



AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE

Silnik do automatycznej kategoryzacji obrazów

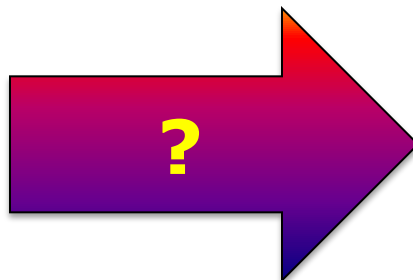
Wykonanie: Michał Cichoń
Opiekun: dr inż. Maciej Śniechowski

Wydział Fizyki i Informatyki Stosowanej, AGH

Kraków, 27.10.2014

Cel pracy

- *Stworzenie silnika kategoryzacyjnego, który ma na celu przypisać dostarczony do niego obraz do odpowiedniej kategorii*
- *Sprawdzenie trafności predykcji kategorii*

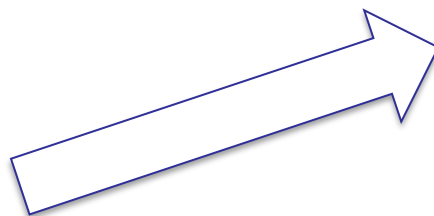


Krajobrazy

Ludzie

Zwierzęta

Budynki



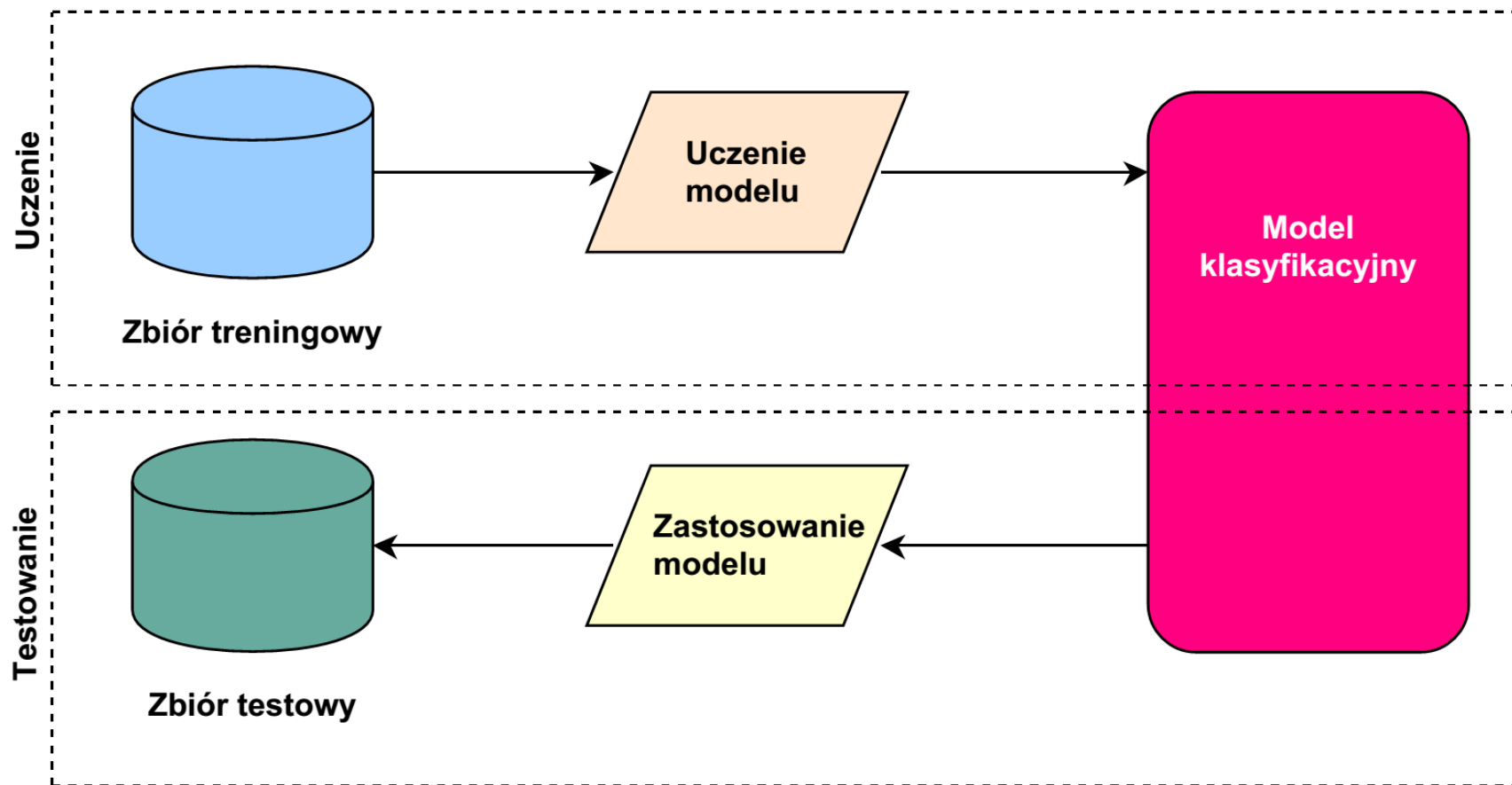
Krajobrazy

Ludzie

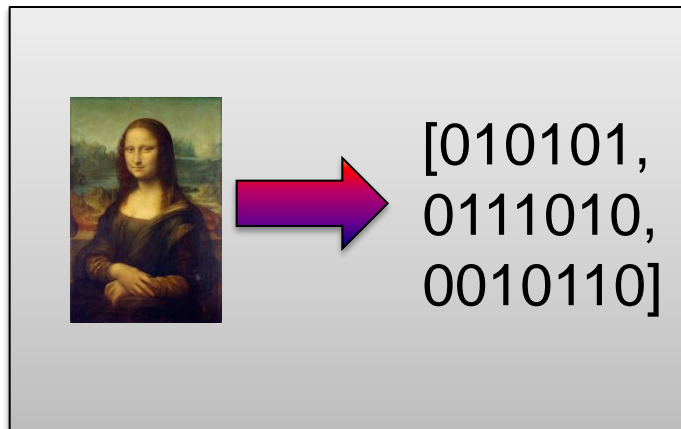
Zwierzęta

Budynki

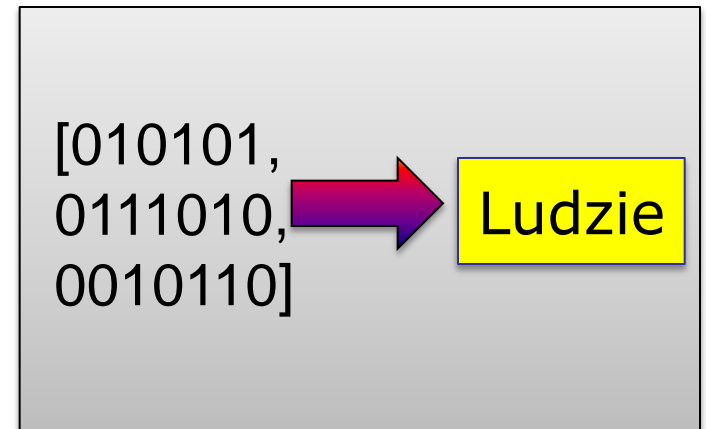
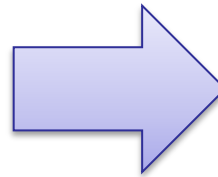
Nadzorowane uczenie maszynowe



Architektura silnika

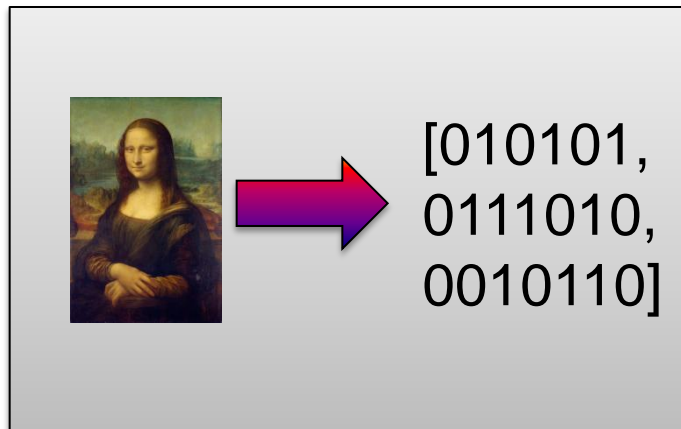


Ekstrakcja cech



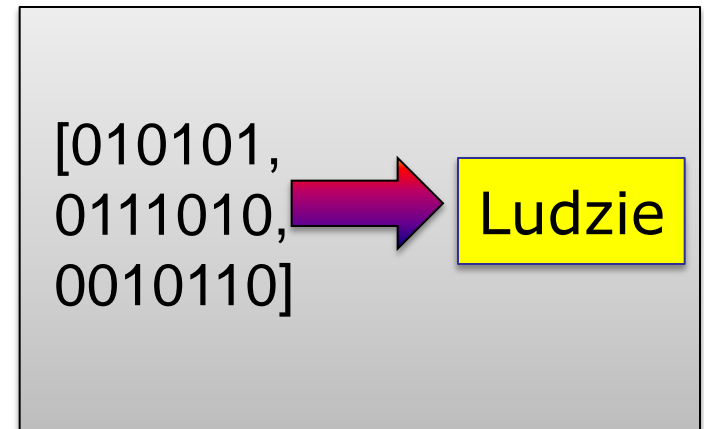
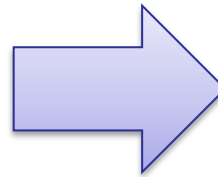
Klasyfikacja

Architektura silnika



Ekstrakcja cech

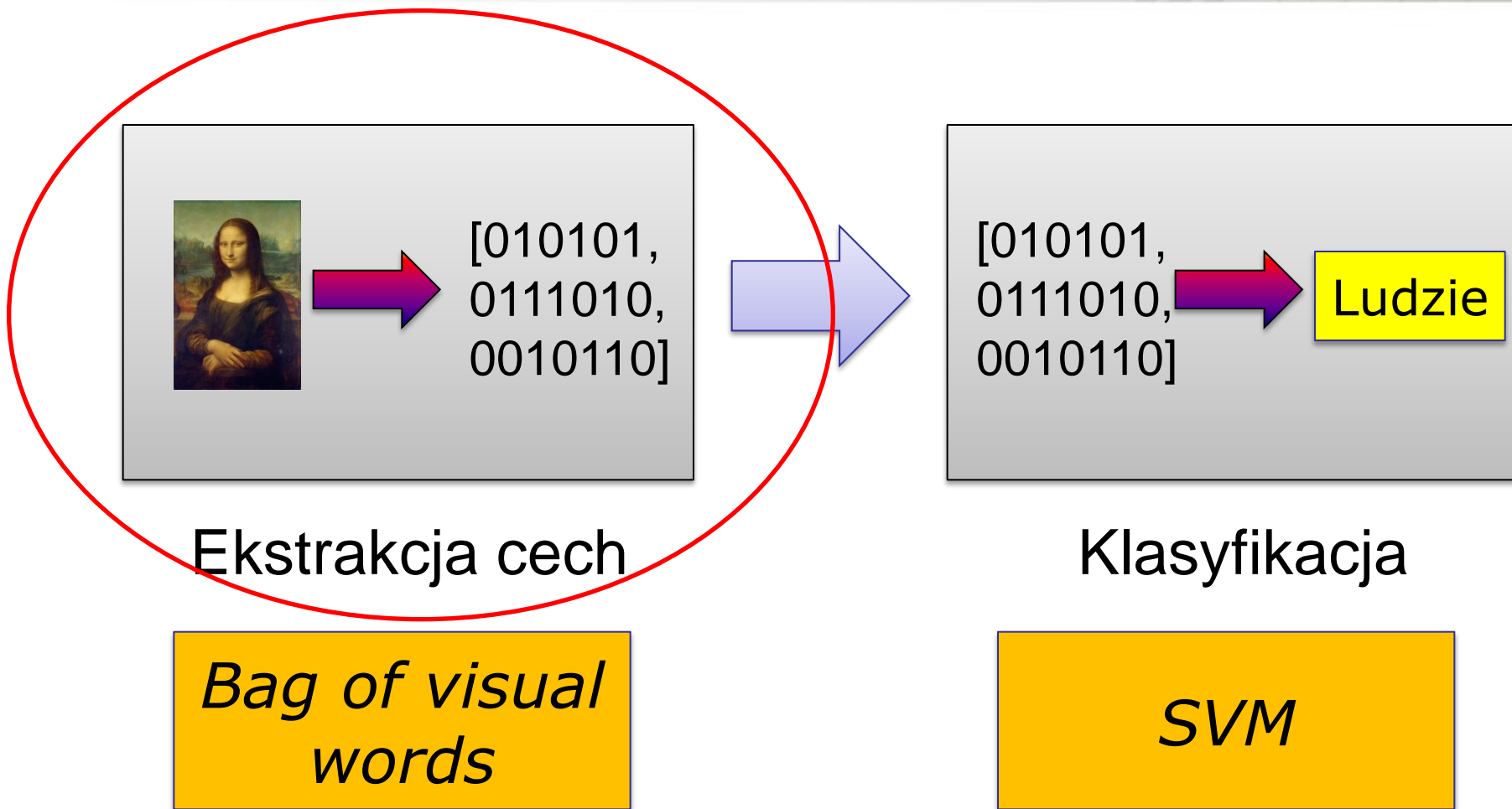
*Bag of visual
words*



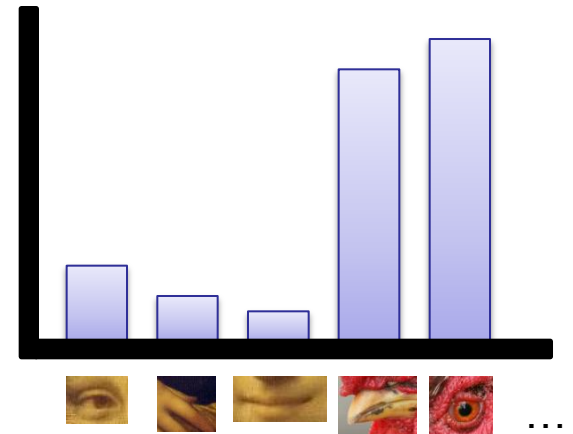
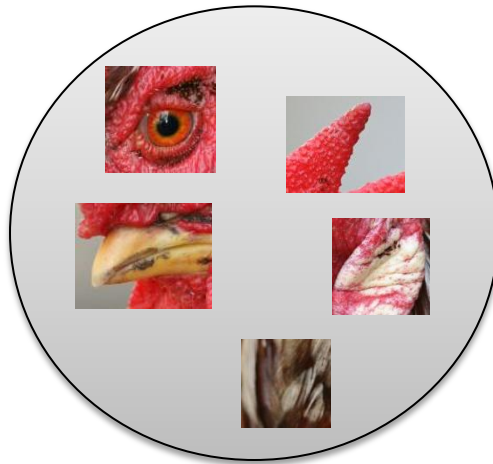
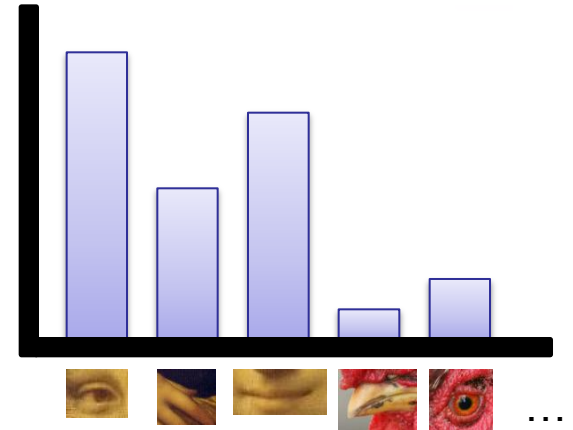
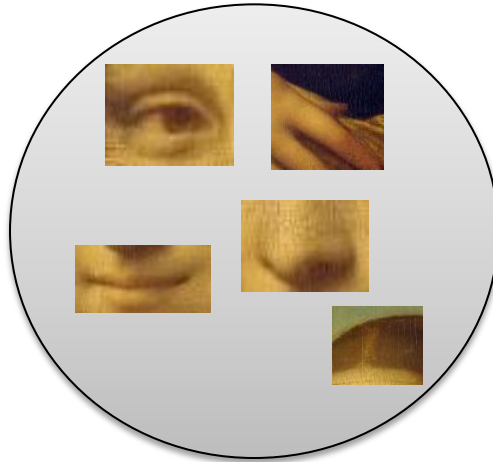
Klasyfikacja

SVM

Architektura silnika



Bag of visual words



BOW z SIFT / SURF

- SIFT i SURF dają dobre wyniki dla rozpoznawania obiektów
- Działają niezmiennie względem rotacji, skali, zmiany oświetlenia oraz położenia



Maszyna wektorów wspierających (SVM)

- + Szeroko stosowana do rozwiązywania problemów klasyfikacji
- + Dobra odporność na przetrenowanie i niedouczenie
- + Dobra generalizacja ze względu na wielowymiarowość przestrzeni cech
- Trening trwa długo
- Rozwiązania generowane przez SVM są skomplikowane

Implementacja

C++

- wieloparadygmatowy język programowania
- statycznie typowany



- biblioteka umożliwiająca przeprowadzanie operacji na obrazach
- zawiera implementację maszyny wektorów wspierających (SVM)



- zawiera bibliotekę Serialization umożliwiającą zapis modelu do pliku

Jak korzystać z silnika

Kategoryzacja pojedynczego obrazu

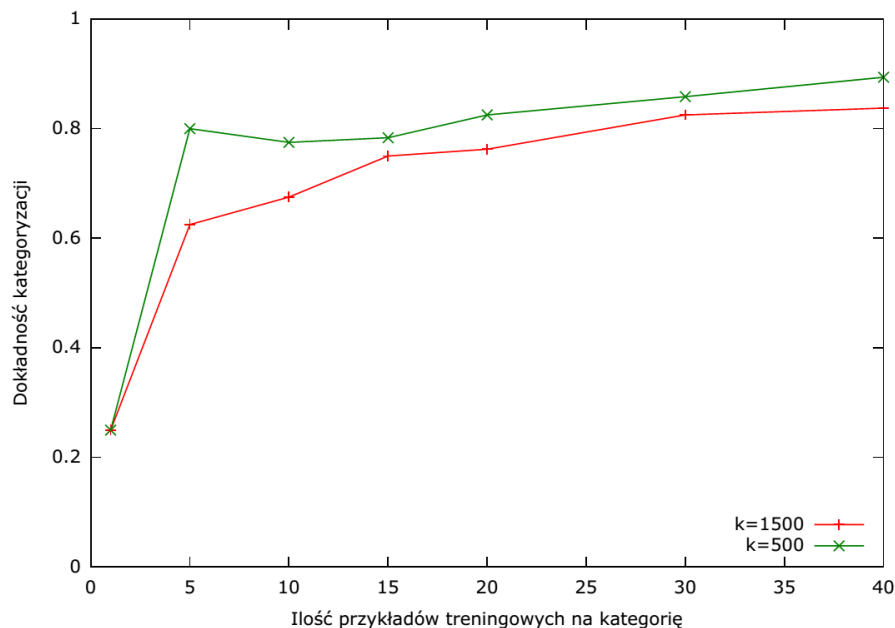
```
ImageCategorization imageCategorization  
    = ImageCategorization("C:\\dataset\\");  
imageCategorization.setNumberOfImagesInCategory(40);  
imageCategorization.refreshModel();  
  
string category = api.getCategory(string("obraz.jpg"));  
cout << category;
```

Uruchomienie testów

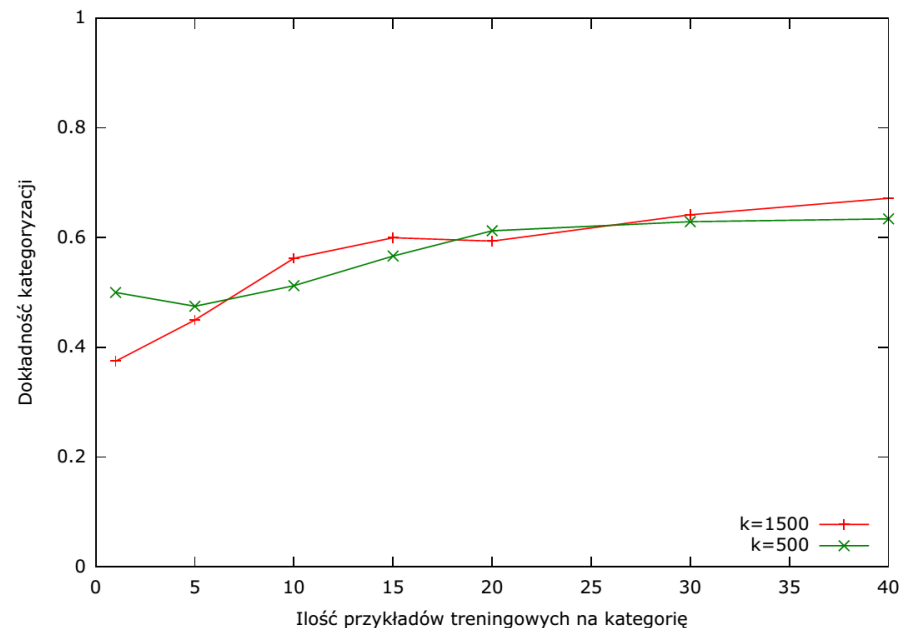
```
imageCategorization.test();
```

Wyniki kategoryzacji

Wynik kategoryzacji dla 4 kategorii (SURF)



Wynik kategoryzacji dla 8 kategorii (SURF)

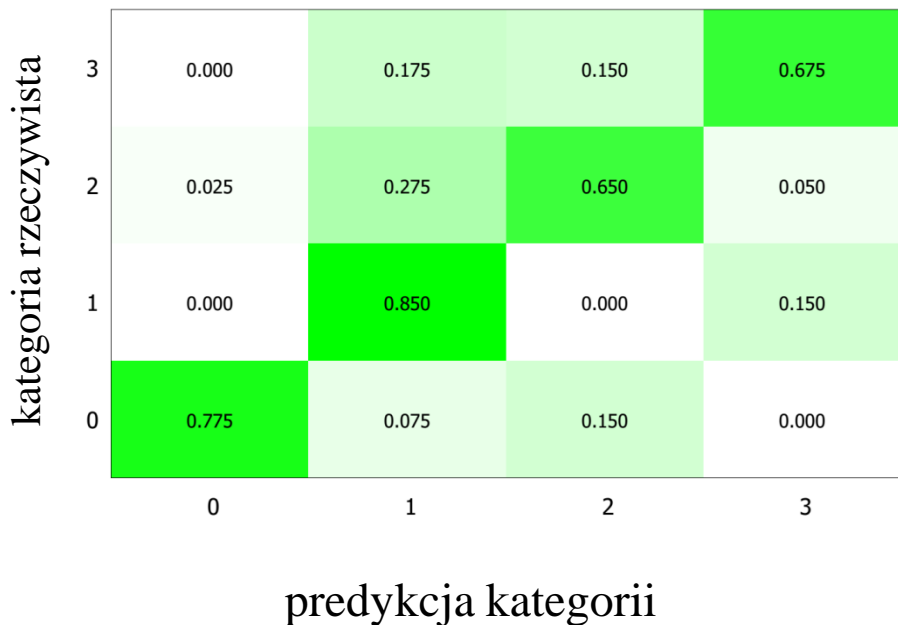


Bardzo wysoka dokładność kategoryzacji dla 4 kategorii
Dobra dokładność dla 8 kategorii

$$\text{Średnia trafność (ang. average accuracy)} = \frac{\sum_{i=1}^l \frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i}}{l}$$

Wyniki kategoryzacji – macierze pomyłek

Macierz pomyłek dla 4 kategorii



Macierz pomyłek dla 8 kategorii



0 – samoloty, 1 – bonsai, 2 – budda, 3 – motyle, 4 – samochody, 5 – żyrandole, 6 – twarze, 7 – pianina

Podsumowanie i wnioski

- Udało się stworzyć API, które pozwala na kategoryzację obrazów,
- Metoda daje dobre wyniki dla kategoryzacji 2 lub 4 kategorii, nieco gorsze dla przypadku 8 kategorii,
- Zastosowanie deskryptorów nieskorelowanych z SIFT/SURF mogłoby poprawić wynik kategoryzacji,
- Warto by było przeprowadzić analizę powtarzalności wyników oraz wpływu zbiorów uczących na osiągnięte rezultaty

Dziękuję za uwagę.