

MICHAŁ DRZAŁ

mgr inż., Sybilla Technologies,
tel: +48 665 892 130, e-mail: michal.
drzal@sybillatechnologies.com

PIOTR OSTASZEWSKI

mgr inż., doktorant w Instytucie
Inżynierii Drogowej, Kolejowej i Trans-
portu, Politechnika Krakowska, Telefon:
+48 669 263 870, e-mail: piotr.
ostaszewski@hotmail.com

Niekonwencjonalne techniki pomiarowe w modelowaniu ruchu¹

Streszczenie: Celem artykułu jest przedstawienie autorskiego projektu wprowadzenia algorytmu opartego o sieci neuronowe w zastosowaniu do pomiarów wykonywanych w transporcie. Jakość, ilość oraz sposób pozyskiwania danych bezpośrednio przekłada się na wyniki tworzonych modeli symulacyjnych. Przeanalizowano różne systemy (zarówno komercyjne, jak i autorskie), które są używane do pozyskania danych do modelowania. W wyniku różnych wątpliwości, niedostosowania systemów lub zbyt wysokich kosztów, zaproponowano alternatywne rozwiązania, które mogą wyeliminować prezentowane problemy. Zaproponowano rozwiązania ograniczające część problemów sygnalizowanych przez autorów w przedmiotowym zakresie. Testowe prace uzasadniły wykorzystanie sieci neuronowych w pomiarach w transporcie. Otrzymano wyniki pomiarów testowych o dostatecznej zgodności z rzeczywistymi obserwacjami oraz porównano je z wynikami systemów dostępnych na rynku. Autorzy poddają analizie dalsze wymagane prace oraz możliwości udoskonalenia stosowanych rozwiązań.

Słowa kluczowe: techniki pomiarowe, modelowanie, pozyskiwanie danych, sieci neuronowe, uczenie maszynowe, automatyczne pomiary ruchu w transporcie

Wprowadzenie

Różnego rodzaju pomiary w obszarze transportu są podstawą analiz istniejących sieci transportowych oraz stanowią integralną część modelowania. Proces wykonywania pomiarów na przestrzeni wielu lat zmieniał się znacząco. Wciąż najczęściej stosowaną metodą są metody manualne. Dotyczy to dokonywania pomiarów na wyznaczonych poligonach oraz manualnego zliczania obiektów z nagranego materiału wideo za pomocą konwencjonalnych kamer. Pozyskiwanie wiarygodnych danych wsadowych na potrzeby tworzonych modeli jest nieodłączną częścią procesu modelowania oraz w znaczący sposób wpływa na jakość wykonywanego modelu. W wielu publikacjach podkreśla się problemy związane z pozyskiwaniem danych w ramach pomiarów lub wręcz z brakiem możliwości pozyskania ich w odpowiedniej jakości. [1, 2, 3]. Wraz z rozwojem modelowania i poprawy dokładności tworzonych modeli, wymagania w zakresie pozyskiwania danych są coraz większe. Zwiększająca się liczba punktów oraz obszarów w ramach prac nad poszczególnymi projektami dyktuje potrzebę

zwiększenia efektywności wykonywania pomiarów oraz skrócenia czasu ich przetwarzania. W komercyjnych oraz naukowych projektach często znaczną część kosztów, zasobów osobowych oraz czasu poświęca się wykonywaniu odpowiednich badań. Na rynku znajduje się wiele rozwiązań mających na celu ułatwienie wykonywania pomiarów oraz skrócenie czasu ich realizacji. Nie wszystkie rozwiązania odpowiadają potrzebom wykonywanych pomiarów w dziedzinie transportu. Stosowane systemy są często nieefektywne lub niedostosowane do realizacji takich zadań, szczególnie, gdy realizowany zakres pomiarów jest krótkookresowy, a wymagania ustawienia stanowisk pomiarowych są bardzo pracochłonne.

W artykule przedstawiono kilka przykładów rozwiązań dostępnych na rynku i zaproponowano własne rozwiązania z zakresu opracowania sprzętu pomiarowego oraz algorytmów wspomagających zliczanie obiektów z materiałów wideo.

Problemy i ograniczenia aktualnie stosowanych technik pomiarowych w transporcie

Najprostszą dostępną techniką pomiarową są manualne pomiary wykonywane przez człowieka. Metoda ta jest silnie uzależniona od warunków atmosferycznych oraz od konkretnej osoby odpowiedzialnej za dany punkt pomiarowy. Jest ona zarówno najbardziej kosztowna, jak i pracochłonna oraz stanowi znaczne wyzwanie organizacyjne, szczególnie w przypadku konieczności wykonania pomiarów w kilkudziesięciu punktach pomiarowych. Pomimo to jak do tej pory jest to najszerzej stosowana technika pozyskiwania danych pomiarowych. Metoda ta nie zapewnia możliwości weryfikacji materiału, a przy rosnących stale kosztach pracy staje się coraz mniej opłacalna. W celu umożliwienia weryfikacji otrzymanego materiału stosuje się nagrywanie wideo danych poligonów lub wyselekcjonowanych miejsc. W większości przypadków stosuje się konwencjonalne kamery lub kamery CCTV. Jeśli chodzi o manualne sposoby pozyskiwania danych z nagranego materiału, to takie rozwiązania są wystarczające. Niestety, chcąc przeprowadzić profesjonalną analizę obrazu oraz wykonywanie pomiarów, to metody te posiadają znaczące wady. Typowe kamery CCTV mają wbudowane algorytmy, które sterują liczbą nagrywanych klatek w danej sekundzie. Wiąże się to z ich przeznaczaniem i najlepszą zdolnością do ograni-

¹ ©Transport Miejski i Regionalny, 2018. Wkład autorów w publikację: M. Drzał 50%, P. Ostaszewski 50%.

czania gromadzenia niepotrzebnego materiału wideo. Jak wynikało z przeprowadzonych testów polegających na rozkładaniu materiału na poszczególne klatki (założono zmienną liczbę klatek w sekundzie zgodną z liczbą wstępnie określoną w oprogramowaniu), straty w czasie nagrania wynikły z działania tych algorytmów wynosiły około 1 min, co przy nagrywaniu materiału całodobowego stanowi znaczącą przeszkodę.

W większości materiałów uzyskanych z konwencjonalnych kamer stosowany jest np. przepłot [4]. Polega on na zapisie materiału wideo w sposób naprzemienny za pomocą linii parzystych bądź nieparzystych (rys. 1)



Rys. 1. Przykład przepłotu w materiale wideo
Źródło: opracowanie własne

Przy analizie obrazu np. w celu pomiarów prędkości, zmienna liczba klatek w danej sekundzie oraz przepłot stosowany w kamerach bardzo komplikuje możliwości prawidłowego pomiaru oraz odpowiedniej obróbki uzyskanego materiału (np. artefakty). Dodatkowym czynnikiem działającym na niekorzyść konwencjonalnych kamer jest obecność filtrów podczerwieni, które są zamontowane w obiektywach (większość kamer CCTV posiada jednak funkcję „zdejmującą” filtr podczerwieni – IR-CUT). Uniemożliwiają one realizację pomiaru w warunkach słabego oświetlenia.

Zastosowanie tradycyjnych metod nagrywania również rodzi pewne wątpliwości ze względu na zasilanie i rozpoznawalność sprzętu dla innych osób niezaangażowanych w pomiar. W większości wymienione wyżej kamery zasilane są za pomocą napięcia 12V lub 240V, co komplikuje rozmieszczenie sprzętu pomiarowego, wydłuża przygotowania i czas jego instalacji oraz znacząco zwiększa jego roz-

miary. Rozpoznawalność sprzętu pomiarowego np. w przypadku wykonywania pomiarów ruchu tranzytowego lub pomiarów prędkości na danym odcinku jest często czynnikiem mającym wpływ na sposób realizacji pomiarów. W przypadku odczytu rejestracji lub nagrania danego skrzyżowania, stosując tradycyjny statyw oraz kamerę, znacząco zwiększamy szansę na zlokalizowanie przez osoby postronne sprzętu pomiarowego oraz szybkie rozpoznanie realizowanych przez niego funkcji. Często osoby rozpoznające kamerę zainstalowaną na poligonie zmieniają swoje zachowanie, a sytuacje przez nich stwarzane zagrażają bezpieczeństwu ruchu w obrębie miejsca wykonywania czynności pomiarowych. Do częstych reakcji uczestników ruchu w takich sytuacjach należą: zwalnianie, celowe omijanie miejsc objętych pomiarem, zatrzymywanie się, blokowanie ruchu, wzywanie służb porządkowych, a nawet groźby pod adresem osób wykonujących pomiar.

Podejmowane przez badaczy próby uzyskiwania dokładnych danych pomiarowych przy zastosowaniu dostępnych systemów są bardzo utrudnione. Szczególnie prace dotyczące pomiarów pieszych narażone są na znaczne błędy oraz utrudnienia w ich realizacji (rys. 2). Sama kalibracja musi zostać wykonana poprawnie i jest ona czasochłonna. Przykłady realizacji pomiarów ruchu pieszego przez badaczy podkreślają pracochłonność wykorzystywanych systemów [5, 6, 7, 8]

Ze względu na opisane wyżej wady stosowania powszechnie dostępnego w sklepach, tradycyjnego sprzętu do nagrywania obrazów wideo, niektóre firmy komercyjne zdecydowały się na opracowanie sprzętu, który w większym zakresie odpowiadałby potrzebom pozyskania wiarygodnych danych (tj. np. Autoscope, Intellio, MioVision, Wavetronix). W ramach prac prowadzonych na Politechnice Krakowskiej, przy realizowanych projektach naukowo-komercyjnych, możliwe było w warunkach praktycznych zweryfikowanie części dostępnego na rynku sprzętu. Niektóre kamery oferowane na rynku (np. firma Autoscope) ze względu na swój wygląd kojarzone są przez kierowców z fotoradarami (rys. 3) oraz muszą być podpięte stale do zasilania zewnętrznego, ponieważ odczyt tablic jest wykonywany w czasie rzeczywistym. Potrzeba użycia zasilania prądem o wysokim napięciu w warunkach polowych znacząco utrudnia realizację pomiaru. Oprogramowanie dostarczane do kamer umożliwia też odczyt tablic z zadanego



Rys. 2.
Przykład zaawansowania kalibracji sprzętu potrzebnego do pomiarów ruchu pieszego
Źródło: [5, 8]



Rys. 3. Autoscope Pn-100 ANPR
Źródło: opracowanie własne

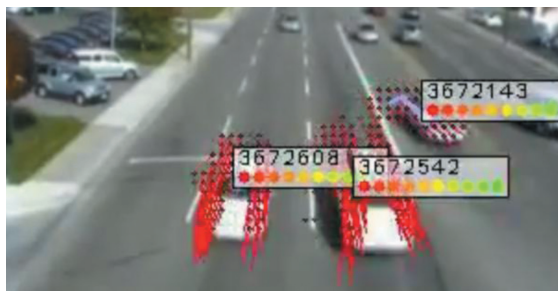
wcześniej nagranych materiału wideo, ale odpowiednie ustawienie konwencjonalnej kamery oraz odpowiednie przygotowanie materiału jest pracochłonne i bardzo wrażliwe na błędy, a w dłuższym okresie nie daje zadowalających wyników. Natomiast ustawienie na obszarze analizy detektorów wykrywających obiekty powiązane jest często ze wstępną kalibracją kamery przed wykonaniem pomiarów, co wymaga czasu i odpowiedniej dokładności oraz umożliwia, w niektórych przypadkach, tylko zapis zdarzeń w czasie rzeczywistym i brak możliwości ponownego zweryfikowania materiału.

Problemy związane z kalibracją sprzętu pomiarowego i poprawnym wykrywaniem obiektów znajdują również potwierdzenie w literaturze [9]. Podczas badań niepoprawne (niewiarygodne) wyniki liczby zliczonych pojazdów na skrzyżowaniu, uzyskano aż na 40% analizowanych skrzyżowań. Autorzy badań podkreślają, że stosowany przez nich sprzęt może mieć zastosowanie do detekcji pojazdów na skrzyżowaniu ale w przypadku dokładnego liczenia pojazdów nie można mu zaufać. W przypadku zastosowania systemu MioVision Scout możemy mówić o wyeliminowaniu problemów dotyczących zasilania oraz instalacji kamery. System posiada specjalnie opracowane własne zasilanie oraz maszt, a analiza obrazu wideo odbywa się po nagraniu materiału wideo. Niestety w przypadku analizy wideo uzyskany obraz z kamery MioVision jest słabej jakości, pozwala on na detekcję pojazdów, ale znacząco wpływa na jej jakość oraz zasięg możliwej detekcji (rys. 4)

Użycie radarów montowanych na mobilnych masztach nie wymaga wielu czynności związanych z kalibracją. Niestety dużą wadą tego typu systemu jest możliwość zliczania obiektów tylko w danym przekroju oraz często brak możliwości wykrycia pojazdu ze względu na wzajemne przesłanianie się pojazdów.

Nie bez znaczenia są koszty pozyskania komercyjnego sprzętu do pomiarów. Są one bardzo duże. Koszt sprzętu firmy Autoscope oszacowany dla skrzyżowania 4 – wlotowego – kształtuje się w okolicach 20 tysięcy dolarów [9], natomiast zakup masztu oraz kamery MioVision to koszt około 3–4 tysiące dolarów. W przypadku systemu MioVision należy także uwzględnić dodatkowe opłaty za każdy wysłany materiał do analiz.

Wykonanie poprawnych pomiarów ruchu eliminujących kompleksowo problemy związane z ich realizacją jest dużym wyzwaniem zarówno dla badaczy, jak i dla osób komercyjnie je realizujących. Użycie amatorskiego sprzętu do



Rys. 4. Przykładowy kadr podczas detekcji pojazdów przez system MioVision
Źródło: [10]

wykonania pomiarów łączy się ze sporymi wadami i utrudnieniami w ich realizacji. Natomiast koszty zastosowania profesjonalnego sprzętu często są bardzo wysokie przy realizacji pomiarów np. w kilkunastu różnych miejscach jednocześnie. Użycie profesjonalnego sprzętu pomiarowego nie daje też gwarancji uzyskania odpowiedniej jakości wyników. Wysokie wydatki na sprzęt pomiarowy nie gwarantują otrzymania wszechstronnego sprzętu nadającego się do pozyskania większości danych pomiarowych potrzebnych w procesach modelowania komputerowego. Bardzo rzadko uwzględnia się możliwość pomiaru ruchu pieszego, a większość producentów skupia swoje rozwiązania na pojazdach kołowych.

Istniejące rozwiązania oraz problemy związane z automatycznymi pomiarami w transporcie

Pomiary ruchu samochodowego oraz ruchu pieszych są dziedziną, w której ciężko było stworzyć rozwiązanie zdolne do automatycznego zliczania. Wynika to z fundamentalnych różnic pomiędzy ludzką analizą obrazu a tą dokonywaną przez maszyny.

Pierwszą różnicą jest różnica sprzętowa. W przeciwieństwie do ludzkiego wzroku, który opiera się na parze oczu zapewniających widzenie stereoskopowe, większość dostępnych na rynku rozwiązań opiera się na pojedynczej kamerze. Tracona w ten sposób informacja o relatywnej odległości pomiędzy obiektami utrudnia rekonstrukcję trójwymiarowej sceny i ustalenie pokonanej przez śledzony obiekt trasy.

Drugą różnicą tkwi w ludzkim mózgu. Podczas gdy maszyny operują na dwuwymiarowych macierzach jasności pikseli w poszczególnych kolorach – czerwonym, niebieskim i zielonym – to ludzie są w stanie automatycznie dzielić trójwymiarową scenę na poszczególne obiekty oraz dokonywać ich klasyfikacji bez wykonywania świadomego wysiłku.

Trzecią istotną różnicą jest zdolność do tworzenia modelu trójwymiarowego sceny. Człowiek, widząc nawet dwuwymiarowy obraz przedstawiający odcinek autostrady jest w stanie bez większych problemów odtworzyć w głowie zależności przestrzenne pomiędzy widocznymi obiektami, w tym istniejące źródła światła. Ludzki mózg posiada również wbudowany model fizyczny świata, który potrafi wyciągać poprawne wnioski na temat obserwowanych obiektów. Przykładem może być osoba, która, przechodząc przez scenę, została na chwilę przesłonięta przez filar lub samochód, który przez chwilę został przesłonięty przez roślinność. Człowiek jest w stanie bez wysiłku zauważyć, że ma

do czynienia z tym samym obiektem. W przypadku automatycznych rozwiązań konieczne jest ponowne zidentyfikowanie obiektu i sprawdzenie jego podobieństwa z obiektem, który zniknął.

W zależności od poziomu sprzętowego rozwiązania może opierać się ono wyłącznie na systemie kamer lub również wykorzystywać dodatkowe informacje o odległości dostarczane przez dodatkowe czujniki. Przy analizie istniejących rozwiązań skupiono się na rozwiązaniach opierających się wyłącznie na obrazie uzyskiwanym z kamer. Są to rozwiązania, które znacznie lepiej nadają się do automatycznego pomiaru ruchu, głównie ze względu na łatwość instalacji oraz demontażu.

Większość istniejących [11], [12], [13] rozwiązań opiera się na rozwiązaniach z zakresu widzenia komputerowego (ang. *computer vision*), w których algorytm, według ustalonych z góry kroków, dokonuje ekstrakcji obiektów. Pierwszym krokiem jest segmentacja obrazu na części, która pozostaje niezmienna lub zmienia się bardzo powoli (tło) oraz poruszające się w ramach sceny obiekty. Dla rozwiązania tego problemu utworzono kilka algorytmów przedstawionych w publikacjach [14], [15], [16].

Na podstawie dokonanej segmentacji wybierane są obszary, które nie zostały zaklasyfikowane jako tło. Dla każdego obszaru obliczany jest zestaw statystyk, na podstawie których odfiltrowywane są obiekty, które nie są pojazdami lub pieszymi. Takie podejście relatywnie dobrze sprawdza się przy pomiarach ruchu, gdzie kamera znajduje się bezpośrednio nad drogą i jedynym jej zadaniem jest pomiar pojazdów przejeżdżających przez wybrany przekrój. Problemy występują, gdy analizowana scena posiada szybko zmieniające się oświetlenie (pomiar o poranku lub późnym wieczorem, reflektory pojazdów z sąsiednich dróg generujące dodatkowe cienie). Dodatkowo większość algorytmów nie dokonuje semantycznej analizy rejestrowanych obiektów. Wnioskowanie opiera się jedynie na rozmiarze zarejestrowanych obszarów i dlatego jest bardzo podatne na błędy. W zależności od ustawienia kamery wymagana jest kalibracja parametrów związanych z poszczególnymi klasami obiektów.

Jakościowym skokiem w widzeniu maszynowym jest zastosowanie sieci neuronowych. Jednym z pierwszych zastosowań było wykorzystanie w rozpoznawaniu odręcznie napisanych liter [17]. Wraz ze wzrostem dostępnej mocy obliczeniowej komputerów sieci neuronowe zaczęły być wykorzystywane do klasyfikacji obrazów niskiej rozdzielczości [18]. Szczególnie konwolucyjne sieci neuronowe [19], inspirowane mechanizmem działania ludzkiej kory wzrokowej [20], okazały się być szczególnie użyteczne w rozwiązaniu kwestii, z którymi ludzki mózg radzi sobie bez większego problemu. Do takich problemów należy detekcja w scenie i klasyfikacja pojazdów oraz pieszych.

Konwolucyjne sieci neuronowe są podstawowym komponentem w konstrukcji autonomicznych pojazdów [21], [22]. Głównym problemem przy konstrukcji tego typu pojazdów jest zdolność do szybkiego wyszukiwania w przychodzącym strumieniu obrazów podstawowych obiektów:

pojazdów, znaków drogowych, pieszych, rowerzystów itd. Dodatkowo algorytmy są w stanie wyznaczać maski odpowiadające poszczególnym obiektom wykrytym na obrazie z kamery. Informacja ta pozyskana z kilku różnych kamer może posłużyć do oceny odległości stanowiska od obiektu.

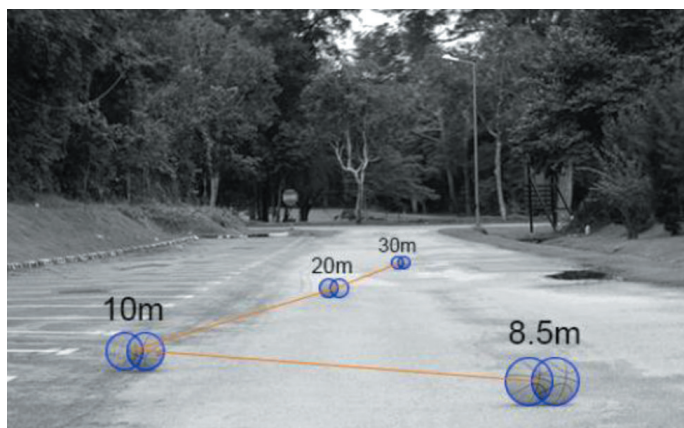
Przy kamerach przeznaczonych do pomiaru ruchu zadanie jest analogiczne. Jediną zasadniczą różnicą jest rozmieszczenie systemu kamer. W przypadku autonomicznych pojazdów względne położenia oraz parametry kamer są dokładnie znane, w przypadku pomiarów ruchu drogowego względne położenie kamer może znacznie się różnić. Przy zastosowaniu systemu odrębnych kamer rodzą się znaczne problemy z możliwościami instalacji sprzętu pomiarowego. Znaczącym ułatwieniem w analizach jest stacjonarne położenie kamer względem analizowanej sceny.

Rozwiązania wdrożone w pracach oraz propozycje możliwe do zastosowania

Rozwiązania techniczne dotyczące rejestrowania danych punktów w przestrzeni są bardzo duże. Ze względu na istotne wady dotychczas stosowanych rozwiązań zostały zdiagnozowane dwa podstawowe elementy z zakresu sprzętowej realizacji pomiarów:

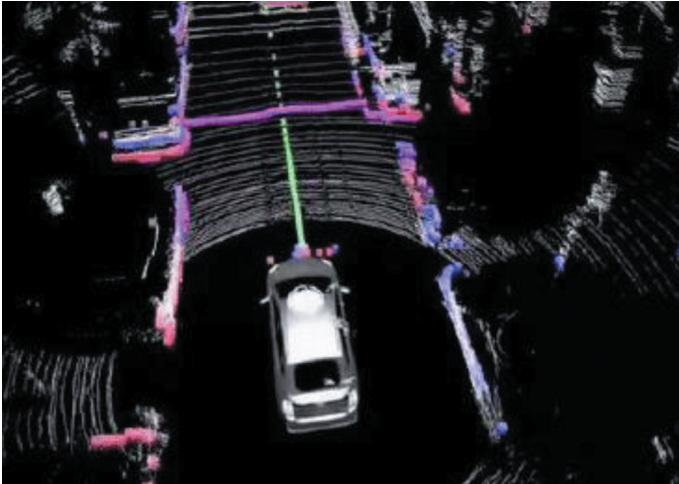
- podstawowe – zastosowanie innych typów kamer z lepszymi przetwornikami obrazu,
- złożone – kombinacja różnych systemów wzajemnie się weryfikujących /uzupełniających.

Podstawowe rozwiązanie może bazować na dotychczas wykorzystywanych technologiach, jednocześnie prezentując nowe funkcje, do tej pory nieużywane w trakcie realizacji pomiarów w transporcie. Jedną z nich jest kamera stereo. Jest to system bazujący na dwóch kamerach oddalonych od siebie o znaną odległość, który jest w stanie, poprzez analizę pozyskanych obrazów, oszacowywać odległość danych przedmiotów od siebie. Badacze znaleźli zastosowanie tego systemu dla pojazdów autonomicznych [23]. Osiągnięto dobrą zgodność wykonanych pomiarów z rzeczywistością, a zaprezentowane błędy są na tyle małe, że takie rozwiązanie może zostać ujęte jako moduł mogący zostać wykorzystany w pomiarach. Przykład odczytu odległości za pomocą kamery stereo zaprezentowano na rysunku 5.



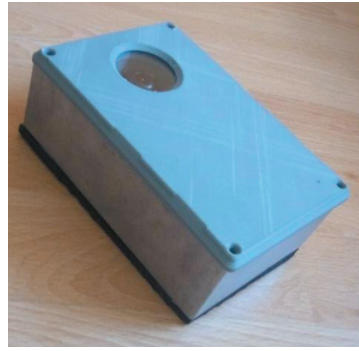
Rys. 5. Przykład odczytu odległości z kamery stereo
Źródło: [23]

Uzupełnieniem systemu kamer mogłyby być urządzenia, które mogą pracować niezależnie od warunków atmosferycznych. Efektywnymi rozwiązaniami mogącymi ułatwić pozyskanie odpowiednich danych są systemy oparte na czujnikach ultradźwiękowych [24] oraz systemy oparte na technice laserowej [25]. Przykład działania systemu opartego na systemie LIDAR został zaprezentowany na rysunku 6. Zastosowanie układów złożonych z kilku technologii ma jednak znaczące wady przy wykonywaniu pomiarów krótkookresowych bądź pomiarów w ciasnych obszarach. Są nimi: potrzeba dostarczenia zasilania całego systemu oraz pracochłonność jego instalacji.



Rys. 6. Przykład działania systemu opartego na technologii LIDAR
Źródło: [26]

W ramach prowadzonych prac badawczych udało się stworzyć kamerę pomiarową (rysunek 7), która eliminuje znaczącą większość problemów sprzętu do nagrań wideo dotychczas stosowanego do pomiarów. Ze względu na testowy charakter opracowanego sprzętu możliwe jest na tym etapie wyszczególnienie tylko jego głównych elementów. Podstawowym czynnikiem brany pod uwagę było dostosowanie kamery do możliwości zdalnego sterowania oraz zarządzania nią za pomocą sieci bezprzewodowej, co eliminowało czynnik ludzki podczas realizacji pomiarów. Udało się to uzyskać za pomocą zaprogramowanego mikrokomputera z odpowiednio dobranymi podzespołami. Moduł elektroniczny kamery został dobrany tak, aby zapewnić wysoką rozdzielczość uzyskiwanego materiału, szeroki kąt oraz możliwość pozyskania materiału w warunkach niedostatecznego oświetlenia. Kolejne wady, które udało się zniwelować, to wymiary sprzętu, jego zasilanie oraz możliwość ciągłego nagrywania. Ze względu na pełne możliwości zaprogramowania mikrokomputera możliwe jest uzyskiwanie ciągłych nagrań bez ograniczeń czasowych. Podczas testów opracowany sprzęt bez problemu realizował 48 godzin ciągłego nagrywania bez potrzeby uzupełniania zasilania. Ze względu na zastosowanie różnych elementów elektronicznych, do opracowania odpowiedniego projektu jej obudowy, użyto drukarki 3D. Dzięki temu uzyskano kształt urządzenia, który nie kojarzy się z tradycyjnymi kamerami.



Rys. 7.
Zaprojektowana kamera
Źródło: opracowanie własne

Autorzy artykułu podjęli pracę na materiale uzyskanym po wykonaniu pomiarów, a nie w ich trakcie, ze względu na oszczędności w zakresie sprzętowym (mniejsze wymagania dla sprzętu) oraz możliwość późniejszej weryfikacji materiału. Biorąc pod uwagę algorytm automatycznie zliczający pojazdy, zdecydowano się na użycie oraz wytrenowanie konwolucyjnej sieci neuronowej do rozpoznawania obiektów kluczowych w pomiarach wykonywanych w transporcie przy jednoczesnym założeniu, że system kamer będący źródłem obrazów jest stacjonarny.

Trenowana sieć neuronowa opiera swoją architekturę na publikacji Facebook AI Research (FAIR) [27]. Proces detekcji obiektów na ramce koncepcyjnie składa się z następujących etapów:

- generowanie prostokątów zawierających regiony z potencjalnymi kandydatami na obiekty;
- klasyfikacja każdego z prostokątów (przyporządkowanie do klasy: samochód osobowy, samochód ciężarowy, pieszy) wraz z dodaniem numerycznej miary określającej pewność klasyfikacji (w zakresie od 0 do 1);
- generacja maski pokrywającej wyłącznie piksele przyporządkowane do obiektu w ramach prostokątów, co do których algorytm jest pewny ich klasyfikacji (są klasyfikowane z pewnością wyższą niż 0.7).

Wytrenowana sieć neuronowa jest w stanie znajdować obiekty w każdej klatce filmu. Osobną kwestią, która zostanie uwzględniona w planach dalszych prac, jest powiązanie obiektów w kolejnych klatkach. Jest to kluczowe przy wyznaczaniu trasy, po której przemieszczał się obiekt w scenie, i pozwala na precyzyjne pomiary ruchu.

Testowe wyniki zastosowanego sprzętu oraz oprogramowania

Opracowany sprzęt został przetestowany w różnych warunkach atmosferycznych oraz sprawdzono różne możliwości nagrywania. Średni błąd, biorąc pod uwagę zakładaną liczbę klatek, wyniósł 0,04% przy rozdzielczości FullHD. Jest to błąd pozwalający na analizy materiału wideo bez przeszkód i umożliwia otrzymanie dobrych wyników zastosowanych algorytmów. W ramach prowadzonych testów wykonano próbne nagrania w kilkudziesięciu różnych lokalizacjach. Do dalszych analiz wyselekcjonowano kilkadziesiąt zróżnicowanych poligonów. Odczytów dokonano zarówno na pojazdach, jak i na pieszych. Zestawienie prezentujące typy materiałów zostało przedstawione w tabeli 1.

Tabela 1

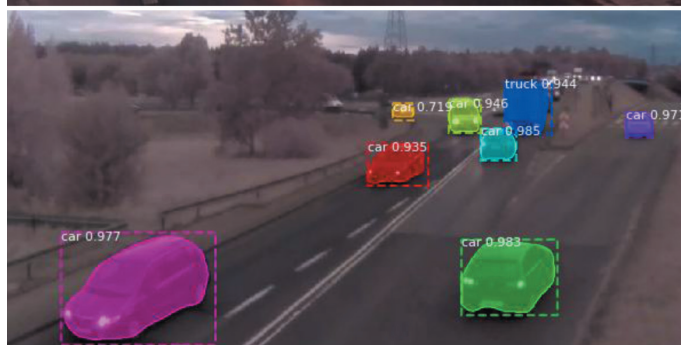
Zestawienie przeanalizowanych typów materiałów wideo	
Typy materiałów	Liczba analiz
Piesi przechodzący przez wąski przekrój	3
Piesi przechodzący przez szeroki przekrój	2
Pojazdy – skrzyżowanie	9
Pojazdy – przekrój drogi	6
Pojazdy – inne (np. pasy wyłączenia/pasy do skrótu)	8
Suma	28

Źródło: opracowanie własne

W ramach prowadzonych testowych prac zdecydowano się na wypróbowanie nie tylko elementów wykrywających pojazdy, ale także sklasyfikowania wykrywanych pojazdów. Wyniki kategoryzacji pojazdów przez algorytm są zadowalające, ale napotkano na znaczące problemy w przypadku odczytu pojazdów dostawczych oraz małych pojazdów ciężarowych. Pojazdy te są często bardzo zbliżone wymiarami, co powoduje błędne odczyty algorytmu. W przypadku dużych różnic wymiarowych kategorii pojazdów, jak pojazdy osobowe, motocykle czy samochody ciężarowe, algorytm działa poprawnie. W przypadku pieszych, podobnie jak w innych pracach cytowanych w artykule, jednym z największych problemów było wzajemne przesłanianie się pieszych i takie sytuacje rodziły najwięcej błędów odczytu. Może to zostać wyeliminowane przez odpowiednie śledzenie pieszych oraz ustawienie obiektywu kamery jak najwyżej nad głowami pieszych. Przykładowe zrzuty ekranu prezentujące rozpoznawanie obiektów (pojazdów oraz postaci ludzkich) zostały przedstawione na rysunkach 8 i 9.

Problemem przy analizie obrazu okazała się pora dnia lub niepożądane zmiany kolorów występujące na obrazie np. cień. Ze względu na różne odbłyски światła, słabą widoczność brył pojazdów oraz ograniczenia sprzętowe algorytm wykrywający obiekty nie był w stanie poprawnie poradzić sobie z taką liczbą zakłóceń. Wykazano, że występują znaczące różnice w wykrywaniu obiektów na wytypowanych materiałach ze względu na warunki pogodowe i oświetleniowe. Drugorzędny element powodującym ewentualne błędy był typ analizowanego materiału – błąd ten był na tyle mały, że zdecydowano się na prezentację błędów dotyczących warunków atmosferycznych/oświetleniowych. Wyniki wykrywania pieszych dotyczą tylko i wyłącznie dobrych warunków pogodowych i oświetleniowych. Na te warunki natężenie ruchu pieszego jest bardzo wrażliwe.

Pomimo błędów odczytu algorytmu w warunkach zmroku oraz cieni na jezdni wyniki pracy algorytmu są obiecujące. Błędy osiągnięte przez działający algorytm w dobrych warunkach atmosferycznych oraz dla ruchu pieszego pozwalają na otrzymanie wyników porównywalnych z wynikami otrzymywanymi w przypadku manualnego zliczania materiału przez ludzi (rys. 10). Biorąc pod uwagę kontynuowanie prac nad algorytmem oraz zakładane polepszenie parametrów sprzętowych, można stwierdzić, że uzyskiwanie poprawnych wyników w większości przypad-



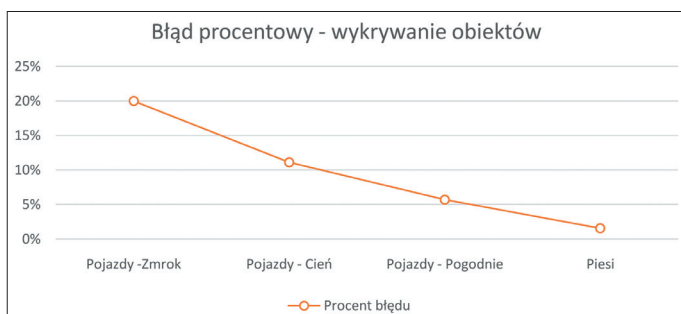
Rys. 8. Przykład wykrywania pojazdów przez przedstawiany algorytm

Źródło: opracowanie własne



Rys. 9. Przykład wykrywania pieszych przez przedstawiany algorytm

Źródło: opracowanie własne



Rys. 10. Wykres błędów wykrywania obiektów w zależności od typu obiektu oraz warunków pogodowych

Źródło: opracowanie własne

ków lokalizacji punktów pomiarów jest możliwe. Pozwoli to znacznie zaoszczędzić czas przeznaczony na opracowanie pomiarów i otrzymać wiarygodne dane, na których można oprzeć dalsze tworzenie i kalibrację modelu.

Podsumowanie i dalsze prace

Opracowany sprzęt oraz nowatorskie zastosowanie sieci neuronowych przyniosły zaskakująco dobre rezultaty. Porównując próbne wyniki z wynikami uzyskiwanymi przez komercyjne firmy z dużym zapleczem finansowym [28], uzyskano lepsze zgodności otrzymywanych danych. Uzyskane wyniki potwierdzają założenie, że konwolucyjne sieci neuronowe mogą z powodzeniem zostać wykorzystane w obszarze pomiarów ruchu drogowego oraz

piesznych. W ramach dalszych prac zostanie zastosowana procedura pozwalająca wyznaczać trasę obiektów wykrywanych na poszczególnych klatkach. Przy pojedynczej kamerze trasa ta będzie jedynie zlokalizowana w dwuwymiarowym układzie współrzędnych projekcji obrazu na matrycy kamery. Przy wykorzystaniu systemu kamer (np. za pomocą kamer stereo) będzie można pokusić się o próbę wyznaczania trójwymiarowej lokalizacji śledzonych obiektów w ramach sceny. Warto zauważyć, że do pomiaru natężenia ruchu przez wybrany przekrój, przy wystarczająco wysokim umieszczeniu, wystarczy pojedyncza kamera, i odtwarzanie trójwymiarowej lokalizacji w scenie nie będzie konieczne.

W przypadku powodzenia dalszych prac badawczych czynnik ludzki wykorzystywany w procesie pozyskania danych/informacji o transporcie będzie ograniczony do minimum. Pomimo dużych możliwości technologicznych oraz coraz bardziej skomplikowanych algorytmów implementowanych w celu uzyskania jak najlepszych wyników automatyzacja pomiarów nie będzie uniwersalnym narzędziem znajdującym zastosowanie w każdym przypadku. Doświadczenie osób oraz czujność przy weryfikacji otrzymanych danych pozwolą uchwycić granice zastosowań systemów, które będą już w pełni rozwinięte. Natomiast znacząco przyspieszy to otrzymywanie danych na potrzeby prowadzonych analiz (testy wdrożonego oprogramowania pozwalają na odpowiedniej stacji roboczej zredukować czas analiz o około 50%). Optymalizacja używanego sprzętu oraz odpowiednie jego wykorzystanie obniży koszty realizacji pomiarów oraz umożliwi wykonywanie ich w większej liczbie.

Literatura

1. Boukhanovsky A., Krzhizhanovskaya V., Melnikov V., Sloot P., *Data-driven modeling of transportation systems and traffic data analysis during a major power outage in the Netherlands*, YSC 2015, 4th International Young Scientists Conference on Computational Science, 2015.
2. Network Rail: Station Capacity Assessment Guidance, maj 2011.
3. Gaca S., Suchorzewski W., Tracz M., *Inżyniera ruchu drogowego. Teoria i praktyka*, Warszawa 2011.
4. Farid H., Wang W., *Exposing Digital Forgeries in Interlaced and De-Interlaced Video*.
5. Hermant L.F.L., *Human movement behavior in South African railway stations: implications for design*.
6. Inamura H., Takeyama Y., Teknomo K., *Measuring Microscopic Flow Performance For Pedestrians*, Tohoku University, Japan, The 9th WCTR Selected Proceedings.
7. Johansson A.F. *Data-Driven Modeling of Pedestrian Crowds*, Doctoral Thesis, 2009.
8. Boltz M., Schadschneider A., Seyfried A., Steffen B., *Automatic Extraction of Pedestrian Trajectories from Video Recordings*, Pedestrian and Evacuation Dynamics 2008.
9. Bonnette J., Codjoe J., Ishak S., Jenkins S., Mousa S., *Traffic Counting Using Existing Video Detection Cameras*, Final Report – November 2015.
10. <https://www.youtube.com/watch?v=m3z79fLczDM>, dostęp w 22.03.2018 r.
11. Cong T., Crouzil A., Khoudour L. Nghi D., Valiere P., *Automatic Vehicle Counting System for Traffic Monitoring*, Journal of Electronic Imaging, Society of Photo-optical Instrumentation Engineers, 2016.
12. Kumar V., Majeti M., *Vehicle Detection, Tracking and Counting Objects For Traffic Surveillance System Using Raspberry-Pi*, 2015.
13. Thool R., Uke N., *Moving Vehicle Detection for Measuring Traffic Count Using OpenCV*, Journal of Automation and Control Engineering.
14. Bowden R., Kaewtrakulpong P., *An Improved Adaptive Background Mixture Model for Realtime Tracking with Shadow Detection*, Proceedings of 2nd European Workshop on Advanced Video-Based Surveillance Systems, 2001.
15. Zivkovic Z., *Improved Adaptive Gaussian Mixture Model for Background Subtraction*, In Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04), Vol. 2.
16. van der Heijden F., Zivkovic Z., *Efficient adaptive density estimation per image pixel for the task of background subtraction*, Pattern Recognition Letters, Volume 27, Issue 7, 2006.
17. LeCun, Yann et al., *Handwritten Digit Recognition with a Back-Propagation Network*, NIPS (1989).
18. Hinton G., Krizhevsky A., Sutskever I., *Imagenet classification with deep convolutional neural networks*, Advances in Neural Information Processing Systems 25, Curran Associates, 2012.
19. Le Cun, Y., Boser B., Denker J.S., Henderson D., Howard R. Hubbard W., Jackel L., *Handwritten Digit Recognition with a Back-Propagation Network*, LeNet-5.
20. Kaneda Y., Masakazu M., Mitari Y., Mori K., *Subject independent facial expression recognition with robust face detection using a convolutional neural network*.
21. Bojarski M., Testa D.D., Dworakowski D., Firner B., Flepp B., Goyal P., Jackel L.D., Monfort M., Muller U., Zhang J., Zhang X., Zhao J., Zieba K., *End to End Learning for Self-Driving Cars*.
22. Fahmizal, Nugraha B., Su S., *Towards self-driving car using convolutional neural network and road lane detector*, 2017 2nd International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology, Jakarta 2017.
23. Kamioka E., Ku-Mahamad K., Salman Y., *Distance measurement for self-driving cars using stereo camera in Zulikba*, Proceedings of the 6th International Conference of Computing & Informatics, 2017.
24. Guerra V., Luna-Rivera M., Perez-Jimenez R., Rabadan J., Rodríguez R., Rufo J., *Hybrid Visible Light and Ultrasound-Based Sensor for Distance Estimation*, SENSORS, Szwajcaria.
25. Barsi Á., Lovas T., Potó V., Árpád Somogyi J., *Laser scanned point clouds to support autonomous vehicles*, 20th EURO Working Group on Transportation Meeting, EWGT 2017.
26. Ifte Khairul Alam Bhuiyan, *LiDAR Sensor for Autonomous Vehicle*, 2017.
27. Dollár P., Gkioxari G., Girshick R., He K., *Mask R-CNN*, 2017.
28. Xin Yu, Prevedouros P.D., Sulijoadikusumo G., *Evaluation of Autoscope, SmartSensor HD and Infra-Red Traffic Logger for Vehicle Classification*, Transportation Research Record: Journal of the Transportation Research Board, No. 2160.