

Marcin Łykowski

numer albumu: 47168

kierunek studiów: Informatyka

specjalność: Inteligencja obliczeniowa

forma studiów: studia niestacjonarne

**ROZPOZNAWANIE TABLIC REJESTRACYJNYCH POJAZDÓW NA
OBRAZACH Z KAMERY SAMOCHODOWEJ**

**RECOGNITION OF VEHICLE LICENSE PLATES ON IMAGES FROM A
CAR CAMERA**

praca dyplomowa magisterska

napisana pod kierunkiem:

dr hab. inż. Przemysław Kłęska

Katedra Metod Sztucznej Inteligencji i Matematyki Stosowanej

Data wydania tematu pracy: 30.03.2022

Data dopuszczenia pracy do egzaminu:
(uzupełnia pisemnie Dziekanat)

Szczecin, 2022

Oświadczenie autora pracy dyplomowej

Oświadczam, że praca dyplomowa magisterska pn. *Rozpoznawanie tablic rejestracyjnych pojazdów na obrazach z kamery samochodowej* napisana pod kierunkiem dr hab. inż. Przemysława Klęska jest w całości moim samodzielnym autorskim opracowaniem sporządzonym przy wykorzystaniu wykazanej w pracy literatury przedmiotu i materiałów źródłowych. Złożona w dziekanacie Wydziału Informatyki treść mojej pracy dyplomowej w formie elektronicznej jest zgodna z treścią w formie pisemnej.

Oświadczam ponadto, że złożona w dziekanacie praca dyplomowa ani jej fragmenty nie były wcześniej przedmiotem procedur procesu dyplomowania związanych z uzyskaniem tytułu zawodowego w uczelniach wyższych.

Podpis autora:

Szczecin, dnia:

Streszczenie

W tym miejscu trzeba napisać streszczenie pracy w języku polskim. Zawiera krótką charakterystykę dziedziny, przedmiotu i wyników zaprezentowanych w pracy. Maksymalnie 1/2 strony.

słowa kluczowe: np. informatyka, sterowanie, grafika komputerowa

Abstract

The abstract's purpose, which should not exceed 150 words, is to provide sufficient information to allow potential readers to decide on the thesis's relevance—a maximum of half the page.

keywords: e.g.: computer science, control, computer graphics

Spis treści

	Wstęp	7
1	Wprowadzenie teoretyczne	9
1.1	Przegląd istniejących metod detekcji tablic rejestracyjnych	11
1.1.1	Metody oparte na krawędziach (ang. <i>Edge based</i>)	11
1.1.2	Metody oparte na kolorach (ang. <i>Color based</i>)	13
1.1.3	Metody oparte na teksturach (ang. <i>Texture based</i>)	14
1.1.4	Klasyfikatory	14
2	Przedstawienie zestawu zdefiniowanych otoczeń możliwych do wykorzystania w pracy	19
2.1	Listy	19
2.2	Tabele i rysunki w pracy	19
2.3	Twierdzenia i definicje	20
2.4	Przykłady	21
2.5	Kod źródłowy	21
	Podsumowanie	23
	Spis literatury	25
	Książki	25
	Artykuły	25
	Źródła internetowe i inne	26
A	Dodatek	27

Wstęp

W dzisiejszych czasach ludzka praca stanowi jeden z największych składników kosztów dla wielu przedsiębiorstw. Taki stan rzeczy prowadzi do poszukiwania rozwiązań mających na celu zautomatyzowanie najbardziej powtarzalnych czynności. Potwierdza to wzrost zainteresowania na przestrzeni ostatnich lat zagadnieniami takimi jak uczenie maszynowe czy widzenie komputerowe. Jedną z gałęzi gospodarki, w której tego rodzaju automatyzacja jest zauważalna nawet dla osób niezwiązanych z branżą, jest transport drogowy. Nieustannie zwiększająca się liczba aut poruszających się po drogach, wzrost sieci dróg i autostrad niejako samoistnie wymusiła próby zautomatyzowania pewnych czynności.

Jednym z najczęściej poruszanych zagadnień jest problem rozpoznawania tablic rejestracyjnych (ang. *Licence Plate Recognition* - LPR). Do zadań takich systemów należy wykrycie na obrazie obszarów, w których znajdują się tablice rejestracyjne. Dokładność uzależniona jest od wielu czynników, takich jak jakość obrazu, prędkość pojazdu, warunki atmosferyczne lub pora dnia.

Celem niniejszej pracy jest przedstawienie tematyki rozpoznawania tablic rejestracyjnych. Wybór takiego zagadnienia w niniejszej pracy dyplomowej, motywuje własnymi zainteresowaniami zarówno na polu motoryzacyjnym jak i również uczenia maszynowego oraz widzenia komputerowego. W skład zakresu pracy wchodzi:

- Omówienie wybranych algorytmów z zakresu przetwarzania obrazów i uczenia maszynowego, potrzebnych do realizacji postawionego zadania.
- Przygotowania odpowiedniego materiału (sekwencje wideo) na potrzeby uczenia maszynowego i testowania.
- Przedstawienie ostatecznego schematu algorytmicznego dla całego procesu.
- Przeprowadzenie eksperymentów, pomiary dokładności i czasów wykonania, wnioski końcowe.

W pierwszej części pracy przedstawione zostaną najczęściej wykorzystywane techniki uczenia maszynowego i widzenia komputerowego do osiągnięcia wysokiej jakości systemu automatycznego rozpoznawania tablic rejestracyjnych. W drugiej części pracy zostanie przedstawiony stworzony program komputerowy do realizacji zadania rozpoznawania tablic rejestracyjnych. Część ta zawiera szczegółowy opis bibliotek wykorzystanych do realizacji przedstawionego zadania oraz implementacji opracowanego algorytmu. Przedstawiono również etapy pozyskania zbioru uczącego dla opracowanego mechanizmu. Po opisaniu opracowanego procesu, zostaną zaprezentowane wyniki dla wyuczonego klasyfikatora.

W przedstawionej pracy udało się zrealizować przedstawione zadanie. Do realizacji programu wykorzystano język programowania Python w wersji 3.8. Klasyfikator oparty

został o cechy Haara (ang. *Haar-like features*). Praca ma charakter eksperymentalny. Z tego powodu oraz z racji ograniczonych zasobów, algorytm ma swoje niedoskonałości. W podsumowaniu pracy zaprezentowano możliwości dalszego rozwoju klasyfikatora.

1. Wprowadzenie teoretyczne

Na przestrzeni ostatnich lat stosowanie Systemów Automatycznego Rozpoznawania Tablic Rejestracyjnych (ARTR)(ang. *Automatic Licence Plate Recognition - ALPR*) stało się znacznie bardziej powszechne. W większości dużych miast istnieją parkingi, gdzie po umieszczeniu opłaty za postój, przy zbliżeniu się do wyjazdu, szlaban otwiera się automatycznie po rozpoznaniu numeru rejestracyjnego pojazdu, w którym się poruszamy. W obecnych czasach wszystkie nowoczesne systemy do zarządzania i sterowania ruchem drogowym oparte są o technologie ARTR. Instytucje takie jak służby drogowe, dzięki rejestrowanym i przetwarzanym w czasie rzeczywistym ogromnym ilościom danych, są w stanie odpowiednio szybko reagować na wydarzenia na drogach takie jak kolizje, korki lub innego rodzaju utrudnienia. Innym z możliwych przykładów zastosowania wspomnianych systemów są odcinkowe pomiary prędkości, opłaty za przejazd płatnymi drogami lub wykrywanie kierowców łamiących przepisy. Dzięki nieustannemu rozwojowi technologii i co raz wydajniejszym komputerom, systemy stają się tańszą i łatwiej dostępną alternatywą dla systemów opartych na RFID (ang. *Radio-frequency identification*), które to wymagają specjalnej etykiety do prawidłowego działania.

Rozpoznawanie tablic rejestracyjnych jest techniką polegającą na wykryciu i odczytaniu znaków z tablicy rejestracyjnych na podstawie zarejestrowanego obrazu. Do tego celu wykorzystywany jest aparat o wysokiej rozdzielczości oraz odpowiedni program komputerowy. Oprogramowanie otrzymuje na wejściu cyfrową reprezentację obrazu. Dla zdjęć kolorowych każdy piksel opisany jest wartościami z palety barw RGB reprezentującymi jego barwę oraz współrzędnymi umiejscowienia w obrazie. Dla zdjęć monochromatycznych barwy opisywane są najczęściej za pomocą wartości luminacji obrazu.

W procesie automatycznego rozpoznawania tablic rejestracyjnych pozyskany obraz jest odpowiednio przetwarzany. Przed przejściem do rozpoznawania, obraz często jest konwertowany do skali szarości i filtrowany za pomocą filtrów np. (Gaussa lub średnio-przepustowego) w celu redukcji szumu. W procesie tym można wyróżnić trzy etapy [1]:

- **detekcję** - określenie położenia tablicy rejestracyjnej w analizowanym obrazie
- **segmentację** - wyodrębnienie pojedynczych znaków na fragmencie obrazu ze zlokalizowaną tablicą
- **identyfikację** - rozpoznanie każdego ze znaków i przedstawienie ich w formie tekstowej, którą można później wykorzystać do dalszych działań w zależności od przeznaczenia systemu

Na rysunku 1.1 przedstawiono graficzną reprezentację powyższego procesu.



Rysunek 1.1: Etapy procesu automatycznego rozpoznawania tablic rejestracyjnych (źródło: opracowanie własne)

Kolejne etapy korzystają z wyników uzyskanych w poprzednich krokach, co oznacza, że błąd powstały we wcześniejszej fazie, będzie rzutował na jakość działania całego systemu. W wielu systemach zanim dojdzie do rozpoznawania tablicy rejestracyjnej, obraz jest w pierwszej kolejności odpowiednio przetwarzany. Jednymi z najpowszechniej stosowanych czynności są skalowanie obrazu, modyfikacje jasności oraz redukcja zakłóceń. W zależności od wymagań stawianych przed danym mechanizmem i środowiskiem jego działania, czynności te mogą znacznie się od siebie różnić. Najbardziej podstawowe systemy wymagają, aby pojazd znajdował się nieruchomo w określonym miejscu. Tego typu rozwiązania najczęściej stosowane są na parkingach, gdzie szlaban otwiera się po odczycie numerów rejestracyjnych pojazdu i potwierdzeniu opłaty za postój w zewnętrznej bazie danych. Takie systemy pracują z reguły w środowisku o niskim poziomie zakłóceń wynikających z warunków atmosferycznych i oświetlenia. Obecnie na rynku znajduje się wiele komercyjnych rozwiązań, które oferują wysoką dokładność (powyżej 95%) dla tego rodzaju detekcji. Taki rodzaj systemów ARTR nazywany systemami statycznymi. Dużo większą

złożonością charakteryzują się systemy dynamiczne, w których znacznie większą rolę odgrywają zakłócenia wynikające ze zmiennych warunków oświetlenia. W obecnych czasach stworzenie dynamicznego systemu ARTR o wysokiej dokładności wciąż stanowi wyzwanie i jest tematem wielu prac naukowych. Celem niniejszej pracy jest analizowanie obrazów pochodzących z kamery samochodowej, co zdecydowanie sprawia, że jest to system dynamiczny. Poniżej przedstawiono najczęściej stosowane metody widzenia komputerowego w systemach ARTR.

1.1 Przegląd istniejących metod detekcji tablic rejestracyjnych

Zgodnie ze słownikiem języka polskiego, definicja tablicy rejestracyjnej brzmi następująco:

Definicja 1.1.1 — Tablica rejestracyjna. Płytką zawierającą numery identyfikacyjne pojazdu, umieszczana z przodu i z tyłu pojazdu.

Dla programu komputerowego powyższe zdanie nie jest zrozumiałe. W zadaniu detekcji tablicy rejestracyjnej, wymagane jest, aby maszyna “zrozumiała” jakich obiektów należy szukać. W kontekście detekcji, za definicję można uznać “prostokątny obszar, z dużym zagęszczeniem horyzontalnych i wertykalnych krawędzi”[13]. W oparciu o powyższe cechy zaprezentowano wiele algorytmów do rozwiązania zadania wykrywania tablic rejestracyjnych. Część z nich wywodzi się z tradycyjnych metod widzenia komputerowego i metod głębokiego uczenia. Każda z metod ma swoje zalety, ale również często ograniczenia. W związku z tym, trudno jednoznacznie stwierdzić, która z metod jest najbardziej efektywna.

Detekcja numerów rejestracyjnych jest wyzywającym zadaniem ze względu na poniższe czynniki:

- tablica rejestracyjna zajmuje niewielki obszar na zdjęciu
- istnienie ogromnej ilości formatów tablic rejestracyjnych (w zależności od kraju rejestracji lub rodzaju pojazdu)
- słabe oświetlenie, rozmazany obraz, refleksy świetlne
- ruch pojazdu, zabrudzone tablice

Tradycyjne metody widzenia komputerowego oparte są na cechach takich jak kształt, kolor, symetria, tekstury itp.[9]. W celu uzyskania lepszych wyników, spotyka się rozwiązania, w których łączy się wiele technik. Poniżej wyróżniono najczęściej stosowane metody w detekcji tablic rejestracyjnych.

1.1.1 Metody oparte na krawędziach (ang. *Edge based*)

Biorąc pod uwagę, że tablica rejestracyjna jest prostokątem o znanych proporcjach, większość badań bazuje na podejściu opartym o wykrywanie krawędzi. W większości przypadków kolor tablicy rejestracyjnej jest różny od koloru pojazdu. Dzięki temu, granice tablicy zostają uznane za krawędzie. Wiele metod wykorzystuje filtr Sobela. Jego działania polega dyskretnym różniczkowaniu i aproksymacji pochodnych kierunkowych intensywności obrazu. Filtr ten składa się z dwóch macierzy o wymiarach 3x3 (1.2) służących do detekcji krawędzi horyzontalnych i wertykalnych.

-1	0	+1
-2	0	+2
-1	0	+1

G_x

+1	+2	+1
0	0	0
-1	-2	-1

G_y

Rysunek 1.2: Macierze detektora krawędzi Sobela (źródło: opracowanie własne)

Zaletą takiego podejścia jest niewątpliwa łatwość użycia, natomiast jedną z głównych wad jest jego wrażliwość na szum.

Często wykorzystywaną metodą do wykrywania krawędzi obiektów w obrazach jest *Binary Image Processing*. Technika ta polega na sprowadzenia obrazu do postaci, w której kolory pikseli przyjmują tylko dwie wartości - czarną lub białą. Osiąga się to za pomocą ustalenia progu, który determinuje kolor piksela. Proóg wyznaczany jest na podstawie histogramu obrazu w odcieniach szarości. Metoda ta jest użyteczna, ze względu na fakt łatwego odseparowania obiektu od tła. Wykorzystuje ona założenie, że krawędzie tablicy są proste i poziome. Przy zdeformowanych lub zabrudzonych tablicach, algorytm ten nie osiąga zadowalających wyników.

Inną stosowaną metodą do wykrywania linii na obrazach binarnych jest transformata Hougha [3]. Motywacją do jej opracowania była metoda siłowa (ang. *Brute force*), która jest jednak znacznie bardziej zasobożerna. Złożoność algorytmu siłowego wynosi $O(n^3)$. Transformata Hougha polega na twierdzeniu, że każda prosta może być jednoznacznie przedstawiona za pomocą dwóch parametrów. Przestrzeń tych parametrów to właśnie przestrzeń Hougha. Najczęściej używanymi parametrami są współczynniki ρ i α z równania prostej w postaci normalnej (1.1).

$$x \cos \alpha + y \sin \alpha = \rho \quad (1.1)$$

W powyższym równaniu ρ jest promieniem wodzącym, natomiast α kątem tworzonym przez ρ z osią X. W związku z powyższym, jest to algorytm o liniowej złożoności obliczeniowej. Można wykazać następujące własności transformacji Hougha:

Twierdzenie 1.1.1 Prostej przestrzeni kartezjańskiej odpowiada w przestrzeni Hougha punkt, natomiast punktowi przestrzeni kartezjańskiej odpowiada w przestrzeni Hougha sinusoidalna krzywa. Punkty leżące na tej samej prostej korespondują z sinusoidami przechodzącymi przez wspólny punkt w przestrzeni Hougha [10].

Zasadę transformacji ilustruje rysunek 1.3.



Rysunek 1.3: Transformata Hougha (źródło: [7])

Innym spotykanym podejściem [5] jest stosowanie dwóch algorytmów. Pierwszy z nich ma za zadanie wyodrębnić odcinki linii i pogrupować je na podstawie wcześniej ustalonego zbioru warunków geometrycznych. Drugi znajduje obszary o najwyższym zagęszczeniu pionowych krawędzi. Dzięki takiemu spojrzeniu na przedstawiony problem, uzyskane wyniki mają wysoką dokładności, szczególnie dla pojazdów znajdujących się w ruchu. Metody oparte na krawędziach są stosowane w wielu rozwiązaniach ze względu na ich szybkość działania i prostotę. Jednakże, rozwiązania te są silnie wrażliwe na niepożądane krawędzie i nie sprawdzają się w rozmytych i złożonych obrazach.

1.1.2 Metody oparte na kolorach (ang. *Color based*)

Metody oparte na kolorach bazują na fakcie, że kolor tablicy jest różny od koloru tła pojazdu. Dla tej grupy rozwiązań, zamiast modelu barw RGB, stosuje się model HSL oparty o nasycenie koloru. Model ten jest jednak wrażliwy na szum.

Często metody wykorzystujące kolor tablicy rejestracyjnej są używane do wyselekcjonowania kandydatów. Innymi słowy, oznacza to wybranie obszarów obrazu, w których może znajdować się tablica rejestracyjna. Technika ta łączona jest z innymi algorytmami, które na kolejnych etapach decydują, czy wskazany obszar rzeczywiście zawiera poszukiwany obiekt. Do tego typu metod wykorzystywany jest m. in. algorytm *Mean shift* [6] i logika rozmyta [12].

Opisywana grupa metod może zostać użyta do detekcji zdeformowanych i pochylonych tablic. Rzadko występują one osobno w metodach detekcji, głównie ze względu na ich dużą czułość na zmiany naświetlenia. Dodatkowo w zależności od kraju oraz przeznaczenia pojazdu, kolory tablic mogą się znacznie różnić. Przykładowo obecnie w Polsce tablice aut elektrycznych mają kolor zielony, a samochody zabytkowe żółty (1.4).



Rysunek 1.4: Tablice rejestracyjne w Polsce dla aut elektrycznych i zabytkowych (źródło: opracowanie własne)

1.1.3 Metody oparte na teksturach (ang. *Texture based*)

Metody oparte na teksturach wykorzystują fakt znajdowania się znaków na tablicach rejestracyjnych. Znaki na tablicy mają z reguły czarny kolor i znajdują się na jasnym tle tworząc duży kontrast. Powyższa grupa algorytmów wykorzystuje wysoką częstość zmiany kolorów w obszarze występowania tablic rejestracyjnych. W [13] autorzy zaproponowali metodę lokalizacji tablic wykorzystując algorytm kwantyzacji wektorów (ang. *Vector Quantization* - VQ). W przeciwieństwie do innych metod, które wykorzystywały krawędzi lub kontrast, metoda VQ wykorzystuje aktualną zawartość tablicy rejestracyjnej. Autorzy wykorzystali bardzo wysoką skuteczność na poziomie 98%.

W analizie tekstur, często stosuje się filtr Gabora. Jest to filtr liniowy, pozwalający na przefiltrowanie obrazu z precyzyjnie dobranym zakresem częstotliwości. Reprezentacje częstotliwości i orientacji filtrów Gabora są uważane przez wielu współczesnych naukowców zajmujących się widzeniem komputerowym za podobne do tych z ludzkiego układu wzrokowego [8]. W [2] zaprezentowano algorytm wykorzystujący filtr Gabora. Jest to jednak metoda czasochłonna i nie znajduje zastosowania dla systemów, w których szybkość działania jest jednym z najistotniejszych czynników.

Wszystkie metody oparte na teksturach są odporne na deformacje tablic. Jest to kluczowa zaleta ich stosowania. Mimo to, metody te wymagają skomplikowanych obliczeń i nie dają zadowalających efektów w złożonych środowiskach z różnymi warunkami oświetlenia.

1.1.4 Klasyfikatory

Wiele badań wykorzystuje cechy Haara razem ze wzmocnieniem adaptacyjnym (ang. *Adaptive Boosting* - AdaBoost) do wyuczenia kaskady klasyfikatorów. Podejście takie zostało zaproponowane po raz pierwszy w [11]. Algorytm Viola-Jones został zaprezentowany w 2001. Pomimo upływu ponad 20 lat, dalej jest on powszechnie stosowany co świadczy o jego ponadczasowości i uniwersalności. Autorzy zaprojektowali go jako detektor twarzy, jednak jego funkcjonalność pozwala na wykrywanie dowolnych obiektów, np. tablic rejestracyjnych, przy odpowiednim wyuczeniu. Aby przedstawić jak działa algorytm Viola-Jones, należy najpierw zrozumieć czym są cechy Haara. Cechy Haara często są przedstawiane jako skalowalne, prostokątne szablony (rysunek 1.5), używane do porównania zależności pomiędzy pikselami, w szczególności jak ciemne są one względem siebie. Wartość cechy obliczana jest jako różnica pomiędzy średnią jasnością pikseli w zbiorze “białym” i średnią jasnością pikseli w zbiorze “czarnym”.

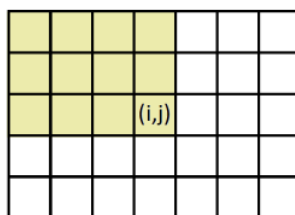


Rysunek 1.5: Cechy Haara używane w algorytmie Viola-Jones (źródło: https://docs.opencv.org/4.x/d2/d99/tutorial_js_face_detection.html)

Podejście takie powoduje jednak, że dla detektora o rozdzielczości 24x24px, algorytm

wymaga obliczenia ponad 180000 cech. Z tego powodu autorzy wprowadzili nowe pojęcie obrazu całkowego (ang. *integral image*).

Twierdzenie 1.1.2 Obrazem całkowym nazywamy tablicę dwuwymiarową o rozmiarze obrazu źródłowego, w której każdy element w i -tym wierszu i j -tej kolumnie przechowuje sumę pikseli z tej części obrazu, której prawym dolnym wierzchołkiem jest piksel (i, j) (rysunek 1.6).



Rysunek 1.6: Obraz całkowy dla punktu (i, j) (źródło: <https://www.spoj.com/WIPING5/problems/WIPING50.pdf>)

Matematycznie obraz całkowy $ii(x, y)$ przedstawiamy jako:

$$ii(x, y) = \sum_{1 \leq j \leq x} \sum_{1 \leq k \leq y} i(j, k) \quad (1.2)$$

Natomiast obliczanie obrazu całkowego z definicji 1.2 dla każdego punktu w obrazie jest czasochłonne i charakteryzuje się złożonością obliczeniową $O(n_x^2 n_y^2)$. Aby osiągnąć złożoność niezależną od rozmiaru okna, sumę oblicza się odejmując od sumy wartości obrazu całkowego dla prawego dolnego i lewego górnego wierzchołka sumę wartości obrazu dla pozostałych wierzchołków analizowanego prostokąta. Poniżej przedstawiono obliczenie sumy dla prostokąta rozpiętego między punktami (x_1, y_1) , a (x_2, y_2) .

$$\sum_{x_1 \leq x \leq x_2} \sum_{y_1 \leq y \leq y_2} i(x, y) = ii(x_2, y_2) - ii(x_1 - 1, y_2) - ii(x_2, y_1 - 1) + ii(x_1 - 1, y_1 - 1) \quad (1.3)$$

Na podstawie powyższego wzoru można zauważyć, że wystarczą operacje tylko na 4 punktach obrazu całkowego. Ten sposób charakteryzuje się liniową złożonością obliczeniową $O(1)$. Dla obliczenia różnicy pomiędzy sumami dwóch dowolnych prostokątów wymagane jest pobranie wartości dla 8 punktów z obrazu całkowego.

Algorytm Viola-Jones wymaga dużej ilości próbek pozytywnych i negatywnych. Do wyuczenia klasyfikatora i wybrania zbioru cech zastosowano algorytm AdaBoost. Autorzy zaproponowali użycie kaskady klasyfikatorów. Takie podejście bazuje na obserwacji, że okna pozytywne stanowią średnio 0.01% wszystkich okien. Rozpatrywane okno w pierwszej kolejności badane jest przez słabsze klasyfikatory, które bazują na mniejszej liczbie cech. Dzięki takiemu podejściu algorytm stał się bardziej wydajny czasowo. Każdy kolejny klasyfikator w kaskadzie oblicza większą liczbę cech. Jeżeli na którymś etapie klasyfikator zwróci odpowiedź negatywną, proces jest przerywany. Do osiągnięcia pozytywnego wyniku, wymagane jest zwrócenie przez wszystkie klasyfikatory odpowiedzi pozytywnej. W algorytmie Viola-Jones kaskada składa się z 32 klasyfikatorów, które

badają od 2 do 200 cech. Łącznie liczone jest 4297 cech, co daje średnio 8 cech na klasyfikator. W trakcie uczenia każdy z klasyfikatorów ma przypisany indywidualny próg decyzyjny. Każdy etap kaskady uczony jest w ramach boostingu (np. poprzez AdaBoost lub RealBoost).

AdaBoost (Adaptive Boosting) jest popularną techniką wzmacniania, której zadaniem jest stworzenie mocnego klasyfikatora na podstawie wielu słabych klasyfikatorów. Słabym klasyfikatorem nazywamy klasyfikator, który osiąga niską dokładność, jednak wyższą od losowych wyników. Za przykład może posłużyć rozpoznawanie płci na podstawie wzrostu. Słaby klasyfikator mógłby bazować na założeniu, że każda osoba o wzroście 175cm lub wyższym jest mężczyzną. Pozostała grupa osób jest kobietami. Wiele osób ze zbioru testowego zostanie określonych błędnie, jednak dokładność klasyfikatora będzie wyższa niż 50%.

AdaBoost jest techniką, która może zostać łączona z dowolnym algorytmem klasyfikującym, jednak nie może zostać wykorzystany jako samodzielny klasyfikator. Po raz pierwszy został on przedstawiony w [4]. AdaBoost występuje w połączeniu z popularnymi wariantami słabych klasyfikatorów:

- AdaBoost + decision stump
- AdaBoost + drzewka decyzyjne
- AdaBoost + klasyfikator liniowy (np. SVM)
- AdaBoost + naiwny Bayes

Głównymi zastosowaniami tej techniki jest wybór najlepszych cech z punktu widzenia klasyfikacji oraz dobór wag dla klasyfikatorów wchodzących w skład kaskady.

Każdy słaby klasyfikator powinien zostać wyuczony na losowym podzbiore zbioru uczącego. Podzbiory mogą się nakładać, natomiast nie powinno się dzielić zbiorów na równe części. AdaBoost przypisuje wagi do jednostek treningowych, które określają prawdopodobieństwo pojawienia się w zbiorze uczącym zgodnie ze wzorem 1.4.

$$D_{t+1}(i) = \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t} \quad (1.4)$$

D_t oznaczono wektor wszystkich wag, natomiast Z_t reprezentuje ich sumę. Indeks i jest numerem kolejnej próbki w zbiorze uczącym. Na początku wszystkie próbki posiadają tą samą wartość wag. Po zakończeniu uczenia, wagi błędnie sklasyfikowanych próbek zostają zwiększone. Dzięki temu, w następnej iteracji uczenia kolejny klasyfikator będzie mógł lepiej rozpoznawać niepoprawnie oznaczone jednostki treningowe w poprzednim kroku. W momencie, gdy każdy klasyfikator zostanie wyuczony, AdaBoost przypisuje do nich wagi na podstawie dokładności każdego z nich. Ostateczny klasyfikator można opisać wzorem 1.5.

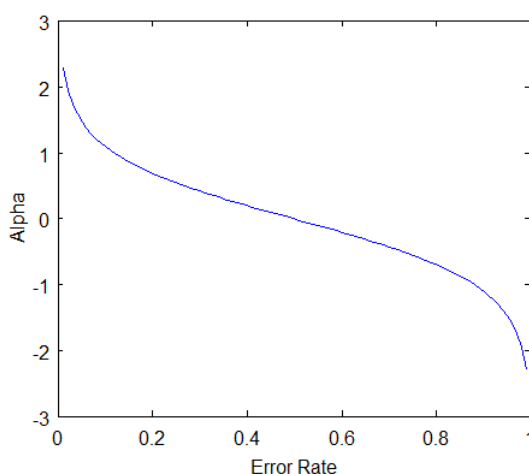
$$H(x) = \text{sign}\left(\sum_{t=1}^T \alpha_t h_t(x)\right) \quad (1.5)$$

Finalny klasyfikator zawiera T słabych klasyfikatorów. Waga klasyfikatora, wyliczona przez AdaBoost, została określona symbolem α_t . Ostateczny rezultat powyższego równania można opisać jako liniowa kombinacja słabych klasyfikatorów.

Pierwszy klasyfikator ($t = 1$) jest trenowany z równym prawdopodobieństwem dla wszystkich próbek. Po zakończeniu uczenia, waga obliczana jest na podstawie wzoru 1.6.

$$\alpha_t = \frac{1}{2} \ln \frac{1 - \epsilon_t}{\epsilon_t} \quad (1.6)$$

Współczynnik α_t jest skorelowany z poziomem błędu klasyfikatora ϵ_t , który to jest liczony jako liczba błędnie sklasyfikowanych próbek dzielona przez rozmiar zbioru uczącego. Rysunek 1.7 przedstawia powyższą zależność.



Rysunek 1.7: Zależność współczynnika α od błędności klasyfikatora (źródło: https://chrisjmccormick.files.wordpress.com/2013/12/adaboost_alphacurve.png)

Z rysunku 1.7 można odczytać, że waga rośnie wykładniczo dla klasyfikatorów o wysokiej dokładności. Dla klasyfikatora o poziomie błędności 0.5 waga jest równa zero. Oznacza to, że taki klasyfikator jest równie dokładny co losowe zgadywanie, dlatego jest on ignorowany. Dla klasyfikatorów o większym poziomie błędności, waga przyjmuje wartość ujemną. Jeśli taki klasyfikator uzna próbkę za negatywną, ostatecznie zostanie ona oznaczona jako pozytywna.

AdaBoost jest techniką uczenia progresywnego. Ważne jest, aby na wejściu próbki uczące były odpowiedniej jakości. Metoda ta jest bardzo czuła na zakłócenia. Jednymi z najczęstszych aplikacji tego typu boostingu jest klasyfikacja tekstu oraz zdjęć.

2. Przedstawienie zestawu zdefiniowanych otoczeń możliwych do wykorzystania w pracy

Szablon definiuje różne otoczenia, z których może korzystać autor pracy dyplomowej. Takowe zostały opisane w niniejszym rozdziale. Jeżeli otoczenie jest zawarte w szablonie, to właśnie takie należy stosować. Jeżeli w ramach pracy zachodzi potrzeba korzystania z elementów nie ujętych w szablonie można je stosować dowolnie.

Test.

2.1 Listy

Strukturę list wyliczeniowych poprawnie wykorzystywanych w języku polskim opisują słowniki. Oto fragment z poradnika sjp.pwn.pl. Podpunkty zdaniowe powinny się zaczynać wielką literą, a niezdaniowe małą. Jeśli podpunkt jest zdaniem, to zamykamy go kropką lub znakiem równoważnym, jeśli nie – przecinkiem lub średnikiem. Nagłówek listy powinien się kończyć dwukropkiem nawet wtedy, gdy następują po nim podpunkty zdaniowe, pisane wielką literą.

Listy numerowane wykorzystuje się, gdy kolejność elementów jest istotna:

1. element 1,
2. element 2,
3. element 34234.

Listy wypunktowane wykorzystuje się, gdy kolejność elementów jest istotna:

- osoba Alicja A,
- zwierzak kot Alicji,
- dom Alicji.

Jeżeli zamiast punktora lub wyliczenia istnieje potrzeba wykorzystania innego słowa jako wyliczenia, to możliwe jest stosowanie listy z nagłówkami (descriptions-definitions),

Nazwa Opis;

Słowo Definicja;

Komentarz Wywód.

2.2 Tabele i rysunki w pracy

Tabele podpisuje się z góry. Ważne, że każdy element osadzony w treści pracy musi zostać przywołany w tekście np. w Tabeli: 2.1 zaprezentowano wyniki eksperymentu

Tabela 2.1: Tabelka z wynikami eksperymentu 1

Lek	Odpowiedź 1	Odpowiedź 2
Lek 1	0.0003262	0.562
Lek 2	0.0015681	0.910
Lek 3	0.0009271	0.296

Rysunki podpisywane są z dołu. Należy tak jak w tabeli w treści pracy odnieść się do etykiety rysunku, np. Rysunek 2.1 przedstawia



Rysunek 2.1: To jest przykład osadzania rysunków (źródło:)

2.3 Twierdzenia i definicje

W pracach o rozbudowanej części teoretycznej może istnieć potrzeba prezentowania twierdzeń i definicji. Poniżej przykłady stosowania odpowiednich otoczeń.

A oto twierdzenie składające się z kilku równań.

Twierdzenie 2.3.1 — Tytuł twierdzenia. In $E = \mathbb{R}^n$ all norms are equivalent. It has the properties:

$$||\mathbf{x}|| - ||\mathbf{y}|| \leq ||\mathbf{x} - \mathbf{y}|| \quad (2.1)$$

$$||\sum_{i=1}^n \mathbf{x}_i|| \leq \sum_{i=1}^n ||\mathbf{x}_i|| \quad \text{where } n \text{ is a finite integer} \quad (2.2)$$

Jest to twierdzenie składające się z tylko jednej linii.

Twierdzenie 2.3.2 Zbiór $\mathcal{D}(G)$ ma gęstość $L^2(G)$, $|\cdot|_0$.

To jest przykład definicji. Definicja może być matematyczna lub może definiować koncepcję. Jeżeli w pracy wprowadza się nowe pojęcia najlepiej ująć je w formie definicji, co będzie spójnie akcentować taki fakt.

Definicja 2.3.1 — Tytuł definicji. Given a vector space E , a norm on E is an application,

denoted $|| \cdot ||$, E in $\mathbb{R}^+ = [0, +\infty[$ such that:

$$||\mathbf{x}|| = 0 \Rightarrow \mathbf{x} = \mathbf{0} \quad (2.3)$$

$$||\lambda \mathbf{x}|| = |\lambda| \cdot ||\mathbf{x}|| \quad (2.4)$$

$$||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}|| \quad (2.5)$$

2.4 Przykłady

W wielu pracach do przedstawionych idei prezentuje się przykłady. Na tą okoliczność przygotowano otoczenie *example*, które pomoże utrzymać spójny wygląd wszystkich przykładów.

Przykład może dotyczyć obliczeń lub prezentacji działania np. metody, algorytmu kodu, itp, itd.

Przykład 2.1 Let $G = \{x \in \mathbb{R}^2 : |x| < 3\}$ and denoted by: $x^0 = (1, 1)$; consider the function:

$$f(x) = \begin{cases} e^{|x|} & \text{si } |x - x^0| \leq 1/2 \\ 0 & \text{si } |x - x^0| > 1/2 \end{cases} \quad (2.6)$$

The function f has bounded support, we can take $A = \{x \in \mathbb{R}^2 : |x - x^0| \leq 1/2 + \epsilon\}$ for all $\epsilon \in]0; 5/2 - \sqrt{2}[$. ■

Przykład 2.2 — Przykład krzyżowania.

Pokolenie rodziców:

$x_1 = (\mathbf{00110}|\mathbf{011})$

$x_2 = (01101|101)$

krzyżowanie
→

Pokolenie potomków:

$(\mathbf{00110}|101)$

$(01101|\mathbf{011})$

■

2.5 Kod źródłowy

Jeżeli w pracy prezentowany jest kod źródłowy, to należy skorzystać z predefiniowanego i skonfigurowanego (w *structure.tex*) otoczenia *lstlisting*. Jeżeli autor chciałby umieścić w pracy pseudokod, to może posłużyć się wybranym przez siebie otoczeniem.

Zasada odwołania w tekście do algorytmu jest taka sama jak do tabeli i rysunku np. algorytm 2.1 został opracowany na potrzeby...

```
1 import numpy as np
2 # change this value for a different result
3 celsius = 37.5
4
5 # calculate fahrenheit
6 fahrenheit = (celsius * 1.8) + 32
```

```

7  print('%0.1f degree Celsius is equal to %0.1f degree
    Fahrenheit' %(celsius , fahrenheit))
8
9  def incmatrix(genl1 , genl2):
10     m = len(genl1)
11     n = len(genl2)
12     M = None # to become the incidence matrix
13     VT = np.zeros((n*m,1), int) # zmienna ąźćł
14
15     # compute the bitwise xor matrix
16     M1 = bitxormatrix(genl1)
17     M2 = np.triu(bitxormatrix(genl2),1)
18
19     for i in range(m-1):
20         for j in range(i+1, m):
21             [r,c] = np.where(M2 == M1[i,j])
22             for k in range(len(r)):
23                 VT[(i)*n + r[k]] = 1;
24                 VT[(i)*n + c[k]] = 1;
25                 VT[(j)*n + r[k]] = 1;
26                 VT[(j)*n + c[k]] = 1;
27
28             if M is None:
29                 M = np.copy(VT)
30             else:
31                 M = np.concatenate((M, VT), 1)
32
33     return M

```

Algorytm 2.1: Fragment algorytmu xxx

Podsumowanie

Podsumowanie pracy powinno na maksymalnie dwóch stronach przedstawić główne wyniki pracy dyplomowej. Struktura zakończenia to:

1. Przypomnienie celu i hipotez
2. Co w pracy wykonano by cel osiągnąć (analiza, projekt, oprogramowanie, badania eksperymentalne)
3. Omówienie głównych wyników pracy
4. Jak wyniki wzbogacają dziedzinę
5. Zamknięcie np. poprzez wskazanie dalszych kierunków badań.

Spis literatury

Książki

Artykuły

- [1] C.N.E. Anagnostopoulos i in. “A License Plate-Recognition Algorithm for Intelligent Transportation System Applications”. W: *IEEE Transactions on Intelligent Transportation Systems* 7.3 (2006), s. 377–392. DOI: 10.1109/TITS.2006.880641.
- [2] Hakan Caner, Hatice Geçim i Ali Alkar. “Efficient Embedded Neural-Network-Based License Plate Recognition System”. W: *Vehicular Technology, IEEE Transactions on* 57 (paź. 2008), s. 2675–2683. DOI: 10.1109/TVT.2008.915524.
- [3] Tran Duc Duan i in. “Building an Automatic Vehicle License-Plate Recognition System”. W:
- [4] Yoav Freund i Robert E. Schapire. “Experiments with a New Boosting Algorithm”. W: *ICML*. 1996.
- [5] Gisu Heo i in. “Extraction of Car License Plate Regions Using Line Grouping and Edge Density Methods”. W: *2007 International Symposium on Information Technology Convergence (ISITC 2007)*. 2007, s. 37–42. DOI: 10.1109/ISITC.2007.79.
- [6] Wenjing Jia i in. “Mean shift for accurate license plate localization”. W: *Proceedings. 2005 IEEE Intelligent Transportation Systems, 2005*. 2005, s. 566–571. DOI: 10.1109/ITSC.2005.1520110.
- [7] Xin Lin i Kazunori Otake. “Hough transform algorithm for real-time pattern recognition using an artificial retina camera”. W: *Opt. Express* 8.9 (kw. 2001), s. 503–508. DOI: 10.1364/OE.8.000503. URL: <http://opg.optica.org/oe/abstract.cfm?URI=oe-8-9-503>.
- [8] Bruno Olshausen i David Field. “Emergence of simple-cell receptive field properties by learning a sparse code for natural images”. W: *Nature* 381 (lip. 1996), s. 607–9. DOI: 10.1038/381607a0.
- [9] Jithmi Shashirangana i in. “Automated License Plate Recognition: A Survey on Methods and Techniques”. W: *IEEE Access* 9 (2021), s. 11203–11225. DOI: 10.1109/ACCESS.2020.3047929.

- [10] Fayed Tarsha-Kurdi. “HOUGH-TRANSFORM AND EXTENDED RANSAC ALGORITHMS FOR AUTOMATIC DETECTION OF 3D BUILDING ROOF PLANES FROM LIDAR DATA”. W: (wrz. 2007).
- [11] P. Viola i M. Jones. “Rapid object detection using a boosted cascade of simple features”. W: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. T. 1. 2001, s. I–I. DOI: 10.1109/CVPR.2001.990517.
- [12] Feng Wang i in. “Fuzzy-based algorithm for color recognition of license plates”. W: *Pattern Recognit. Lett.* 29 (2008), s. 1007–1020.
- [13] R. Zunino i S. Rovetta. “Vector quantization for license-plate location and image coding”. W: *IEEE Transactions on Industrial Electronics* 47.1 (2000), s. 159–167. DOI: 10.1109/41.824138.

Źródła internetowe i inne

A. Dodatek

W dodatkach mogą znaleźć się większe fragmenty kodów źródłowych, instrukcje działania programów, specyfikacje opcji, z którymi wywołuje się program, większe dane tabelaryczne, specyfikacje oprogramowania, dłuższe opisy teoretyczne, itp.

What is Lorem Ipsum? Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged. It was popularised in the 1960s with the release of Letraset sheets containing Lorem Ipsum passages, and more recently with desktop publishing software like Aldus PageMaker including versions of Lorem Ipsum.

Why do we use it? It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editors now use Lorem Ipsum as their default model text, and a search for 'lorem ipsum' will uncover many web sites still in their infancy. Various versions have evolved over the years, sometimes by accident, sometimes on purpose (injected humour and the like).

Where does it come from? Contrary to popular belief, Lorem Ipsum is not simply random text. It has roots in a piece of classical Latin literature from 45 BC, making it over 2000 years old. Richard McClintock, a Latin professor at Hampden-Sydney College in Virginia, looked up one of the more obscure Latin words, consectetur, from a Lorem Ipsum passage, and going through the cites of the word in classical literature, discovered the undoubtable source. Lorem Ipsum comes from sections 1.10.32 and 1.10.33 of "de Finibus Bonorum et Malorum" (The Extremes of Good and Evil) by Cicero, written in 45 BC. This book is a treatise on the theory of ethics, very popular during the Renaissance. The first line of Lorem Ipsum, Lorem ipsum dolor sit amet..", comes from a line in section 1.10.32.

The standard chunk of Lorem Ipsum used since the 1500s is reproduced below for those interested. Sections 1.10.32 an

What is Lorem Ipsum? Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic

typesetting, remaining essentially unchanged. It was popularised in the 1960s with the release of Letraset sheets containing Lorem Ipsum passages, and more recently with desktop publishing software like Aldus PageMaker including versions of Lorem Ipsum.

Why do we use it? It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editors now use Lorem Ipsum as their default model text, and a search for 'lorem ipsum' will uncover many web sites still in their infancy. Various versions have evolved over the years, sometimes by accident, sometimes on purpose (injected humour and the like).

Where does it come from? Contrary to popular belief, Lorem Ipsum is not simply random text. It has roots in a piece of classical Latin literature from 45 BC, making it over 2000 years old. Richard McClintock, a Latin professor at Hampden-Sydney College in Virginia, looked up one of the more obscure Latin words, *consectetur*, from a Lorem Ipsum passage, and going through the cites of the word in classical literature, discovered the undoubtable source. Lorem Ipsum comes from sections 1.10.32 and 1.10.33 of "*de Finibus Bonorum et Malorum*" (The Extremes of Good and Evil) by Cicero, written in 45 BC. This book is a treatise on the theory of ethics, very popular during the Renaissance. The first line of Lorem Ipsum, *Łorem ipsum dolor sit amet..*", comes from a line in section 1.10.32.

The standard chunk of Lorem Ipsum used since the 1500s is reproduced below for those interested. Sections 1.10.32 and 1.10.33 of "*de Finibus Bonorum et Malorum*" (The Extremes of Good and Evil) by Cicero, written in 45 BC. This book is a treatise on the theory of ethics, very popular during the Renaissance. The first line of Lorem Ipsum, *Łorem ipsum dolor sit amet..*", comes from a line in section 1.10.32. Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsum has been the industry's standard dummy text ever since the 1500s, when an unknown printer took a galley of type and scrambled it to make a type specimen book. It has survived not only five centuries, but also the leap into electronic typesetting, remaining essentially unchanged. It was popularised in the 1960s with the release of Letraset sheets containing Lorem Ipsum passages, and more recently with desktop publishing software like Aldus PageMaker including versions of Lorem Ipsum.

Why do we use it? It is a long established fact that a reader will be distracted by the readable content of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-or-less normal distribution of letters, as opposed to using 'Content here, content here', making it look like readable English. Many desktop publishing packages and web page editors now use Lorem Ipsum as their default model text, and a search for 'lorem ipsum' will uncover many web sites still in their infancy. Various versions have evolved over the years, sometimes by accident, sometimes on purpose (injected humour and the like).

Where does it come from? Contrary to popular belief, Lorem Ipsum is not simply random text. It has roots in a piece of classical Latin literature from 45 BC, making it over 2000 years old. Richard McClintock, a Latin professor at Hampden-Sydney College in Virginia, looked up one of the more obscure Latin words, *consectetur*, from a Lorem Ipsum passage, and going through the cites of the word in classical literature, discovered the undoubtable source. Lorem Ipsum comes from sections 1.10.32 and 1.10.33 of "*de Finibus Bonorum et Malorum*" (The Extremes of Good and Evil) by Cicero, written in 45

BC. This book is a treatise on the theory of ethics, very popular during the Renaissance. The first line of Lorem Ipsum, Lorem ipsum dolor sit amet..", comes from a line in section 1.10.32.

The standard chunk of Lorem Ipsum used since the 1500s is reproduced below for those interested. Sections 1.10.32 an