

Akademia Górniczo-Hutnicza  
im. Stanisława Staszica w Krakowie  
Wydział Elektrotechniki, Automatyki, Informatyki i  
Inżynierii Biomedycznej



Projekt realizowany w ramach przedmiotu  
Techniki Obrazowania Medycznego

Tytuł projektu:

**Automatyczna segmentacja nerek i  
nowotworów nerek w obrazach tomografii  
komputerowej przy użyciu konwolucyjnych sieci  
neuronowych**

Zespół:

Michał Motak  
Kamil Hosaja  
Maciej Wowra

Kraków, 22.06.2020

# Spis treści

<b>1. Wprowadzenie</b>	<b>3</b>
<b>2. Metody</b>	<b>3</b>
2.1 Zbiór danych	3
2.2 Preprocessing	3
2.3 Architektura sieci	3
2.4 Ewaluacja	4
<b>3. Wyniki</b>	<b>4</b>
3.1 Sieć segmentująca nerkę	4
3.2 Sieć segmentująca guza	7
<b>4. Podsumowanie i dyskusja</b>	<b>10</b>
<b>Bibliografia</b>	<b>11</b>

# 1. Wprowadzenie

Celem projektu było zaimplementowanie w pełni automatycznego algorytmu pozwalającego na segmentację nerki oraz guza na skanach brzucha pochodzących z tomografu komputerowego.

## 2. Metody

### 2.1 Zbiór danych

W związku z dużą ilością danych zawężono liczbę wykorzystywanych przypadków ze zbioru kits19 z 210 do 60. Przypadki te podzielono na zbiór treningowy oraz testowy w stosunku 0.7/0.3 za pomocą funkcji `train_test_split()` z biblioteki `sklearn`. W każdym przypadku testowym obecnych było bardzo dużo przekrojów nie zawierających ani nerki ani guza, w związku z czym, w celu optymalizacji procesu uczenia oraz przyspieszenia obliczeń zapisywano tylko przekroje oraz maski na których była widoczna jedna z dwóch klas. Otrzymano ostatecznie 3254 przekrojów w zbiorze treningowym oraz 3254 przekrojów w zbiorze testowym. Ilość przekrojów dla obu modeli jest taka sama z racji na zastosowanie tego samego ziarna (*seed*) przy

### 2.2 Preprocessing

Skany wolumetryczne i maski każdego przypadku otrzymano w postaci macierzy za pomocą funkcji `get_fdata()` z biblioteki `NiBabel`. Dla każdego przekroju z każdego przypadku zmieniono rozmiar obrazu z oryginalnego rozmiaru 512x512 na 256x256. W przypadku masek zerowano wartości odpowiadające obecności guza a wartościom odpowiadającym nerce przypisywano wartości 1 (w przypadku modelu segmentującego guza, zrobiono na odwrót). Następnie przekroje obrazów i masek wczytano do macierzy typu `ndarray` (`X_train`, `Y_train`, `X_test`, `Y_test`).

### 2.3 Architektura sieci

W celu segmentacji nerki oraz guza wykorzystano metody uczenia głębokiego. Zaimplementowano architekturę w pełni konwolucyjnej sieci neuronowej UNet (zmniejszonej w stosunku do oryginalnej architektury), złożonej z 5 bloków dekodujących (liczba filtrów w kolejności 16, 32, 64, 128, 256), 4 bloków enkodujących (liczba filtrów w kolejności 128, 64, 32, 16) oraz warstwy wyjściowej z sigmoidalną funkcją aktywacji. Jako funkcję aktywacji w warstwach konwolucyjnych zastosowano Rectified Linear Unit (`relu`). Architektura taka pozwala na segmentację, gdy w masce występuje tylko jedna klasa (np. obecny jest tylko guz), w związku z tym stworzono dwa modele, jeden segmentujący nerkę oraz drugi segmentujący guz.

Podczas treningu modelu zastosowano optymalizator Adam z wartością `learning_rate` równą  $1e-4$ . Jako funkcję kosztu wykorzystano `DiceLoss`, natomiast śledzonymi metrykami były `F1Score` (`Dice`) oraz `IOUScore` (`Jaccard index`) (funkcja kosztu oraz metryki pochodziły z biblioteki `segmentation_models`). Model w obu przypadkach miał ustawiony parametr

EPOCHS na 15, z instancją EarlyStopping o wartości parametru patience = 1 (trening w przypadku nerki zakończył się po 11 epokach, natomiast w przypadku guza po 5 epokach).

## 2.4 Ewaluacja

Ewaluacja zaproponowanej metody polegała na porównaniu otrzymanych masek (R) z maskami wzorcowymi (G). Wybrano i opisano trzy metryki.

- 1) Dice coefficient (DICE) - stosunek dwukrotnej części wspólnej do sumy masek otrzymanych i wzorcowych. Wartość zawiera się w przedziale [0,1]. Im wyższa wartość tego współczynnika tym większe podobieństwo maski otrzymanej do maski wzorcowej.

$$DICE(R, G) = \frac{2|R \cap G|}{|R| + |G|}$$

- 2) Intersection over Union (IOU, Jaccard index) - stosunek części wspólnej do sumy rezultatu i maski wyrażona w procentach. Wartość zawiera się w przedziale [0,1], wyższa wartość oznacza lepszą segmentację.

$$IOU(R, G) = \frac{|R \cap G|}{|R \cup G|}$$

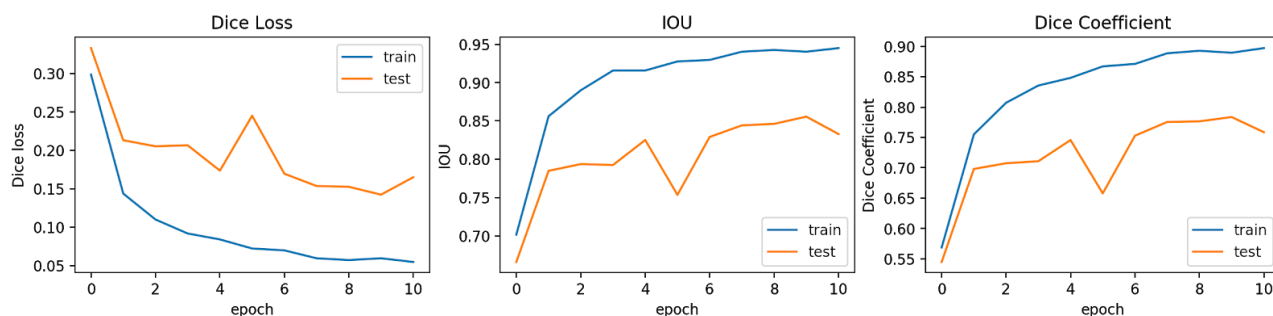
- 3) relative volume difference (VD) - wskazuje na podobieństwo objętości masek. Nie uwzględnia ich odwzorowania, jedynie objętość. Pokazuje czy metoda ma tendencję do przesegmentowania (wartość dodatnia) albo podsegmentowania (wartość ujemna). Wartości bliskie 0 wskazują na najlepszą segmentację.

$$VD(R, G) = \frac{|R| - |G|}{|G|} \cdot 100 \%$$

## 3. Wyniki

### 3.1 Sieć segmentująca nerkę

Na rysunek 1 przedstawiono zmiany wartości funkcji kosztu, współczynnika Dice oraz współczynnika IOU na przestrzeni epok dla modelu segmentującego nerkę.



Rys. 1 Przebieg procesu uczenia dla modelu segmentującego nerkę.

Model został zatrzymany przez EarlyStopping w 11 epoce uczenia. Zapisany został najlepszy model uzyskany podczas procesu, dla którego wartości funkcji kosztu, współczynnika Dice oraz IOU wyniosły:

$$DiceLoss_{train} = 0.046, DiceScore_{train} = 0.95, IOUScore_{train} = 0.91$$

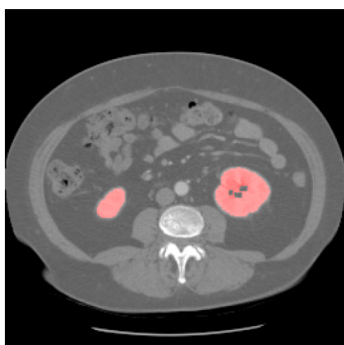
$$DiceLoss_{test} = 0.142, DiceScore_{test} = 0.86, IOUScore_{test} = 0.78$$

Dodatkowo wyznaczono wartości współczynnika VD dla zbioru treningowego oraz testowego, które wyniosły:

$$VD_{train} = 9.38\% \text{ i } VD_{test} = 7.65\%$$

Na poniższych rysunkach (2, 3, 4) zestawiono oryginalne maski z maskami przewidzianymi (nałożone na oryginalny obraz) w kolejności od najlepszych do najgorszych (maski wybierane były ze względu na wartości współczynnika Dice).

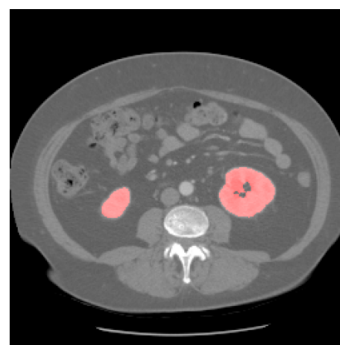
Original image  
with original mask



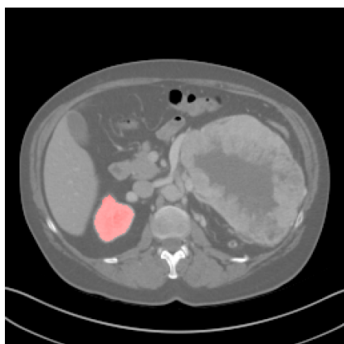
Predicted mask



Predicted mask  
on image



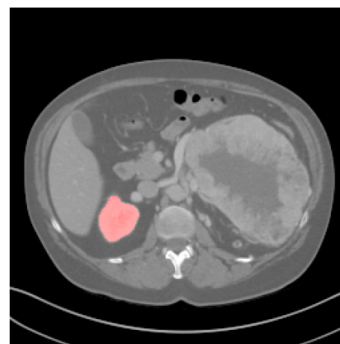
Original image  
with original mask



Predicted mask



Predicted mask  
on image

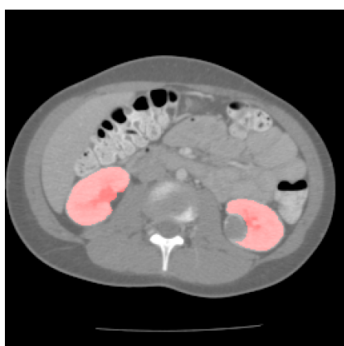


Rys. 2 Zestawienie najlepiej dopasowanych masek dla segmentacji nerek.

Tabela 1. Zestawienie wartości dla powyższych (najlepszych) przekrojów.

Dice	IoU	VD [%]
0.97	0.94	5.3
0.98	0.96	5.7

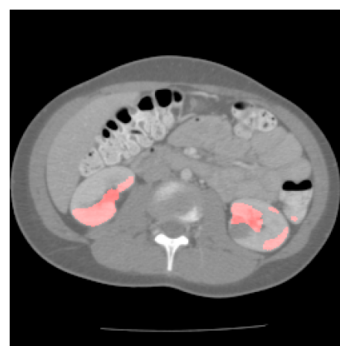
Original image  
with original mask



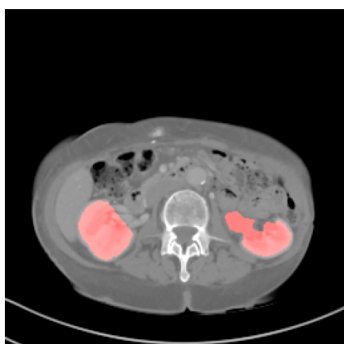
Predicted mask



Predicted mask  
on image



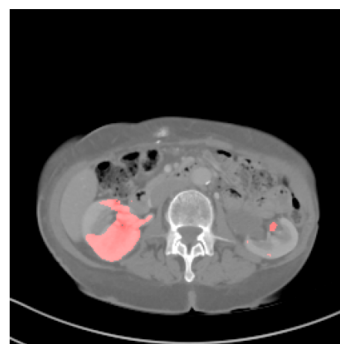
Original image  
with original mask



Predicted mask



Predicted mask  
on image

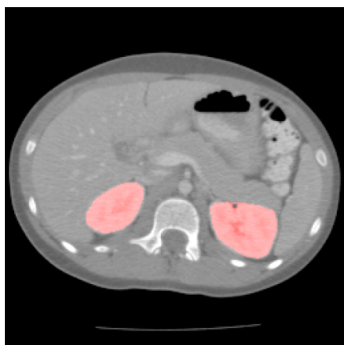


Rys. 3 Zestawienie przeciętnych dopasowanych masek dla segmentacji nerek.

Tabela 2. Zestawienie wartości dla powyższych (przeciętnych) przekrojów.

Dice	IoU	VD [%]
0.56	0.38	-52.3
0.51	0.34	-48.3

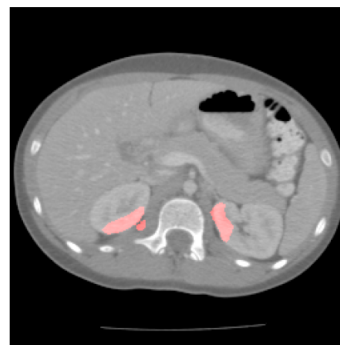
Original image  
with original mask



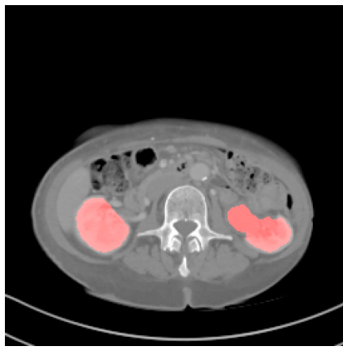
Predicted mask



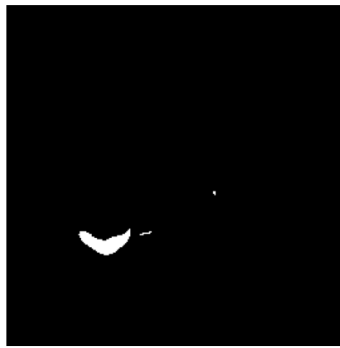
Predicted mask  
on image



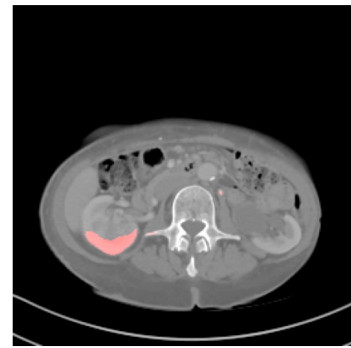
Original image  
with original mask



Predicted mask



Predicted mask  
on image



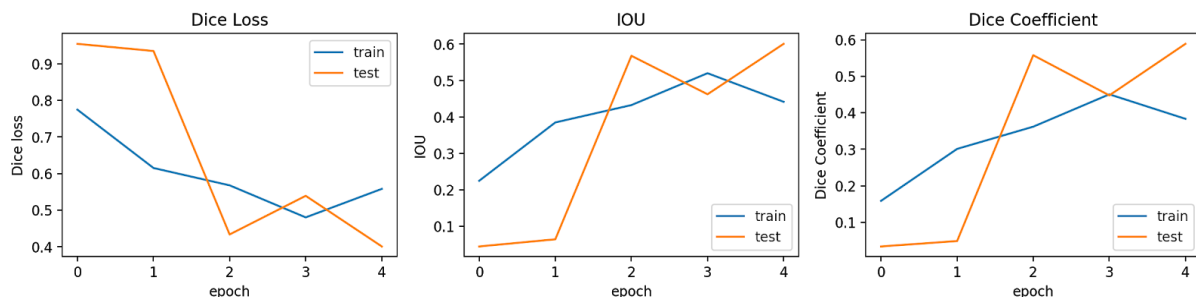
Rys. 4 Zestawienie najgorzej dopasowanych masek dla segmentacji nerek.

Tabela 3. Zestawienie wartości dla powyższych (najgorszych) przekrojów.

Dice	IoU	VD [%]
0.30	0.18	-77.9
0.24	0.14	-81.7

### 3.2 Sieć segmentująca guza

Na rysunku 5 przedstawiony został proces uczenia modelu segmentującego guza. Model został zastopowany w



Rys. 5 Przebieg procesu uczenia dla modelu segmentującego guza.

Model został zatrzymany przez EarlyStopping w 5 epoce uczenia. Zapisany został najlepszy model uzyskany podczas procesu, dla którego wartości funkcji kosztu, współczynnika Dice oraz IOU wyniosły:

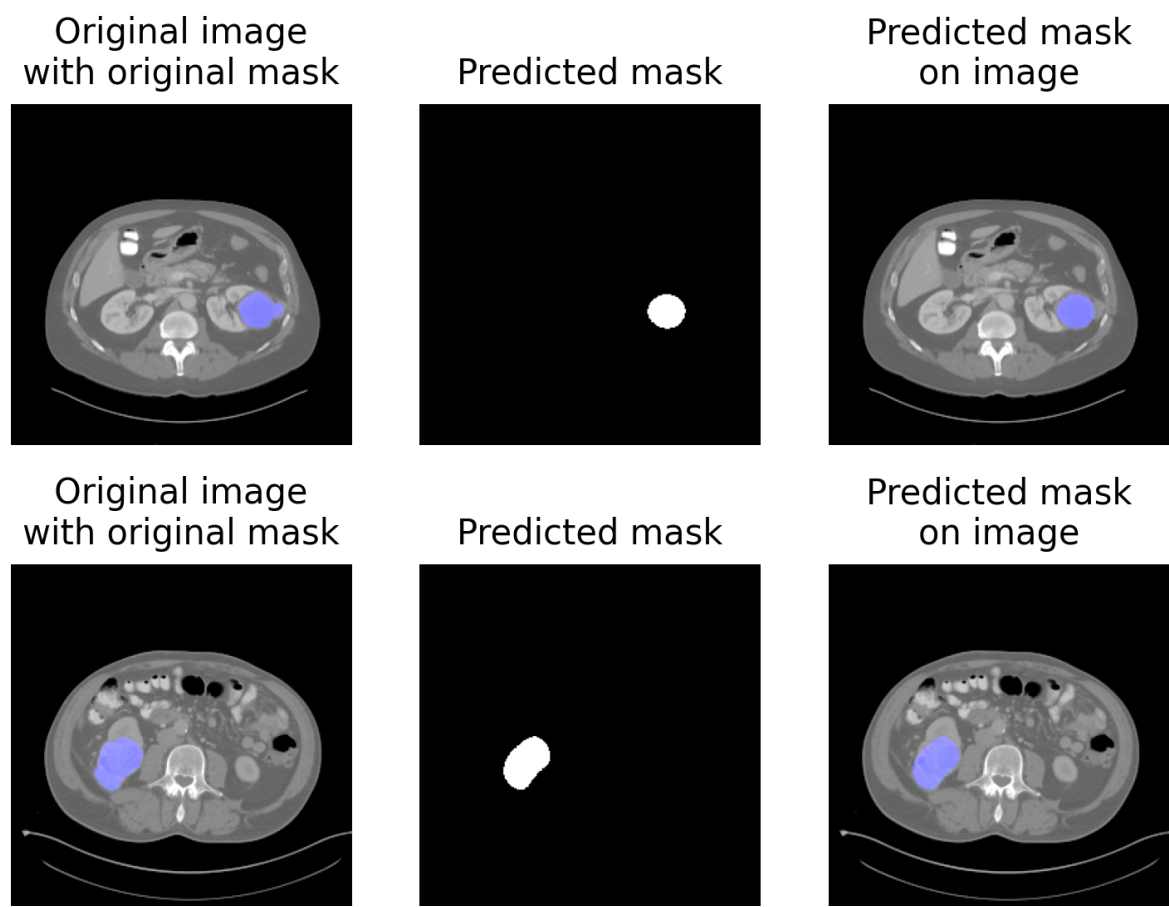
$$DiceLoss_{train} = 0.415, DiceScore_{train} = 0.59, IOUScore_{train} = 0.51$$

$$DiceLoss_{test} = 0.401, DiceScore_{test} = 0.60, IOUScore_{test} = 0.59$$

Dodatkowo wyznaczono wartości współczynnika VD dla zbioru treningowego oraz testowego, które wyniosły:

$$VD_{train} = 17.57\% \text{ i } VD_{test} = 22.08\%$$

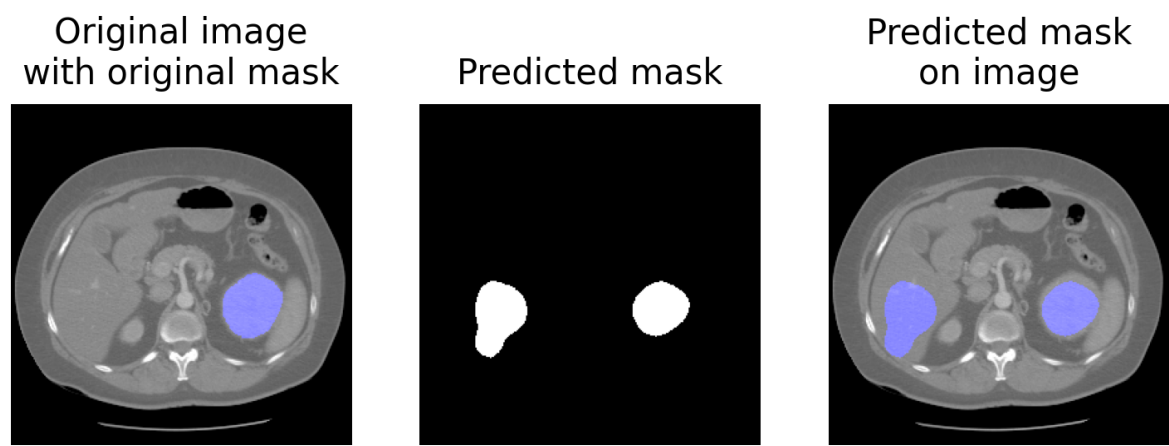
Na poniższych rysunkach (6, 7, 8) zestawiono oryginalne maski z maskami przewidzianymi (nałożone na oryginalny obraz) w kolejności od najlepszych do najgorszych (maski wybierane były ze względu na wartości współczynnika Dice).



Rys. 6 Zestawienie najlepiej dopasowanych masek dla segmentacji guza.

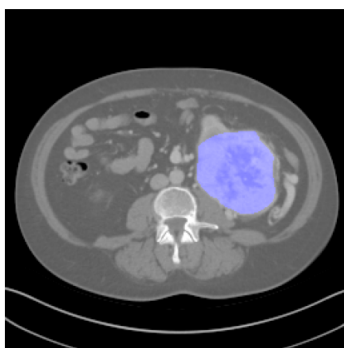
Tabela 4. Zestawienie wartości dla powyższych (najlepszych) przekrojów.

Dice	IoU	VD [%]
0.90	0.80	-9.3
0.94	0.87	-4.3

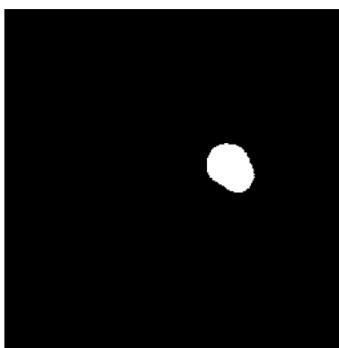




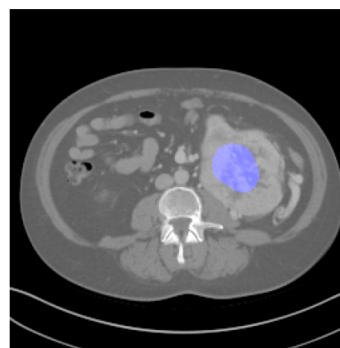
Original image  
with original mask



Predicted mask



Predicted mask  
on image

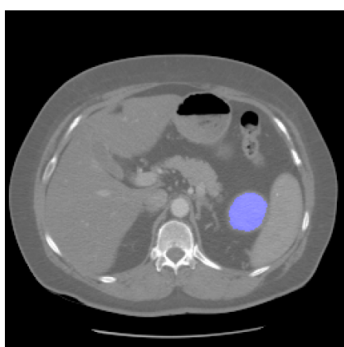


Rys. 7 Zestawienie przeciętnie dopasowanych masek dla segmentacji guza.

Tabela 5. Zestawienie wartości dla powyższych (przeciętnych) przekrojów.

Dice	IoU	VD [%]
0.57	0.39	77.1
0.50	0.32	-66.3

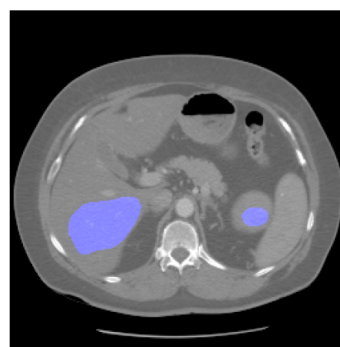
Original image  
with original mask



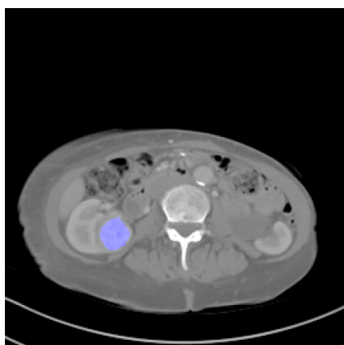
Predicted mask



Predicted mask  
on image



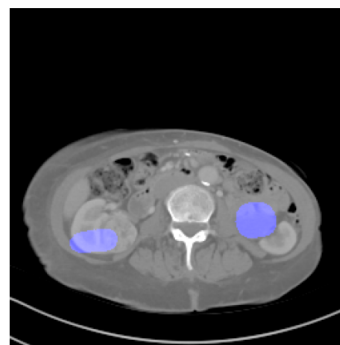
Original image  
with original mask



Predicted mask



Predicted mask  
on image



Rys. 8 Zestawienie najgorzej dopasowanych masek dla segmentacji guza.

Tabela 6. Zestawienie wartości dla powyższych (najgorszych) przekrojów.

Dice	IoU	VD [%]
0.18	0.10	198.3
0.13	0.07	205.0

## 4. Podsumowanie i dyskusja

Wyniki otrzymane dla modelu segmentującego nerkę są zadowalające: współczynnik Dice na poziomie 0.95 podczas treningu (najlepszy zapisany model) oraz 0.85 dla walidacji, natomiast wartość IoU 0.91 podczas treningu oraz 0.78 dla walidacji. Na podstawie takich wartości można wnioskować o małym przetrenowaniu modelu (wartości treningowe są większe od wartości walidacyjnych). Wyznaczone wartości VD dla zbiorów treningowego i testowego były do siebie zbliżone (różnica ~2%) oraz były na niskim poziomie, co dla współczynnika VD jest dobrym znakiem. Wartości VD są dodatnie co wskazuje na delikatne przesegmentowanie masek. Zaprezentowane na rysunkach 2-4 zestawienia masek obrazuje zdolności segmentacyjne modelu.

W przypadku sieci segmentującej nowotwór otrzymano znacznie gorsze wyniki. Dice podczas treningu wyniósł ~0.58 oraz 0.60 dla walidacji, natomiast IoU to odpowiednio 0.50 i 0.59. Wartości współczynnika VD również nie osiągnęły zadowalających wyników. Wartości te nie są wystarczająco wysokie, żeby poprawnie segmentować guza nerki. Wybór prezentowanych masek w obu przypadkach był subiektywny, pominięto zdjęcia całkowicie puste oraz takie, na których segmentowana nerka była bardzo mała.

Analizując rysunki 2-8 oraz tabele 1-6 można dojść do wniosku, iż najlepiej wysegmentowane maski cechują się wysokim współczynnikiem Dice i IoU, przy czym VD jest bliskie 0. Z kolei jedne z najgorszych otrzymanych segmentacji cechują się małą wartością Dice i IoU, natomiast współczynnik VD jest bardzo wysoki.

Zdolności segmentacyjne trenowanych modeli mogą być poprawione kilkoma metodami, których nie zastosowano ze względu na długi czas obliczeń oraz brak zasobów sprzętowych. Mowa tu o zastosowaniu generatorów, które pozwoliłyby na wczytanie większej ilości bardziej zróżnicowanych danych. Możliwe byłoby wtedy również zastosowanie augmentacji danych, przez co zbiór treningowy byłby jeszcze bardziej rozbudowany. Dodatkowo, dobrym pomysłem może być zastosowanie wag dla klas, ponieważ pełny zbiór treningowy jest niezbalansowany (biblioteka segmentation-models umożliwia wprowadzenie wag w metrykach oraz funkcjach kosztu).

Próbowano zaimplementować większą architekturę sieci UNet, jednak w tym przypadku nie dawała ona zadowalających rezultatów (głównie ze względu na dużą ilość parametrów, która powodowała powolne uczenie modelu oraz podwyższone zużycie pamięci RAM). Dodatkową zaletą modelu o mniejszej ilości parametrów jest mniejsza zdolność do przetrenowania.

Segmentację przeprowadzono na podstawie dwóch modeli (osobno dla nerki oraz guza), uzyskując dla jednego obrazu dwie maski wejściowe. Kolejnym krokiem byłoby dopisanie

fragmentu kodu nakładającego dwie maski na siebie, który brałby pod uwagę prawdopodobieństwa zwracane z funkcji predict() modelu.

Podział prac pomiędzy członków zespołu zgodny był z tym co napisane zostało w początkowym planie projektu oraz dodatkowo uwidacznia to aktywność na GitHubie.

## **Bibliografia**

[1] O. Ronneberger, P. Fischer, i T. Brox, „U-Net: Convolutional Networks for Biomedical Image Segmentation”, *arXiv:1505.04597 [cs]*, maj 2015, Udostępniono: cze. 07, 2020. [Online]. Dostępne na: <http://arxiv.org/abs/1505.04597>.

[2] Q. Dou i in., „3D deeply supervised network for automated segmentation of volumetric medical images”, *Medical Image Analysis*, t. 41, s. 40–54, paź. 2017, doi: 10.1016/j.media.2017.05.001.

[3] <https://segmentation-models.readthedocs.io/en/latest/>