

Raport

Random Forest

Aleksandra Stafiej
Dominika Ciupek
Patrycja Kwiek
Teodor Pstrusiński

Z racji braku wcześniejszego pomysłu dotyczącego automatyzacji kodu, który został napisany do tej pory, podjęliśmy próbę wykorzystania lasów losowych.

Przedstawienie metody:

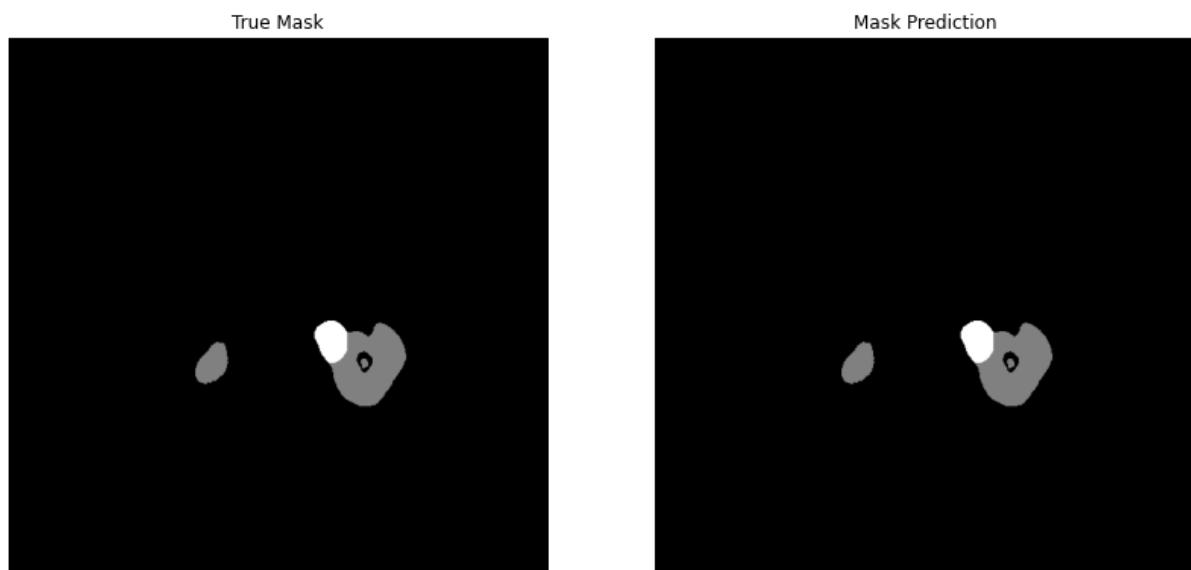
W artykule pt. “Wykrywanie nieprawidłowości nadnerczy za pomocą lasów losowych” przedstawiono wykorzystanie lasów losowych w celu segmentacji z obrazów CT nadnerczy. Aby to osiągnąć, klasyfikację przeprowadzono piksel po pikselu. [1]

Postanowiliśmy w podobny sposób przeprowadzić segmentację nerek. W tym celu po wczytaniu każdego case’a obrazy zwektoryzowaliśmy. Następnie zastosowaliśmy metodę lasu losowego.

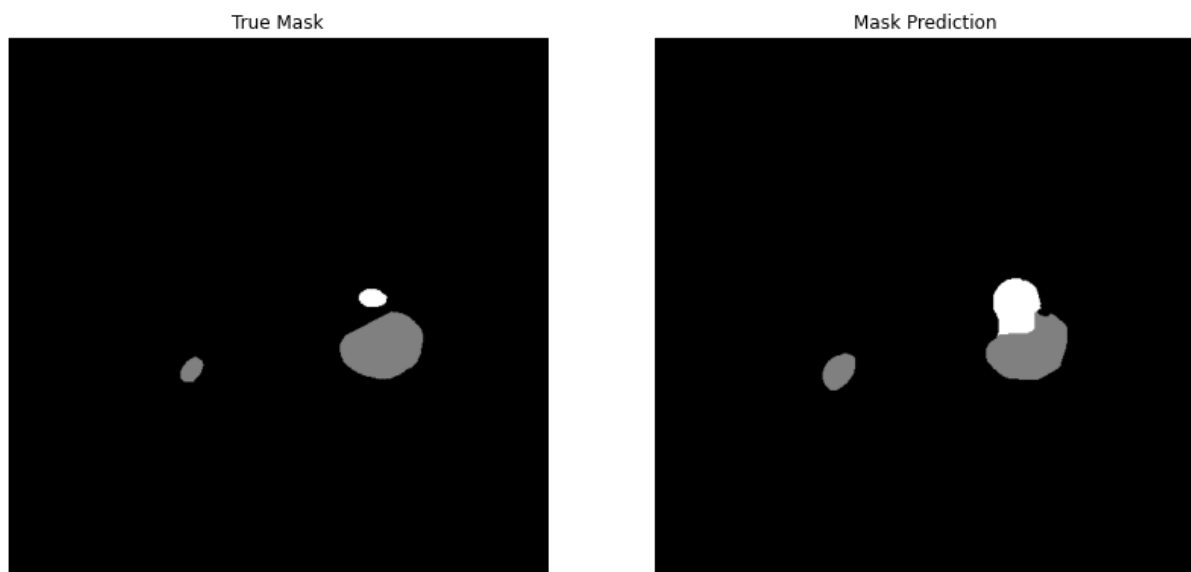
Wstępnie próbę przeprowadziliśmy na 3 case’ach o numerach 10, 11 i 12.

Wyniki:

Następnie w celu wizualnej oceny wyników naszego modelowania zobrazowano dwie maski wyliczone ze zbioru testowego przez nasz model oraz (dla porównania) maskę oryginalną, która była jako prawdziwa podana w danych.



Rys.1. Zobrazowanie masek przedstawiających nerki (szary) i guzy nerek (biały) Dla maski, którą dostaliśmy w zestawie danych (pierwszy obraz) i wyliczonej przez model (drugi obraz).



Rys. 2. Zobrazowanie masek przedstawiających nerki (szary) i guzy nerek (biały) Dla maski, którą dostaliśmy w zestawie danych (pierwszy obraz) i wyliczonej przez model (drugi oraz).

Jak widać dokładność w niektórych przypadkach jest bardzo duża, zaś w innych pozostawia wiele do życzenia. Dlatego też model trzeba poprawić.

Wiadomo, że w celu poprawienia wyników należy zwiększyć ilość danych treningowych i testowych, a także zoptymalizować parametry modelu.

Co dalej:

Dlatego też następnymi krokami będą znalezienie jak najlepszego sposobu ewaluacji wyników oraz optymalizacja parametrów modelu, aby uzyskiwać możliwie jak najlepsze wyniki. Na koniec zwiększymy ilość użytych danych. Jest to spowodowane tym, iż obliczenia dużych ilości danych wykonują się bardzo długo, a więc postanowiono najpierw zoptymalizować parametry modelu, który następnie zastosuje się do większej ilości danych.

Bibliografia:

[1] Saiprasad, Ganesh, et al. "Adrenal gland abnormality detection using random forest classification." *Journal of digital imaging* 26.5 (2013): 891-897.