

Projekt 4 - raport

Badanie skuteczności narzędzia trace.moe w wyszukiwaniu tytułów kreskówek anime na podstawie pojedynczych scen

Opis problemu

Wśród miłośników filmów i seriali nierzadko zdarzają się sytuacje, kiedy widząc pojedynczą scenę, mają oni trudność ze znalezieniem źródła kinematograficznego, z którego ona pochodzi. Z pomocą mogą przyjść wtedy różne narzędzia CBIR, które z mniejszą lub większą skutecznością pomogą osiągnąć zamierzony cel. Warto się jednak zastanowić, czy taką skuteczność da się podnieść, stosując narzędzia wyspecjalizowane w wyszukiwaniu konkretnych danych w bardzo specyficznych zasobach. Jako wieloletniemu entuzjaście japońskich kreskówek **anime** udało mi się znaleźć stronę internetową **trace.moe**, specjalizującą się w wyszukiwaniu tytułu takiego serialu na podstawie pojedynczej sceny przesłanej jako plik graficzny przez użytkownika. W niniejszym projekcie sprawdzę skuteczność tego narzędzia na pewnym zbiorze danych i spróbuję wyciągnąć wnioski, w jakich sytuacjach trace.moe spełnia stawiane przez użytkownika wymagania, a kiedy radzi sobie gorzej.

Cel projektu

Poniższe działania mają na celu zbadanie skuteczności narzędzia trace.moe w zależności od stopnia podobieństwa obrazu używanego do wyszukiwania i obrazu oryginalnego (fragmentu kreskówki).

Dodatkowo podjęta zostanie próba zdefiniowania, jaki typ danych wejściowych może potencjalnie spowodować spadek wydajności.

Konstrukcja doświadczenia

Zbiór trzydziestu obrazów, znajdujących się w katalogu „**Dane**”, przesłano do trace.moe, które zwracało tytuł serialu, który jego zdaniem pasuje do uwiecznionej w danej grafice sceny. Oprócz tego, narzędzie zwraca listę czterech innych scen, które również (choć w mniejszym stopniu) wydały mu się podobne. Co więcej, użytkownik jest także informowany o stopniu podobieństwa swojego zapytania do odpowiedzi, wyrażonym skali 0-100%, gdzie 100% oznacza identyczność. Dodatkowo, w celu weryfikacji tych wyników wykorzystano jeszcze jeden serwis, który pozwala uzyskać taką informację - **IMGonline.com.ua**.

Charakterystyka danych wejściowych

Badane obrazy odzwierciedlają różne sytuacje, w których użytkownik mógłby chcieć skorzystać z testowanego narzędzia:

- **I1 – I5**: grafiki możliwie bliskie oryginałowi, „czyste”, pozbawione napisów i innych modyfikacji
- **I6 – I10**: na obrazach znajdują się napisy, odzwierciedlające częste sytuacje, gdy użytkownik nie ogląda kreskówki z dubbingiem

- **I11 – I30**: obrazy zmodyfikowane, np. poprzez umieszczenie na nich napisów o nietypowym formacie, innych grafik oraz przekształceniu ich w memy lub podobne formy twórczości, ale w taki sposób, by człowiek, który wie skąd pochodzą, nadal nie miał problemu z ich identyfikacją. Można podejrzewać, że będą mieć negatywny wpływ na skuteczność narzędzia

Punktacja

Przy ocenie skuteczności narzędzia wykorzystałem własną wiedzę na temat tego, co rzeczywiście przedstawiają obrazy, a także przyjąłem następujące kryteria punktowe:

- 1 - trafne wskazanie tytułu anime, z którego pochodzi dana scena
- 0,5 – złe wskazanie, ale poprawna odpowiedź znajduje się na dalszym miejscu na liście
- 0 – złe wskazanie, a poprawnego nie ma nawet na dalszym miejscu na liście

Wyniki

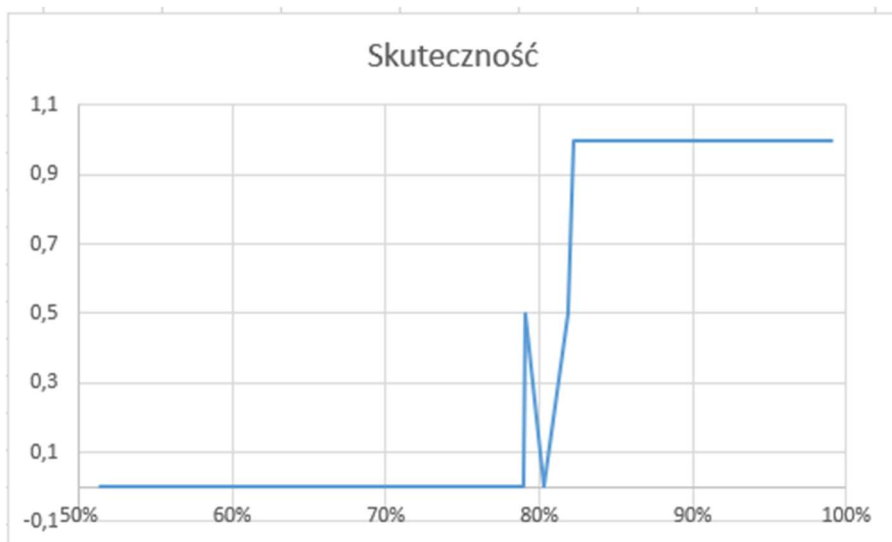
Wszystkie uzyskane rezultaty znajdują się w pliku „**Wyniki.xlsx**”.

W tabeli obok przedstawiono wyniki uzyskane przez wszystkie obrazy, posortowane malejąco według ich podobieństwa do oryginalnego fragmentu anime, z którego pochodzą.

Obserwacje

- Nietrudno zauważyć, że obrazy o wyższym podobieństwie do oryginału przekładają się na wyższą skuteczność algorytmu
- Jeśli spojrzeć, które to konkretnie obrazy, to zgodnie z przewidywaniami okaże się, że obrazy niezmodyfikowane lub zmodyfikowane nieznacznie (pierwsza dziesiątka w numeracji I) są łatwiej identyfikowane przez trace.moe
- Średnia ocena wynosi ok. 0,67, co znaczy, że mimo iż 20 z 30 scen zostało w zauważalny sposób zniekształconych, to narzędzie zachowuje skuteczność wyraźnie powyżej połowy

Poniższy wykres przedstawia zależność oceny uzyskiwanej przez obraz od stopnia podobieństwa

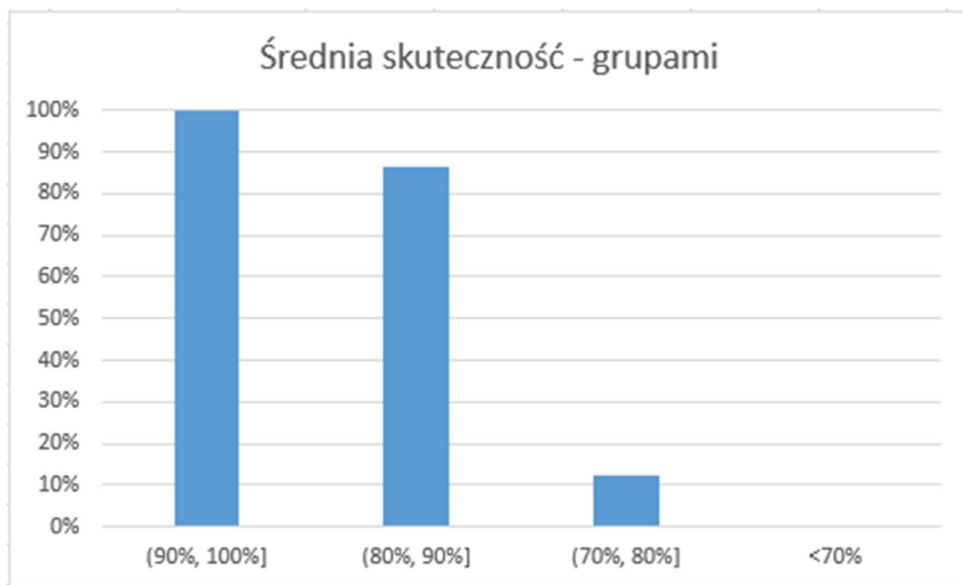


Obraz	Podobieństwo	Ocena
I2	99%	1
I4	98,00%	1
I6	98,00%	1
I11	96,59%	1
I1	96,00%	1
I3	95,35%	1
I8	93,16%	1
I23	93,07%	1
I16	92,03%	1
I7	91,60%	1
I14	89,28%	1
I5	88,94%	1
I17	88,92%	1
I10	88%	1
I26	87,09%	1
I20	85,93%	1
I30	85,71%	1
I15	84,99%	1
I21	82,22%	1
I24	81,91%	0,5
I13	80,31%	0
I9	79%	0,5
I19	78,92%	0
I12	72,29%	0
I25	70,13%	0
I28	63,65%	0
I27	61,88%	0
I18	57,03%	0
I29	56,24%	0
I22	51,38%	0
		Średnia
		0,666667

Obserwacje

- Powyżej 82,22% podobieństwa algorytm zachował 100% skuteczność
- W okolicach 80% podobieństwa następuje gwałtowny spadek wydajności. Występuje tam też jednak pewna nieregularność na wykresie, co sugeruje, że w tych okolicach skuteczność serwisu trace.moe może być trudna do przewidzenia
- Poniżej tej granicy znajdują się obrazy mocno zmodyfikowane – przede wszystkim przez naniesienie na nie innych grafik (np. przy tworzeniu z danych scen memów). Algorytm praktycznie nie generuje tu poprawnych wyników

Powyższe wnioski może pomóc zobrazować także kolejny wykres. Tutaj pogrupowano poszczególne wyniki ze względu na % podobieństwa. Ponownie, widać tu duży spadek wydajności w okolicach 80%.



Wnioski końcowe

Po przestudiowaniu powyższych wyników można spróbować odpowiedzieć na postawione wcześniej pytania.

- W przyjętej skali punktowej średni wynik dla badanego zbioru danych to ok. 0,67, co znaczy, że generalnie narzędzie trace.moe częściej niż nie zwraca poprawne wyniki.
- Dolną granicą podobieństwa obrazów, dla których można oczekiwać dobrych rezultatów zdaje się być ok. 80%.
- Można spróbować zidentyfikować, jakie konkretne cechy grafik sprawiają, że są one mniej podobne do oryginału:
 - Obrazy, których modyfikacja polegała jedynie na dodaniu napisów (zarówno typowych napisów zastępujących dubbing, jak i własnych elementów tekstowych dodawanych np. przy produkcji memów), generalnie nie wpływała bardzo mocno na stopień podobieństwa. Był on niższy, jednak z reguły wciąż akceptowalny z punktu widzenia algorytmu wyszukiującego.
 - Obrazy zmodyfikowane mocno, w szczególności poprzez naklejenie na nie innych grafik, w dużym stopniu traciły podobieństwo do oryginału, nawet jeśli, z perspektywy człowieka, modyfikacje nie zasłaniały kluczowych informacji zawartych w danej scenie.

Innowacje wniesione przez projekt

Przedstawiona we wnioskach przewaga modyfikacji tekstowych przy edytowaniu obrazów zdaje się mieć nieoczywiste podłoże.

- Z jednej strony może wynikać z czysto graficznych właściwości czcionek i faktu, że naniesienie ich na obraz w gruncie rzeczy pozostawia dużą jego część niezmodyfikowaną.
- Być może taka właściwość jest także celowym skutkiem doboru deskryptorów, faworyzującego tekst na obrazach. Istnieje realne prawdopodobieństwo, że twórcy serwisu trace.moe przewidzieli, że duża część użytkowników ogląda anime z napisami i może chcieć w ten sposób korzystać z wyszukiwarki.

W związku z powyższym na przyszłość warto się zastanowić nad innymi praktycznymi zastosowaniami tego typu deskryptorów w wyszukiwaniu źródeł wielu innych form graficznych, np. memów, które często zdominowane są przez tekst, choć tym razem już niekoniecznie pochodzących z anime.

Wykorzystane narzędzia

- trace.moe – badany serwis do wyszukiwania scen z anime
- IMGonline.com.ua – badanie stopnia podobieństwa obrazów
- GIMP – modyfikacja niektórych obrazów
- MS Excel – przygotowanie tabel i wykresów