

WYDZIAŁ PODSTAWOWYCH PROBLEMÓW TECHNIKI
POLITECHNIKA WROCŁAWSKA

ROZPOZNAWANIE EMOCJI TŁUMÓW METODAMI UCZENIA MASZYNOWEGO

ERYK KRUPA

NR INDEKSU: 244993

Praca inżynierska napisana
pod kierunkiem
dr inż. Jakuba Lemiesza



Politechnika
Wrocławska

WROCŁAW 2020

Spis treści

1	Wstęp	1
2	Analiza problemu	3
2.1	Analiza i porównanie istniejących rozwiązań	3
2.2	Przegląd literatury	3
3	Projekt systemu	5
3.1	Dane	5
3.2	Model	5
3.3	Wyniki	5
4	Implementacja systemu	7
4.1	Opis technologii	7
4.2	Omówienie kodów źródłowych	7
5	Instalacja i wdrożenie	9
6	Podsumowanie	11
	Bibliografia	13

Wstęp

Sieci neuronowe pozwalające na analizę emocji tłumów to wciąż jeszcze słabo przebadany temat, który jednak zyskuje zainteresowanie naukowe w ciągu ostatnich kilku lat, głównie ze względu na ogromną ilość danych dostępnych na serwisach społecznościowych zawierających zdjęcia grup osób uczestniczących w różnych wydarzeniach towarzyskich.

Poniższa praca obejmuje swoim zakresem:

- głębokie konwolucyjne sieci neuronowe, będące w stanie analizować dominujące emocje panujące na zdjęcia dużych grup i tłumów ludzi,
- bazy danych przechowujące zdjęcia wraz z odpowiadającymi im oznaczeniami.

Celem pracy jest próba zaprojektowania, zaimplementowania i wyszkolenia sztucznej sieci neuronowej zdolnej do analizy emocji tłumów na podstawie zdjęć. Sieć ma być w stanie określić emocje panującą na fotografii przydzielając ją do jednej z trzech kategorii: emocje pozytywne, neutralne i negatywne.

Praca składa się z czterech rozdziałów. W rozdziale pierwszym szczegółowo przedstawiono i przeanalizowano rozpatrywany problem, opisano i uzasadniono zastosowane podejście oraz dokładnie sprecyzowano cele pracy.

W rozdziale drugim zaprezentowano informacje dotyczące bazy danych jak również wyjaśniono, dlaczego właśnie ta baza została wybrana. Ponadto, dokładnie omówiono architekturę zaimplementowanego modelu konwolucyjnej sieci neuronowej. Na końcu rozdziału przedstawiono wyniki analizy obrazów przeprowadzanej przez zaimplementowany model.

W rozdziale trzecim opisano technologie implementacji projektu: wybrany język programowania oraz frameworki. Przedstawiono dokumentację techniczną najistotniejszych fragmentów kodów źródłowych oraz sygnatury dostępnych dla użytkownika metod publicznych.

W rozdziale czwartym przedstawiono sposób instalacji i wdrożenia modelu w środowisku docelowym.

Końcowy rozdział stanowi podsumowanie stanu zakończonych prac projektowych i implementacyjnych wraz z potencjalnymi, dalszymi kierunkami ich rozwoju.



Analiza problemu

W tym rozdziale zostanie przedstawiony problem analizy emocji na zdjęciach tłumów i grup ludzi, jak również dotychczasowe podejścia do tego problemu, moje podejście, oraz to, czego dokładnie oczekujemy od modelu.

2.1 Analiza i porównanie istniejących rozwiązań

Jednym z przykładów pracy poruszającej ten temat jest Group Emotion Recognition Using Machine Learning autorstwa Samanyou Garg [1]. Autor wykorzystuje hybrydowy system, który składa się z głębokiej konwolucyjnej sieci neuronowej (Deep Convolutional Neural Network, CNN) i klasyfikatora bayesowskiego. Na początku wydobywa zdjęcia twarzy z obrazu przedstawiającego tłum lub grupę ludzi, a następnie przekazuje je do sieci neuronowej, która analizuje emocje na poszczególnych zdjęciach twarzy zwracając uśredniony wynik. Z kolei klasyfikator bayesowski wnioskuje o globalnej emocji związanej z obrazem, integrując wizualne cechy zawartości zdjęcia uzyskane za pomocą deskryptora sceny. Ostatecznie sieć neuronowa przekazuje swój wynik do klasyfikatora, a ten zwraca ostateczną decyzję odnośnie emocji panującej na zdjęciu.

2.2 Przegląd literatury

Tutaj znajdują się pozycje ogólnie związane z implementacją konwolucyjnych sieci neuronowych.



Projekt systemu

Tutaj ogólne wprowadzenie w architekturę systemu.

3.1 Dane

Tutaj zostaną przedstawione wszystkie informacje dotyczące bazy danych Group Affect Database, odpowiedź na pytanie dlaczego została wybrana oraz aspekty prawne.

3.2 Model

Tutaj zostanie omówiona architektura sieci neuronowej, oraz spróbuję odpowiedzieć na pytanie czemu akurat taka architektura spełnia swoją rolę.

3.3 Wyniki

Tutaj zostaną przedstawione wyniki (również wydajnościowe) dla testów.



Implementacja systemu

4.1 Opis technologii

Tutaj będzie opis technologii użytych przy implementacji systemu (Python, Keras).

4.2 Omówienie kodów źródłowych

Tutaj najważniejsze/ najciekawsze kody źródłowe i sygnatury metod publicznych.



Instalacja i wdrożenie

Tutaj znajdują się informacje o zawartości paczki, założenia co do środowiska, w którym realizowany system będzie instalowany, procedura instalacji i wdrożenia systemu, tj. instalacja Pythona i środowiska wirtualnego, sposób uruchomienia aplikacji i podstawy jej użytkowania (przekazanie zdjęcia do modelu by otrzymać wynik).



Podsumowanie

Informacje o stanie zakończonych prac projektowych i implementacyjnych, co udało się zrealizować, dalsze możliwe kierunki rozwoju systemu.



Bibliografia

- [1] S. Garg. Group emotion recognition using machine learning. 2019.

