



Politechnika Świętokrzyska
Kielce University of Technology

Systemy rozpoznawania mowy i obrazu - Projekt

System rozpoznawania logo i marek

Skład zespołu:

Łukasz Wiktorowicz, 094136

Kacper Gawin, 094034

Kierunek/specjalność: Informatyka/Grafika

komputerowa

Studia: stacjonarne

Numer grupy: 4ID13B

Jaki problem rozwiązuje projekt?

W wielu sytuacjach istnieje potrzeba automatycznego rozpoznawania marek widocznych na zdjęciach lub nagraniach wideo. Ręczna analiza takich materiałów jest czasochłonna, kosztowna i podatna na błędy, szczególnie gdy liczba danych jest bardzo duża.

Przykładowe sytuacje:

- analiza tysięcy zdjęć publikowanych w internecie i mediach społecznościowych
- monitoring materiałów reklamowych i promocyjnych
- nagrania z kamer monitoringu lub wydarzeń publicznych

Dlaczego to problem:

- człowiek musi poświęcić dużo czasu na ręczne sprawdzanie
- łatwo o pomyłki przy dużej liczbie danych
- proces nie skaluje się przy rosnącej liczbie obrazów

Dla kogo i po co jest to rozwiązanie?

System rozpoznawania logotypów może być użyteczny dla różnych grup użytkowników, nie tylko osób technicznych.

- **Firmy marketingowe i analityczne**
 - automatyczna analiza widoczności marek na zdjęciach i nagraniach
- **Media i agencje reklamowe**
 - monitoring materiałów promocyjnych i reklamowych
- **Firmy dbające o ochronę marek**
 - wykrywanie nieautoryzowanego użycia logotypów
- **Zwykli użytkownicy**
 - szybkie rozpoznanie restauracji na podstawie zdjęcia lub obrazu z kamery

Dlaczego to potrzebne:

- użytkownik nie musi znać logotypów ani marek
- system działa automatycznie i szybko
- eliminuje błędy wynikające z ręcznej analizy
- może pracować na dużej liczbie danych

Zastosowane rozwiązanie

W ramach projektu zaprojektowano i zaimplementowano system automatycznego rozpoznawania logotypów marek restauracyjnych na podstawie obrazów. System wykorzystuje metody uczenia maszynowego do analizy obrazu i klasyfikacji widocznych logotypów.

Jak działa system:

- użytkownik wczytuje obraz lub korzysta z kamery
- obraz jest automatycznie przetwarzany
- system rozpoznaje logotyp i wskazuje markę
- wyświetlany jest najbardziej prawdopodobny wynik

Zastosowane podejście:

- wykorzystano konwolucyjną sieć neuronową CNN
- model był trenowany na zbiorze obrazów logotypów

Stack technologiczny

Język programowania

- Python – główny język użyty do implementacji projektu, trenowania modelu oraz stworzenia aplikacji demonstracyjnej.

Framework / biblioteki do uczenia maszynowego

- TensorFlow – framework do budowy i trenowania modeli uczenia maszynowego.
- Keras – wysokopoziomowe API wchodzące w skład TensorFlow, ułatwiające tworzenie sieci neuronowych.

Biblioteki do przetwarzania danych i obrazu

- OpenCV – biblioteka do wczytywania i wstępnego przetwarzania obrazów oraz obsługi kamery.
- NumPy – biblioteka do obliczeń numerycznych i operacji na macierzach.

Biblioteki do wizualizacji wyników

- Matplotlib – tworzenie wykresów i prezentacja wyników.
- Seaborn – wizualizacja macierzy pomyłek (confusion matrix).

Interfejs użytkownika

- Tkinter – biblioteka do stworzenia graficznego interfejsu aplikacji demonstracyjnej.

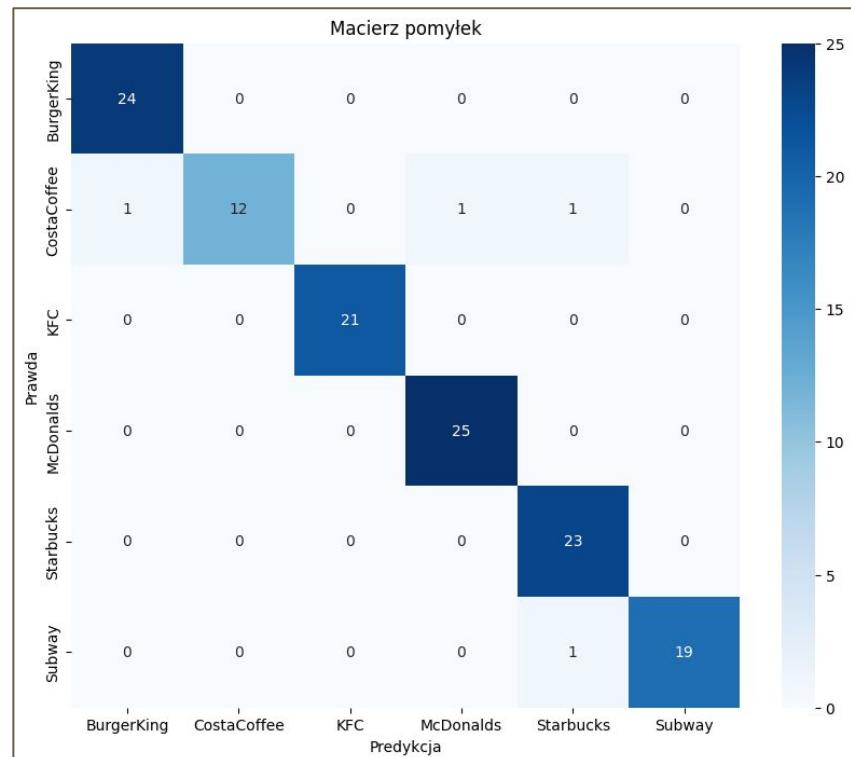
Wyniki i skuteczność systemu

Obserwacje:

- wysoka skuteczność rozpoznawania dla większości klas
- brak poważnych problemów z całkowicie błędna klasyfikacją
- pojedyncze pomyłki występują głównie między wizualnie podobnymi logotypami

Analiza:

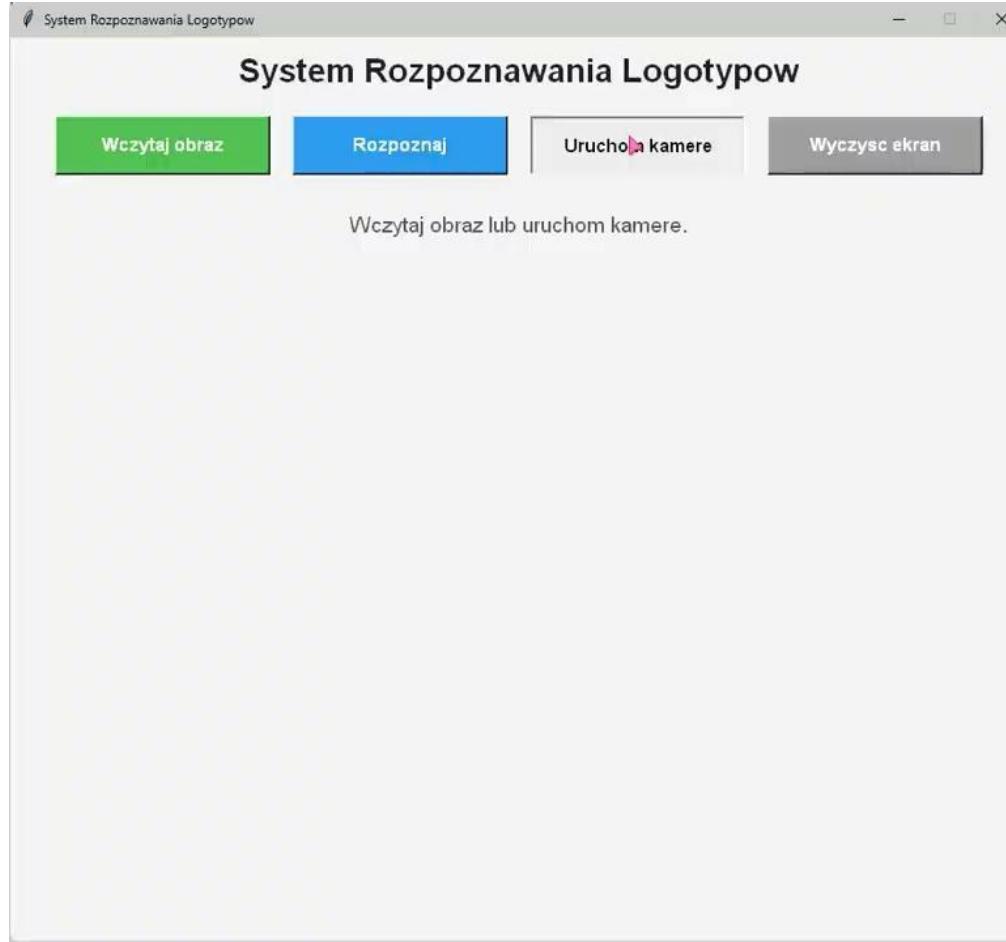
- model dobrze nauczył się cech charakterystycznych logotypów
- system w większości przypadków poprawnie generalizuje na dane wcześniej niewidziane, choć w przypadku klas słabiej reprezentowanych w zbiorze treningowym dokładność jest niższa,
- uzyskane wyniki potwierdzają zasadność zastosowanego podejścia, jednak wskazują również, że dalsze zwiększenie liczby danych oraz ich lepsze zbalansowanie mogłyby dodatkowo poprawić skuteczność modelu i ograniczyć liczbę pomyłek.



Rozpoznawanie logo ze zdjęcia



Rozpoznawanie logo w czasie rzeczywistym



Wnioski

Wnioski:

- zrealizowany system poprawnie rozpoznaje logotypy marek restauracyjnych
- zastosowane podejście oparte na sieciach neuronowych okazało się skuteczne
- projekt spełnił założone cele funkcjonalne i jakościowe
- rozwiązanie może być wykorzystywane w praktycznych zastosowaniach

Ograniczenia:

- liczba rozpoznawanych marek jest ograniczona
- skuteczność zależy od jakości i różnorodności danych uczących
- pojedyncze pomyłki pojawiają się głównie dla wizualnie podobnych logotypów oraz w przypadkach, gdy tło obrazu jest złożone lub zakłóca widoczność samego znaku.

Przyszłość

Możliwe kierunki rozwoju:

- dodanie większej liczby marek i danych treningowych
- dalsze dostrajanie parametrów modelu
- stworzenie aplikacji mobilnej
- zastosowanie systemu w szerszej analizie obrazów