanych rozmiarach. Ustalenie kroku przesuwania maski filtru o wartości większej od 1 pozwala na redukcję wymiaru map cech. Oryginalna sieć AlexNet przetwarzała obrazy o wymiarach 224 x 224 x 3, natomiast wykorzystywany w badaniach model sieci AlexNet wymaga podania na wejście obrazów o rozmiarach 227 x 227 x 3.

2. BAZA DANYCH I METODOLOGIA BADAŃ

Wykorzystanie sieci AlexNet w zadaniu weryfikacji tożsamości wymaga adaptacji jej struktury. Konieczne jest, aby ostatnia warstwa pełnego połączenia posiadała tyle wyjść ile klas ma rozpoznawać dana sieć. Dlatego we wstępnie wytrenowanym modelu sieci AlexNet należy zastąpić trzy ostatnie warstwy nowymi. Wagi ostatniej warstwy pełnego połączenia muszą zostać wyznaczone w procesie tzw. douczenia sieci (ang. *fine-tuning*) z wykorzystaniem odpowiedniej bazy danych.

Przeprowadzone badania miały na celu porównanie efektywności działania dwóch zaproponowanych metod weryfikacji tożsamości opartych na wstępnie wytrenowanym modelu sieci AlexNet. Pierwsze podejście (metoda nr 1) zakładało wykorzystanie jednej douczonej sieci AlexNet. W tym rozwiązaniu na wejście sieci podawane są jednocześnie dwa obrazy twarzy tej samej osoby lub dwóch różnych osób. Zadaniem sieci jest stwierdzenie, że na obrazach wejściowych znajduje się ta sama osoba lub dwie różne. Drugie podejście (metoda nr 2) zakładało wykorzystanie jednej sieci do weryfikacji tożsamości konkretnej osoby. Zatem konieczne jest przeprowadzenie pewnej procedury kalibracyjnej polegającej na douczeniu tylu sieci ile jest osób, których tożsamość może zostać podana w procesie weryfikacji. Zadaniem sieci jest wypracowanie decyzji potwierdzającej lub zaprzeczającej, że obraz wejściowy przedstawia osobę, do rozpoznawania której sieć jest dedykowana.

W badaniach wykorzystana została baza danych obrazowych, zgromadzona w celu opracowania systemu monitoringu stanu psychofizycznego kierowców [12]. Baza danych zawiera obrazy twarzy 19 osób prezentujących normalny wyraz twarzy oraz symulujących wybrane oznaki zmęczenia, takie jak ziewanie czy przymykanie oczu. Wszystkie obrazy zostały zarejestrowane w zakresie biskiej



Rys. 1. *Poglądowa ilustracja przetwarzania zachodzącego w strukturze neuronowej sieci konwolucyjnej*

podczerwieni z wykorzystaniem kamery Basler acA2000-165umNIR wyposażonej w odpowiedni obiektyw szerokopasmowy oraz filtr pasmowo-przepustowy dopasowany do długości fali oświetlacza. Wybór tego zakresu spektralnego wynika z potrzeby monitorowania stanu kierowcy zarówno w dzień jak i w nocy, przy braku dostatecznej ilości światła widzialnego uniemożliwiającego rejestrację obrazu tradycyjną kamerą. Zarejestrowane obrazy zostały poddane procedurze detekcji twarzy z wykorzystaniem algorytmu Violi-Jonesa. Przykładowe obrazy twarzy dwóch osób z bazy danych zostały zamieszczone na rys. 2. Wszystkie zarejestrowane obrazy były obrazami w skali szarości.

Rys. 2. *Przykładowe obrazy twarzy dwóch osób ze zgromadzonej bazy danych*

Metodologia badań różniła się dla obu zaproponowanych metod weryfikacji tożsamości. Na potrzeby metody nr 1 ze zgromadzonych obrazów twarzy 19 osób utworzone zostały pary obrazów przedstawiających te same oraz różne osoby. Na wejście sieci podawane były oba obrazy z pary jednocześnie w następujący sposób: pierwszy obraz umieszczony był w pierwszej warstwie obrazu wejściowego, drugi obraz w drugiej warstwie, natomiast w trzeciej warstwie obrazu wejściowego znajdował się obraz utworzony z wartości średnich poszczególnych pikseli obu obrazów. Stwierdzono doświadczalnie, że taki sposób podawania danych na wejście sieci dawał lepsze wyniki niż różnica obrazów albo obraz średni. Proces uczenia i testowania sieci został przeprowadzony z wykorzystaniem 19 zbiorów danych uczących i testujących. Za każdym razem zbiór testowy tworzyły pary obrazów, z których przynajmniej jeden przedstawiał zadaną osobę. W ten sposób uzyskano wyniki uczenia i testowania sieci dla każdej z 19 osób, przy czym obrazy danej osoby nie zasilały zbioru danych uczących. Liczebność obu klas w zbiorach uczących i testujących była taka sama. Ponadto, za każdym razem procedura uczenia i testowania sieci była powtarzana kilkakrotnie a w artykule przedstawione zostały najlepsze uzyskane wyniki. Jako miara jakości klasyfikatora wykorzystana została dokładność na etapie testowania, wyznaczana jako iloraz liczby poprawnych identyfikacji do liczby wszystkich analizowanych przypadków. Procedura uczenia sieci była wykonywana z wykorzystaniem 64 obrazów w jednej iteracji. W celu uniknięcia efektu przeuczenia sieci proces uczenia był przerywany w momencie gdy wartość średnia dokładności uzyskanych w ostatnich 20 iteracjach była większa niż 95 %.

W metodzie nr 2 na wejście podawane były pojedyncze obrazy twarzy. Zadaniem sieci było określenie czy analizowany obraz przedstawia osobę, do rozpoznawania której sieć została wytrenowana czy nie. Zbiór uczący stanowiło 75 % losowo wybranych obrazów danej osoby oraz taka sama liczba obrazów połowy pozostałych osób. Reszta obrazów danej osoby oraz obrazy osób nie biorących udziału w uczeniu sieci stanowiły zbiór testujący, przy czym liczebność klas również była zrównywana. Dla każdej osoby procedura uczenia i testowania sieci była powtarzana kilka razy tak jak w przypadku metody nr 1. Takie same były również parametry uczenia sieci i warunek stopu.

Do oceny stopnia wytrenowania sieci oraz jej zdolności generalizacji na danych testujących wykorzystana została macierz błędów, zwana również macierzą pomyłek. Jest to macierz kwadratowa, której wiersze odpowiadają zadanym klasom, natomiast kolumny klasom wskazanym w wyniku rozpoznania. Wartości bezwzględne lub względne znajdujące się na przekątnej macierzy odpowiadają poprawnym wynikom, natomiast pozostałe elementy macierzy wskazują na błędy klasyfikacji. W przypadku rozpoznawania dwóch klas macierz błędów jest macierzą czteroelementową. Znaczenie poszczególnych elementów macierzy przedstawiono w tabelach 1 i 2. Wykorzystanie macierzy pomyłek w sposób rzetelny uwzględnia przypadki występowania poszczególnych błędów klasyfikacji.

Tab. 1. *Bezwzględna macierz pomyłek w problemie weryfikacji tożsamości*

Klasa zadana	**Wynik klasyfikacji**	
	Brak zgodności	*Ta sama osoba*
Brak zgodności	TN (ang. True Negative)	FP (ang. False Positive)
Ta sama osoba	FN (ang. False Negative)	TP (ang. True Positive)

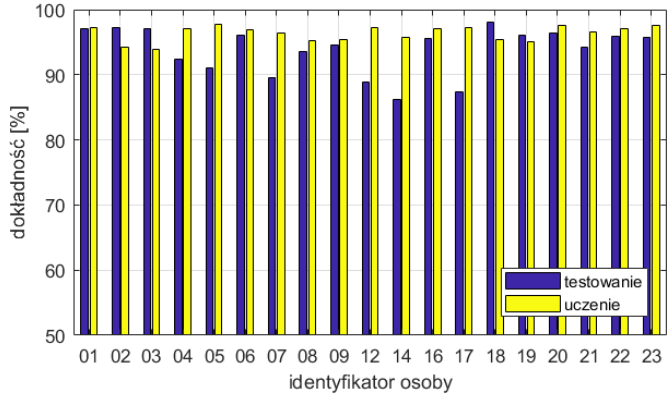
Tab. 2. *Względna macierz pomyłek w problemie weryfikacji tożsamości*

Klasa zadana	**Wynik klasyfikacji**	
	Brak zgodności	*Ta sama osoba*
Brak zgodności	TNR = TN / (TN + FP) (ang. True Negative Rate)	FPR = FP / (TN + FP) (ang. False Positive Rate)
Ta sama osoba	FN = F / (FN + TP) (ang. False Negative Rate)	TPR = TP / (FN + TP) (ang. True Positive Rate)

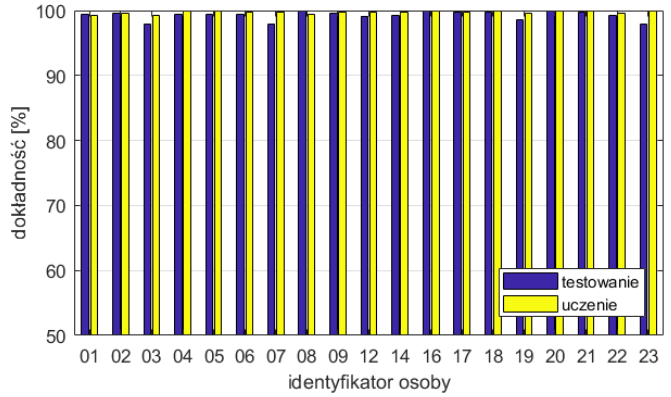
3. WYNIKI

Jak już wcześniej wspomniano, dla obu zaproponowanych metod uzyskano wyniki odpowiadające poszczególnym osobom, których obrazy znajdowały się w wykorzystanej bazie danych. Warto przypomnieć, że w przypadku metody nr 1 żaden obraz wybranej do testów osoby nie występował w bazie danych uczących. Obrazy pozostałych 18 osób zasilały zarówno zbiór uczący jak i testujący. Natomiast w metodzie nr 2 obrazy wybranej osoby, dla której dana sieć była dedykowana, występowały zarówno w zbiorze uczącym jak i testującym i stanowiły klasę przeciwną do reszty osób. Jednak pozostałe 18 osób stanowiło zbiory rozłączne w procesie uczenia i testowania, tzn. obrazy osób zasilających zbiór uczący nie występowały w zbiorze testującym i na odwrót. Takie podejście umożliwiło sprawdzenie działania sieci w przypadku konieczności analizy osoby nowej, nieznanej na etapie uczenia, czyli próbującej oszukać system. Wykresy zamieszczone na rysunkach 3 i 4 przedstawiają dokładności uzyskane dla poszczególnych osób na etapie uczenia i testowania sieci przy wykorzystaniu obu zaproponowanych metod weryfikacji tożsamości. W celu ułatwienia analizy wartości obu osi Y pokrywają zakres od 50 % (poziom klasyfikatora naiwnego) do 100 %. Bez problemu można zauważyć, że uzyskane dokładności testowania są wyraźnie większe w przypadku metody nr 2 i wynoszą ponad 97 %. Natomiast przy zastosowaniu metody nr 1 wartości te wykazują dość duże zróżnicowanie dla poszczególnych osób i wahają się w zakresie od około 87 % do nieco ponad 97 %. Porównując dokładności dla danych uczących można zauważyć, że mniejsze wartości uzyskano dla metody nr 1. Oznacza to, że wykorzystany w trakcie uczenia warunek stopu pozwolił na lepsze dopasowanie się sieci do danych uczących w przypadku metody nr 2. Ponadto, należy zaznaczyć, że spełnienie warunku zakończenia procesu uczenia w przypadku metody nr 2 wymagało wykonania kilkakrotnie większej liczby iteracji. Może to wynikać z faktu, że podejście zasto-

sowane w metodzie nr 1 wydaje się być trudniejszym zadaniem klasyfikacyjnym w porównaniu z metodą nr 2.



Rys. 3. *Dokładności uzyskane dla poszczególnych osób na etapie uczenia i testowania dla metody nr 1*



Rys. 4. *Dokładności uzyskane dla poszczególnych osób na etapie uczenia i testowania dla metody nr 2*

Zdefiniowana powyżej dokładność jest globalną miarą działania klasyfikatora. Dlatego w celu sprawdzenia poprawności rozpoznawania poszczególnych klas w każdym przypadku wyznaczone zostały macierze błędów. W tab. 3 zestawione zostały wartości wskaźników poprawnych identyfikacji uzyskane w procesie testowania dla poszczególnych osób przy wykorzystaniu obu zaproponowanych metod weryfikacji tożsamości. Współczynnik TPR wyraża prawdopodobieństwo otrzymania poprawnej pozytywnej decyzji klasyfikatora, czyli wskazania, że badana osoba jest tą, za którą się podaje. Duża jego wartość oznacza, że przypadki błędnego negatywnego wyniku weryfikacji, czyli stwierdzenia, że analizowana osoba podała nieprawdziwą tożsamość, są relatywnie rzadkie. Natomiast, wskaźnik TNR określa prawdopodobieństwo wystąpienia poprawnego negatywnego wyniku weryfikacji a jego duża wartość wskazuje na rzadkie przypadki otrzymania błędnego pozytywnego wyniku klasyfikacji. Błąd tego typu wydaje się być dużo poważniejszy niż przeciwny, gdyż oznacza zaakceptowanie przez system osoby podającej się za kogoś innego.

Na podstawie bezwzględnych macierzy błędów uzyskanych dla poszczególnych osób z bazy danych wyznaczone zostały wynikowe macierze błędów będące sumą wszystkich 19 macierzy składowych. Następnie obliczono macierze względne. Procedura ta została wykonana zarówno dla danych uczących jak i testujących. Wynikowe względne macierze błędów dla obu zaproponowanych metod weryfikacji tożsamości zamieszczono w tabelach 4 i 5. Wartości obu błędów klasyfikacji na etapie testowania dla metody nr 2 są mniejsze niż 1 %. Zastosowanie metody nr 1 skutkuje czteroprocentowym prawdopodobieństwem uzyskania błędnego pozytywnego wyniku weryfikacji, czyli akceptacji osoby podającej fałszywą tożsamość. Natomiast prawdopodobieństwo popełnienia przez system błędu przeciwnego, czyli odrzucenia osoby podającej prawdziwą tożsamość, wynosi około 10 %. Z punktu widzenia systemu monitorowania kierowcy częstsze przypadki błędnego wykrycia próby oszustwa niż akceptacji fałszywej osoby wydają się być bardziej akceptowalne niż sytuacja odwrotna.

Tab. 3. *Wartości TPR i TNR uzyskane w procesie testowania dla poszczególnych osób*

Identyfikator osoby	Metoda nr 1		Metoda nr 2	
	TPR [%]	*TNR [%]*	*TPR [%]*	*TNR [%]*
01	93,6	98,9	98,8	100,0
02	95,3	98,2	100,0	99,2
03	94,7	98,3	95,8	100,0
04	81,8	97,8	100,0	98,7
05	82,7	95,2	100,0	98,9
06	93,3	97,6	99,4	99,4
07	83,8	92,6	98,8	97,1
08	99,6	90,7	100,0	100,0
09	92,4	95,6	100,0	99,2
12	72,2	97,2	100,0	98,3
14	66,2	96,2	99,6	98,8
16	96,7	95,0	100,0	100,0
17	72,4	94,9	100,0	99,6
18	98,0	98,0	100,0	99,6
19	94,9	96,7	99,5	97,5
20	95,8	96,7	100,0	100,0
21	99,1	91,7	100,0	99,5
22	98,4	94,7	98,6	100,0
23	95,8	95,8	98,8	96,9

Tab. 4. *Wynikowa macierz błędów uzyskana dla metody nr 1*

Klasa zadana	Wynik rozpoznania (testowanie)		Wynik rozpoznania (uczenie)	
	Brak zgodności	*Ta sama osoba*	*Brak zgodności*	*Ta sama osoba*
Brak zgodności	95,9 %	4,1 %	98,6 %	1,4 %
Ta sama osoba	10,2 %	89,8 %	5,9 %	94,1 %

Tab. 5. *Wynikowa macierz błędów uzyskana dla metody nr 2*

Klasa zadana	Wynik rozpoznania (testowanie)		Wynik rozpoznania (uczenie)	
	Brak zgodności	*Ta sama osoba*	*Brak zgodności*	*Ta sama osoba*
Brak zgodności	99,1 %	0,9 %	99,9 %	0,1 %
Ta sama osoba	0,6 %	99,4 %	0,5 %	99,5 %

PODSUMOWANIE

Przedstawiony materiał wskazuje na potencjalną możliwość poprawy funkcjonalności tachografów za pomocą opcji weryfikacji tożsamości z wykorzystaniem współczesnych sieci neuronowych. Poddana badaniom metoda nr 1 wydaje się być bardziej praktycznym rozwiązaniem. Po pierwsze, przygotowanie systemu opartego na takim podejściu wymaga douczenia tylko jednej sieci, natomiast w rozwiązaniu wykorzystującym metodę nr 2 konieczne jest wytrenowanie wielu sieci, po jednej dla każdej osoby, co do pewnego

stopnia komplikuje proces przygotowania systemu. Po drugie, w metodzie nr 1 do procedury uczenia sieci nie są potrzebne obrazy osób, których tożsamości mogą być później podawane w procesie weryfikacji. Procedura weryfikacji wymaga tylko zgromadzenia ich wzorcowych obrazów twarzy w referencyjnej bazie danych. W związku z powyższym gotowy system może zostać dostarczony ewentualnemu odbiorcy, co stanowi olbrzymią zaletę tego rozwiązania. W metodzie nr 2 do przygotowania systemu konieczne jest zgromadzenie odpowiedniej liczby obrazów każdej osoby. Po trzecie, dołączenie kolejnej osoby do systemu weryfikacji w metodzie 1 nie wymaga ponownego douczania sieci, wystarczy tylko do referencyjnej bazy danych dodać obraz twarzy nowej osoby. W metodzie nr 2 należy zebrać odpowiednią liczbę obrazów twarzy nowej osoby a następnie wytrenować dedykowaną sieć do jej rozpoznawania. Niemniej jednak, metoda nr 2 wydaje się być związana z łatwiejszym zadaniem klasyfikacyjnym i cechuje się lepszymi wynikami, co niewątpliwie jest istotną jej zaletą.

BIBLIOGRAFIA

1. Krueger G. P., *Sustained Work, Fatigue, Sleep Loss and Performance: A Review of the Issues, Work and Stress*, An International Journal of Work, Health and Organisations, Volume 3, 1989 - Issue 2.
2. Thiffault P., Bergeron J., *Monotony of road environment and driver fatigue: a simulator study*, Accident Analysis and Prevention 35 (2003), ss. 381–391.
3. http://motoryzacja.interia.pl/raporty/raport-polskie-drogi/wiadomosci/news-pomyslowy-sposob-na-oszukanie-tachografu,nId,2366900 (dostępny kwiecień 2018)
4. https://www.tvn24.pl/wiadomosci-z-kraju,3/tachograf-oszukac-bardzo-latwo-zmuszaja-do-tego-pracodawcy,262569.html (dostępny kwiecień 2018)
5. Stocerz M., wypowiedź w artykule pt. *Kontrole kierowców zawodowych bez zatrzymywania pojazdów*, Specjalistyczny Kwartalnik Informacyjny „Czas na transport", Nr 1 (12), 2018, ss. 68-89.
6. Rozporządzenie Parlamentu Europejskiego i Rady (UE) Nr 165/2014 z dnia 4 lutego 2014r w sprawie tachografów stosowanych w transporcie drogowym.
7. Steele F. J., *Improved digital tachograph system*, patent Nr WO 2006008527 A2, 2006.
8. Taigman Y., Yang M., Ranzato M. A., Wolf L., *DeepFace: Closing the Gap to Human-Level Performance in Face Verification*, 2014 IEEE Conference on Computer Vision and Pattern Recognition, materiały konferencyjne, ss. 1701-1708.
9. Krizhevsky A., Sutskever I., Hinton G. E., *Imagenet classification with deep convolutional neural networks*, Neural Information Processing Systems Conference (NIPS), 2012.
10. Bengio Y., Courville A., Goodfellow I., *Deep Learning. Systemy uczące się*, Wyd. PWN, Warszawa 2018.
11. Glorot X., Bordes A., Bengio Y., *Deep Sparse Rectifer Neural Networks*, Proceedings of the 14th International Conference on Artifcial Intelligence and Statistics (AISTATS), ss. 315-323, 2011.
12. Chmielińska J., Jakubowski J., *Zastosowanie sieci konwolucyjnej do wykrywania wybranych symptomów zmęczenia kierowcy*, Przegląd Elektrotechniczny, vol. 93, no. 10, 2017, ss. 6-10.

Biometrical driver face verification

The paper discusses the problem of face verification in a driver monitoring system for the purpose of traffic safety. Two different methods of face verification were proposed. Both of them are based on a convolutional neural network and were developed with the use of a transfer learning technique. In the paper, the results produced by both proposed method have been presented and compared. Moreover, their advantages and disadvantages have been discussed.

Autorzy:

mgr inż. **Jolanta Chmielińska** – Wojskowa Akademia Techniczna, Wydział Elektroniki, e-mail: jolanta.pacan@wat.edu.pl

dr hab. inż. **Jacek Jakubowski**, prof. ndzw. WAT – Wojskowa Akademia Techniczna, Wydział Elektroniki,
e-mail: jacek.jakubowski@wat.edu.pl

JEL: L92 **DOI**: 10.24136/atest.2018.039
Data zgłoszenia: 2018.05.17 **Data akceptacji**: 2018.06.15