Rozpoznawanie flag Europejskich

1. Jakub Bednarski 235829
2. Weronika Kuś 235909
3. Hubert Woziński 236015
4. Daniel Chróst 235843
5. Wstęp

Jak dotąd na rynku pojawiały się różne podobne rozwiązania. Niektóre opierają się na tym samym schemacie, ale wykorzystuje się je w innych okolicznościach. System rozpoznający znaki drogowe jest wykorzystywany między innymi przy dystrybucji luksusowych samochodów. Podobnym rozwiązaniem możemy nazwać stronę flagid.org, która identyfikuje flagę na podstawie filtrów wybranych przez użytkownika.

Nasz program otrzymuje zdjęcie, po tym jak użytkownik zaznaczy flagę na tym zdjęciu potrafi określić jakiego Państwa jest ta flaga. W dalszej części będą omówione: Użyte technologie, Budowa i funkcje programu, Schemat sieci neuronowej.

W odróżnieniu od innych, podobnych rozwiązań nasz program potrafi określić Państwo z różnych zdjęć. Jeśli na zdjęciu znajduje się więcej niż jedna flaga, program określi te Państwa. Program wyznacza 3 najwyższe prawdopodobieństwa.

1. Materiały i metody
   1. Dane

Program otrzymuje zdjęcie, na którym jest widoczna co najmniej jedna flaga Europejskiego Państwa. Obrazy, których użyto przy tworzeniu programu przedstawiały flagi Państw Europejskich w różnych widokach. Niektóre były proste bez skazy, inne natomiast były cieniowane, krzywe. Obrazy te pozyskane zostały ze strony https://www.google.com/imghp

* 1. Technologie i narzędzia

Program w całości został napisany w języku Python. Podczas tworzenia zostały użyte następujące biblioteki:

- Tensorflow

- Keras

- Matplotlib, numpy

- opencv

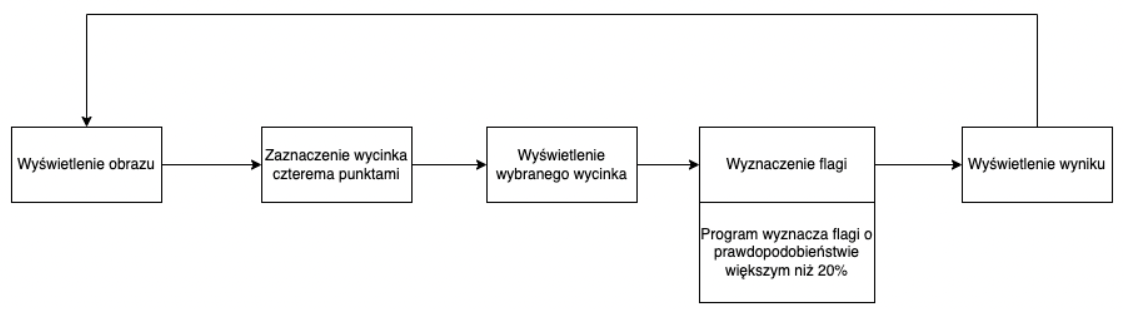
-scikit-learn

Biblioteki Tensorflow oraz Keras były niezbędne do utworzenia sieci neuronowej. Numpy został użyty do przetwarzania danych. Matplotlib wykorzystano do przedstawienia danych wynikowych oraz do utworzenia interfejsu użytkownika.

* 1. Metody

Rozpoznawanie obrazów flag Państw Europy z interfejsem użytkownika.

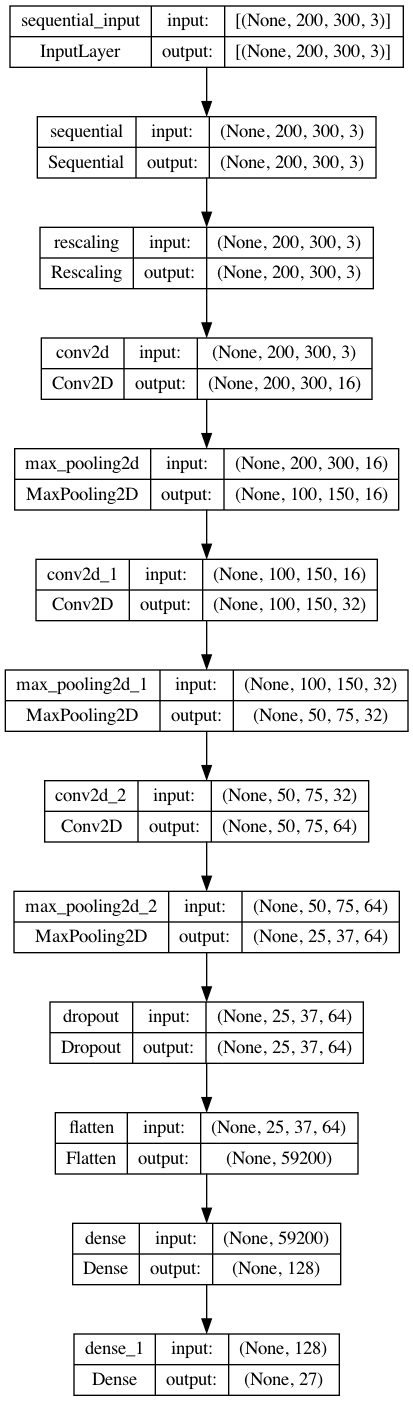
Wycinanie wybranego fragmentu – użytkownik zaznacza na wybranym obrazie 4 punkty. Punkty te tworzą obszar, a następnie program wyznacza flagi na tym obszarze i wyświetla wynik końcowy w postaci flagi i nazwy Państwa.



Metoda predykcji flagi - po wybraniu części obrazu jest ona przetwarzana przez sieć neuronową. Dzięki modelowi sekwencyjnemu, który składa się z samodzielnie dobranych warstw jesteśmy w stanie określić prawdopodobieństwo wystąpienia odpowiedniej z flag. Każda z warstw ma pojedyncze wejście oraz wyjście co skutkuje uproszczeniem całej sieci neuronowej.

* 1. Sieć Neuronowa

Schemat sieci neuronowej użytej w programie:



Omówienie budowy sieci:

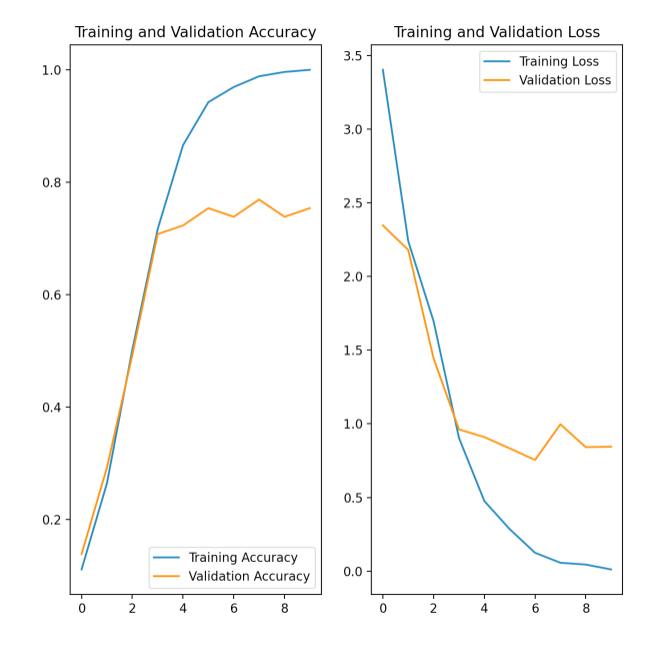
Sieć neuronowa pokazana na powyższym schemacie jest siecią sekwencyjną. Stwierdzenie to określa sieć, która na każdej z warstw ma jedno wyjście oraz wejście. W pierwszym z bloków widzimy pobieranie danych wejściowych będących niezbędnymi do poprawnego uczenia sieci. Kolejny z bloków określa typ sieci, który został już omówiony.

W pierwszej z warstw obraz jest przeskalowany do wielkości 200x300. Kolory obrazu zostały zachowane co można zauważyć poprzez trzeci parametr określający głębokość. Kolejna warstwą jest Conv2d co oznacza działanie konwolucji czyli wyodrębnianie dodatkowych filtrów dla danego obrazu. Maxpooling2D jest warstwą sieci neuronowej, która jest używana do zmniejszania rozmiaru obrazu podczas uczenia modelu. Działa ona poprzez wybieranie największych wartości z każdego obszaru (zwanego "oknem") i zastępowanie ich jedną maksymalną wartością. To pozwala modelowi skupić się na najważniejszych cechach obrazu i ogranicza ilość danych, z którymi musi się uczyć.

Operacje z dwóch ostatnio opisanych warstw jest wykonywana trzykrotnie, aby wygenerować jak najwięcej różnic między obrazami. Następną warstwą jest warstwa *Dropout* usprawniająca(?) działanie sieci neuronowej. Operacja ta została wprowadzona, aby zapobiec przeuczaniu się sieci neuronowej. Warstwa flatten spłaszcza tensor do postaci jednowymiarowej, dzięki temu możemy traktować cały tensor jak wektor filtrów. Dwie kolejne warstwy tworzą tensor odpowiadający **ilości** przetwarzanych typów danych.

1. Wyniki

Model sieci neuronowej bez warstw dropout1 i augmentacji

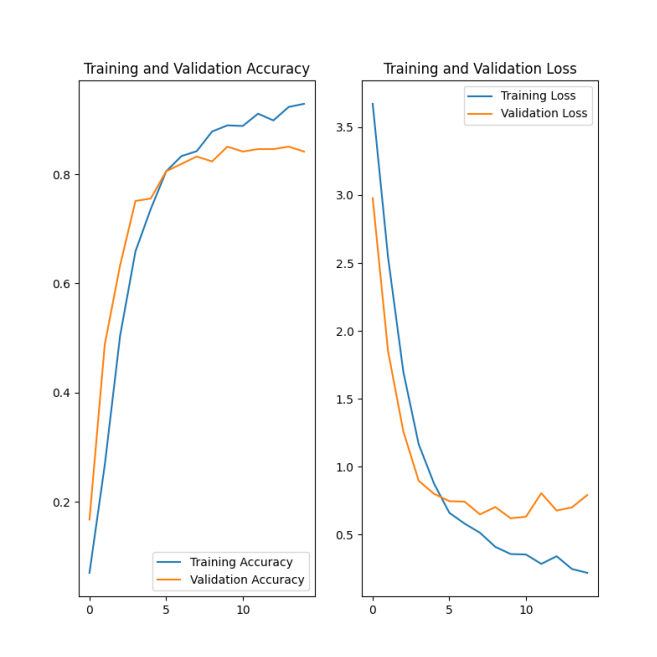


Lewy wykres przedstawia zależność zgodności i walidacji do liczby epok. Widoczne jest to, że walidacja osiąga maksymalnie nie więcej niż 80% zgodności.

Na prawym wykresie można zauważyć, że strata związana z walidacją jest większa niż strata treningowa. Oznacza to, że model jest niedostosowany, czyli model nie jest w stanie dokładnie modelować danych szkoleniowych przez co generuje duże błędy.

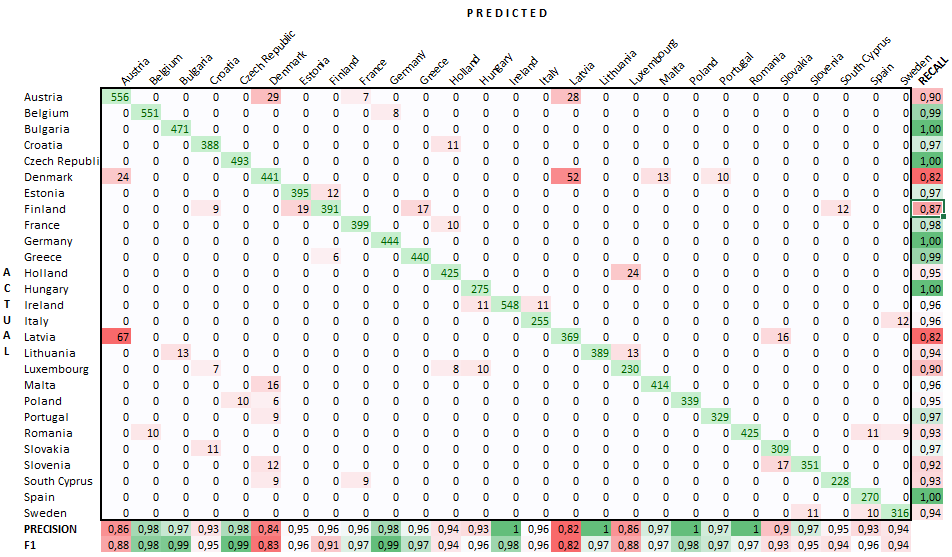
Dropout1 – Warstwa minimalizująca efektu przeuczenia

Model sieci neuronowej z warstwami dropout i augmentacją



Na lewym wykresie można zauważyć, że zgodność walidacji wynosi ponad 80% przy 15 epokach.

Na prawym wykresie widoczny jest spadek straty związanej z walidacją

Macierz konfuzji

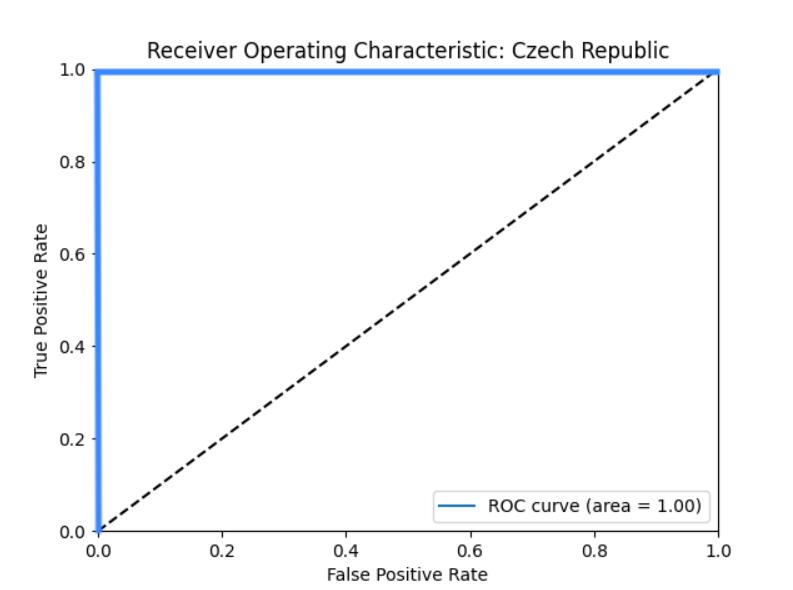
Macierz konfuzji jest narzędziem, które służy do oceny jakości klasyfikatora (np. modelu machine learning). Macierz konfuzji pokazuje, ile razy klasyfikator pomylił się w swoich predykcjach oraz ile razy był poprawny.

Precision to ułamek prawdziwych pozytywnych predykcji dokonanych przez klasyfikatora z wszystkich pozytywnych predykcji dokonanych przez klasyfikatora

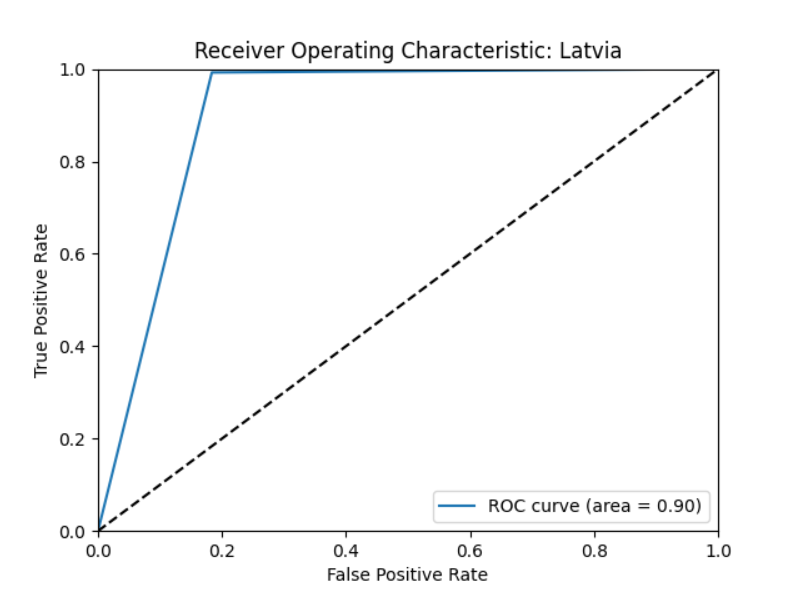
Recall to procent prawdziwych pozytywnych predykcji dokonanych przez klasyfikatora z wszystkich prawdziwych pozytywnych predykcji

F1 to średnia harmoniczna recall i precision, i jest pojedynczym wynikiem, który równoważy zarówno recall, jak i precision. Jest obliczane jako średnia harmoniczna recall i precision, przy czym wyższy wynik F1 oznacza lepsze wyrównanie między recall a precision.





Dla Czeskiej flagi zgodność wynosi 100%, jest to idealny przypadek



Dla Łotewskiej flagi zgodność wynosi 90% i jest najniższym odnotowanym wynikiem

1. Wnioski

* Po dodaniu do sieci neuronowej warstw dropout i augmentacji, zgodnośc walidacji wzrosła o ponad 10%. W ten sposób poprawiliśmy problem z przeuczaniem się sieci.
* Dokładność wyników otrzymywanych z sieci neuronowej wyniosła 94,5%
* Najwyższą odnotowaną dokładność otrzymaliśmy z flagi Republiki Czeskiej i wyniosła 100%
* Najniższą odnotowaną dokładność otrzymaliśmy z flagi Łotwy i wyniosła 90%