

# Segmentacja aorty z obrazów tomografii komputerowej przy użyciu głębokich sieci neuronowych

Piotr Paniec<sup>1</sup> , Maciej Stanuch<sup>2</sup> , Andrzej Skalski<sup>3</sup> 

<sup>1</sup> Akademia Górniczo-Hutnicza, Wydział Inżynierii Metali i Informatyki Przemysłowej, Kraków

<sup>2</sup> MedApp S.A., Kraków, Poland

<sup>3</sup> Akademia Górniczo-Hutnicza, Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej, Kraków

---

**Streszczenie:** W artykule zaproponowano wykorzystanie architektury nnU-Net do segmentacji aorty z obrazów tomografii komputerowej. Przetwarzanie wstępne obejmowało transformację różnych zakresów wartości do zakresu odpowiadającemu skali jednostek Hounsfielda oraz procedurę okienkowania. Na etapie treningu sieci przeprowadzono pięciokrotną walidację krzyżową oraz wytrenowano pięć modeli, których predykcje na zbiorze testowym zostały połączone. Dla modeli uzyskano średni współczynnik Dice'a na poziomie 0,921 oraz IoU równe 0,858, oznaczające wysoką jakość predykcji. Wartość odległości Hausdorffa ujawniła istnienie elementów odstających, a analiza trójwymiarowych rekonstrukcji wskazała obszary o niższej dokładności segmentacji.

**Słowa kluczowe:** uczenie głębokie, sieci neuronowe, segmentacja obrazów, obrazowanie medyczne, tomografia komputerowa, informatyka, aorta

## AORTA SEGMENTATION FROM CT IMAGES USING DEEP NEURAL NETWORKS

**Abstract:** The manuscript proposes employing the nnU-Net architecture for the segmentation of the aorta from computed tomography images. Preprocessing involved transforming various value ranges to align with the Hounsfield unit scale, in addition to implementing a windowing procedure. During the network training phase, a five-fold cross-validation protocol was utilized, resulting in the training of five models, whose predictions on the test set were aggregated. The models achieved an average Dice coefficient of 0.921 and an Intersection over Union of 0.858, signifying high prediction quality. The value of the Hausdorff distance indicated the presence of outliers, while the analysis of three-dimensional reconstructions highlighted regions of diminished segmentation accuracy.

**Keywords:** deep learning, neural networks, image segmentation, medical imaging, computed tomography, computer science, aorta

## 1. Wprowadzenie

Tomografia komputerowa (CT, ang. *computed tomography*) jest jedną z najważniejszych technik stosowanych w diagnostyce obrazowej (Hsieh i Flohr 2021). Aorta, największa tętnica organizmu, bywa dotknięta patologiami w budowie i strukturze, które bez diagnozy i odpowiedniego leczenia mogą zagrażać zdrowiu i życiu. W konsekwencji niezwykle istotna jest jej odpowiednia ocena, ustalana na podstawie obrazowania tomografii komputerowej.

Jednym z narzędzi informatycznych, które mogą zostać wykorzystane w tym celu, jest automatyczna segmentacja aorty w obrazach CT, która umożliwia między innymi wyodrębnienie jej światła, trójwymiarową rekonstrukcję geometrii naczynia, a w konsekwencji precyzyjne pomiary geometrii. Celem pracy jest stworzenie narzędzia, bazującego na sieci neuronowej, które umożliwi segmentację w pełni automatycznie, niezależnie od protokołu akwizycji.

Pod względem danych obrazowych wartości powinny być reprezentowane w jednostkach Hounsfielda będących względnym współczynnikiem osłabienia promieniowania rentgenowskiego odniesionego do wody. Woda destylowana reprezentowana jest przez wartość 0 HU (Gücük i Üyetürk 2014, Jhala i Thaker 2015). Ze względu na sposób zapisu wartości w danych, oznaczenia miejsc w obrazach, które były poza obszarem akwizycji, typu tomografu, jednostki czy protokołu akwizycji zakres wartości może być inny.

## 2. Segmentacja aorty

W pracy zaproponowano na potrzeby segmentacji wykorzystanie narzędzia bazującego na uczeniu głębokim. Jest to podejście często stosowane w analizie obrazów medycznych. Powodem popularności jest automatyczna ekstrakcja cech, oznaczająca, że model samodzielnie, w trakcie procesu uczenia, na podstawie dostarczonych danych, określa istotne cechy, które wpływają na predykcję wyniku.

W niniejszej pracy zastosowano framework nnU-Net (Isensee i in. 2019b, 2021, 2024) w wersji V2. Jest to system stworzony na potrzeby dopasowania rodziny architektur U-Net (Ronneberger i in. 2015) do domeny modalności stosowanych w obrazowaniu medycznym. Framework nnU-Net analizuje statystykę zbioru (rozmiary danych, rozmiary woksela, liczbę klas) i na tej podstawie generuje wariant architektury U-Net oraz odpowiadający mu zestaw parametrów treningu. W konsekwencji dobór topologii sieci czy hiperparametrów jest konfigurowany automatycznie. Jest to narzędzie ułatwiające pracę, przy jednoczesnym zachowaniu najwyższych wyników dla uczenia nadzorowanego (Huang i in. 2022, Isensee i in. 2019b, 2024).

W przypadku nnU-Net możliwe jest uczenie modelu zarówno przekrojami dwuwymiarowymi, jak i danymi 3D, a w obrębie danych 3D występuje kilka domyślnie zdefiniowa-

nych sposobów budowania sieci. Nazywane są one konfiguracjami (ang. *configurations*), z których domyślnie dostępne są cztery: *2d*, *3d\_fullres*, *3d\_lowres*, *3d\_lowres\_cascade*. Do treningu wybrano konfigurację *3d\_fullres*, która tworzy architekturę opartą na konwolucji trójwymiarowej. Pozwala to sieci analizować ciągłe zależności przestrzenne w obrębie przetwarzanej objętości obrazowania, co zwiększa ilość przestrzennego kontekstu i może podnieść skuteczność predykcji (Isensee i in. 2019b). Zastosowana została pięciokrotna walidacja krzyżowa, a każdy model trenowany był przez 200 epok.

W przypadku współczynnika uczenia (ang. *learning rate*) początkowo używana jest ustalona stała wartość wynosząca 0,01. Domyślnie wartość maleje wraz z postępem treningu, jest zależna od epoki treningu według wzoru (1):

$$lr(e) = lr_0 \left(1 - \frac{e}{E}\right)^{0,9} \quad (1)$$

gdzie:

- $lr_0$  – wartość początkowa współczynnika uczenia (domyślnie 0,01),
- $e$  – bieżąca epoka,
- $E$  – maksymalna liczba epok (domyślnie 1000),
- $lr(e)$  – wartość współczynnika uczenia dla danej epoki.

Przed rozpoczęciem treningu użyty framework domyślnie projektuje zestaw augmentacji jako element planu eksperymentu, natomiast stosuje je na bieżąco w trakcie uczenia, generując losowo przekształcone wersje obrazów wejściowych. Domyślnie stosowana augmentacja obejmuje losową rotację, losowe skalowanie, szum gaussowski, rozmycie gaussowskie, losową jasność, losowy kontrast, symulowanie niższej rozdzielczości, korekcję gamma, odbicia lustrzane (Isensee i in. 2019a), co ma na celu zwiększenie różnorodności obrazów i poprawę uogólniania. W prezentowanym podejściu zrezygnowano z dodania losowego skalowania, losowej jasności, losowego kontrastu czy korekcji gamma ze względu na fakt, że te transformacje nie odzwierciedlają możliwych zmian w danych. W ostatecznym treningu zastosowano losową rotację, szum gaussowski oraz odbicie lustrzane. Te zabiegi odzwierciedlają możliwość transformacji, które mogą wystąpić w rzeczywistych danych, i jednocześnie poprawiają generalizację modelu.

Do rozwiązania problemu przedstawionego w pracy wykorzystano publicznie dostępny zbiór zanonimizowanych danych tomograficznych. Obszar obrazowania obejmuje odcinek od szyi do miednicy. Zbiór danych zawiera tomografie komputerowe pochodzące z trzech ośrodków (Radl i in. 2022), wraz z referencyjnymi obrysami wykonanymi przez lekarzy. W konsekwencji uzyskano zbiór 56 trójwymiarowych obrazowań tomografii komputerowej i masek, określających położenie aorty i jej odgałęzień.

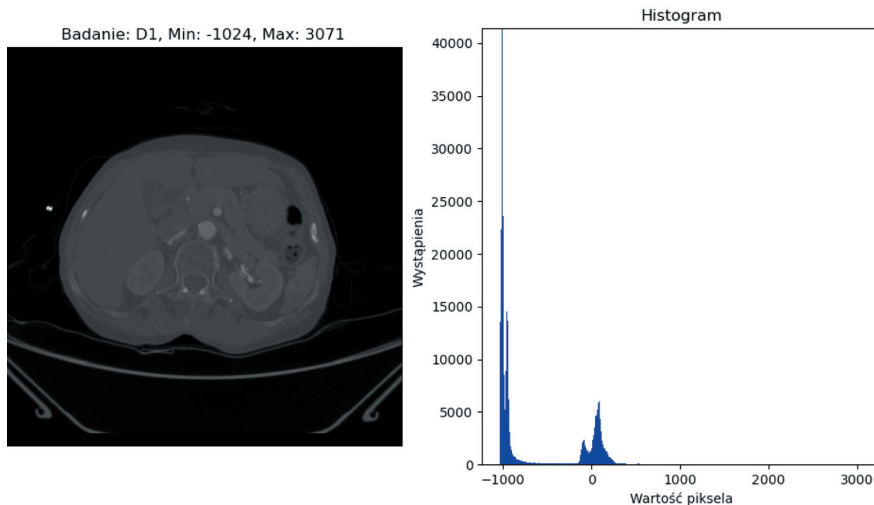
Maski zostały przygotowane w sposób półautomatyczny. Każdy przypadek znajduje się w osobnym folderze. Foldery zawierają plik *.nrrd* z danymi z przeprowadzonego

badania tomografii komputerowej, a także plik *.seg.nrrd* z maską z zaznaczoną aortą, używaną dalej w procesie uczenia. W tabeli 1 przedstawiono podstawowe informacje na temat danych tomograficznych wykorzystanych w pracy.

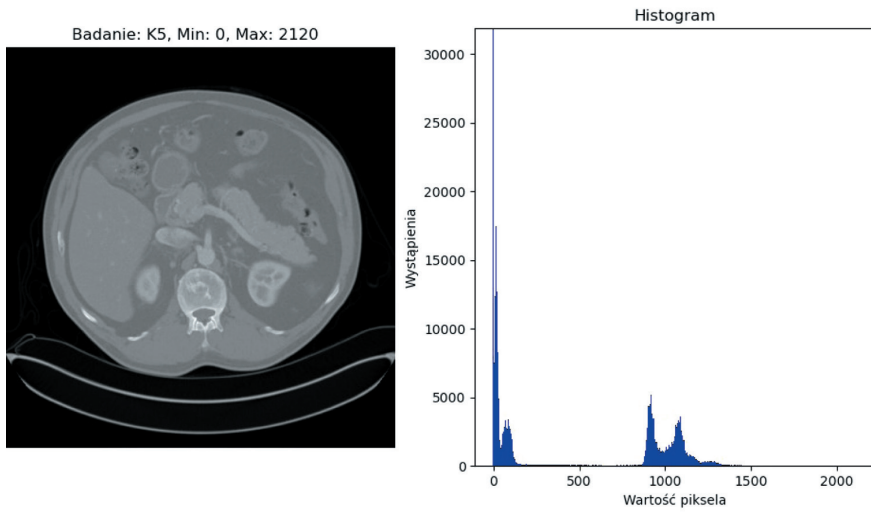
**Tabela 1**  
Zestawienie informacji o zbiorze danych (Radl i in. 2022)

Informacje o obrazach	KiTS	RIDER	Dongyang
Rozdzielczość x/y	512 × 512 pikseli	512 × 512 pikseli	512 × 666 pikseli
Liczba przekrojów względem osi	94/146/1059	260/1008/1140	122/149/251
Grubość przekroju	0,5/5/5 mm	0,625/0,625/2,5 mm	2/3/3 mm
Patologie	brak	rozwarstwienie aorty, tętniaki aorty brzusznej	brak
Liczba przypadków	20	18	18

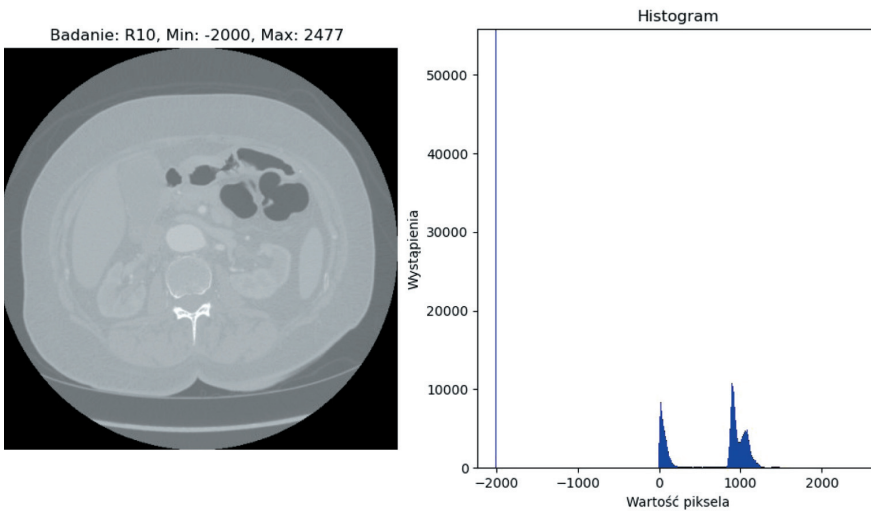
Dane z tomografii komputerowej powinny być zapisane w skali Hounsfielda. Na podstawie histogramów (rys. 1–4) wyznaczonych dla obrazowań zauważono, że przypadki w zbiorze różnią się wartością minimalną, a wartości dla tych samych struktur (np. aorty) różnią się w zależności od pochodzenia danych. Ponadto w przypadku zbioru RIDER obszar zawierający piksele spoza walca obrazowania miał przypisaną wartość  $-2000$  (poza skalą Hounsfielda) jako objętość nieobjętą akwizycją, co przedstawia rysunek 3.



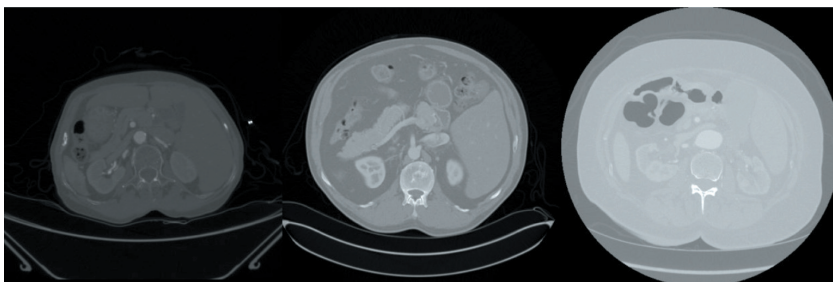
**Rys. 1.** Przekrój badania D1 oraz jego histogram



Rys. 2. Przekrój badania K5 oraz jego histogram



Rys. 3. Przekrój badania R10 oraz jego histogram



Rys. 4. Przekroje badań D1 (Dongyang), K5 (KiTS) oraz R10 (RIDER) przed przetwarzaniem wstępnym

W konsekwencji dla wartości wokseli obrazowań ze zbioru RIDER, o początkowej wartości  $-2000$ , zaproponowano zastąpienie jej wartością minimalną z obrazowania znalezionej poza tym obszarem.

Rozbieżność sposobu zapisu wartości wokseli wymagała ujednoczenia zakresu względem skali Hounsfielda przed rozpoczęciem uczenia sieci. W związku z tym przyjęto liniową transformację:

$$I_{HU} = \alpha I_{raw} + \beta \quad (2)$$

gdzie:

$I_{HU}$  – wartość woksela w skali Hounsfielda,

$I_{raw}$  – pierwotna, zapisana w pliku wartość woksela,

$\alpha$  – współczynnik skalujący (odpowiada polu *RescaleSlope* w standardzie DICOM),

$\beta$  – przesunięcie (odpowiada polu *RescaleIntercept* w standardzie DICOM).

W nagłówkach plików *.nrrd* brakowało metadanych *RescaleSlope* i *RescaleIntercept*, dlatego obie wartości oszacowano ręcznie. Współczynniki  $\alpha$  oraz  $\beta$  dobrano tak, by charakterystyczne zakresy sygnału odpowiadały typowym poziomom HU dla aorty, płuc i trzonów kręgow. Dzięki temu obrazy ze zbiorów KiTS i RIDER zostały przeskalowane do rzeczywistych jednostek HU (dla zbioru RIDER po zamianie wartości  $-2000$  na wartość minimalną), w przypadku obu zbiorów  $\alpha = 1$ ,  $\beta = -1024$ , natomiast dane z Dongyang były już poprawnie reprezentowane w skali Hounsfielda.

Otrzymane po tym ujednoczeniu wartości wokseli dla aorty stały się zbliżone, co nie tylko umożliwiło późniejsze okienkowanie, ale także poprawiło jakość procesu uczenia sieci.

W pracy wykorzystano trzy zbiory danych różniące się m.in. rozdzielczością przestrzenną. W konsekwencji konieczne było dostosowanie rozmiarów obrazów do jednej przyjętej wartości. Zdecydowano się ujednoczyć rozdzielczość przekrojów do wymiarów  $512 \times 512$  pikseli. Jest to standardowa wartość stosowana w tomografii komputerowej (Balmer i in. 2022). Zbiory danych RIDER oraz KiTS od początku miały docelową rozdzielczość, zbiór Dongyang należało dopasować.

Obrazy ze zbioru Dongyang mają wymiary  $512 \times 666$  pikseli. W celu dopasowania rozmiarów danych zdecydowano się wykadrować tomografie z powyższego zbioru tak, aby ustandaryzować rozdzielczość wszystkich obrazowań. Upewniono się, że w usuwanym fragmencie nie znajdują się istotne dane. Jeśli w obszarze, który ma zostać usunięty, występują wartości inne niż 0, wówczas program o tym informuje. W obszarze nie znajdowały się informacje istotne pod względem procesu uczenia, wobec czego można było wykadrować obrazy bez utraty informacji.

Do danych wyrażonych w skali Hounsfielda zastosowano okienkowanie wyrażone wartościami domykającymi przedział. Badanie wykazało istotny wpływ okienkowania

na trening modelu, a także usunięcie artefaktów powstałych w trakcie wykonywania badania. Na podstawie wizualizacji danych, po zastosowaniu różnych wartości okienkowania do ostatecznego rozwiązania użyto przedziału  $[-50; 600]$ .

Dane używane do trenowania sieci powinny być znormalizowane. W nnU-Net domyślna normalizacja dla tomografii komputerowej polega na progowaniu intensywności do percentyli 0,5 i 99,5 wyznaczonych na podstawie danych ze zbioru treningowego oraz na standaryzacji (z-score) z użyciem globalnej średniej i odchylenia standardowego. W tej pracy wyłączono domyślnie oferowaną normalizację intensywności, a zastosowano normalizację polegającą na transformacji wejściowego przedziału po okienkowaniu do przedziału  $[0; 1]$ , na podstawie wzoru (3).

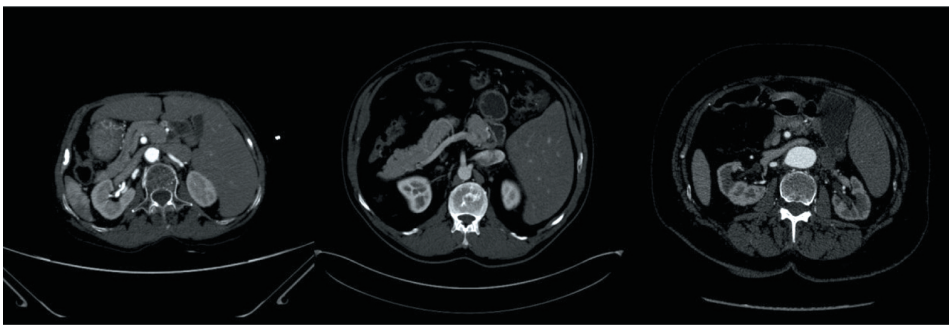
$$I'(x) = \frac{I(x) - I_{\min.}}{I_{\max} - I_{\min.}} \quad (3)$$

gdzie:

- $I(x)$  – oryginalna wartość woksela w punkcie  $x$ ,
- $I_{\min.}$  – wartość minimalna w całym obrazie po okienkowaniu,
- $I_{\max}$  – wartość maksymalna w całym obrazie po okienkowaniu,
- $I'(x)$  – wartość normalizowana, należąca do przedziału  $[0; 1]$ .

Normalizacja pozwala na ograniczenie problemu eksplozji gradientu (ang. *gradient explosion*) podczas procesu uczenia, a tym samym na jego stabilizację. Normalizacja danych, zgodnie z oczekiwaniami, przyniosła poprawę wyników treningu (Cabana i Lago-Fernández 2021).

Rysunek 5 przedstawia przykładowe przekroje z każdego ze zbioru danych po zastosowaniu opisanych metod przetwarzania wstępnego.



Rys. 5. Przekroje badań D1, K5 oraz R10 po przetworzeniu wstępnym

Domyślnie stosowanym w nnU-Net optymalizatorem jest SGD (ang. *stochastic gradient descent*). Innym, bardzo często spotykanym, jest Adam (ang. *adaptive moment estimation*). Chociaż obie metody dążą do tego samego celu, jakim jest minimalizacja

funkcji kosztu, ich szybkość zbieżności może być różna, zależnie od zadania (Dogo i in. 2018). Z tego powodu warto przeprowadzić test dla konkretnych danych.

Na podstawie zebranych danych stwierdzono, że wynik dla obu metod optymalizacji jest zbliżony, różnica uzyskanych oboma metodami współczynników Dice'a wyniosła 0,01. W opisywanym przypadku nie zdecydowano się na dalsze testowanie, ponieważ obie metody dają zbliżone wyniki. W konsekwencji wybrano domyślnie proponowany algorytm SGD, wraz z domyślną początkową wartością współczynnika uczenia wynoszącą 0,01.

Obrazy tomograficzne składają się z przekrojów (ang. *slices*), w których wartości najczęściej przypisane są do kwadratowych pikseli. Dodatkowym parametrem jest grubość przekroju (ang. *thickness*), która umożliwia skonstruowanie przestrzennego woksela danych. Wymiary przestrzenne woksela reprezentują rzeczywisty rozmiar obszaru, którego uśredniona wartość jest zapisywana jako wartość woksela (Tino i in. 2021). Taki sposób zapisu danych pozwala przekształcić wielkość woksela na rzeczywiste jednostki odległości, a tym samym dokonywać pomiarów wielkości zadanych struktur. Objętość przestrzenna, którą zajmuje woxsel, jest różna w zależności od sposobu przeprowadzania badania. Oznacza to, że obrazy mające tę samą rozdzielczość mogą prezentować zupełnie różne rzeczywiste wymiary, skąd wynika potrzeba ujednoczenia wymiarów woksela przed rozpoczęciem treningu.

Rozdzielczość obrazowania oraz wymiary woksela są używane do obliczenia rzeczywistych rozmiarów, przez co są ze sobą powiązane według wzoru (4).

$$L_k = N_k \cdot \Delta_k \quad (4)$$

gdzie:

$L_k$  – fizyczna długość obrazu wzdłuż osi  $k$  (mm),

$N_k$  – liczba wokseli (rozdzielczość) wzdłuż osi  $k$ ,

$\Delta_k$  – wielkość pojedynczego woksela (ang. *spacing*) wzdłuż tej osi (mm).

Framework nnU-Net zachowuje oryginalną rozdzielczość obrazów, natomiast dla każdego dopasowuje wymiary woksela. Dla danych treningowych trójwymiarowych, niezależnie dla każdej osi, wyszukiwana jest mediana wymiarów woksela. Wartość ta staje się nową wartością docelową dla wszystkich obrazowań. W przypadku treningu dla sieci stosującej dane dwuwymiarowe (przekroje) wybierane są przekroje względem tych osi, które mają najwyższą rozdzielczość. Jeśli wszystkie osie są izotropowe (dla każdej osi rozdzielczość jest taka sama), przekroje wybierane są względem dwóch ostatnich osi (Huang i in. 2022, Isensee i in. 2021). Interpolacja wykonywana jest z wykorzystaniem funkcji sklepanych trzeciego stopnia (Isensee i in. 2019b).

Dobór optymalnej architektury sieci do problemu stanowi duże wyzwanie, dlatego też twórcy frameworku nnU-Net zautomatyzowali to zadanie. Zaimplementowano sys-

tem ekspertowy, który najpierw wyznacza *dataset fingerprint* (kluczowe właściwości zbioru), a następnie na podstawie zestawu reguł automatycznie konfiguruje sposób przetwarzania danych, w tym wstępne przetwarzanie, architekturę sieci U-Net, parametry treningu oraz końcowe przetwarzanie.

Ręcznie podzielono wszystkie 56 obrazowań na zbiór treningowy oraz testowy w stosunku 0,9 względem przypadków, co dało 48 przypadków treningowych (85,7%) oraz 8 testowych (14,3%). Do zbioru treningowego dodano badania R1, R7 oraz R8, pochodzące ze zbioru RIDER, zawierające obrazowania z rozwarstwienia aorty, a także badanie R6, zawierające obrazowanie z tętniakami aorty. Do zbioru testowego przydzielono badania R2 oraz R5, również należące do zbioru RIDER, zawierające obrazowania z rozwarstwienia aorty. Algorytm tworzący podziały służące do realizacji walidacji krzyżowej podzielił zbiór treningowy względem liczby obrazowań w sposób przedstawiony w tabeli 2.

**Tabela 2**  
Podział walidacji krzyżowej względem liczości zbiorów

	Obrazowania treningowe	Obrazowania walidacyjne
Model 1	38	10
Model 2	38	10
Model 3	38	10
Model 4	39	9
Model 5	39	9

Do oceny stworzonego rozwiązania użyto metryk Dice'a Score (5), IoU (6) oraz odległości Hausdorffa (7). Dzięki zastosowaniu różnorodnych metryk oraz oceny jakościowej możliwa jest dokładna i kompleksowa ocena modelu. Każda z metryk uwypukla inne aspekty, co przekłada się na pełniejsze zrozumienie możliwości stworzonego narzędzia oraz identyfikację jego mocnych i słabych stron, a więc umożliwiła określenie dalszego kierunku rozwoju.

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \quad (5)$$

gdzie:

$A$  – obraz maski wyliczony jako predykcja modelu,

$B$  – obraz prawidłowej maski, dostarczony wraz ze zbiorem,

$|A \cap B|$  – część wspólna obu masek,

$|A| + |B|$  – suma arytmetyczna pikseli należących do masek  $A$  oraz  $B$ .

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (6)$$

gdzie:

$|A \cap B|$  – część wspólna obu masek,

$|A \cup B|$  – suma logiczna pikseli należących do masek  $A$  oraz  $B$ .

$$H(A, B) = \max\left(\sup_{a \in A} \inf_{b \in B} d(a, b), \sup_{b \in B} \inf_{a \in A} d(b, a)\right) \quad (7)$$

gdzie:

$d(a, b), d(b, a)$  – odległość euklidesowa pomiędzy punktami  $a$  oraz  $b$ ,

$\inf_{b \in B} d(a, b)$  – infimum odległości (najmniejsza możliwa odległość) pomiędzy punktem  $a$  ze zbioru  $A$  a wszystkimi punktami  $b$  ze zbioru  $B$  (analogicznie dla  $\inf_{a \in A} d(b, a)$ ),

$\sup_{a \in A} \inf_{b \in B} d(a, b)$  – supremum odległości (największa możliwa odległość) dla każdego punktu  $a$  ze zbioru  $A$  ze zbioru wcześniej obliczonych infimów (analogicznie dla  $\sup_{b \in B} \inf_{a \in A} d(b, a)$ ),

$\max()$  – funkcja zwracająca maksimum z obu obliczonych wartości.

Oprócz klasycznej odległości Hausdorffa raportuje się również metrykę Hausdorff95, czyli 95. percentyl rozkładu odległości między granicami (powierzchniami) dwóch masek segmentacji. Metryka ta ogranicza wpływ pojedynczych wartości odstających.

### 3. Wyniki

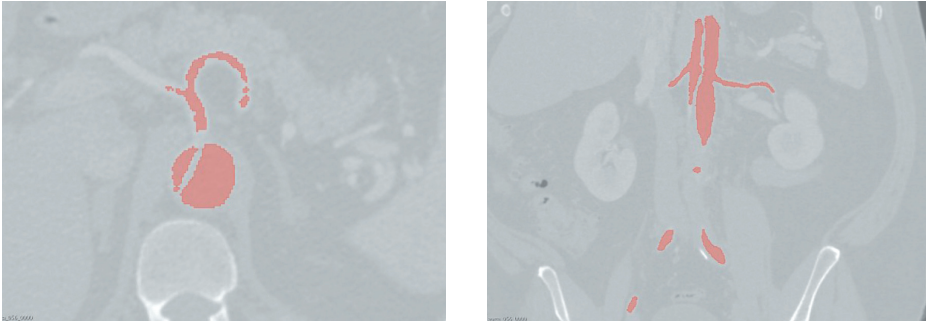
Wyniki przewidywań zbioru testowego z pięciu modeli oraz wyniki uśrednione przez nnU-Net przedstawia tabela 3.

**Tabela 3**  
Metryki dla poszczególnych modeli

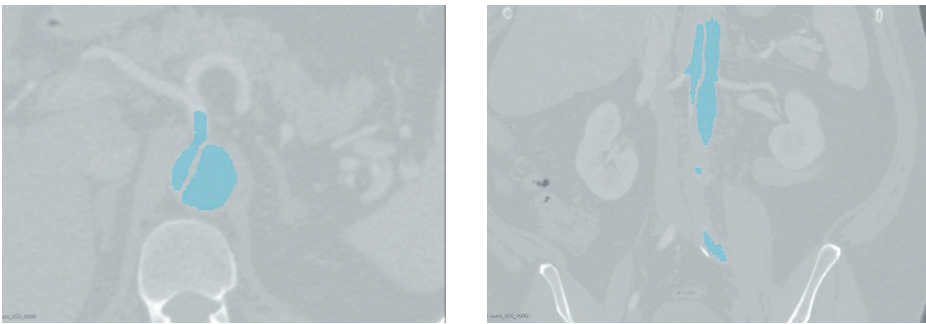
	Model 1	Model 2	Model 3	Model 4	Model 5	Model uśredniony
Dice	0,919	0,895	0,921	0,916	0,908	0,921
IoU	0,855	0,824	0,858	0,849	0,837	0,858
Hausdorff95 [mm]	6,233	12,467	6,013	13,784	12,770	5,870
Hausdorff [mm]	152,876	131,122	133,5	117,431	130,914	114,208

W celu ewaluacji końcowego rozwiązania zwizualizowano powstałe segmentacje oraz porównano je z prawidłowymi maskami, co umożliwiło ocenę problematycznych miejsc.

Jednym ze znanych problemów rodziny architektur U-Net jest trudność z odwzorowaniem szczegółów (Ahmed i in. 2025, Valanarasu i in. 2020). W przypadku używanego zbioru danych sytuacja występuje przy mniejszych odgałęzieniach, co pokazują rysunek 6 oraz rysunek 7.



**Rys. 6.** Prawidłowa segmentacja odgałęzienia

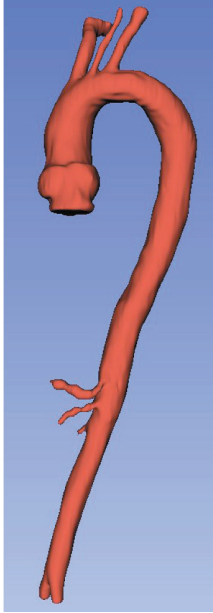
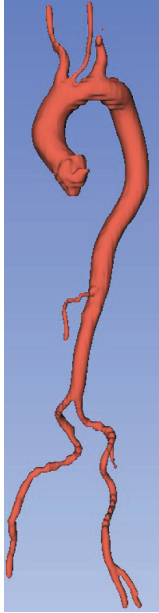
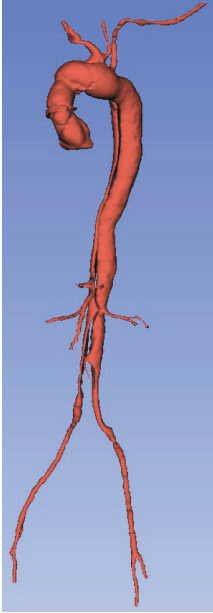

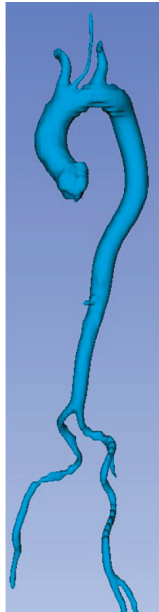
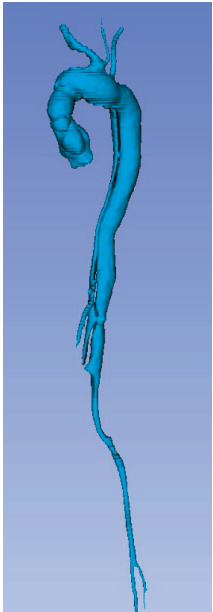


**Rys. 7.** Segmentacja utworzona przez sieć

Tabela 4 przedstawia trójwymiarowe modele, stworzone na podstawie dostarczonych ze zbiorem danych (czerwone) oraz wygenerowanych przez sieć (niebieskie) masek segmentacji. Na ich podstawie można potwierdzić przypuszczenie, że model dobrze odnajduje duże struktury, natomiast znacznie gorzej przewiduje odgałęzienia.

Wartość współczynnika Dice'a na poziomie 0,921 oraz IoU wynoszące 0,858 oznaczają stosunkowo wysoką dokładność przewidywań, co zweryfikowano również na podstawie wizualizacji wyników jako trójwymiarowych modeli. Odległość Hausdorffa wskazuje na obecność elementów odstających, natomiast 95. percentyl o wartości 5,870 mm oznacza stosunkowo wysoką jakość przewidywań w dominującej objętości struktury. Stworzone rozwiązanie dobrze realizuje zadanie, jakim jest segmentacja aorty. Jeśli celem predykcji jest również wyodrębnienie odgałęzień oraz mniejszych naczyń, zachodzi potrzeba zweryfikowania zastosowanego podejścia. Ograniczeniem metody jest brak możliwości odwzorowania małych naczyń odchodzących od aorty.

**Tabela 4**  
Porównanie modeli 3D prawidłowych oraz predykcji

	Dongyang	KiTS	RIDER
Model 3D stworzony na podstawie dostarczonych masek			
Model 3D stworzony na podstawie predykcji sieci			

#### 4. Podsumowanie oraz wnioski

Zaprezentowane rozwiązanie miało za zadanie realizować segmentację aorty z obrazów tomografii komputerowej. Otrzymane wyniki potwierdzają skuteczność metody w zakresie segmentacji aorty, nawet w trudniejszych przypadkach, takich jak jej rozwarstwienie. Można stwierdzić, że segmentacja większych, jednolitych struktur działa zadowalająco, w przeciwieństwie do odgałęzień i drobniejszych naczyń.

Framework nnU-Net jest odpowiednim narzędziem, jeżeli celem segmentacji jest jednolita, dużych rozmiarów struktura. Ze względu na uproszczenie wielu etapów tworzenia rozwiązania opartego na sieciach neuronowych, a także automatyczne wykrywanie rodzaju danych treningowych i dopasowywanie do nich architektury, nnU-Net jest przydatnym i wygodnym w użyciu rozwiązaniem. Kluczowym elementem jest prawidłowe przeprowadzenie przetwarzania wstępnego umożliwiającego generalizację modelu dla danych z różnych ośrodków.

Dalszym obszarem badań może być opracowanie rozwiązania, które wychwytuje globalny kontekst i jednocześnie dba o szczegóły. Możliwe jest również wykorzystanie mechanizmu uwagi, zastosowanie augmentacji poprawiających kontrast drobnych struktur czy zastosowanie funkcji kosztu minimalizującej błędy topologiczne.

#### Literatura

- Ahmed R.M., Iltaf A., Elmann M., Zhao G., Li H., Du Y., Li B., Zhou S., 2025, *MSA-UNet3+: Multi-scale attention UNet3+ with new supervised prototypical contrastive loss for coronary DSA image segmentation*, arXiv:2504.05184v3. <https://arxiv.org/pdf/2504.05184v3>.
- Balmer B.D., Blüthgen C., Bäessler B., Martini K., Huber F.A., Ruby L., Schönenberger A., Frauenfelder T., 2022, *Influence of CT image matrix size and kernel type on the assessment of HRCT in patients with SSC-ILD*, *Diagnostics*, vol. 12(7), 1662. <https://doi.org/10.3390/diagnostics12071662>.
- Cabana A., Lago-Fernández L.F., 2021, *Backward gradient normalization in deep neural networks*, arXiv:2106.09475. <https://doi.org/10.48550/arXiv.2106.09475>.
- Dogo E.M., Afolabi O.J., Nwulu N.I., Twala B., Aigbavboa C.O., 2018, *A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks*, [w:] 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), IEEE, s. 92–99. <https://doi.org/10.1109/CTEMS.2018.8769211>.
- Gücük A., Üyetürk U., 2014, *Usefulness of Hounsfield unit and density in the assessment and treatment of urinary stones*, *World Journal of Nephrology*, vol. 3(4), s. 282–286. <https://doi.org/10.5527/wjn.v3.i4.282>.
- Hsieh J., Flohr T., 2021, *Computed tomography recent history and future perspectives*, *Journal of Medical Imaging*, vol. 8(5), 052109. <https://doi.org/10.1117/1.JMI.8.5.052109>.

- Huang Z., Wang H., Ye J., Niu J., Tu C., Yang Y., Du S., Deng Z., Gu L., He J., 2022, *Revisiting nnU-Net for iterative pseudo labeling and efficient sliding window inference*, [w:] Ma J., Wang B. (Eds.), *Fast and Low-Resource Semi-supervised Abdominal Organ Segmentation*, FLARE 2022. Lecture Notes in Computer Science, vol. 13816, s. 178–189. [https://doi.org/10.1007/978-3-031-23911-3\\_16](https://doi.org/10.1007/978-3-031-23911-3_16).
- Isensee F., Jäger P.F., Kohl S.A.A., Petersen J., Maier-Hein K.H., 2019a, *Automated design of deep learning methods for biomedical image segmentation*, arXiv:1904.08128v2. <https://doi.org/10.48550/arXiv.1904.08128>.
- Isensee F., Petersen J., Kohl S.A.A. Jäger P.F., Maier-Hein K.H., 2019b, *nnU-Net: Breaking the spell on successful medical image segmentation*, arXiv:1904.08128v1. <https://arxiv.org/abs/1904.08128v1>.
- Isensee F., Jaeger P.F., Kohl S.A.A., Petersen J., Maier-Hein K.H., 2021, *nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation*, *Nature Methods*, vol. 18(2), s. 203–211. <https://doi.org/10.1038/s41592-020-01008-z>.
- Isensee F., Wald T., Ulrich C., Baumgartner M., Roy S., Maier-Hein K., Jäger P., 2024, *nnU-Net revisited: A call for rigorous validation in 3D medical image segmentation* arXiv:2404.09556v2. <https://doi.org/10.48550/arXiv.2404.09556>.
- Jhala V.M., Thaker V.S., 2015, *X-ray computed tomography to study rice (Oryza sativa L.) panicle development*, *Journal of Experimental Botany*, vol. 66(21), s. 6819–6825. <https://doi.org/10.1093/jxb/erv387>.
- Radl L., Jin Y., Pepe A., Li J., Gsaxner C., Zhao F., Egger J., 2022, *AVT: Multicenter aortic vessel tree CTA dataset collection with ground truth segmentation masks*, *Data in Brief*, vol. 40, 107801. <https://doi.org/10.1016/j.dib.2022.107801>.
- Ronneberger O., Fischer P., Brox T., 2015, [w:] Navab N., Hornegger J., Wells W.M., Frangi A.F. (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI. 18<sup>th</sup> International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18*, Lecture Notes in Computer Science, vol. 9351, s. 234–241, arXiv:15.05.04597v1. <https://doi.org/10.48550/arXiv.1505.04597>.
- Tino R., Yeo A., Brandt M., Leary M., Kron T., 2021, *The interlace deposition method of bone equivalent material extrusion 3D printing for imaging in radiotherapy*, *Materials & Design*, vol. 199, 109439. <https://doi.org/10.1016/j.matdes.2020.109439>.
- Valanarasu J.M.J., Sindagi V.A., Hacihaliloglu I., Patel V.M., 2020, *Kiu-net: Towards accurate segmentation of biomedical images using over-complete representations*, [w:] Martel A.L., Abolmaesumi P., Stoyanov D., Mateus D., Zuluaga M.A., Zhou S.K., Racoceanu D., Joskowicz L. (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020, 23<sup>rd</sup> International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part IV*, Lecture Notes in Computer Science, vol. 12263, s. 363–373.