

Chciałbym podziękować moim rodzicom za to, że wspierali mnie w trakcie studiów. Dodawali mi otuchy w momentach w których byłem zrezygnowany, wspierali ciepłym słowem i że pamiętali o mnie w modlitwach. Chciałbym również podziękować Panu Doktorowi Tomaszowi Kajdanowiczowi, dzięki któremu ta praca mogła powstać.

Spis treści

Streszczenie pracy	4
Abstract	4
1 Wstęp	5
1.1 Określenie problemu i motywacja	6
1.2 Cel	6
1.3 Konspekt pracy	6
2 Przegląd Literatury	7
3 Wstęp teoretyczny	11
3.1 Ekstrakcja danych z Instagrama	11
3.1.1 Pozyskiwanie obrazów i metadanych z wykorzystaniem API Instagrama	11
3.2 Skalowiezmiennicze przekształcenie cech - SIFT	13
3.2.1 Przestrzeń skalowa	14
3.2.2 Lokalizacja punktów kluczowych	15
3.2.3 Kierunki punktów kluczowych	15
3.2.4 Deskryptor punktu kluczowego	16
3.2.5 Dopasowywanie punktów kluczowych	17
4 Opis metody rozwiązania problemu	18
4.1 Metoda grafów	18
4.1.1 Tworzenie grafów	18
4.1.2 Wybieranie hashtagów	21
4.1.3 Sprawdzanie ilości poprawnie wybranych hashtagów	21
4.2 Metoda oparta o algorytm SIFT	21
4.2.1 Przygotowanie danych	21
4.2.2 Wybieranie hashtagów	22
4.2.3 Sprawdzanie ilości poprawnie wybranych hashtagów	23
5 Konfiguracja eksperymentu	24
5.1 Pochodzenie danych eksperymentalnych	24
5.2 Miary ewaluacji	25
5.3 Zapytania do API i ograniczenia początkowe	27

6	Wyniki ekspertymentu	28
6.1	Metoda grafów	28
6.1.1	Wyniki dla metody grafów	28
6.1.2	Wnioski dla metody grafów	31
6.2	Metoda oparta o algorytm SIFT	32
6.2.1	Wyniki dla metody opartej o algorytm SIFT	32
6.2.2	Wnioski dla metody opartej o algorytm SIFT	36
6.3	Porównanie wyników	37
7	Podsumowanie	39

Streszczenie pracy

Internetowe serwisy takie jak Instagram czy Flickr umożliwiają użytkownikom dzieleniem się swoimi zdjęciami z rodziną, przyjaciółmi i innymi osobami w sieci. Ważnym elementem tych serwisów jest to, że autorzy zdjęć mogą samodzielnie dodać adnotację do swojego zdjęcia, tak zwany hashtag. W tej pracy zostały zaproponowane dwie metody klasyfikacji zdjęć lub innymi słowy zostały stworzone systemy rekomendowania hashtagów do grafik na Instagramie. Nie ma prac naukowych, które zajmowałyby się rekomendowaniem hashtagów na serwisie Instagram. Praca oparta jest na grafikach z informacjami, gdzie zdjęcie zostało wykonane. Dane były pobierane przy pomocy API Instagrama. Pierwsza metoda opiera się na budowaniu grafu współwystępujących ze sobą hashtagów. Zaproponowany sposób ma predyspozycje do skutecznego rekomendowania hashtagów ze względu na szybkość zwracanych wyników. Druga metoda korzysta z algorytmu skaloniezmienniczego przekształcania cech (SIFT). Rozwiązanie te zwraca lepsze wyniki niż wcześniej zaproponowany algorytm. Obliczenia w tej metodzie trwają bardzo długo przez co nie można mówić o efektywnym wykorzystywaniu tego rozwiązania w praktyce. Porównując metody zaproponowane w artykułach naukowych z tymi, które są opisane w tej pracy można stwierdzić, że rozwiązanie oparte o budowanie grafu nie zwraca tak dobrych wyników jak inne metody. Rozwiązanie oparte o algorytm SIFT zwraca podobne rezultaty jak w innych pracach naukowych. Dodatkowym aspektem tej pracy było przeanalizowanie jak zmieniają się wyniki (pochodzące z z) w zależności od odległości pomiędzy punktem, gdzie znajduje się atrakcja turystyczna, a miejscem skąd było wykonane zdjęcie i wgrane na Instagram.

Abstract

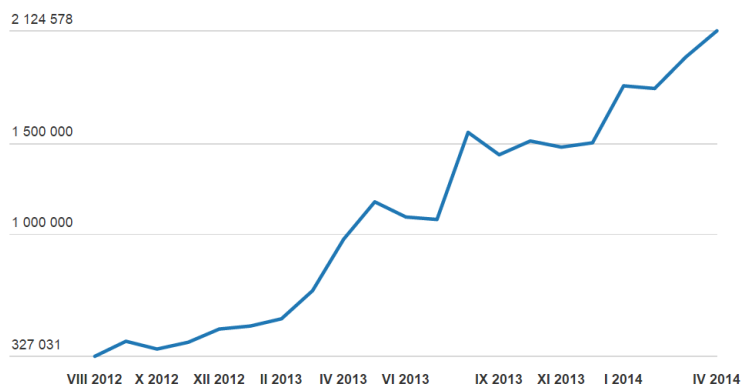
Online services such as Instagram or Flickr allow users to share their photos with family, friends and another people on Internet. An important element of these services is that authors can single-handedly add an annotation to their photos, which are called hashtag. In this work two methods to classify pictures have been presented or in another words there have been invented two systems which recommend hashtags to images on Instagram. There are no scientific papers which describe methods to recommend hashtags on Instagram. This work bases on graphic with information where the photo was taken. The first method is based on building a graph. The graph is created base on co-occures hashtags. The proposed method has a predisposition to effectively recommend hashtags because of the speed of results returned. The second method uses an algorithm scale-invariant feature transform (SIFT). This solution returns better results than previously proposed algorithm. The calculations in this method take a long time so this algorithm cannot be effectively used in practice. Comparing the methods proposed in the scientific articles with those that are described in this work it can be concluded that a solution based on building a graph does not return as effective results as other methods. The solution based on SIFT algorithm returns similar results as in other scientific papers. Another aspect of this work was to analyze how the results are changing (from the tag recommendation system) depending on the distance between the point where is a tourist attraction and the place where the photo was taken and uploaded to the Instagram.

Rozdział 1

Wstęp

Serwis społecznościowy to medium w którym ludzie dzielą się swoimi przeżyciami, zdjęciami, filmami z innymi osobami. Serwisy te zyskują bardzo szybko na popularności. Jeszcze w 2010 roku z takich stron korzystało 43% Polaków [1], a w roku 2013 już blisko 90% [2].

Praca ta jest poświęcona serwisowi społecznościowemu Instagram. Z Instagrama można korzystać dzięki aplikacji zainstalowanej na komórce. Na stronie google play, skąd można pobrać dedykowaną aplikację do tego serwisu, opisywany jest on w następujący sposób: *Instagram is a simple way to capture and share the world's moments*. Twórcy chwala się również, że ich aplikacji została pobrana ponad 300 milionów razy przez ludzi na całym świecie. Badania przeprowadzone przez wirtualnedia.pl wyraźnie pokazują (Obrazek 1.1), że popularność tej aplikacji w Polsce cały czas wzrasta.



Rysunek 1.1: Wykres popularności Instagrama. Źródło: [3]

Na Instagrama można wysyłać zdjęcia przy pomocy urządzeń mobilnych. Zdjęcia, które są przesyłane można ohashtagować, czyli nadać im hasztag (*ang. hashtag*). Hasztag jest to pojedyncze słowo lub wyrażenie bez spacji, które jest poprzedzone znakiem "#", na przykład: "#sky", "#love", "#PracaMagisterska". Hasztagi nadawane są przez autora obrazka jako pewnego rodzaju opis zdjęcia.

1.1 Określenie problemu i motywacja

Dziennie Instagram przyjmuje około 70 miliona zdjęć [4]. Nie wszystkie obrazki się ohashtagowują. Ta praca magisterska ma pomóc w rozwiązaniu problemu nieohashtagowanych zdjęć na Instagramie. W momencie kiedy każda grafika na internecie byłaby opisana to wyszukiwarki internetowe potrafiłyby zwracać więcej trafniejszych wyników.

1.2 Cel

Celem tej pracy jest zaproponowanie metod dzięki którym komputer będzie potrafił sklasyfikować, które hasztagi trafnie opisują zdjęcie, które zostało dodane do portalu Instagram z podaniem długości i szerokości geograficznej w której zdjęcie zostało wykonane. Obrazki, które będą brane pod uwagę wraz z ich hasztagami będą pochodzić wyłącznie z serwisu Instagram. Innymi słowy celem tej pracy jest zaproponowanie metod rekomendacji hasztagów do zdjęć wykorzystując informacje o położeniu geograficznym.

1.3 Konspekt pracy

W rozdziale 2 przedstawiony jest przegląd literatury dotyczący tematu hasztagowania zdjęć w portalach społecznościowych. W rozdziale 3 przedstawiony jest wstęp teoretyczny w którym opisana jest metoda wydobywania danych z Instagrama oraz wytłumaczenie zasady działania algorytmu SIFT na podstawie którego, działa jedna z zaproponowanych w tej pracy metod. W rozdziale 4 wytłumaczone są nowe zaproponowane metody klasyfikacji zdjęć. W rozdziale 5 podane są informacje skąd zebrane są dane do eksperymentu, wytłumaczone miary ewaluacji oraz opisane ograniczenia. W rozdziale 6 opisane są wyniki eksperymentu i porównane z innymi metodami. W rozdziale 7 znajduje się podsumowanie tej pracy oraz propozycje rozwijania zaproponowanych metod w przyszłości.

Rozdział 2

Przegląd Literatury

Pomimo starań nie udało się znaleźć prac naukowych, które dotyczyłyby tematu klasyfikacji zdjęć na Instagramie. Instagram w artykułach był wspominany w *Abstrakt* bądź *Introduction* jako jedno w wielu mediów społecznych. Bardzo podobnym serwisem do Instagrama, który był przytaczany przez większość prac naukowych dotyczył serwisu Flickr. Czasem zdarzały się prace dotyczące serwisu Delicious.

W artykułach naukowych dotyczących klasyfikacji zdjęć na serwisach społecznościowych widać, że początkowo były one nastawione na analizowanie zawartości obrazu pod względem wizualnym i tworzenia relacji pomiędzy podobnymi grafikami. Z upływem czasu zaczęto zmieniać podejście. Starano się zwrócić uwagę podczas rekomendacji hashtagów na relacje, które panują pomiędzy użytkownikami serwisów medialnych. Zaczęto analizować hashtagi, które użytkownik i jego znajomi używali, aby później trafnie dobrać wyrazy, które opisywałyby zdjęcie.

W artykule [5] autorzy postanowili porównać 16 różnych algorytmów do rekomendacji hashtagów. Głównie opierały się one na algorytmie kNN, Naiwnego Bayesa i algorytmach klasteryzacji. W celu ewaluacji jakości działania algorytmów stworzyli 5 różnych zbiorów danych. Autorzy stworzyli aplikacje dzięki której przeprowadzili badania z udziałem ludzi, którzy oceniali działanie algorytmów. 80% respondentów odpowiedziało, że wszystkie algorytmy nadają się do rekomendowania przykładowych hashtagów. Jedno z pytań brzmiało: "Czy w Twojej opinii, system jest zdolny do zarekomendowania Tobie 1-3 / 4-6 / 7-10 pasujących hashtagów?" Najczęściej ankietowani odpowiadali: 4-6 hashtagów.

W artykule [6] jest napisane, że na serwisie Flickr tylko około 50% hashtagów ma związek z tym co jest pokazane na obrazku. Innym problem z rekomendacją hashtagów, który wymienia autor jest to, że użytkownicy używają skrótów, bądź robią literówkę w napisanym wyrazie.

W artykule [7] autorzy zaproponowali metodę rekomendacji hashtagów dla serwisu Flickr bazując na hashtagach, których używał sam użytkownik, jego znajomi, osoby w grupach do których ta osoba należała i ogólnie występujących hashtagach na Flickerze.

W artykule [8] autorzy postanowili stworzyć metodę, która sugeruje hashtagi na podstawie geolokalizacji, czasu, cech wizualnych zrobionego zdjęcia oraz preferencji używania hashtagów przez użytkownika. W celu sprawdzenia tego algorytmu dane były pobrane z strony Flickr. Co najmniej jedno zdjęcie z hashtagiem i danymi geolokalizacjami wgrało na serwer 89,09% użytkowników. Zdjęcia, które zawierają hashtagi i geolokalizację stanowią 28,34% wszystkich dostępnych zdjęć na serwisie Flickr. Ilość trafnych rekomendacji była

zależna od ilości proponowanych hashtagów. Autorzy zauważyli, że średnia ilość hashtagów, którą użytkownicy przypisywali do zdjęć rozkłada się w następujący sposób: [1,5] hashtagów - 52,64% użytkowników, [6,10] hashtagów - 30,31% użytkowników, >10 hashtagów - 17,05% użytkowników.

W artykule [9] autorzy zauważają, że w ostatnich latach analiza danych multimedialnych to bardzo popularny temat, który jest poruszany w badaniach naukowych. Szczególnie dużo uwagi w badaniach poświęca się na wyszukiwanie grafiki w Internecie. Piszą, że najważniejszym czynnikiem w tego rodzaju badaniach jest stworzenie algorytmu, który potrafi dopasować hashtagi do wizualnej zawartości obrazów. Autorzy zwracają uwagę, że w ostatnich latach media społecznościowe takie jak Flickr lub Youtube zyskały potężną popularność. Użytkownicy tych serwisów mogą hashtagować media (czyli zdjęcia lub filmy) które tam wgrali. Jednak wyszukiwanie danych cyfrowych w tych serwisach na podstawie hashtagów daje mało satysfakcjonujące wyniki dla użytkowników. Przyczyny tego stanu rzeczy upatrują w tym, że użytkownicy nie piszą hashtagów, które w oczywisty sposób odnoszą się do zawartości (dane są zaszumione) oraz że nie ma wystarczająco dobrych algorytmów, które rozwiązywałyby problem skutecznego wyszukiwania informacji.

Autorzy proponują nową metodę na rozwiązanie tego problemu. Każdy obrazek jest reprezentowany dzięki zbiorowi słów oraz zbiorowi wizualnych słów (punktów wyznaczonych przez algorytm SIFT). Następnie tworzone są grupy bazując na wcześniej wymienionych cechach. Powstaje hipergraf, który reprezentuje relacje pomiędzy wszystkimi obrazkami. Każdy wierzchołek hipergrafu oznacza obrazek, a krawędź jest wizualnym albo teksturalnym słowem (hashtagiem). Krawędź łączy wiele wierzchołków. Waga każdej krawędzi zdefiniowana jest bazując na wizualnych podobieństwa obrazka. Złożoność obliczeniowa tej metody wynosi $O(n^2)$. Dane z których korzystali pochodzą z serwisu Flickr. Sprawdzenie jak dobrze działa metoda zostało przeprowadzone dla 52 hashtagów.

W artykule [10] autorzy piszą, że liczba obrazków, które można znaleźć w sieci jest wyzwaniem dla systemów komputerów, które je przechowują, przesyłają, indeksują i zarządzają. Artykuł ten proponuje nową metodę ekstrakcji cech charakterystycznych obrazka, w celu znalezienia najbardziej podobnych grafik. Zaznaczają, że sposób ten jest bardzo łatwy do zaimplementowania i że zaproponowana przez nich metoda daje akceptowalne rezultaty w szybki i prosty sposób.

Według autorów ważne jest, aby stworzyć efektywną technikę indeksowania obrazów w celu szybkiego ich później wyszukiwania. Są wyróżnione trzy metody, które mogą posłużyć jako zapytania w celu wyszukiwania grafik. Zapytanie przez obraz, przez zarys grafiki oraz przez słowa. Wyszukiwanie mediów przez słowa było kiedyś popularnym tematem naukowych. Aktualnie prym wiedzie wyszukiwanie, w którym to zapytaniem jest medium, na przykład film. W przeciwieństwie do tekstu, obrazy składają się z pikseli, które nie mają żadnego znaczenia dla komputera.

Zdjęcia są opisywane przez ludzi. Jednakże to implikuje pewne problemy. Po pierwsze, adnotacje zależą od osoby, które je zrobiła. Ludzka percepcja obrazów jest bardzo zróżnicowana i różne osoby mogą przypisać różne znaczenia do tego samego obrazka. Nawet jeżeli dwoje ludzi postrzega dany obiekt w ten sam sposób, to mogą użyć innych słów, aby go opisać. Drugą wadą jest to, że ręczne opisywanie zdjęć kosztuje bardzo dużo czasu. Ostatnim minusem jest to, że użytkownicy, którzy chcieliby wyszukiwać dane musieliby używać tych samych słów do opisanie grafiki.

Autorzy chcą użyć danych tekstowych i cech wizualnych obrazka. Metoda ta opiera się na histogramach kanałów RGB. Dla każdego kanału tworzone są osobne histogramy. Następnie histogramy są wyrównywane. W ten sposób powstają cechy charakterystyczne obrazków. Podobieństwo obrazków obliczane jest bazując na odległości euklidesowej. Testy zostały przeprowadzone na grupie 500 obrazków. Obrazy były podzielone na 5 kategorii (autobusy, dinozaury, słonie, konie, kwiaty), po 100 grafik na kategorię. Zdjęcia były różnej wielkości. Porównanie jak podobne są do siebie obrazki były obliczane na podstawie odległości Euklidesa.

W artykule [11] autor pisze, że ludzie aktualnie używają do robienia zdjęć urządzeń mobilnych, które zbierają dużą ilość metadanych. Dzielenie się zdjęciami, które zostały przed chwilą zrobione jest bardzo popularne. Ludzie bardzo chętnie robią zdjęcia, ponieważ chcą zapisać każde swoje doświadczenie. Najpopularniejszymi platformami, gdzie ludzie dzielą się zdjęciami są: Facebook - dziennie wysyłanych jest 350 milionów zdjęć, Instagram - 55 milionów, Snapchat i Flickr - 1,6 milionów (artykuł pochodzi z 2014 roku).

Urządzenia mobilne są intensywnie używane do wysyłania zdjęć do Internetu. Hasztagowanie zdjęć jest bardzo czasochłonne, a robienie tego na urządzeniu mobilnym jest mniej wygodniejsze niż na komputerze. Autor pisze, że powstały strony, które oferują możliwość automatycznego hasztagowania zdjęć na przykład autotag.me¹ lub iqengines.com/glow/².

Ludzie hasztagują swoje zdjęcia w dwóch celach: aby były bardziej rozpropagowane wśród innych użytkowników Internetu oraz aby łatwiej potrafili je potem samemu odnaleźć.

Autor przeprowadził badania na serwisie Flickr. Proponuje on metodę automatycznego hasztagowania zdjęć przy użyciu metadanych. Wśród 42,000 zdjęć, które były losowo wybrane to 90% z nich nie były ohasztagowane. Ohasztagowane zdjęcia zawierały głównie od 1 do 3 tagów. Tylko kilku użytkowników przypisywało większą ilość hasztagów. Autor wyróżnia trzy powody dlaczego, aż 90% zdjęć jest nie ohasztagowanych: ludzie mogą używać tych platform jako darmową przestrzeń do przechowywania zdjęć, gdzie mogą przetrzymać zdjęcia; użytkownicy nie wiedzą, że mogą zwiększyć popularność swojego zdjęcia dzięki hasztagowaniu; ludzie są zbyt leniwi, aby hasztagować zdjęcia.

Autor zaznacza, że nie można stwierdzić czy zdjęcie jest dobrze lub źle ohasztagowane. To zawsze zależy od opinii użytkownika.

Metoda do rekomendacji hasztagów, którą proponuje autor ma używać następujących metadanych: długość i szerokość geograficzna w którym zostało zrobione zdjęcie, data, czas, orientacja zdjęcia, czy została użyta lampa błyskowa przy robieniu zdjęcia oraz poziom dźwięku w trakcie robienia zdjęcia. Te dane wraz z zdjęciem są wysyłane na serwer. Serwer zwraca proponowane hasztagi. Kiedy użytkownik odrzuci jakiś hasztag, ta informacja jest przesyłana do serwera i klasyfikator jest aktualizowany. Naukowiec pisze, że na Instagramie występują *social tags*, na przykład: *l4l*, *f4f*, *love*, *make-up*. Autor podaje tylko przykłady *social tags*, bez ich wyjaśnienia. Na Flickrze hasztagi mają większe znaczenie niż na Instagramie. Autor stworzył bazę 142 najbardziej popularnych hasztagów na Flickrze. Oto jak rozkłada się w niej tematyka hasztagów:

¹Na dzień 2015-03-14 na tej stronie nie można się zarejestrować. Rejestracja jest konieczna w celu korzystania z usługi, która proponuje hasztagi do zdjęć. Po skontaktowaniu się z twórcami tej strony dostałem odpowiedź, że „strona jest dalej nieaktywna”.

²Na dzień 2015-05-26 strona nie istnieje.

- 29% nazwy miejsc (np. Nowy Jork, Europa)
- 14% żywe rzeczy (np. para ludzi, kot)
- 13% fizyczne obiekty (np. samochód, woda)
- 10% meta informacje (np. iphone, samsung)
- 8% przedziały czasowe (np. wieczór, lato)
- 8% wydarzenia społeczne (np. wakacje, ślub)
- 7% miejsca (np. plaża, ogród)
- 6% kolory (np. niebieski, czerwony)
- 5% emocje, przymiotniki (np. kocham, starość)

Artykuł ten dzieli dane na niski i wysoki poziom. Niskiego poziomu są to informacje wyciągnięte z metadanych, na przykład: pogoda, lokalizacja miejsca, przedział czasowy. Wysokiego poziomu to: główny kolor na obrazku, stwierdzenie czy na obrazku są ludzie.

Baza danych składa się z 300 zdjęć, które są opisane danymi geolokalizacyjnymi i posiadają datę. Każde zdjęcie jest opisane co najmniej trzema hasztagami.

W artykule [12] autorzy zauważyli, że w ostatnich latach strony takie jak: Flickr, YouTube, Slashdot zyskały dużą popularność. Strona Flickr służy entuzjastom fotograficznym do dzielenia się zdjęciami, które zrobili i otrzymywaniu od innych uwag na temat swojej pracy. Zdjęcia można dodać do odpowiedniej kategorii. Jednak istnieje wiele podobnych grup, na przykład: "*Travel / Travel Photography / Travel in Asia*". W tym artykule autorzy starali się przydzielić zdjęcia do odpowiednich grup.

Autorzy zauważają, że w innych pracach naukowcy więcej uwagi przykładają na automatyczne hasztagowanie zdjęć na podstawie cech wizualnych obrazka, mniej natomiast bazując na interakcjach społecznych, które są kluczowe dla mediów społecznościowych. Autorzy uważają, że ich artykuł jest pierwszym, który łączy cechy wizualne grafiki z komunikacją społecznościową na portalu w celu odpowiedniego sklasyfikowania zdjęć.

Autorzy proponują nowy sposób hasztagowania zdjęć w media społecznościowych. Zostały stworzone trzy grupy cech na podstawie których następowała klasyfikacja obrazków: cechy obrazka, hasztagów oraz aktywności użytkownika. Cechami obrazka są: kolor, histogram koloru oraz moment koloru; tekstura; kształt; punkty charakterystyczne SIFT. Cechy hasztagów powstają w oparciu o najczęściej używane hasztagi przez użytkownika oraz do jakich grup przydzielał użytkownik swoje obrazki. Cechy aktywności użytkownika są grupy w których to dana osoba jest aktywna. Na podstawie tych trzech cech powstaje wektor. Następnie uczony jest algorytm kNN i sprawdzana jest jakość algorytmu. Model ten był testowany na serwisie Flickr. W badaniu wzięło udział 15689 zdjęć skategoryzowanych do 925 grup (średnio 17 obrazków na grupę).

Autorzy uważają, że wykorzystywanie danych, które oferują sieci społecznościowe stwarza możliwości lepszego opracowywania metod do rekomendowania hasztagów, niż biorąc tylko pod uwagę cechy wizualne obrazków.

Rozdział 3

Wstęp teoretyczny

W tym rozdziale zostanie omówiony sposób w jaki można dokonać ekstrakcji danych z portalu Instagram (podrozdział 3.1) oraz zostanie opisany algorytm SIFT na podstawie którego działa jedna z zaproponowanych w tej pracy metod w celu klasyfikacji zdjęć (podrozdział 3.2).

3.1 Ekstrakcja danych z Instagrama

API Instagrama daje możliwość wyciągnięcia wielu informacji z tego serwisu, na przykład: informacji o użytkownikach, ich komentarzach, relacjach z innymi użytkownikami serwisu, znajdowanie zdjęć, które są opisane dowolnym hasztagiem i wiele innych¹.

API nie zwraca żadnych wyników z kont, które są oznaczone jako prywatne na Instagramie.

Aby móc korzystać w pełni z API Instagrama trzeba uzyskać token². Token ten musi być podany jako jeden z parametrów przy korzystaniu z API Instagrama. Na dzień 2015-04-03 API daje możliwość wykonaniu 5000 zapytań na godzinę na jeden token³.

W celu przeprowadzenia badań były potrzebne zdjęcia, miejsca w których zostały one wykonane oraz ich hasztagi.

3.1.1 Pozyskiwanie obrazów i metadanych z wykorzystaniem API Instagrama

W tej pracy wykorzystywany był punkt końcowy „media/search” z API Instagrama. Ten punkt końcowy zwraca jako wynik informacje o zdjęciach, które spełniają wymagania określone przez parametry wejściowe.

¹Więcej informacji jakie możliwości daje API Instagrama opisanych jest na stronie <https://instagram.com/developer/endpoints/>.

²<http://bobmckay.com/web/simple-tutorial-for-getting-an-instagram-clientid-and-access-token/>. Dostęp: 2015-04-03.

³W przykładowych zapytaniach podany jest nieaktywny token.

Parametry wejściowe

Parametry wejściowe określają jakie warunki muszą spełniać media, aby mogłyby być zwrócone jako wynik. Parametry jakie przyjmuje ten punkt końcowy to:

- **lat** - szerokość geograficzna z której zostało wgrane zdjęcie na serwer.
- **lng** - długość geograficzna z której zostało wgrane zdjęcie na serwer.
- **min_timestamp** - czas w formacie UNIX określający, że zostaną wrócone wszystkie media, które znalazły się na serwerze po tym czasie.
- **max_timestamp** - czas w formacie UNIX określający, że zostaną wrócone wszystkie media, które znalazły się na serwerze przed tym czasem.
- **count** - określa ilość mediów, która ma być zwrócona w jednym zapytaniu. Wartość domyślna wynosi 20. Wartość maksymalna wynosi 100.
- **distance** - określa długość promienia okręgu w metrach od wcześniej wskazanej długości i szerokości geograficznej, z którego to rejonu zostaną zwrócone zdjęcia. Wartość domyślna wynosi 1000 metrów. Wartość maksymalna wynosi 5000 metrów.

Zapytanie 4.1 zwróci zdjęcia, które były wykonane w promieniu 50 metrów od szerokości 41.403585 i długości 2.174191 geograficznej w czasie pomiędzy 2014-03-14 00:00:00, a 2014-03-24 00:00:00. API zwróci 100 wyników lub mniej jeżeli nie znajdzie aż tylu mediów odpowiadającym tym kryteriom.

Zapytanie 4.1: Przykładowe zapytanie do punktu końcowego.

```
https://api.instagram.com/v1/media/search?lat=41.403585&lng=2.174191&max_timestamp=1403395200&min_timestamp=1394755200&count=100&distance=50&access_token=1384643294.1fb234f.824821f952274b0b9b32577f6a440dbb
```

Dane wyjściowe

Wynik jest zwracany w formacie JSON. Zawiera on informacje o mediach, które spełniają wymagania parametrów wejściowych. API do Zapytania 4.1 zwraca wiele informacji. Główne informacje, które w pliku JSON są szerzej opisane to: kto dodał zdjęcie, kto polubił zdjęcie, jakie hasztagi ma zdjęcie, gdzie zostało zrobione zdjęcie, komentarze do zdjęcia, kiedy zdjęcie zostało dodane do serwisu, stosowany filtr na zdjęciu, użytkownicy, którzy znajdują się na zdjęciu, adresy URL do: miniaturki, małego zdjęcia i standardowej wielkości zdjęcia.

W dalszej części pracy będą wykorzystywane informacji: jakie hasztagi ma zdjęcie oraz URL do standardowej wielkości zdjęcia.

Problemy związane z ekstrakcją danych

Chcąc uzyskać dane z okresu 2014-01-01 00:00:00 do 2015-01-01 00:00:00 i podając te wartości jako **min_timestamp** oraz **max_timestamp** API zwróci maksymalnie w odpowiedzi 100 zdjęć lub tyle ile jest ich dostępnych. W celu uzyskania większej ilości wyników programowo została wprowadzona zmienna **P**. Określa ona różnicę czasu w dniach między kolejnymi zapytaniami do API. Dla $P = 10$, **min_timestamp** = 2014-01-01, **max_timestamp** = 2015-01-01, zapytania do API pod względem momentu wgrania na serwer, będą przybierać następujące wartości czasu: <2014-01-01; 2014-01-11>, (2014-01-11; 2014-01-21>, (2014-01-21; 2014-01-31>, (2014-01-31; 2014-02-10> i tak dalej.

Wprowadzenie zmiennej P prezentuje Algorytm 3.1

Algorytm 3.1: Wprowadzenie zmiennej P

```

1 while (min_timestamp <= max_timestamp)
2 {
3   if (min_timestamp + P >= max_timestamp)
4     Pobierz dane z okresu od min_timestamp do max_timestamp;
5   else
6     Pobierz dane z okresu od min_timestamp do min_timestamp + P
7
8   min_timestamp += P;
9 }
```

Podczas badań można było zauważyć, że API Instagrama nie zawsze zwracało identyczne wyniki dla tego samego zapytania. W przypadku zapytania, które różniło się tylko parametrem **distance** były zwracane czasem nielogiczne wyniki, np. przy parametrze **distance** = 50 było zwracanych 40 obrazków, a dla **distance** = 25 API zwróciło 44 obrazki.

3.2 Skalowiezmiennicze przekształcenie cech - SIFT

Pozyskane obrazy w procesie ekstrakcji danych z portalu Instagram są następnie przetwarzane w celu uzyskania punktów kluczowych do zdjęcia. Dzięki punktom kluczowym można porównywać czy zdjęcia są do siebie wizualnie podobne. Następnie obliczana jest ilość wspólnych punktów kluczowych pomiędzy dwoma grafikami. Jest to kluczowy element dla metody zaproponowanej w podrozdziale 4.2.

Algorytm ten został stworzony w 1999 roku przez Davida G. Lowe [13]. Aktualnie algorytm ten jest szeroko stosowany w różnych dziedzinach, np. rozpoznawania obiektów, rekonstrukcjach 3D, śledzenia obiektów.

Algorytm SIFT daje możliwość detekcji i opisu lokalnych cech (inaczej nazywanych punktami kluczowymi) na obrazie, które są niezależne od położenia, rotacji, skali, zmiany oświetlenia oraz w pewnym zakresie od zmiany kąta widzenia i są odporne na zakłócenia [14]. W momencie kiedy są już opisane punkty kluczowe zdjęcia, to mogą być one porównane z punktami kluczowymi obrazu wzorcowego.

Algorytm SIFT można podzielić na cztery główne kroki:

1. Budowanie przestrzeni skalowej (podrozdział 3.2.1). Na tym etapie wykrywane są ekstrema obrazka za pomocy różnic rozmyć Gaussowskich. Na tych ekstremach znajdują się potencjalne punkty kluczowe.
2. Lokalizacja punktów kluczowych (podrozdział 3.2.2). Wykrywane są punkty kluczowe wśród potencjalnych punktów kluczowych. Zostają również odrzucone te punkty kluczowe, które znajdują się blisko krawędzi zdjęcia.
3. Wyznaczanie kierunku punktu kluczowego (podrozdział 3.2.3). Dzięki temu stadium zapewniona jest niezależność względem obrotu zdjęcia.
4. Tworzenie deskryptora punktu kluczowego (podrozdział 3.2.4). W tym kroku tworzony jest wektor, który opisuje punkt kluczowy. Wektor ten zawiera dane, które były wyznaczone we wcześniejszych etapach.

W momencie kiedy dwa zdjęcia mają wyznaczone punkty kluczowe wraz z ich deskryptorami można znaleźć cechy wspólne dla obydwu obrazków (podrozdział 3.2.5).

3.2.1 Przestrzeń skalowa

W celu wykrycia cech lokalnych budowana jest piramida przestrzenna skali. Piramida jest podzielona na oktawy. Oktawa zawiera zdjęcia o tej samej rozdzielczości, ale wzrastającym odchyleniu standardowym filtru Gaussa. W kolejnych oktawach (poziomach piramidy) zmniejsza się wielkość obrazów.

Obraz wejściowy przekształcany jest do trójwymiarowej przestrzeni skali, w której do dwóch wymiarów przestrzennych dodawany jest trzeci - skala, poprzez splot z funkcją Gaussa o zmiennym odchyleniu standardowym:

$$L(x, y, \sigma) = G(x, y, \sigma) \cdot I(x, y) \quad (3.1)$$

gdzie $I(x, y)$ to obraz wejściowy, $G(x, y, \sigma)$ to operator Gaussa, ze współczynnikiem rozmycia σ , $L(x, y, \sigma)$ to nowy obraz rozmyty [15].

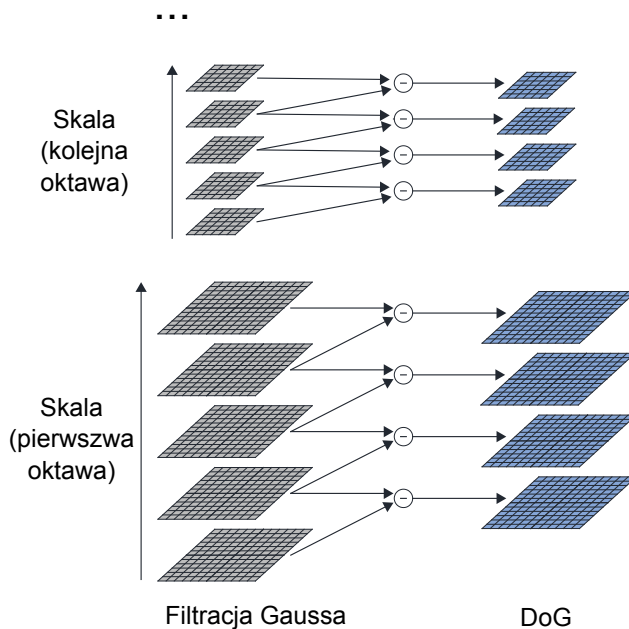
W kolejnym kroku obliczane są różnice rozmyć Gaussowskich dwóch kolejnych obrazków, tak zwany Difference of Gaussians (skrót: DoG). Powstaje nowa macierz, w której wartości 0 przyjmują piksele, gdzie obrazy mają taką samą wartość po rozmyciu (Równanie 3.2).

$$[G(x, y, k\sigma) - G(x, y, \sigma)] \cdot I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) = D(x, y, \sigma) \quad (3.2)$$

W artykule [13] autor wyznaczył empirycznie optymalne wartości parametrów: liczba oktaw = 4, liczba poziomów skali = 5, początkowa wartość $\sigma = 1,6$, $k = \sqrt{2}$.

Operacje opisane powyżej można przedstawić w uproszczonej wersji obrazkowej (Rysunek 3.1).

W momencie kiedy DoG jest wyznaczone, zaczyna się etap wyszukiwania lokalnych ekstremów. Odbywa się to przez porównanie punktu z obrazu DoG z ich sąsiadami. Jeden piksel porównywany jest z 8 sąsiadami na tym samym poziomie skali oraz 9 na poprzednim i następnym poziomie skali (Rysunek 3.2). Jeżeli dany punkt jest lokalną ekstremą to jest to potencjalny punkt kluczowy.



Rysunek 3.1: Obliczanie DoG. Źródło: własne.
Opracowane na podstawie: [16].

3.2.2 Lokalizacja punktów kluczowych

Punkty kluczowe to niektóre elementy wśród potencjalnych punktów kluczowych. Potencjalnych punktów kluczowych jest bardzo dużo. Spowodowane jest to między innymi szumami na zdjęciu. Zostaną odrzucone te punkty potencjalne, które mało wyróżniają się wśród innych. W tym celu obliczany jest szereg Taylora w przestrzeni skali (Równanie 3.3).

$$D(\mathbf{x}) = D + \frac{\partial D}{\partial \mathbf{x}} \mathbf{x} + \frac{1}{2} \mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2} \mathbf{x} \quad (3.3)$$

gdzie $\mathbf{x} = [x, y, \sigma]^T$.

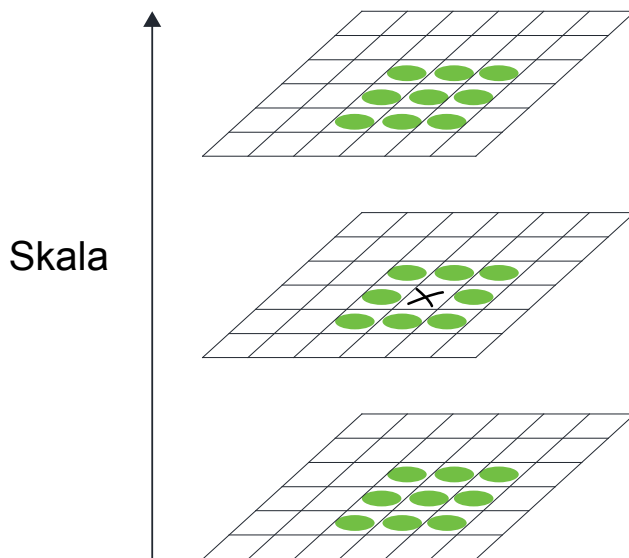
Następnie wyznaczane jest położenie ekstremum, dzięki pochodnej funkcji $D(x)$ (Równanie 3.4) i zostaje ona przyrównana do wartości 0.

$$\hat{\mathbf{x}} = -\frac{\partial^2 D^{-1}}{\partial \mathbf{x}^2} \frac{\partial D}{\partial \mathbf{x}} \quad (3.4)$$

Jeżeli \hat{x} jest większe niż 0.5 w którymkolwiek wymiarze, to zostaje odrzucony ten kandydat z możliwych punktów kluczowych [16]. Jeżeli intensywność punktu w ekstremum jest mniejsza niż 0.03 to punkt ten jest odrzucony. Dodatkowo dużo punktów przy krawędzi obrazka może być traktowane jako punkty kluczowe. Punkty te są odrzucone, gdy spełniają warunek $\det(\mathbf{H}) \leq 0$ lub $\frac{\text{Tr}(\mathbf{H})^2}{\det(\mathbf{H})} \geq \frac{(r+1)^2}{r}$, gdzie $\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$ oraz $r = 10$.

3.2.3 Kierunki punktów kluczowych

Na tym etapie wyznaczana jest orientacja punktu kluczowego. Wyznacza się ją dzięki lokalnym gradientom. Przy pomocy takiego zabiegu uzyskuje się niezależność punktów klu-



Rysunek 3.2: Wyznaczenie ekstremów. Źródło: własne. Opracowane na podstawie: [16].

czowych od ich rotacji. W celu wyznaczenia orientacji punktu kluczowego obliczany jest kierunek i wielkość (moc) gradientu. Wyniki przedstawiane są w formie histogramu. Histogram posiada 36 przedziałów. Każdy przedział odpowiada zakresowi 10° . Moc i wielkość gradientu obliczana jest w sposób następujący:

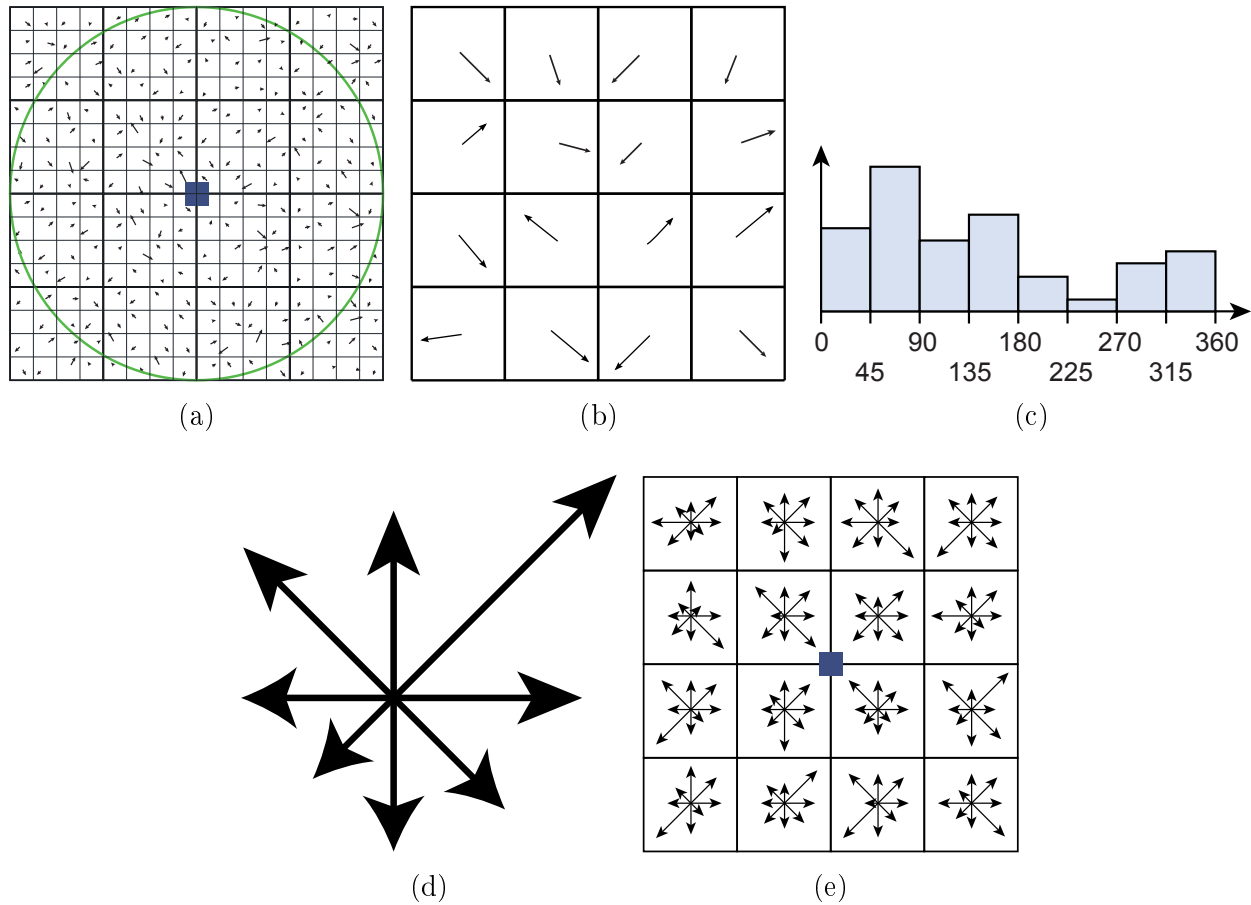
$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (3.5)$$

$$\theta(x, y) = \arctg\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \quad (3.6)$$

gdzie $m(x, y)$ to wielkość gradientu, a $\theta(x, y)$ to kierunek gradientu wyrażony w stopniach. Następnie odcina się wartości na histogramie, które nie przekraczają 80% wartości najwyższego słupka. Może istnieć kilka kierunków dla jednego punktu na obrazie. Operacja ta jest powtarzana dla wszystkich punktów kluczowych na obrazie.

3.2.4 Deskryptor punktu kluczowego

W celu stworzenia deskryptora punktu kluczowego brane jest jego sąsiedztwo o wymiarach 16×16 (Rysunek 3.3a), a następnie dzielone na okna o wielkości 4×4 (Rysunek 3.3b). W tych oknach wyliczany jest histogram orientacji w sposób podobny jak w 3.2.3. Histogram ten ma jednak w tym przypadku 8 przedziałów, co odpowiada jednej kolumnie 45° (Rysunek 3.3c). W kolejnym kroku tworzy się orientację punktu bazując na powstałym wcześniej histogramie (Rysunek 3.3d). Następnie nakładana jest waga na próbki w postaci dwuwymiarowej funkcji Gaussa. Powstaje 16 okien (Rysunek 3.3e). Każde okno posiada 8 wartości. W ten sposób tworzony jest wektor cech o 128 wartościach dla punktu kluczowego.



Rysunek 3.3: Deskryptor punktu kluczowego. Źródło: własne. Opracowane na podstawie: [16].

3.2.5 Dopasowywanie punktów kluczowych

Punkt kluczowy opisany jest wektorem. Jeżeli wymaga się porównania dwóch obrazów i znalezienia podobieństw pomiędzy nimi należy porównywać ze sobą deskryptory punktów kluczowych tych zdjęć. Im bardziej punkty kluczowe są do siebie podobne, tym wektory te powinny znajdować się bliżej siebie to znaczy, że odległość Euklidesowa pomiędzy wektorami powinna być mała. Porównywanie wielowymiarowych wektorów jest znanym problemem związanym ze złożonymi obliczeniami. W celu przyspieszenia obliczeń rekomendowany jest algorytm Najlepszy Koszyk Pierwszy (*ang. best-bin-first*), który określa najbliższe położone wektora z wysoką skutecznością i z użyciem ograniczonej ilości obliczeń. W pewnych sytuacjach zdarza się, że drugi najbliższy leżący wektor jest bardzo blisko pierwszego. W takim przypadku oblicza się stosunek odległości pierwszego wektora do drugiego. Jeżeli stosunek ten jest większy niż 0.8 to te punkty są odrzucane. Dzięki temu zabiegowi eliminowanych jest około 90% fałszywych dopasowań [17].

Rozdział 4

Opis metody rozwiązania problemu

W tym rozdziale zostaną zaprezentowane dwie metody rozwiązania problemu klasyfikacji zdjęć na portalu Instagram. Pierwsza z zaproponowanych metod (podrozdział 4.1) polega na tworzeniu grafów, gdzie wierzchołkami są hasztagi występujące w zbiorze uczącym. Druga z zaproponowanych metod (podrozdział 4.2) korzysta z algorytmu SIFT. Znajdowane są wspólne punkty kluczowe na zdjęciu testowym, a następnie liczona jest ilość wspólnych punktów kluczowych pomiędzy grafiką testową, a obrazkami ze zbioru uczącego.

4.1 Metoda grafów

Pierwszą z proponowanych metod będzie nazywana metodą grafów. Rozwiązanie to polega na stworzeniu grafu, gdzie wierzchołkami są hasztagi, a krawędzie opisane są liczbą, która określa ilość współwystępujących ze sobą hasztagów. Liczba ta niech będzie nazywana *siłą* krawędzi. Dwa wierzchołki z krawędzią i siłą krawędzi będzie nazywane *połączeniem*.

4.1.1 Tworzenie grafów

W celu stworzenia grafu pobierane są dane z API (Podrozdział 3.1). Następnie brane są pod uwagę tylko te informacje, gdzie ilość występujących hasztagów do jednego zdjęcia jest równa bądź większa niż dwa.

Niech \mathbf{U} będzie zbiorem zdjęć wraz ze skojarzonymi z nimi hasztagami. Dla każdego $\mathbf{u}_{\mathbf{N} \times \mathbf{M}} \in \mathbf{U}$, \mathbf{N} oznacza kolejne zdjęcie, \mathbf{M} oznacza kolejny hasztag zdjęcia. Dla przykładu $\mathbf{U} = \{\{\text{love, like4like, l4l}\}; \{\text{serce, miłość, wrocław}\}; \{\text{boyfriend, girlfriend, instagood, love}\}\}$. Dla $\mathbf{u}_{1 \times 1} = \text{love}$, dla $\mathbf{u}_{3 \times 3} = \text{instagood}$.

W celu stworzenia grafu dla każdego zdjęcia z zbioru \mathbf{U} przeprowadzane są następujące operacje: tworzone są pary hasztagów z hasztagów, które należą do tego obrazka. Poprzez parę będzie rozumiany zbiór dwuelementowy. Tworzone są pary w taki sposób, aby każdy hasztag współwystępował z wszystkimi innymi hasztagami. Przykład: dla jednego zdjęcia, które ma hasztagi: $\{\text{love, sky, l4l}\}$ powstają pary: $\{\text{love, sky}\}$, $\{\text{love, l4l}\}$, $\{\text{sky, l4l}\}$. Następnie analizowana jest każda kolejna para hasztagów. Sprawdzane jest czy istnieją w grafie wierzchołki, które mają taką samą nazwę jak hasztagi z analizowanej aktualnie pary. Jeżeli graf nie posiada jednego bądź obydwu wierzchołków to są one tworzone. Zostanie teraz opisany przypadek tworzenia wierzchołków grafu. Założenie jest takie, że graf jest pusty - nie

ma żadnego wierzchołka. Są pary hasztagów $\{\text{love}, \text{sky}\}$, $\{\text{love}, \text{l4l}\}$, $\{\text{sky}, \text{l4l}\}$ pochodzące z pierwszego zdjęcia. Sprawdzana jest pierwsza para $\{\text{love}, \text{sky}\}$. W grafie nie ma wierzchołka, który jest opisany słowem *love*. Dodawany jest nowy wierzchołek, który będzie mieć nazwę *love*. Drugi hasztag z pary: *sky* również nie istnieje w grafie. Zostaje dodany nowy wierzchołek z nazwą *sky*. Analizowany jest kolejna para hasztagów $\{\text{love}, \text{l4l}\}$. Słowo *love* już istnieje w grafie, więc nie zostaną przeprowadzone żadne operacje. Wyrażenie *l4l* nie istnieje w grafie. Zostaje więc dodany nowy wierzchołek, który będzie opisany jako *l4l*. Sprawdzona teraz jest ostatnia para hasztagów: $\{\text{sky}, \text{l4l}\}$. Słowa *sky* i *l4l* istnieją już w grafie, więc nie zostanie przeprowadzona żadna operacja. Ostatnim krokiem podczas tworzenia grafu jest nadanie *siły* krawędzią pomiędzy wierzchołkami. W przypadku opisanym wcześniej pomiędzy każdą parą hasztagów powstaje krawędź, a jej *siła* ma wartość jeden. W przypadku kiedy kolejne zdjęcie brane pod uwagę podczas tworzenia grafu będzie posiadało tylko jedną parę hasztagów $\{\text{love}, \text{sky}\}$ to nie zostaną stworzone nowe wierzchołki ponieważ wierzchołki *love* i *sky* już istnieją. Zostanie zwiększona wartość *siły* krawędzi pomiędzy tymi dwoma wierzchołkami o jeden i w ten sposób wartość *siły* krawędzi pomiędzy wierzchołkami *love* oraz *sky* jest równa 2.

Zoptymalizowany kod tworzenia grafu przedstawia Algorytm 4.1.

Algorytm 4.1: Zoptymalizowany kod tworzenia grafu \mathbf{G}

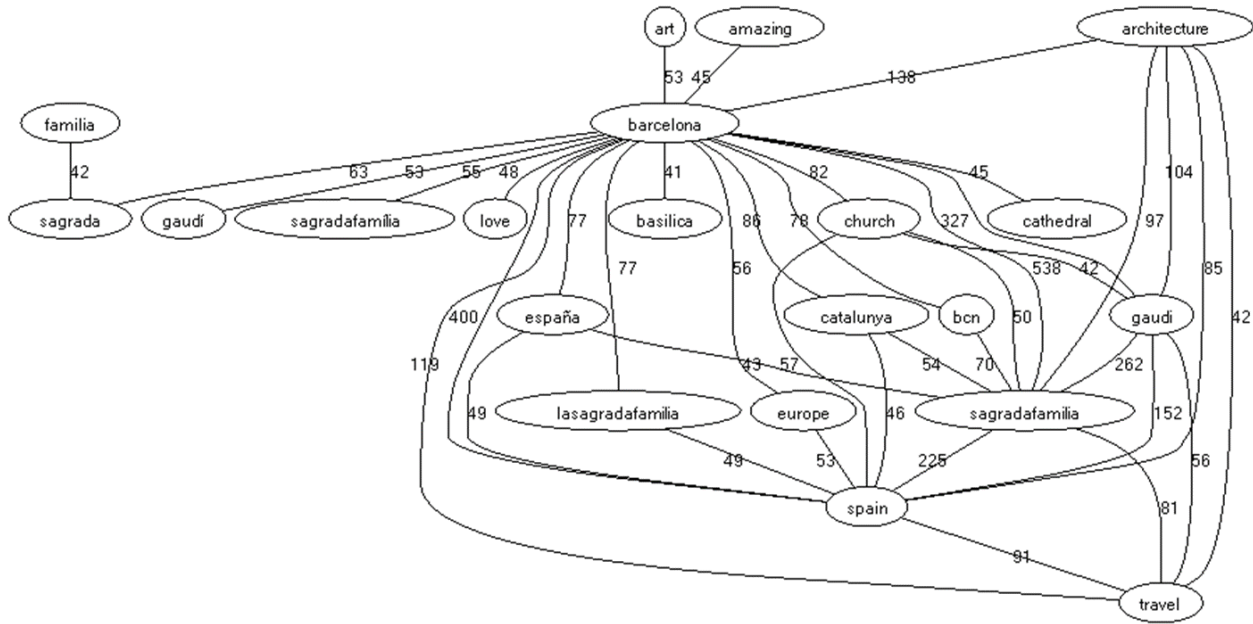
```

1 for (int N = 0; N < LiczbaZdjęćWZbiorze(U); N++)
2 {
3     for (int M = 0; M < LiczbaHasztagówPrzyZdjęciu(UN); M++)
4     {
5         if (CzyWystępujeWGracieHasztag(UNxM))
6         {
7             for (int I = M + 1; I < LiczbaHasztagówPrzyZdjęciu(UN); I++)
8             {
9                 if (CzyIstniejeKrawędźSkładającaSięZHasztagów(UNxM, UNxI))
10                    ZwiększOJedenSiłęKrawędzi(UNxM, UNxI);
11                else
12                    StwórzNowąKrawędźIUstawJegoSiłęNaJeden(UNxM, UNxI);
13            }
14        }
15        else
16        {
17            for (int I = M + 1; I < LiczbaHasztagówPrzyZdjęciu(UN); I++)
18                StwórzNowąKrawędźIUstawJegoSiłęNaJeden(UNxM, UNxI);
19        }
20    }
21 }

```

Na Rysunku 4.1 przedstawiona jest przykładowa wizualizacja grafu dla 20 największych wartości siły połączenia. Graf powstał dla przypadku, gdzie zmienne przyjmowały następujące wartości: **lat** = 41.403585, **lng** = 2.174191, **min_timestamp** = 2014.03.14, **max_timestamp** = 2015.03.14, **P** = 10, **distance** = 50, **count** = 100.

Niech otrzymany w wyniku działania algorytmu graf będzie oznaczony jako \mathbf{G} , $\mathbf{G}(tag_1)$ będzie zwracać zbiór wszystkich hasztagów, które są połączone na grafie \mathbf{G} z hasztagiem tag_1 , $\mathbf{G}(tag_1, tag_2)$ będzie zwracać siłę połączenia pomiędzy tagami tag_1, tag_2 . Dla przy-



Rysunek 4.1: Przykładowa wizualizacja wycinka grafu \mathbf{G} . Opracowanie własne.

kładu zakładając, że \mathbf{G} jest odwzorowaniem grafu z Rysunku 4.1 to $\mathbf{G}(„architecture”) = \{„barcelona”; „sagradafamilia”; „gaudi”; „spain”; „travel”\}$, $\mathbf{G}(„architecture”, „barcelona”) = 138$. Aby przejść do kolejnego etapu graf \mathbf{G} musi mieć co najmniej jedno połączenie.

W kolejnym kroku pobierane są zdjęcia z hasztagami, które pochodzą z innego okresu czasu niż te, które służyły do stworzenia grafu \mathbf{G} . Na ich podstawie będzie przeprowadzona analiza jak dobrze algorytm radzi sobie z klasyfikacją zdjęć. Zbiór ten będzie oznaczony jako \mathbf{K} . W zbiorze \mathbf{K} są tylko zdjęcia, które mają co najmniej 10 hasztagów. Sposób klasyfikacji jest zbliżony do zaprezentowanego w artykule [5]. Z zbioru \mathbf{K} pobierane jest zdjęcie k . Z k pobierane są losowe dwa hasztagi. Powstaje zbiór T , który zawiera dwa wcześniej wylosowane hasztagi. Następnie tworzony jest graf \mathbf{G}_T . Graf \mathbf{G}_T powstaje z skopiowania fragmentów z grafu \mathbf{G} . Poprzez fragmenty rozumiane są połączenia, które zawierają w sobie hasztagi z zbioru T . Dla przykładu niech $T = \{„classic”; „picoftheday”\}$. Wierzchołek „classic” nie istnieje w grafie \mathbf{G} , więc żadne połączenia związane z wyrazem „classic” nie są kopiowane. Wierzchołek „picoftheday” istnieje w grafie \mathbf{G} dlatego są przenoszone wszystkie połączenia, które zawierają w sobie hasztag „picoftheday” z grafu \mathbf{G} do \mathbf{G}_T . Nowo powstały graf \mathbf{G}_T reprezentuje Rysunek 4.2.

W przypadku kiedy dwa pierwsze losowo wybrane hasztagi nie znajdują swojego odwzwiedlenia na grafie \mathbf{G} i powstanie pusty graf \mathbf{G}_T to obrazek nie jest brany pod uwagę w trakcie dalszej analizy.

Algorytm 4.2 przedstawia metodę tworzenia grafu \mathbf{G}_T .

Algorytm 4.2: Tworzenie grafu \mathbf{G}_T

```

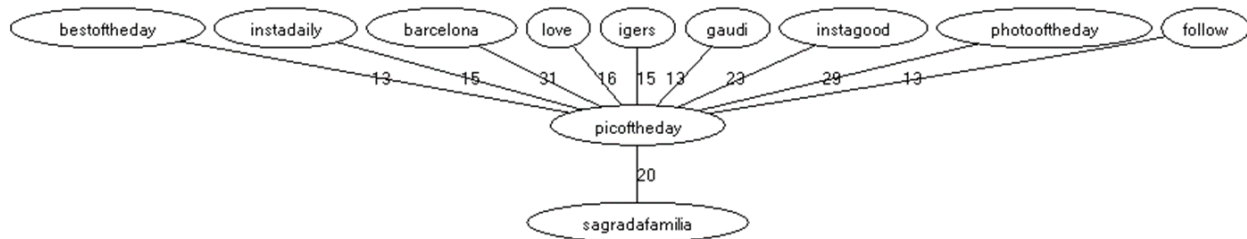
1 for (int index = 0; index < LiczbaHashtagówWZbiorze(T); index++)
2 {
3     ZnalezionePołączenia = ZnajdźWszystkiePołączenieWGracieGWktó
        rychWystępujeHashtag( $T_{index}$ );

```

```

4
5   WszystkieZnalezionePołączeniaDodajDoGrafuGT(ZnalezionePołączenia);
6 }

```



Rysunek 4.2: Przykładowa wizualizacja G_T . Opracowanie własne.

4.1.2 Wybieranie hashtagów

Niech W_{tag} oznacza sumę siły wszystkich krawędzi skierowanych do wierzchołka opisanym wartością „tag” w grafie G_T . Dla przypadku opisanego na Rysunku 4.2 to $W_{\text{follow}} = 13$, $W_{\text{picoftheday}} = 188$. W celu znalezienia hashtagu, który będzie opisywać obrazek wyszukuje się największą wartość sumy siły krawędzi należącej do wierzchołka wyłączając hashtagi należące do zbioru T . Dla przypadku przedstawionego na Rysunku 4.2 zostanie wybrany hashtag „barcelona”. Wybrany wierzchołek dołączany jest do zbioru T . Na nowo tworzony jest graf G_T , bazując na zbiorze T , który w tym momencie zawiera trzy hashtagi: „classic”, „picoftheday” i „barcelona”. Następuje wyszukiwanie kolejnego nowego hashtagu w sposób opisany powyżej. Parametr k będzie określał ile hashtagów ma wyznaczyć ten algorytm.

4.1.3 Sprawdzanie ilości poprawnie wybranych hashtagów

W celu ewaluacji ile hashtagów poprawnie klasyfikuje obrazek pochodzący ze zbioru K obliczana jest ilość wspólnych przecięć dwóch zbiorów. Zbiór hashtagów który, opisuje obrazek k będzie wyrażony przez Z_1 . Drugi zbiór to są hashtagi pochodzące ze zbioru T , wyłączając te które zostały wybrane w sposób losowy. Zbiór ten będzie oznaczany jako Z_2 . Następnie obliczana jest liczba poprawnie wybranych hashtagów, która powstała z działania $Z_1 \cap Z_2$.

4.2 Metoda oparta o algorytm SIFT

Metoda ta opiera się na liczeniu punktów wspólnych pomiędzy dwoma obrazkami przy pomocy algorytmu SIFT (Podrozdział 3.2).

4.2.1 Przygotowanie danych

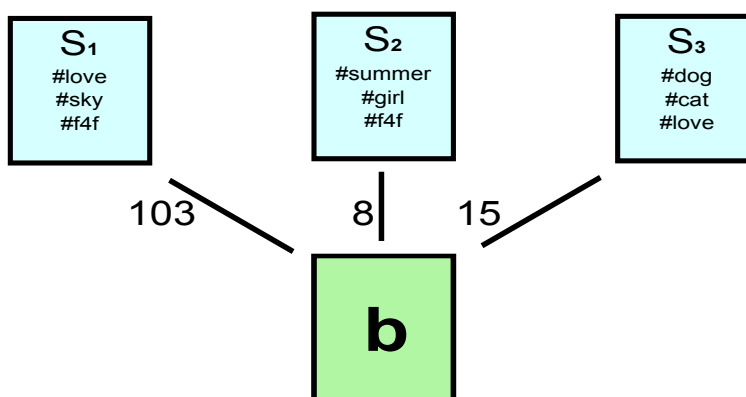
Przy pomocy API (Podrozdział 3.1) pobierany jest zbiór zdjęć na podstawie których będą klasyfikowane zdjęcia. Wśród tych zdjęć zostają odrzucone wszystkie te które nie posiadają żadnego hashtagu. Niech S będzie zbiorem zdjęć wraz ze skojarzonymi z nimi hashtagami.

Dla każdego $s_{N \times M}$, N oznacza kolejne zdjęcie, M oznacza kolejny hashtag zdjęcia. Dla każdego zdjęcia ze zbioru S wykrywane są punkty kluczowe za pomocą algorytmu SIFT (Podrozdział 3.2). Następnie pobierane są zdjęcia, które pochodzą z innego okresu czasu niż obrazy ze zbioru S . Wśród nowo pobranych zdjęć wybierane są te, które mają co najmniej 10 hashtagów. Zbiór tych grafik niech będzie oznaczony przez zmienną B . Dla wszystkich zdjęć ze zbioru B również obliczane są punkty kluczowe.

4.2.2 Wybieranie hashtagów

Niech jedno dowolne zdjęcie z zbioru B będzie oznaczone jako b . Porównywane jest b z wszystkimi pozostałymi grafikami ze zbioru S . Dla każdego zdjęcia ze zbioru S obliczana jest liczba wspólnych punktów kluczowych z b . Liczba punktów wspólnych staje się właściwością zdjęć ze zbioru S . N -te zdjęcie ze zbioru S z wartością właściwości równej x , będzie oznaczane jako $P(S_N) = x$. Tworzona jest tablica struktur. Struktura zawiera dwa pola. Pierwsza zawiera nazwę hashtagu, która będzie opisywać zdjęcie b , drugie pole posiada wartość, która opisuje jak bardzo możliwe jest to, że hashtag z pierwszego pola będzie opisywać zdjęcie b . Struktury są unikatowe względem pierwszego pola. Zdjęcie s_N zawiera hashtagi. Hashtagi te będą przypisywane do pierwszego pola struktury. Do drugiego pola w strukturze przypisywana jest wartość $P(S_N)$. W momencie kiedy hashtag pojawia się przy kilku różnych zdjęciach to nie są tworzone dodatkowe struktury. Wyszukiwana jest już istniejąca struktura, która w pierwszym polu posiada już taką samą nazwę hashtagu. W tej strukturze do drugiego pola dodawana jest wartość funkcji P ze zdjęć, które posiadają ten sam hashtag. Na koniec struktury w tablicy są szeregowane malejąco względem wartości drugiego pola. W tym momencie tablica jest uszeregowana od najbardziej prawdopodobnych hashtagów, które poprawnie będą klasyfikować obrazek do najmniej prawdopodobnych.

Przykład: są trzy zdjęcia w zbiorze S . Każde zdjęcie z zbioru S posiada trzy hashtagi. Obrazek b porównywany jest z wszystkimi grafikami ze zbioru S . Dla $P(S_1) = 103$, $P(S_2) = 8$, $P(S_3) = 15$. Wizualizacja przykładu znajduje się na Rysunku 4.3. Następnie powstaje lista uszeregowanych struktur, która pokazana jest w Tabelcy 4.1. Algorytm 4.3 przedstawia kod wybierania hashtagów.



Rysunek 4.3: Wizualizacja przykładu dla metody działania algorytmu SIFT. Opracowanie własne.

Hashtag	Wartość
love	118
f4f	111
sky	103
dog	15
cat	15
summer	8
girl	8

Tablica 4.1: Przykładowa uszeregowana lista struktur.

Algorytm 4.3: Wybieranie hashtagów w metodzie opartej o algorytm SIFT.

```

1 for (int N = 0; N < LiczbaZdjęćWZbiorze(S); N++)
2 {
3     x = ObliczIlośćWspólnychPunktówKluczowychPomiędzyZdjęciami(SN, b);
4     S(PN) = x;
5 }
6
7 for (int N = 0; N < LiczbaZdjęćWZbiorze(S); N++)
8 {
9     for (int Tag = 0; Tag < IlośćHashtagówPrzyZdjęciu(SN); Tag++)
10    {
11        if (SprawdźCzyIstniejeJuzTakiHashtagWTablicyStruktur(SNxTag))
12        {
13            DodajDoStrukturyZHashtagiemWartość(SNxTag, S(PN))
14        }
15        else
16        {
17            StwórzNowaStrukturęIDodajJąDoTablicyStruktur(SNxTag, S(PN));
18        }
19    }
20 }
21
22 PosortujTabliceStrukturMalejącoWzględemDrugiegoPola();

```

4.2.3 Sprawdzanie ilości poprawnie wybranych hashtagów

W celu ewaluacji ile hashtagów poprawnie klasyfikuje obrazek pochodzący ze zbioru \mathbf{B} obliczana jest ilość wspólnych przecięć dwóch zbiorów. Zbiór hashtagów który, opisuje obrazek b będzie wyrażony przez W_1 . Drugi zbiór to są hashtagi pochodzące z posortowanej tablicy struktur. Wybierane są kolejne hashtagi, która będą klasyfikować obrazek b zaczynając od pierwszej struktury. Parametr k będzie określał ile kolejnych hashtagów ma zostać pobranych. Zbiór z hashtagami, który powstał z wybranych struktur będzie oznaczany jako W_2 . Następnie obliczana jest liczba poprawnie wybranych hashtagów, która powstała z działania $W_1 \cap W_2$.

Rozdział 5

Konfiguracja eksperymentu

W podrozdziale 5.1 zostały opisane zbiory danych testowych i treningowych. Dobór odpowiednich miejsc jest kluczowy, aby algorytmy mogły zwracać wyniki dzięki którym można wyprowadzać wnioski. W podrozdziale 5.2 wytłumaczone są miary ewaluacji na podstawie których była przeprowadzana analiza zagadnienia. W podrozdziale 5.3 omówione są ograniczenia początkowe dla metod, które są ewaluowane.

5.1 Pochodzenie danych eksperymentalnych

Badania zostały przeprowadzone dla popularnych miejscach turystycznych. Selekcja tych miejsc jest ważnym czynnikiem dla badań. Miejsca zostały wybrane na podstawie rankingów z dwóch stron internetowych:

- <http://www.fly4free.pl/uwaga-na-tlumy-oto-najliczniej-odwiedzane-atrakcje-turystyczne-na-swiecie/>. Z tej witryny zostały wybrane miejsca z Ameryki Północnej i Południowej. Zostały odrzucone te atrakcje turystyczne, które występowały w Azji Wschodniej ze względu na duże prawdopodobieństwo występowania azjatyckich znaków językowych.
- http://metro.gazeta.pl/Lifestyle/5,127257,13001774,0skary_podroznicze_2012_Najlepsza_atrakcja_turystyczna.html. Z tej strony zostały wybrane europejskie atrakcje turystyczne.

Dodatkowo został dodany polski akcent: Pałac Kultury i Nauki w Warszawie.

Czynnikami, które motywowały do wybrania miejsc turystycznych do dalszego badania metod były:

- Możliwość zrobienia zdjęcia w popularnym miejscu za darmo i wgrania go na Instagram. Wyjątkiem jest tutaj Disneyland Drive, gdzie za wstęp trzeba zapłacić.
- W najbliższym otoczeniu nie znajdują się inne popularne obiekty turystyczne.
- Charakterystyczne elementy wizualne w danych miejscach. Wyjątkiem jest tutaj Plac Zocalo, który jest znany z dużej pustej przestrzeni w centrum miasta Meksyk.

Dodatkowo został dodany Wodospad Niagara jako znane miejsce, które nie jest dziełem rąk ludzkich.

Dokładne zestawienie miejsc, które były brane pod uwagę znajduje się w Tabelicy 5.1. Długość i szerokość geograficzna została zebrana z strony <http://www.latlong.net/>.

Miejsce	Długość geograficzna	Szerokość Geograficzna
Pałac Kultury i Nauki w Warszawie	21.005931	52.231831
Tower Bridge	-0.075086	51.505840
Wieża Eiffel	2.294481	48.858370
Katedra Notre Dame	2.349693	48.853016
Wodospad Niagara	-79.068634	43.085906
Sagrada Familia	2.174191	41.403585
Times Square	-73.984863	40.758837
Grand Central Terminal	-73.977056	40.752998
Disneyland Drive	-117.918980	33.812719
Plac Zocalo	-99.133205	19.432602

Tablica 5.1: Wybrane miejsca, który były brane pod uwagę w trakcie eksperymentu.

W dalszej części pracy używane będą nazwy skrócone miejsc wyżej wymienionych. Skróty są przedstawione w Tabelicy 5.2.

Miejsce	Nazwa skrócona
Pałac Kultury i Nauki w Warszawie	PKiNW
Tower Bridge	TB
Wieża Eiffel	WE
Katedra Notre Dame	KND
Wodospad Niagara	WN
Sagrada Familia	SF
Times Square	TS
Grand Central Terminal	GCT
Disneyland Drive	DD
Plac Zocalo	PZ

Tablica 5.2: Skrócone nazwy miejsc.

5.2 Miary ewaluacji

W celu ewaluacji jakie wyniki zwracają zaproponowane metody została zaadoptowana metryka z artykułów [18], [7]. **Precyzja dla k rekomendowanych hashtagów ($P@k$)** jest opisana wzorem:

$$P@k = \frac{S}{k} \quad (5.1)$$

gdzie S jest liczbą opisująca ilość poprawnie wybranych hashtagów dla pojedynczego obrazka ze zbioru testowego (podrozdział 4.1.3 lub podrozdział 4.2.3), k jest liczbą, która opisuje ile hashtagów ma opisywać dany obrazek.

W tej pracy również będą używane następujące metryki:

$$AP@k_{m,d} = \frac{\sum_{i=1}^n P@k_{m,d,i}}{n} \quad (5.2)$$

gdzie m określa z jakiej lokalizacji były pobrane zdjęcia (Tabela 5.1), d określa parametr **distance**, n to liczba wszystkich zdjęć, które zostały pobrane z Instagrama do jednej kategorii (ustalonych parametrów **distance**, **lat** i **lng**), i określa kolejne zdjęcie z wybranej kategorii. $AP@k_{m,d}$ będzie oznaczać średnią arytmetyczną $P@k$ dla wszystkich zdjęć należących do jednej kategorii.

W celu obliczenia średniej wartości $P@k$ dla wszystkich zdjęć z kilku kategorii, gdzie stałymi współczynnikami są d oraz k , a zmiennym parametrem będzie miejsce dla którego algorytm działa, stosowany jest wzór:

$$AP@k_d = \frac{\sum_m^M AP@k_{m,d}}{M} \quad (5.3)$$

gdzie M to liczba wszystkich miejsc z Tablicy 5.1, a m to kolejne miejsca z Tablicy 5.1.

Aby obliczyć średnią wartość $P@k$ dla wszystkich zdjęć z kilku kategorii, gdzie stałymi współczynnikami są m oraz d , a zmiennym parametrem będzie k , stosowany jest wzór:

$$AP@m_d = \frac{\sum_k^K AP@k_{m,d}}{K} \quad (5.4)$$

gdzie K to liczba wszystkich parametrów k dla których był wykonywany eksperyment.

W celu obliczenia średniej wartości $P@k$ dla wszystkich zdjęć z kilku kategorii, gdzie stałym współczynnikiem jest m , a zmiennymi parametrami są k oraz d , stosowany jest wzór:

$$AP@m = \frac{\sum_d^D AP@m_d}{D} \quad (5.5)$$

gdzie D to liczba parametrów **distance**, dla których był przeprowadzony eksperyment, d to kolejne wartości parametru **distance**. Dzięki temu wzorowi można wyznaczyć dla jakiej lokalizacji najlepiej działa algorytm.

Za obliczenie średniej wartości $P@k$ dla stałego parametru k , niezależnie od parametru d i m odpowiada wzór:

$$AP@k = \frac{\sum_d^D AP@k_d}{D} \quad (5.6)$$

Dzięki temu wzorowi można wnioskować dla jakiego parametru k algorytmy zwracały najlepsze wyniki.

Aby dowiedzieć się dla jakiej wartości stałego parametru **distance** algorytmy najlepiej klasyfikują hashtagi do zdjęcia został użyty następujący wzór:

$$AP@d = \frac{\sum_k^K AP@m_d}{K} \quad (5.7)$$

5.3 Zapytania do API i ograniczenia początkowe

Dane testowe są pobierane z okresu od 2013.03.14 do 2014.03.14. Dane treningowe pochodzą z okresu od 2014.03.14 do 2015.03.14. Stałymi parametrami są $\mathbf{P} = 10$ oraz $\mathbf{count} = 100$. Parametr **distance** przyjmuje kolejno wartości: 10, 25, 50, 100. Zbieranych jest łącznie 88 zestawów danych; dla 11 miejsc, 4 różnych wartości parametru **distance** oraz 2 różnych okresów czasowych z których pochodzą dane.

Na etapie ewaluacji dla zbioru testowego brane pod uwagę są tylko te zdjęcia, które posiadają co najmniej 10 hashtagów. Dodatkowo dane testowe i treningowe dla pojedynczej lokalizacji muszą mieć więcej niż 15 zdjęć.

Wszystkie wyniki w kolejnym rozdziale są wypisane po zaokrągleniu do dwóch miejsc po przecinku.

Rozdział 6

Wyniki ekspertymentu

W celu uzyskania wyników dla metody grafów (Podrozdział 4.1) były przeprowadzone obliczenia na komputerze z taktowaniem 2,5 GHz. Obliczenia trwały 4 godziny. Obliczenia dla metody opartej o algorytm SIFT (Podrozdział 4.2) były przeprowadzone na maszynie wirtualnej która posiadała 16 procesorów z taktowaniem 1,9 GHz. Procesy były wykorzystywane niemal przez cały czas w 100%. Obliczenia trwały 5 dni. W podrozdziale 6.3 są opisane wyniki z innych artykułów.

6.1 Metoda grafów

W podrozdziale 6.1.1 są przedstawione tablice z wynikami dla metody grafów, a w podrozdziale 6.1.2 są wyciągnięte wnioski dla tej metody.

6.1.1 Wyniki dla metody grafów

W Tablicach 6.1, 6.2, 6.3, 6.4 przedstawione są wyniki $AP@k_{m,d}$ dla różnych kategorii. W Tablicy 6.4 widać, że wartości $AP@k_{m,d}$ dla miejsc PK, TB, WE, SF, GCT przyjmują wartości 0. Jest to spowodowane tym, że nie zostały osiągnięte warunki początkowe opisane w Podrozdziale 5.3. Z tego względu nowa Tablica 6.5 została stworzona po usunięciu miejsc z Tablicy 6.4, które nie spełniają wymagań początkowych. W Tablicy 6.6 są skondensowane wyniki dla wszystkich wcześniejszych tabelic w tym podrozdziale. Tablica 6.7 zawiera porównanie wartości $AP@m$. Na Rysunku 6.1 są przedstawione wyniki w formie graficznej.

Maksymalna wartość parametru k wynosi 8, ponieważ zdjęcia testowe muszą posiadać co najmniej 10 hasztagów, z czego w metodzie grafów 2 hasztagi są losowo pobierane z tego obrazka do dalszej analizy (Podrozdział 4.1.1).

		Lokalizacja (skrót)										$AP@k_d$
		PK	TB	WE	KND	WN	SF	TS	GCT	DD	PZ	
Parametr k	1	0,53	0,63	0,7	0,65	0,45	0,56	0,33	0,55	0,72	0,42	0,55
	2	0,38	0,46	0,49	0,47	0,31	0,43	0,26	0,37	0,48	0,32	0,4
	3	0,31	0,36	0,4	0,39	0,27	0,37	0,24	0,32	0,38	0,28	0,33
	4	0,27	0,31	0,35	0,33	0,23	0,33	0,22	0,29	0,32	0,25	0,29
	5	0,24	0,28	0,31	0,3	0,21	0,3	0,21	0,26	0,28	0,22	0,26
	6	0,21	0,25	0,29	0,27	0,2	0,27	0,2	0,24	0,26	0,2	0,24
	7	0,2	0,23	0,26	0,25	0,19	0,25	0,19	0,22	0,24	0,19	0,22
	8	0,18	0,22	0,25	0,23	0,18	0,23	0,18	0,21	0,22	0,18	0,21
$AP@k_m$		0,29	0,34	0,38	0,36	0,26	0,34	0,23	0,31	0,36	0,26	

Tablica 6.1: Wyniki dla metody grafów. Parametr $d = 100$.

		Lokalizacja (skrót)										$AP@k_d$
		PK	TB	WE	KND	WN	SF	TS	GCT	DD	PZ	
Parametr k	1	0,5	0,65	0,73	0,65	0,46	0,65	0,44	0,6	0,78	0,35	0,58
	2	0,35	0,45	0,47	0,45	0,32	0,47	0,33	0,41	0,52	0,27	0,4
	3	0,31	0,36	0,38	0,38	0,28	0,37	0,29	0,36	0,41	0,23	0,34
	4	0,27	0,31	0,33	0,32	0,24	0,33	0,26	0,32	0,35	0,21	0,29
	5	0,24	0,28	0,3	0,29	0,21	0,29	0,24	0,29	0,31	0,2	0,26
	6	0,21	0,25	0,27	0,26	0,2	0,27	0,23	0,27	0,29	0,19	0,24
	7	0,2	0,24	0,26	0,24	0,18	0,25	0,22	0,25	0,26	0,17	0,23
	8	0,18	0,22	0,24	0,23	0,17	0,24	0,21	0,23	0,24	0,17	0,21
$AP@k_m$		0,28	0,34	0,37	0,35	0,26	0,36	0,28	0,34	0,4	0,22	

Tablica 6.2: Wyniki dla metody grafów. Parametr $d = 50$.

		Lokalizacja (skrót)										$AP@k_d$
		PK	TB	WE	KND	WN	SF	TS	GCT	DD	PZ	
Parametr k	1	0,49	0,65	0,69	0,68	0,47	0,65	0,49	0,54	0,78	0,4	0,58
	2	0,36	0,45	0,46	0,52	0,29	0,46	0,35	0,4	0,55	0,29	0,42
	3	0,3	0,36	0,37	0,44	0,26	0,4	0,31	0,34	0,45	0,26	0,35
	4	0,26	0,31	0,33	0,38	0,22	0,34	0,28	0,31	0,38	0,23	0,3
	5	0,24	0,28	0,29	0,33	0,2	0,31	0,25	0,28	0,34	0,21	0,27
	6	0,22	0,25	0,27	0,3	0,19	0,29	0,24	0,26	0,31	0,19	0,25
	7	0,2	0,24	0,25	0,28	0,17	0,26	0,23	0,25	0,29	0,18	0,24
	8	0,18	0,22	0,24	0,26	0,17	0,25	0,23	0,23	0,27	0,17	0,22
$AP@k_m$		0,28	0,34	0,36	0,4	0,25	0,37	0,3	0,33	0,42	0,24	

Tablica 6.3: Wyniki dla metody grafów. Parametr $d = 25$.

		Lokalizacja (skrót)									$AP@k_d$	
		PK	TB	WE	KND	WN	SF	TS	GCT	DD		PZ
Parametr k	1	0	0	0	0,75	0,47	0	0,27	0	0,7	0,27	0,25
	2	0	0	0	0,55	0,29	0	0,28	0	0,52	0,25	0,19
	3	0	0	0	0,44	0,26	0	0,24	0	0,44	0,23	0,16
	4	0	0	0	0,38	0,22	0	0,23	0	0,37	0,2	0,14
	5	0	0	0	0,34	0,2	0	0,24	0	0,33	0,19	0,13
	6	0	0	0	0,31	0,19	0	0,23	0	0,3	0,18	0,12
	7	0	0	0	0,28	0,17	0	0,22	0	0,28	0,17	0,11
	8	0	0	0	0,26	0,17	0	0,21	0	0,26	0,16	0,11
$AP@k_m$		0	0	0	0,41	0,25	0	0,24	0	0,4	0,21	

Tablica 6.4: Wyniki dla metody grafów. Parametr $d = 10$.

		Lokalizacja (skrót)					$AP@k_d$
		KND	WN	TS	DD	PZ	
Parametr k	1	0,75	0,47	0,27	0,7	0,27	0,49
	2	0,55	0,29	0,28	0,52	0,25	0,38
	3	0,44	0,26	0,24	0,44	0,23	0,32
	4	0,38	0,22	0,23	0,37	0,2	0,28
	5	0,34	0,2	0,24	0,33	0,19	0,26
	6	0,31	0,19	0,23	0,3	0,18	0,24
	7	0,28	0,17	0,22	0,28	0,17	0,23
	8	0,26	0,17	0,21	0,26	0,16	0,21
$AP@k_m$		0,41	0,25	0,24	0,4	0,21	

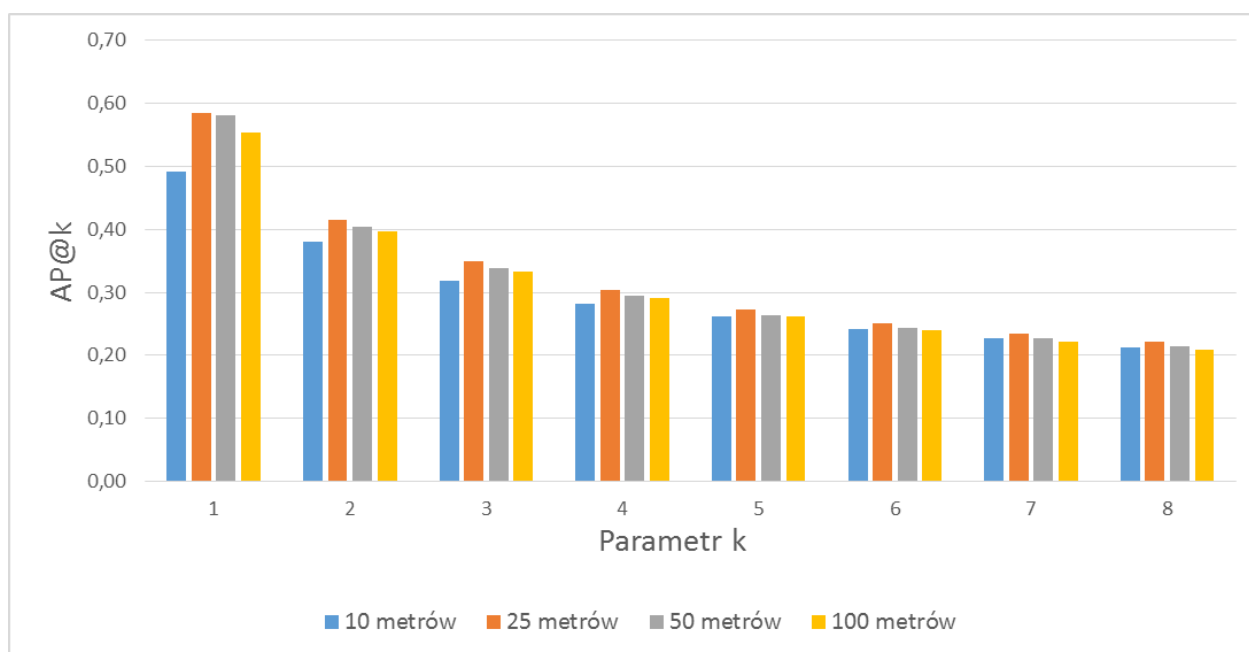
Tablica 6.5: Wyniki dla metody grafów. Parametr $d = 10$. Bez wartości 0.

		Parametr distance				$AP@k$
		10	25	50	100	
Parametr k	1	0,49	0,58	0,58	0,55	0,55
	2	0,38	0,42	0,40	0,40	0,40
	3	0,32	0,35	0,34	0,33	0,33
	4	0,28	0,30	0,29	0,29	0,29
	5	0,26	0,27	0,26	0,26	0,27
	6	0,24	0,25	0,24	0,24	0,24
	7	0,23	0,24	0,23	0,22	0,23
	8	0,21	0,22	0,21	0,21	0,21
$AP@d$		0,30	0,33	0,32	0,31	

Tablica 6.6: Wyniki dla metody grafów.

		Lokalizacja (skrót)									
		PK	TB	WE	KND	WN	SF	TS	GCT	DD	PZ
distance	10	0,00	0,00	0,00	0,41	0,25	0,00	0,24	0,00	0,40	0,21
	25	0,28	0,34	0,36	0,40	0,25	0,37	0,30	0,33	0,42	0,24
	50	0,28	0,34	0,37	0,35	0,26	0,36	0,28	0,34	0,40	0,22
	100	0,29	0,34	0,38	0,36	0,26	0,34	0,23	0,31	0,36	0,26
AP@m		0,21	0,26	0,28	0,38	0,25	0,27	0,26	0,24	0,4	0,23

Tablica 6.7: Wyniki dla metody grafów. Porównanie miejsc.



Rysunek 6.1: Wyniki eksperymentu przy użyciu metody grafów. Opracowanie własne.

6.1.2 Wnioski dla metody grafów

Metoda ta zwraca szybko wyniki, jeżeli jest stworzony wcześniej graf na podstawie którego klasyfikowane jest zdjęcie.

Z eksperymentu wynika, że najlepszą wartością dla parametru **distance** jest 25 metrów (Tablica 6.6). Najgorszą natomiast 10 metrów, co przeczy intuicji, że im bliżej atrakcji turystycznej tym powinny być lepsze wyniki. $AP@k_d$ dla parametru $k = 1$ wyraźnie odstaje porównując z innymi parametrami d . Wartości $AP@k$, w momencie kiedy **distance** przyjmuje wartości 25 metrów, 50 metrów i 100 metrów są bardzo podobne. Jednakże czas obliczeń w przypadku 50 metrów i 100 metrów jest dłuższy, ponieważ jest więcej zdjęć z hashtagami, które muszą być przetworzone w celu stworzenia grafu.

Wartości $AP@k_d$ spadają wraz ze wzrostem parametru d co jest zgodne z intuicją (Tablica 6.6).

Najlepszy współczynnik $AP@m$ ma Disneyland Drive oraz Katedra Norte Dame (Tablica 6.7). Tylko Sagrada Familia ma trend, że wraz ze wzrostem parametru d wartość $AP@k_m$ maleje (wyłączając $d = 10$, ponieważ za mało danych było w eksperymencie dla tego miejsca). Najgorsze wartości $AP@m$ mają Pałac Kultury i Nauki oraz Plac Zocalo, które są głównymi miejscami turystycznymi w stolicach Polski i Meksyku.

Wyniki jakie zostały uzyskane w Tablicy 6.7 i przeprowadzona analiza danych daje pewien pogląd dlaczego najgorszy wynik uzyskał Pałac Kultury i Nauki oraz Plac Zocalo. Hasztagi, które współwystępują ze sobą i mają największą siłę połączenia znaczeniowo są takie same. Przykład pary hasztagów, które współwystępują ze sobą: {poland, warsaw}; {poland, warszawa}; {polska; warszawa}. W przypadku Placu Zocalu sytuacja jest taka sama. Klasyfikacja zdjęcia w takiej sytuacji jest trudna.

Najlepsze rezultaty uzyskał Disneyland Drive. Jest to miejsce w którym, aby zrobić zdjęcie trzeba było zapłacić za wstęp. Polski akcent uzyskał najgorsze rezultaty. Wodospad Niagara natomiast nie wyróżniała się wśród innych miejsc pod względem miary $AP@m$.

6.2 Metoda oparta o algorytm SIFT

W podrozdziale 6.2.1 są przedstawione tablice z wynikami dla metody opartej o algorytm SIFT, a w podrozdziale 6.2.2 są wyciągnięte wnioski dla tej metody.

6.2.1 Wyniki dla metody opartej o algorytm SIFT

W Tablicach 6.8, 6.9, 6.10, 6.11 przedstawione są wyniki $AP@k_{m,d}$ dla różnych kategorii. W Tablicy 6.11 widać, że wartości $AP@k_{m,d}$ dla miejsc PK, TB, WE, SF, GCT przyjmują wartości 0. Jest to spowodowane tym, że nie zostały osiągnięte warunki początkowe opisane w Podrozdziale 5.3. Z tego względu została nowa Tablica 6.12 nie zawiera miejsc, które nie spełniają wymagań początkowych. W Tablicy 6.13 są skondensowane wyniki dla wszystkich wcześniejszych tabel w tym podrozdziale. Tablica 6.14 zawiera porównanie wartości $AP@m$. Na Rysunku 6.2 są przedstawione wyniki w formie graficznej.

		Lokalizacja (skrót)										$AP@k_d$
		10	25	50	100	WN	SF	TS	GCT	DD	PZ	
Parametr k	1	0,65	0,76	0,88	0,84	0,68	0,7	0,46	0,67	0,91	0,49	0,7
	2	0,52	0,64	0,65	0,7	0,5	0,59	0,39	0,49	0,76	0,38	0,56
	3	0,48	0,51	0,6	0,64	0,47	0,55	0,37	0,48	0,54	0,36	0,5
	4	0,39	0,45	0,54	0,52	0,42	0,5	0,31	0,43	0,45	0,32	0,43
	5	0,34	0,4	0,51	0,45	0,36	0,44	0,29	0,39	0,41	0,28	0,38
	6	0,31	0,37	0,46	0,4	0,32	0,39	0,27	0,36	0,38	0,25	0,35
	7	0,28	0,34	0,43	0,37	0,3	0,35	0,25	0,33	0,36	0,24	0,33
	8	0,26	0,32	0,4	0,34	0,28	0,33	0,24	0,31	0,34	0,22	0,3
	9	0,24	0,3	0,38	0,32	0,28	0,3	0,22	0,29	0,32	0,21	0,28
	10	0,22	0,28	0,36	0,3	0,28	0,28	0,22	0,27	0,29	0,2	0,27
$AP@k_m$		0,37	0,44	0,52	0,49	0,39	0,44	0,3	0,4	0,47	0,29	

Tablica 6.8: Wyniki dla metody opartej o algorytm SIFT. Parametr $d = 100$.

		Lokalizacja (skrót)										$AP@k_d$
		10	25	50	100	WN	SF	TS	GCT	DD	PZ	
Parametr k	1	0,67	0,76	0,86	0,8	0,77	0,7	0,53	0,78	0,9	0,49	0,73
	2	0,55	0,65	0,63	0,64	0,52	0,6	0,45	0,57	0,8	0,4	0,58
	3	0,5	0,52	0,58	0,59	0,46	0,54	0,45	0,56	0,64	0,34	0,52
	4	0,41	0,44	0,54	0,48	0,42	0,48	0,38	0,49	0,56	0,32	0,45
	5	0,36	0,39	0,5	0,41	0,38	0,42	0,35	0,42	0,49	0,28	0,4
	6	0,32	0,37	0,45	0,39	0,33	0,38	0,33	0,39	0,44	0,25	0,36
	7	0,29	0,34	0,42	0,36	0,31	0,34	0,29	0,37	0,41	0,22	0,34
	8	0,27	0,32	0,39	0,33	0,29	0,32	0,28	0,35	0,4	0,19	0,31
	9	0,25	0,3	0,37	0,31	0,28	0,29	0,26	0,32	0,38	0,18	0,29
	10	0,23	0,28	0,35	0,29	0,28	0,28	0,26	0,3	0,36	0,18	0,28
$AP@k_m$		0,38	0,44	0,51	0,46	0,4	0,44	0,36	0,46	0,54	0,29	

Tablica 6.9: Wyniki dla metody opartej o algorytm SIFT. Parametr $d = 50$.

		Lokalizacja (skrót)										$AP@k_d$
		10	25	50	100	WN	SF	TS	GCT	DD	PZ	
Parametr k	1	0,68	0,76	0,83	0,87	0,67	0,75	0,53	0,73	0,94	0,51	0,73
	2	0,55	0,66	0,59	0,73	0,43	0,63	0,4	0,54	0,81	0,42	0,58
	3	0,5	0,53	0,57	0,66	0,4	0,59	0,44	0,53	0,71	0,38	0,53
	4	0,41	0,47	0,52	0,53	0,36	0,53	0,37	0,45	0,6	0,34	0,46
	5	0,36	0,42	0,48	0,45	0,34	0,45	0,34	0,41	0,53	0,3	0,41
	6	0,32	0,38	0,43	0,42	0,33	0,42	0,32	0,39	0,48	0,26	0,37
	7	0,3	0,35	0,41	0,38	0,3	0,38	0,31	0,37	0,44	0,22	0,35
	8	0,27	0,32	0,38	0,35	0,28	0,35	0,3	0,35	0,42	0,21	0,32
	9	0,25	0,31	0,36	0,33	0,26	0,33	0,29	0,31	0,4	0,19	0,3
	10	0,23	0,29	0,34	0,32	0,25	0,3	0,27	0,3	0,38	0,18	0,29
$AP@k_m$		0,39	0,45	0,49	0,5	0,36	0,47	0,36	0,44	0,57	0,3	

Tablica 6.10: Wyniki dla metody opartej o algorytm SIFT. Parametr $d = 25$.

		Lokalizacja (skrót)										$AP@k_d$
		10	25	50	100	WN	SF	TS	GCT	DD	PZ	
Parametr k	1	0	0	0	0,9	0,78	0	0,59	0	0,93	0,57	0,38
	2	0	0	0	0,77	0,53	0	0,45	0	0,8	0,47	0,3
	3	0	0	0	0,69	0,48	0	0,5	0	0,65	0,42	0,27
	4	0	0	0	0,55	0,44	0	0,46	0	0,61	0,36	0,24
	5	0	0	0	0,47	0,39	0	0,42	0	0,52	0,31	0,21
	6	0	0	0	0,42	0,38	0	0,39	0	0,47	0,29	0,2
	7	0	0	0	0,39	0,34	0	0,36	0	0,43	0,26	0,18
	8	0	0	0	0,36	0,31	0	0,34	0	0,41	0,24	0,17
	9	0	0	0	0,34	0,28	0	0,33	0	0,4	0,22	0,16
	10	0	0	0	0,33	0,28	0	0,31	0	0,38	0,21	0,15
$AP@k_m$		0	0	0	0,52	0,42	0	0,42	0	0,56	0,33	

Tablica 6.11: Wyniki dla metody opartej o algorytm SIFT. Parametr $d = 10$.

		Lokalizacja (skrót)					$AP@k_d$
		KND	WN	TS	DD	PZ	
Parametr k	1	0,9	0,78	0,59	0,93	0,57	0,76
	2	0,77	0,53	0,45	0,8	0,47	0,6
	3	0,69	0,48	0,5	0,65	0,42	0,55
	4	0,55	0,44	0,46	0,61	0,36	0,48
	5	0,47	0,39	0,42	0,52	0,31	0,42
	6	0,42	0,38	0,39	0,47	0,29	0,39
	7	0,39	0,34	0,36	0,43	0,26	0,36
	8	0,36	0,31	0,34	0,41	0,24	0,33
	9	0,34	0,28	0,33	0,4	0,22	0,31
	10	0,33	0,28	0,31	0,38	0,21	0,30
$AP@k_m$		0,52	0,42	0,42	0,56	0,33	

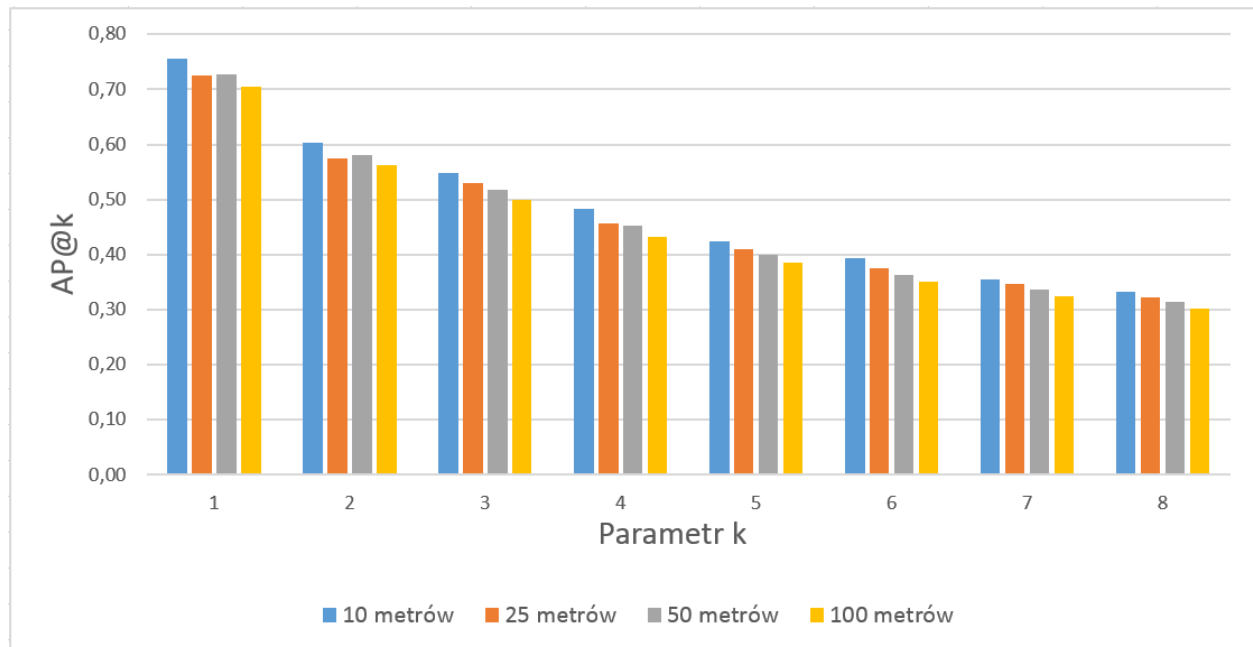
Tablica 6.12: Wyniki dla metody opartej o algorytm SIFT. Parametr $d = 10$. Bez wartości 0.

		Parametr distance				AP@k
		10	25	50	100	
Parametr k	1	0,76	0,73	0,73	0,70	0,73
	2	0,60	0,58	0,58	0,56	0,58
	3	0,55	0,53	0,52	0,50	0,52
	4	0,48	0,46	0,45	0,43	0,46
	5	0,42	0,41	0,40	0,38	0,40
	6	0,39	0,37	0,36	0,35	0,37
	7	0,36	0,35	0,34	0,33	0,34
	8	0,33	0,32	0,31	0,30	0,32
	9	0,31	0,30	0,29	0,28	0,30
	10	0,30	0,29	0,28	0,27	0,28
AP@d		0,45	0,43	0,43	0,41	

Tablica 6.13: Wyniki dla metody SIFT.

		Lokalizacja (skrót)									
		PK	TB	WE	KND	WN	SF	TS	GCT	DD	PZ
distance	10	0,00	0,00	0,00	0,52	0,42	0,00	0,42	0,00	0,56	0,33
	25	0,39	0,45	0,49	0,50	0,36	0,47	0,36	0,44	0,57	0,30
	50	0,38	0,44	0,51	0,46	0,40	0,44	0,36	0,46	0,54	0,29
	100	0,37	0,44	0,52	0,49	0,39	0,44	0,30	0,40	0,47	0,26
AP@m		0,29	0,33	0,38	0,49	0,39	0,34	0,36	0,32	0,54	0,3

Tablica 6.14: Wyniki dla metody opartej o algorytm SIFT. Porównanie miejsc.



Rysunek 6.2: Wyniki eksperymentu przy użyciu metody opartej o algorytm SIFT. Opracowanie własne.

6.2.2 Wnioski dla metody opartej o algorytm SIFT

Metoda ta jest bardzo czasochłonna obliczeniowo. Złożoność obliczeń dla jednej kategorii danych wynosi: $O(n^K)$, gdzie n to liczba wszystkich zdjęć testowych, a K to liczba wszystkich zdjęć treningowych.

Analizując wyniki z Tablicy 6.14 wyraźnie widać, że najgorsze wyniki są dla Pałacu Kultury i Nauki w Warszawie oraz Placu Zocalo. W przypadku PK można powiedzieć, że wynik jest zaniżony ponieważ dla $d = 10$ jest wartość 0. Natomiast Plac Zocalo ma uzupełnione wszystkie wartości $AP@k_m$ i ma niski współczynnik $AP@m$. Może to być spowodowane tym, że na zdjęciach wykonanych w tym miejscu jest bardzo dużo szumu. Jest mało punktów charakterystycznych na tych obrazkach. Najlepsze wyniki ma Disneyland. Szerokość i długość geograficzna Disneylandu wskazuje na zamek księżniczki. Jest on bardzo charakterystyczny pod względem wizualnym i w najbliższej okolicy nie ma bardziej innej tak bardzo wyróżniającej się atrakcji turystycznej. Wodospad Niagara natomiast nie wyróżniała się wśród innych miejsc pod względem $AP@m$.

Rysunek 6.2 pokazuje, że najlepsze wyniki $AP@d$ są zwracane dla parametru $d = 10$. Różnica pomiędzy innymi wartościami parametru d jest niewielka, jednak czas obliczeń dla $d = 10$ jest znacznie mniejszy, ponieważ API Instagrama zwraca mniej wyników dla mniejszego parametru d .

Z Tablicy 6.13 można odczytać, że wraz ze wzrostem parametru k wartość $AP@k$ maleje, co jest zgodne z intuicją.

Czytając wyniki z Tablicy 6.14 można zauważyć trend, że wraz ze wzrostem parametru d to $AP@k_m$ maleje.

6.3 Porównanie wyników

Porównując dwie zaproponowane metody w tej pracy można stwierdzić, że metoda oparta o algorytm SIFT zwraca w każdym aspekcie lepsze wyniki od metody opartej o budowanie grafów, jednak czas zwracania wyników jest nieporównywalnie dłuższy.

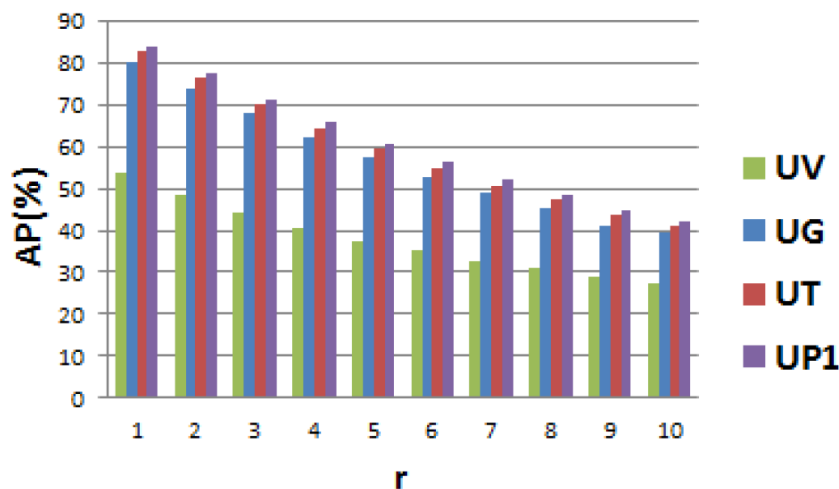
W artykule [5] autorzy porównali 16 różnych algorytmów do rekomendacji hashtagów. W artykule nie jest dokładnie opisane jak zostały pozyskane dane do badania. W celu ewaluacji wyniku algorytm zwracał 10 hashtagów ($TopN$), które najbardziej pasują do zdjęcia I . Następnie porównywane są one z oryginalnymi hashtagami ze zdjęcia I . $Precision = \frac{TopN \cap I}{TopN}$. Najlepszy wynik jaki został osiągnięty to $precision = 66.95\%$. Kolejny algorytm osiągnął $precision = 25.41\%$. Następne algorytmy miały wskaźnik $precision$ poniżej 17%.

W artykule [9] autorzy używali bazy danych pochodzącą z [19]. Zestaw danych składał się z zdjęć wyszukanych przy pomocy 52 różnych hashtagów. Problemem, który postawili przed sobą autorzy było prawidłowe sklasyfikowanie zdjęć do jednego z 52 hashtagów. Metryka którą użyli, aby ocenić skuteczność algorytmów nazywa się *Mean Average Precision (MAP)*. Pierwszy algorytm osiągnął $MAP = 0.576$, drugi $MAP = 0.593$.

W artykule [12] autorzy sami stworzyli bazę danych, która zawiera 16,689 obrazków, należących do 925 grup (kategorii zdjęć). Średnia ilość hashtagów na zdjęcie to 6. Autorzy postawili sobie za cel, aby prawidłowo dopasować zdjęcie do odpowiedniej grupy. Wynik jaki osiągnęli to $precision = 0.62$.

W artykule [18] autorzy stworzyli bazę danych z 331 obrazkami z hashtagami. Obrazki pochodziły z kilku kategorii. 131 zdjęć było treningowych, a reszta testowa. Algorytm zwracał 10 rekomendowanych hashtagów do zdjęcia. Następnie poproszono ludzi, aby ocenili czy algorytm dobrze rekomenduje hashtagi do zdjęcia testowego. Ich oceny tworzyły *ground truth*. Najlepsze wyniki jakie osiągnął algorytm to $P@1 = 0.6750$ oraz $P@5 = 0.5420$.

W artykule [8] baza danych zawierała ponad 1.9 mln zdjęć z hashtagami. Metoda ewaluacji jest zbieżna do tej jaka jest zaprezentowana w tej pracy magisterskiej. Algorytm zwraca k rekomendowanych hashtagów, a potem są one porównywane z oryginalnymi hashtagami należącymi do obrazka. Wyniki znajdują się na Rysunku 6.3.



Rysunek 6.3: Wyniki eksperymentu z artykułu [8]. Źródło: [8].

W artykule [7] autorzy stworzyli bazę danych z 3000 zdjęciami z ich hasztagami. Metoda ewaluacji jest taka sama jak w tej pracy magisterskiej. Wynik, który został osiągnięty w tym artykule to $P@5 \approx 25\%$.

Porównanie wyników w formie tabeli znajduje się w Tablicy 6.15. W tablicy są umieszczone tylko najlepsze wyniki z artykułów lub te dla których parametr k był równy 5. Wszystkie artykuły odnoszą się do serwisu Flickr.

Artykuł / Metoda	Rezultat	Uwagi
Metoda oparta o algorytm SIFT	$AP@5 = 0.42$	BRAK
Metoda grafów	$AP@5 = 0.27$	BRAK
[5]	$precision = 0.67$	Brak informacji w artykule w jaki sposób zostały pozyskiwane dane od badania
[9]	$MAP = 0.59$	Wąski zakres dostępnych hasztagów, które mają opisać zdjęcie
[12]	$precision = 0.62$	Tworzone były profile użytkowników
[18]	$P@5 = 0.54$	Ludzie sprawdzali czy algorytm trafie rekomenduje hasztagi
[8]	$AP@5 = 0.60$	Tworzone były profile użytkowników
[7]	$P@5 \approx 0.25$	Mała baza danych

Tablica 6.15: Porównanie wyników z różnymi artykułami.

Jak widać w Tablicy 6.15 wyniki dla artykułów naukowych są prawie zawsze lepsze. Tylko artykuł [7] miał gorsze wyniki od tych, które zaprezentowane są w tej pracy. Każda z prac naukowych inaczej podchodziła do tematu klasyfikacji zdjęć. Użytkownikami serwisu Flickr są głównie osoby, które profesjonalnie zajmują się robieniem zdjęć. Serwis Flickr posiada kategorie do których można przypisywać zdjęcia. Grupą docelową serwisu Instagram są głównie osoby młode. Różnice w strukturze oraz grupie docelowej tych portali ma znaczący wpływ na jakość hasztagowania zdjęć, a to implikuje lepsze rezultaty klasyfikacji zdjęć.

Rozdział 7

Podsumowanie

Klasyfikacja mediów jest popularnym tematem naukowym. W dniu 2015-06-11 Google Scholar zwracał 4140 wyników dla okresu od 2015 roku do zapytania „tag recommendation”. Świadczy to o tym, że jest duże zapotrzebowanie na stworzenie algorytmu, który trafnie będzie potrafił rekomendować odpowiednie adnotacje do zdjęć na portalach społecznościowych. Kiedy zostaną już stworzone skuteczne systemy rekomendacji hashtagów w mediach społecznościowych, będzie możliwość automatycznego indeksowania danych w portalach takich jak na przykład Instagram.

Hashtagowanie zdjęć to bardzo popularna metoda indeksowania zdjęć. W tej pracy zostały zaprezentowane dwie metody klasyfikacji zdjęć do odpowiednich hashtagów. Zagadnienie hashtagowania mediów jest trudnym tematem, ponieważ ludzie różnie odbierają rzeczywistość i często różnie opisują ten sam obiekt. Kolejnym problem z opisywaniem rzeczywistości jest język. W Polsce ludzie używają głównie hashtagów w języku polskim. Systemy mogą mieć problemy z przetworzeniem różnych języków, na przykład ten sam obiekt może być opisany angielskim słowem „summer” lub polskim „lato”. Przez taki sposób nadawania hashtagów, algorytmy zwracają gorsze rezultaty. Najczęściej systemy tworzone są z myślą o języku angielskim.

W Internecie można znaleźć wiele różnych wskazówek jak stosować hashtagi, między innymi: używaj jeden lub dwa hashtagi w opisie, używaj angielskich liter w hashtagach (zamiast „#PiłkaNożna” lepiej napisać „#PilkaNozna”), hashtagi powinny być krótkie, używaj popularnych hashtagów, hashtag nie powinien się składać z więcej niż trzech połączonych wyrazów (na przykład „#NiechWygraFutbol”). W momencie kiedy użytkownicy portali społecznościowych zaczną stosować się do wymienionych wcześniej reguł hashtagowania będą powstać metody, które będą bardzo skutecznie rekomendować hashtagi.

Dzięki algorytmom, które z dobrym rezultatem potrafią klasyfikować zdjęcia można zarabiać pieniądze. Strona <http://imagg.com/> pokazuje, że cena za próbę odpowiedniego opisanie pięciu zdjęć na sekundę może wynieść nawet 349 dolarów za miesiąc użytkowania z takiej usługi. Jednak strona ta nie była sprawdzona pod względem jak trafnie potrafi hasztaogwać zdjęcia na serwisach społecznościowych.

Porównując serwis Flickr oraz Instagram można stwierdzić, że na Instagramie użytkownicy luźniej podchodzą do tematu hashtagowania mediów. Często używają hashtagów, które nie opisują rzeczywistości (tak zwane: *social tags*), na przykład: „l4l”, „f4f”. Spowodowane może być to tym, że Instagram jest skierowany do wszystkich osób zainteresowanych dzieleniem

się zdjęciami ze swojego życia. Nikt nie nakładam tam ludziom ograniczeń pod względem dodawania adnotacji do zdjęć.

Praca ta pokazała, że użytkownicy Instagrama nie używają podobnych hashtagów w celu opisania mediów pomimo tego, że miejsca skąd były wykonane zdjęcia znajdowały się w bliskiej odległości siebie i wizualnie reprezentowały to samo. Każdy człowiek indywidualnie odbiera rzeczywistość i na swój sposób dobiera słowa, które by ten świat opisywały. Analizując artykuły naukowe z rozdziału 2 można zaproponować stworzenie systemu opierającego się na indywidualnej aktywności użytkownika w portalu społecznościowym Instagram. Prace naukowe dotyczące Flickr pokazały lepsze klasyfikowanie zdjęć w momencie, kiedy były tworzone „profile użytkowników”. Hashtagi, które są przypisywane do zdjęć zależą również od użytkownika, który ma te grafiki opisać. W celu stworzenia skutecznej metody rekomendacji hashtagów na portalu Instagram można zaproponować algorytm, który będzie brać pod uwagę takie informacje jak: wcześniej używane hashtagi przez tego użytkownika; hashtagi używane przez znajomych danej osoby, pogrupować w kategorii używane hashtagi (na przykład stwierdzić, że użytkownik często używa *social tagów*).

Metoda grafów zaprezentowana w tej pracy w przyszłości mogłaby stanowić załączek do dobrego systemu rekomendowania hashtagów dla użytkowników Instagrama, ponieważ wyniki są zwracane relatywnie szybko. Dodatkowo można byłoby rozszerzyć algorytm o zbieranie informacji o używanych hashtagach przez osobę zarejestrowaną w serwisie i stworzenia „profilu użytkownika”.

W przypadku metody opartej o algorytm SIFT w celu uzyskania lepszych wyników można zastosować metody statystyczne dla ilość punktów wspólnych dla dwóch porównywanych obrazków, aby część obrazków nie brać po uwagę w trakcie rekomendowania hashtagów. Dodatkowo można stworzyć „profile użytkowników” jak było to opisane wcześniej.

Spis rysunków

1.1	Wykres popularności Instagrama. Źródło: [3]	5
3.1	Obliczanie DoG. Źródło: własne. Opracowane na podstawie: [16].	15
3.2	Wyznaczenie ekstremów. Źródło: własne. Opracowane na podstawie: [16].	16
3.3	Deskryptor punktu kluczowego. Źródło: własne. Opracowane na podstawie: [16].	17
4.1	Przykładowa wizualizacja wycinka grafu \mathbf{G} . Opracowanie własne.	20
4.2	Przykładowa wizualizacja \mathbf{G}_T . Opracowanie własne.	21
4.3	Wizualizacja przykładu dla metody działania algorytmu SIFT. Opracowanie własne.	22
6.1	Wyniki eksperymentu przy użyciu metody grafów. Opracowanie własne.	31
6.2	Wyniki eksperymentu przy użyciu metody opartej o algorytm SIFT. Opracowanie własne.	36
6.3	Wyniki eksperymentu z artykułu [8]. Źródło: [8].	37

Spis tablic

4.1	Przykładowa uszeregowana lista struktur.	23
5.1	Wybrane miejsca, który były brane pod uwagę w trakcie eksperymentu.	25
5.2	Skrócone nazwy miejsc.	25
6.1	Wyniki dla metody grafów. Parametr $d = 100$	29
6.2	Wyniki dla metody grafów. Parametr $d = 50$	29
6.3	Wyniki dla metody grafów. Parametr $d = 25$	29
6.4	Wyniki dla metody grafów. Parametr $d = 10$	30
6.5	Wyniki dla metody grafów. Parametr $d = 10$. Bez wartości 0.	30
6.6	Wyniki dla metody grafów.	30

6.7	Wyniki dla metody grafów. Porównanie miejsc.	31
6.8	Wyniki dla metody opartej o algorytm SIFT. Parametr $d = 100$	33
6.9	Wyniki dla metody opartej o algorytm SIFT. Parametr $d = 50$	33
6.10	Wyniki dla metody opartej o algorytm SIFT. Parametr $d = 25$	34
6.11	Wyniki dla metody opartej o algorytm SIFT. Parametr $d = 10$	34
6.12	Wyniki dla metody opartej o algorytm SIFT. Parametr $d = 10$. Bez wartości 0.	35
6.13	Wyniki dla metody SIFT.	35
6.14	Wyniki dla metody opartej o algorytm SIFT. Porównanie miejsc.	35
6.15	Porównanie wyników z różnymi artykułami.	38

Bibliografia

- [1] Dawid Długosz. 43% polaków korzysta z serwisów społecznościowych. <http://www.komputerswiat.pl/artykuly/redakcyjne/2010/10/43-procent-polakow-korzysta-z-serwisow-spolesznosciowych.aspx>. Dostępność: 2015-05-26.
- [2] Polska Agencja Prasowa. Blisko 90 proc. internautów korzysta z serwisów społecznościowych. <http://naukawpolsce.pap.pl/aktualnosci/news,394820,blisko-90-proc-internautow-korzysta-z-serwisow-spolesznosciowych.html>. Dostępność: 2015-05-26.
- [3] TW. 2 mln polskich użytkowników instagramu. przeważają kobiety, młodzi i ludzie z miast. <http://www.wirtualnemedi.pl/artykul/2-mln-polskich-uzytkow-nikow-instagramu-przewazaja-kobiety-mlodzi-i-ludziez-miast>. Dostępność: 2015-05-26.
- [4] Instagram. Strona dla prasy. <https://instagram.com/press/>. Dostępność: 2015-05-26.
- [5] M.A. Chatti, S. Dakova, H. Thus, and U. Schroeder. Tag-based collaborative filtering recommendation in personal learning environments. *Learning Technologies, IEEE Transactions on*, 6(4):337–349, Oct 2013.
- [6] Hyun-Ki Hong, Kyung-Wook Park, and Dong-Ho Lee. Tag recommendation system for multimedia retrieval in mobile environment. In *Consumer Electronics (ISCE 2014), The 18th IEEE International Symposium on*, pages 1–2, June 2014.
- [7] Adam Rae, Börkur Sigurbjörnsson, and Roelof van Zwol. Improving tag recommendation using social networks. In *Adaptivity, Personalization and Fusion of Heterogeneous Information*, RIAO '10, pages 92–99, Paris, France, France, 2010.
- [8] Xiaoxiao Liu, Xueming Qian, Dan Lu, Xingsong Hou, and Liejun Wang. Personalized tag recommendation for flickr users. In *Multimedia and Expo (ICME), 2014 IEEE International Conference on*, pages 1–6, July 2014.
- [9] Yue Gao, Meng Wang, Huanbo Luan, Jialie Shen, Shuicheng Yan, and Dacheng Tao. Tag-based social image search with visual-text joint hypergraph learning. In *Proceedings of the 19th ACM International Conference on Multimedia*, MM '11, pages 1517–1520, New York, NY, USA, 2011. ACM.

- [10] N.S.T. Sai and R.C. Patil. New feature vector for image retrieval: Sum of value of histogram bins. In *Advances in Computing, Control, Telecommunication Technologies, 2009. ACT '09. International Conference on*, pages 550–554, Dec 2009.
- [11] Nina Runge, Dirk Wenig, and Rainer Malaka. Keep an eye on your photos: Automatic image tagging on mobile devices. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services, MobileHCI '14*, pages 513–518, New York, NY, USA, 2014. ACM.
- [12] M. De Choudhury, H. Sundaram, Yu-Ru Lin, A. John, and D.D. Seligmann. Connecting content to community in social media via image content, user tags and user communication. In *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*, pages 1238–1241, June 2009.
- [13] D.G. Lowe. Object recognition from local scale-invariant features. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1150–1157 vol.2, 1999.
- [14] Marcin Kielczewski. Rozpoznawanie wzorców za pomocą cech lokalnych. <http://etacar.put.poznan.pl/marcin.kielczewski/P0iSW11.pdf>. Dostępność: 2015-05-26.
- [15] Paweł Fronczak. Łączenie obrazów barwnych z dopasowaniem. <http://home.elka.pw.edu.pl/~pfroncza/Praca.pdf>. Dostępność: 2015-05-26.
- [16] Andrzej Wójtowicz. Opis obrazu i rozpoznawanie obiektów. http://andre.faculty.wmi.amu.edu.pl/pdf/opis_obrazu_rozp_obiektow.pdf. Dostępność: 2015-05-26.
- [17] Alexander Mordvintsev. Introduction to sift. http://opencv-python-tutroals.readthedocs.org/en/latest/py_tutorials/py_feature2d/py_sift_intro/py_sift_intro.html. Dostępność: 2015-05-26.
- [18] Börkur Sigurbjörnsson and Roelof van Zwol. Flickr tag recommendation based on collective knowledge. In *Proceedings of the 17th International Conference on World Wide Web, WWW '08*, pages 327–336, New York, NY, USA, 2008. ACM.
- [19] Meng Wang, Kuiyuan Yang, Xian-Sheng Hua, and Hong-Jiang Zhang. Towards a relevant and diverse search of social images. *Multimedia, IEEE Transactions on*, 12(8):829–842, Dec 2010.
- [20] Ming Cheung and James She. Bag-of-features tagging approach for a better recommendation with social big data. In *IMMM 2014, The Fourth International Conference on Advances in Information Mining and Management*, pages 83–88, 2014.
- [21] Piotr Grzegorski. Wyszukiwanie obrazów na podstawie zawartości. Master's thesis, Politechnika Warszawska, 2015.