



Politechnika
Wrocławska

Zastosowanie algorytmów uczenia maszynowego do oceny degeneracyjnych zmian kręgosłupa lędźwiowego na podstawie obrazów MRI

Autor:

Jakub Skwierczyński

Promotor:

Dr inż. Łukasz Jeleń



HR EXCELLENCE IN RESEARCH

Plan prezentacji

- Cel i zakres pracy
- Przegląd literatury
- Opis zbioru danych
- Wstępne przetwarzanie danych
- Implementacja modeli
- Walidacja i ocena modeli
- Podsumowanie

Cel pracy

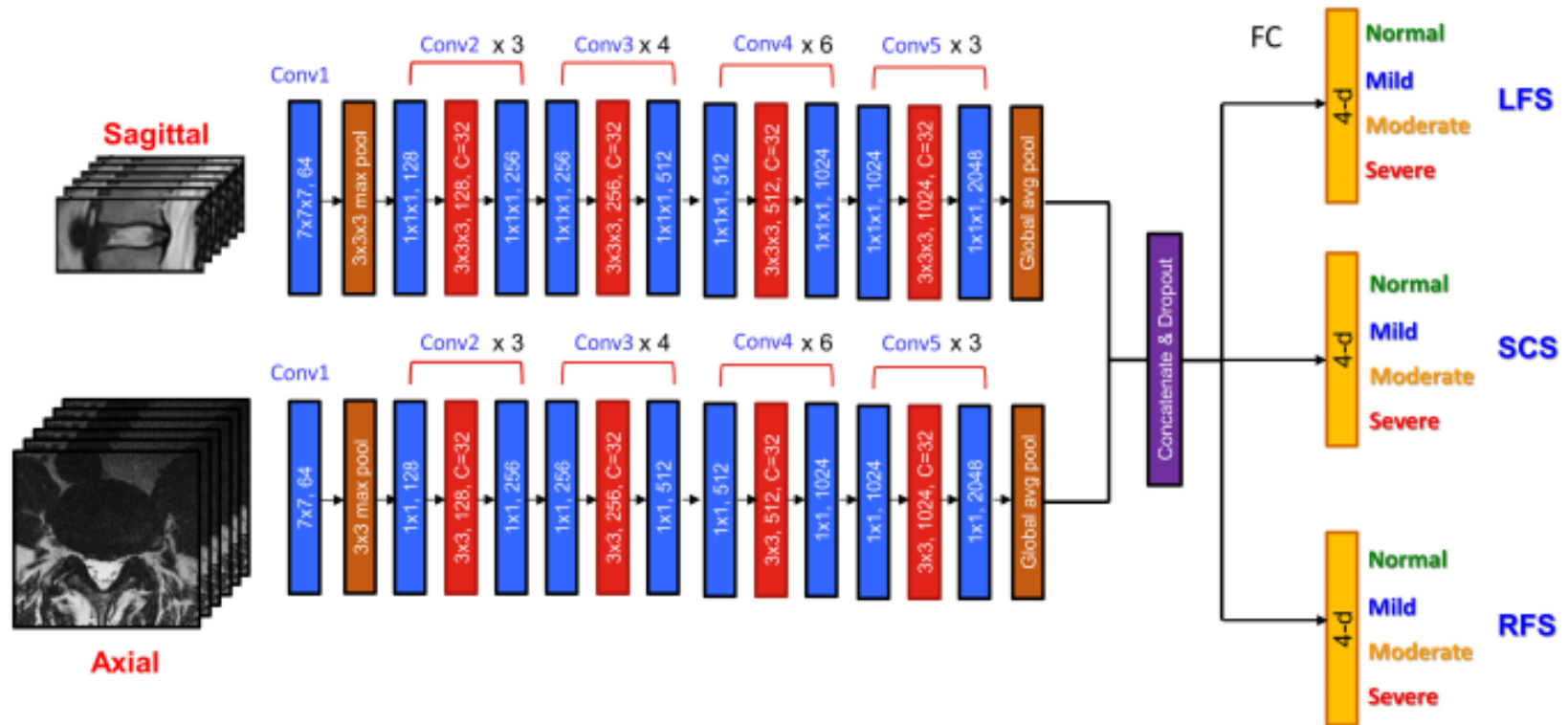
Celem pracy jest opracowanie i walidacja modeli uczenia maszynowego do automatycznej identyfikacji i oceny degeneracyjnych zmian kręgosłupa lędźwiowego na podstawie obrazów z rezonansu magnetycznego. Proponowane rozwiązanie ma szansę znacząco wspierać proces diagnostyczny zarówno w ortopedii jak i reumatologii.

Praca obejmuje eksplorację metod głębokiego uczenia i ich potencjału diagnostycznego w interpretacji obrazów z rezonansu magnetycznego.

Zakres pracy

1. Przegląd literatury dotyczącej degeneracyjnych zmian kręgosłupa oraz metod analizy obrazów medycznych.
2. Zebranie i analiza odpowiednich danych obrazowych.
3. Wstępna obróbka danych:
 - Segmentacja ROI
 - Przygotowanie zbioru danych treningowych, walidacyjnych i testowych.
4. Implementacja i strojenie parametrów wybranych modeli.
5. Walidacja modeli oraz analiza wyników.
6. Dyskusja i podsumowanie badań.
7. Opracowanie wniosków i synteza pracy.

Przegląd literatury



Multi-input, multi-task, and multi-class version of ResNeXt-50

Jen-Tang Lu, Stefano Pedemonte, Bernardo Bizzo, Sean Doyle, Katherine P. Andriole, Mark H. Michalski, R. Gilberto Gonzalez, Stuart R. Pomerantz W DeepSPINE: Automated Lumbar Vertebral Segmentation, Disc-level Designation, and Spinal Stenosis Grading Using Deep Learning

Przegląd literatury

Table 1: Class accuracy for stenosis grading.

Spinal Canal Stenosis (% , mean \pm std)				
Normal	Mild	Mod.	Severe	Class Avg.
78.7 \pm 5.4	59.6 \pm 2.0	61.3 \pm 2.1	82.7 \pm 5.6	70.6 \pm 2.1
Normal	Mild / Moderate		Severe	Class Avg.
79.7 \pm 3.3	83.7 \pm 3.4		77.7 \pm 1.5	80.4 \pm 1.6
Foraminal Stenosis (% , mean \pm std)				
Normal	Mild	Mod.	Severe	Class Avg.
80.5 \pm 0.3	61.3 \pm 5.8	52.0 \pm 6.0	74.8 \pm 3.1	67.1 \pm 2.2
Normal	Mild / Moderate		Severe	Class Avg.
79.6 \pm 0.8	84.2 \pm 0.7		70.5 \pm 0.8	78.1 \pm 0.4

Table 2: Comparison of models trained with axial input only, sagittal input only, and both inputs in class average accuracy (% , mean \pm std).

	Axial Only	Sagittal Only	Axial + Sagittal
Spinal Canal Stenosis	78.6 \pm 2.7	78.6 \pm 2.4	80.4 \pm 1.6
Foraminal Stenosis	76.6 \pm 2.5	74.3 \pm 1.7	78.1 \pm 0.4

Przegląd literatury

Table 3: Comparison of the proposed algorithm with the best published results on binary classification of spinal canal and foraminal stenosis. Performance metric is overall accuracy.

	Zhang et al. (2017)	Jamaludin et al. (2017a)	Ours
Type of Scan	Axial	Sagittal	Axial + Sagittal
Spinal canal stenosis (% , mean \pm std)			
L3-L4	87.2 \pm 3.2	94.7	94.5 \pm 0.7
L4-L5	85.1 \pm 3.4	85.9	95.3 \pm 0.2
L5-S1	87.5 \pm 3.3	93.7	99.1 \pm 0.5
Foraminal stenosis (% , mean \pm std)			
L3-L4	84.3 \pm 3.9	N/A	94.0 \pm 0.7
L4-L5	84.0 \pm 4.0	N/A	89.0 \pm 1.4
L5-S1	87.1 \pm 3.4	N/A	91.2 \pm 1.6

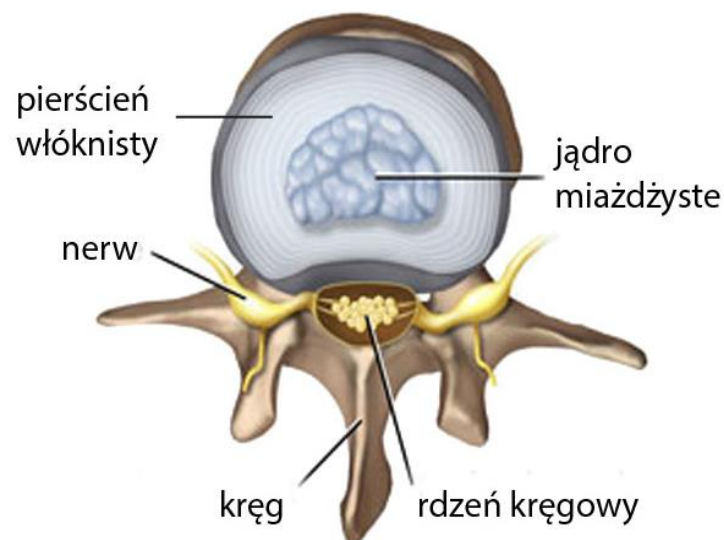
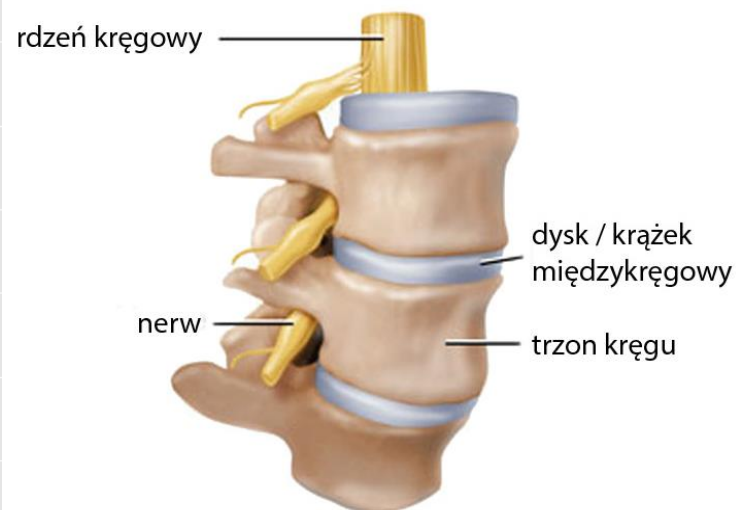
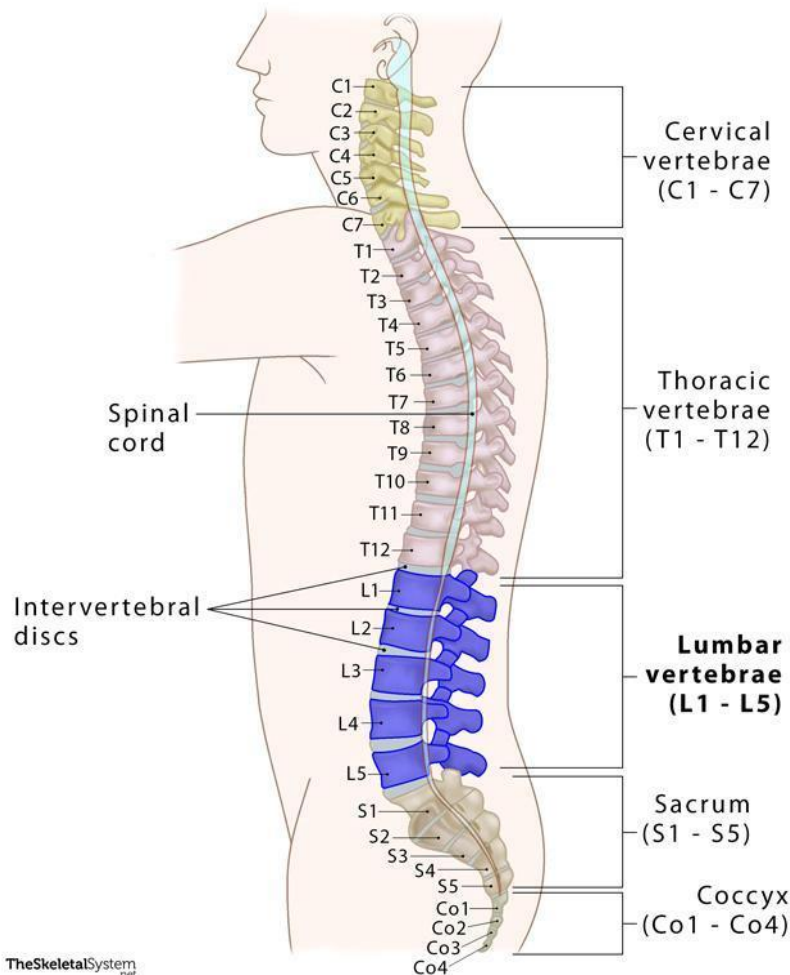
Opis zbioru danych

Zbiór danych pochodzi z konkursu RSNA 2024, w którym wykorzystano obrazy MRI odcinka lędźwiowego kręgosłupa w formacie DICOM. Dane obejmują:

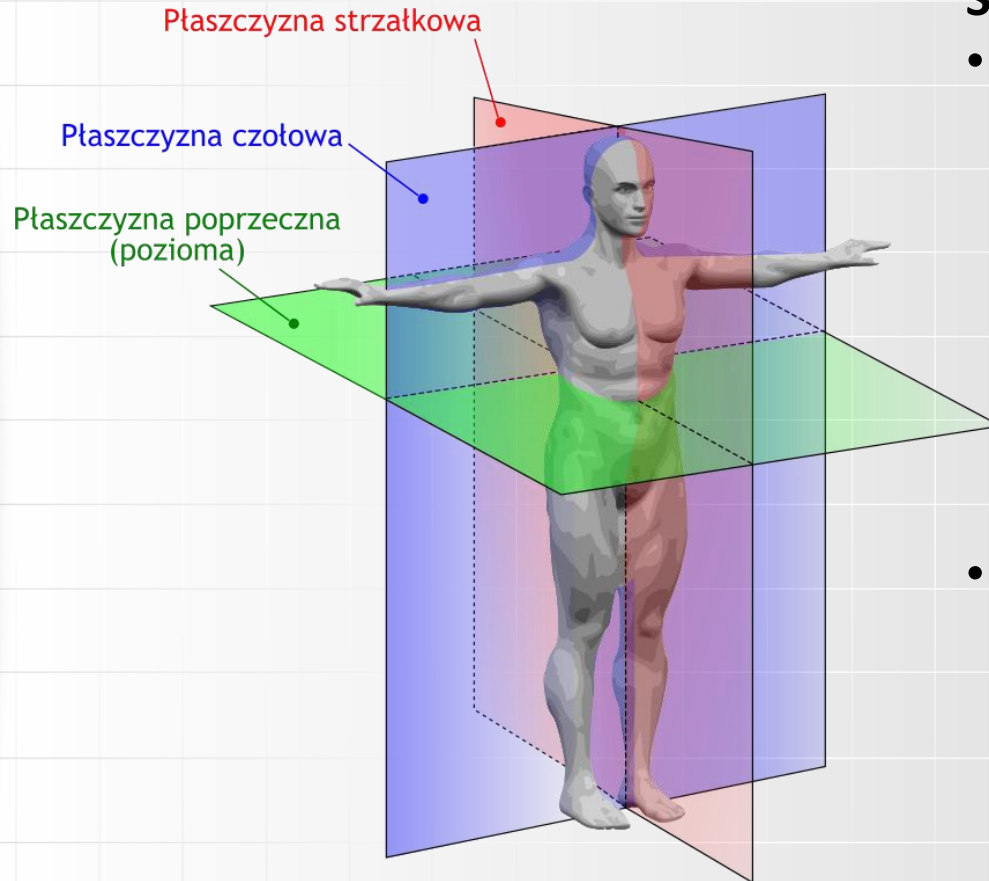
- **1975 badań**, łącznie 147 218 obrazów MRI,
- **5 klas** zmian degeneracyjnych:
 - **Zwężenie kanału kręgowego** (Spinal Canal Stenosis),
 - **Zwężenie otworów międzykręgowych** (Foraminal Narrowing) - lewego i prawego,
 - **Stenoza podchrzęstna** (Subarticular Stenosis) - lewa i prawa.
- Każde badanie zawiera oceny nasilenia zmian na poziomach od L1/L2 do L5/S1:
 - **Normalne/łagodne, Umiarkowane, Ciężkie.**

Odcinek lędźwiowy kręgosłupa

Lumbar Spine (Lumbar Vertebrae)



Sekwencje MRI i płaszczyzny



Sekwencje MRI:

- **T1-zależne** – sekwencje, które dają obrazy o wysokim kontraście między tkankami o różnej zawartości wody i tłuszczu. Na obrazach T1-zależnych tkanki tłuszczowe są jasne, a obszary z dużą ilością wody są ciemne.
- **T2-zależne** – w tych sekwencjach obszary z dużą ilością wody są jasne, a tkanki tłuszczowe są ciemne. Są one szczególnie przydatne w wykrywaniu patologii, takich jak obrzęki czy guzy.

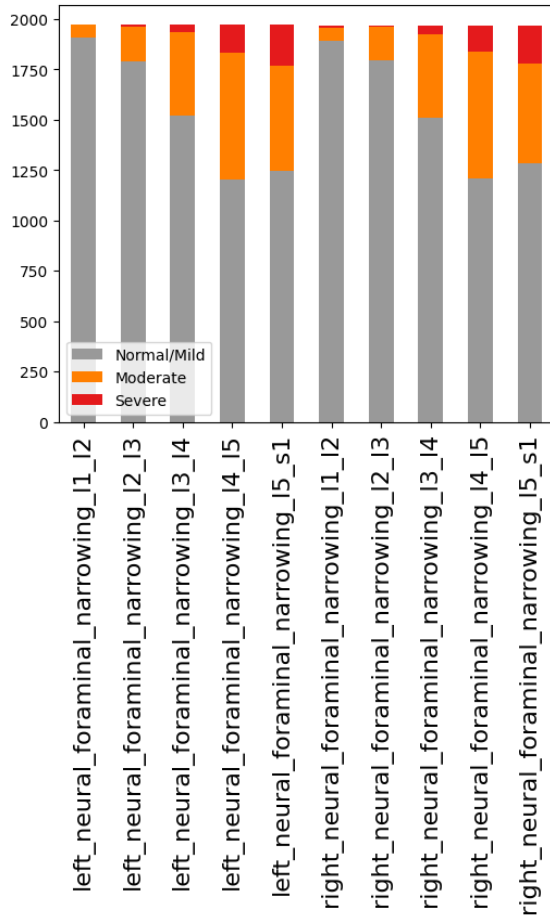
Sekwencje MRI i płaszczyzny



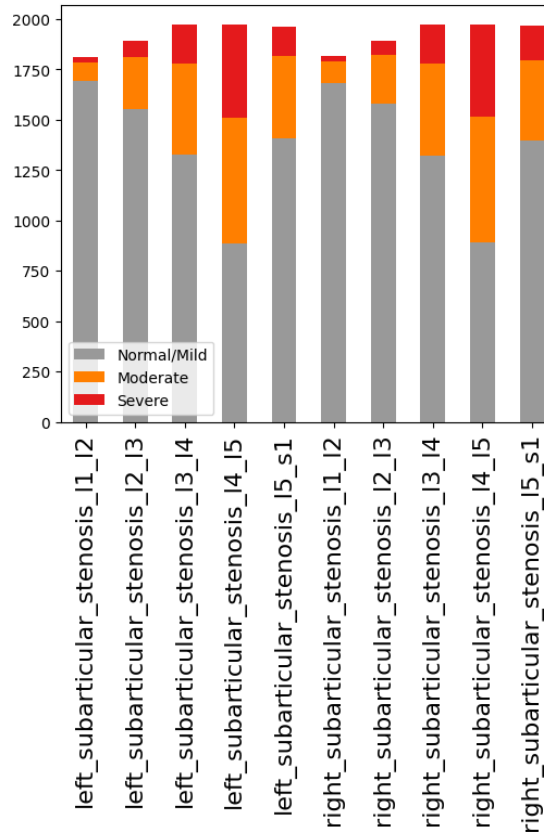
https://www.researchgate.net/publication/353949715_Feasibility_and_Implementation_of_a_Deep_Learning_MR_Reconstruction_for_TSE-Sequences_in_Musculoskeletal_Imaging/figures?lo=1&utm_source=google&utm_medium=organic

Rozkład zbioru danych

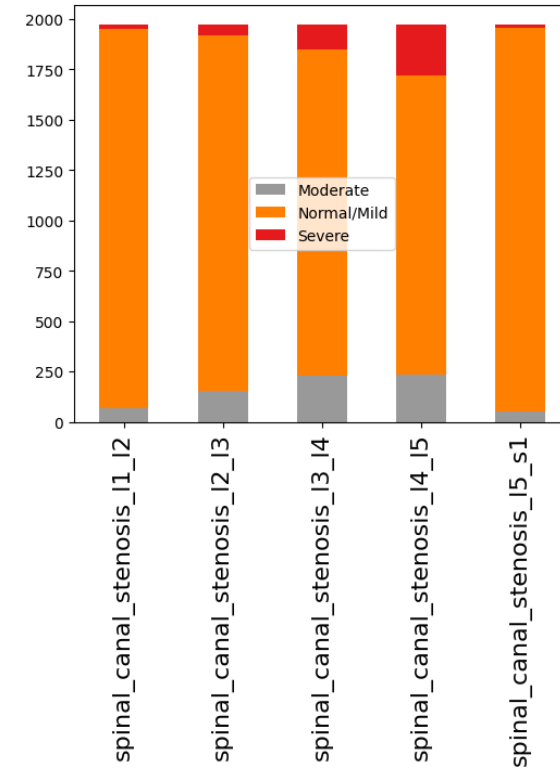
foraminal distribution



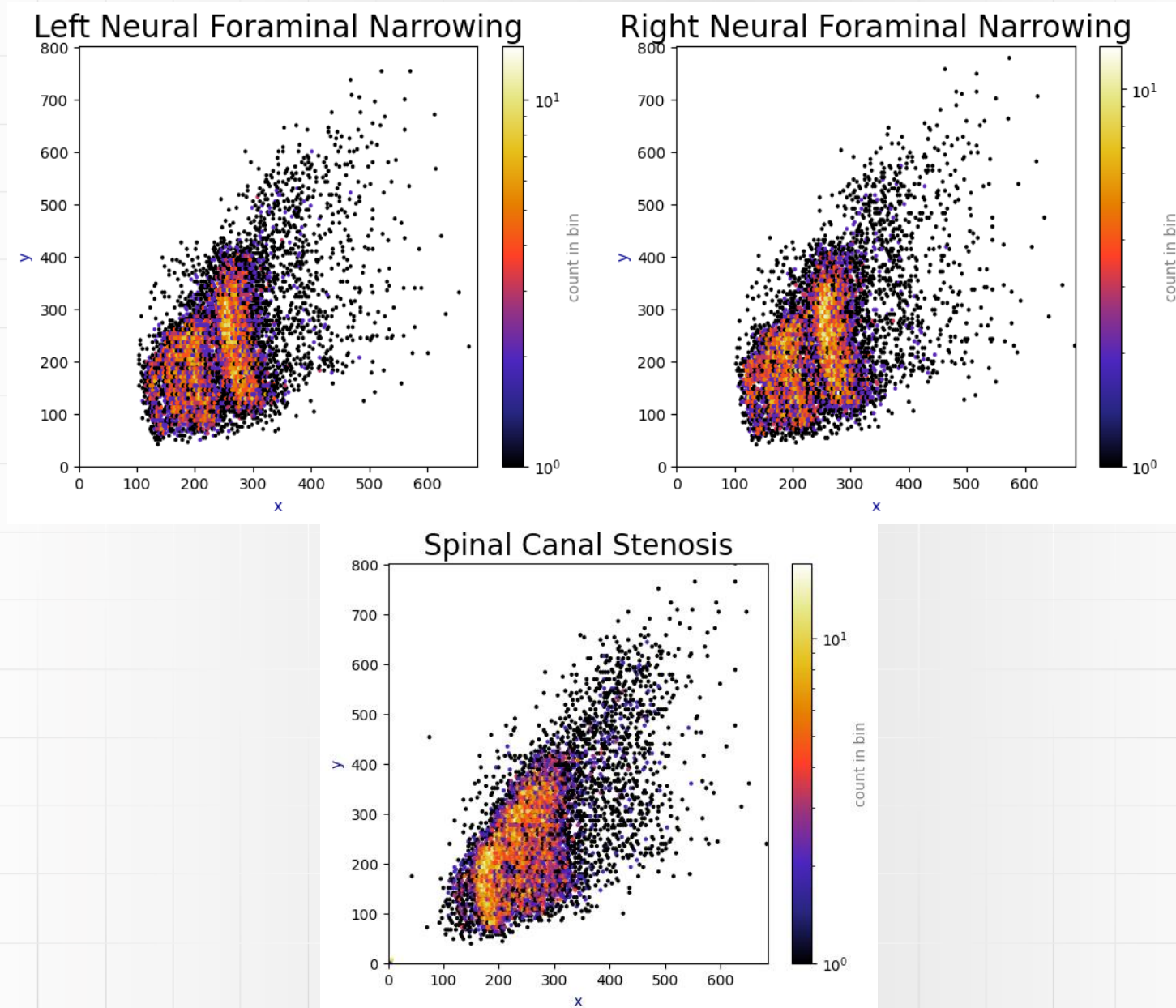
subarticular distribution



canal distribution

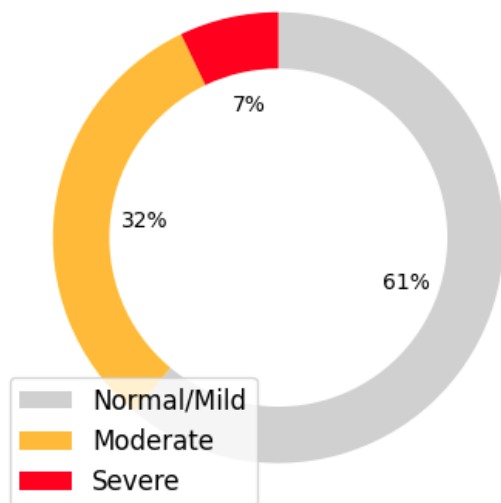


Lokalizacja schorzenia

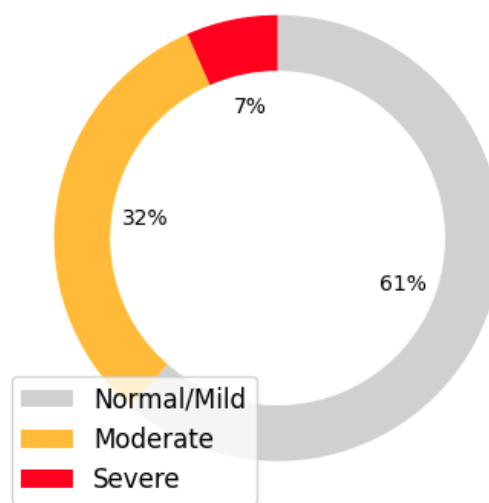


Rozkład nasilenia schorzenia

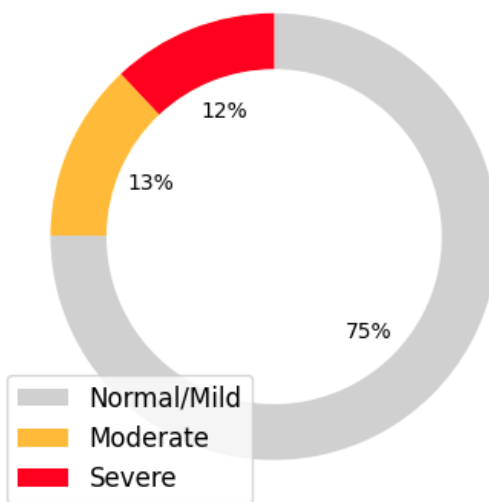
left_neural_foraminal_narrowing_I4_I5



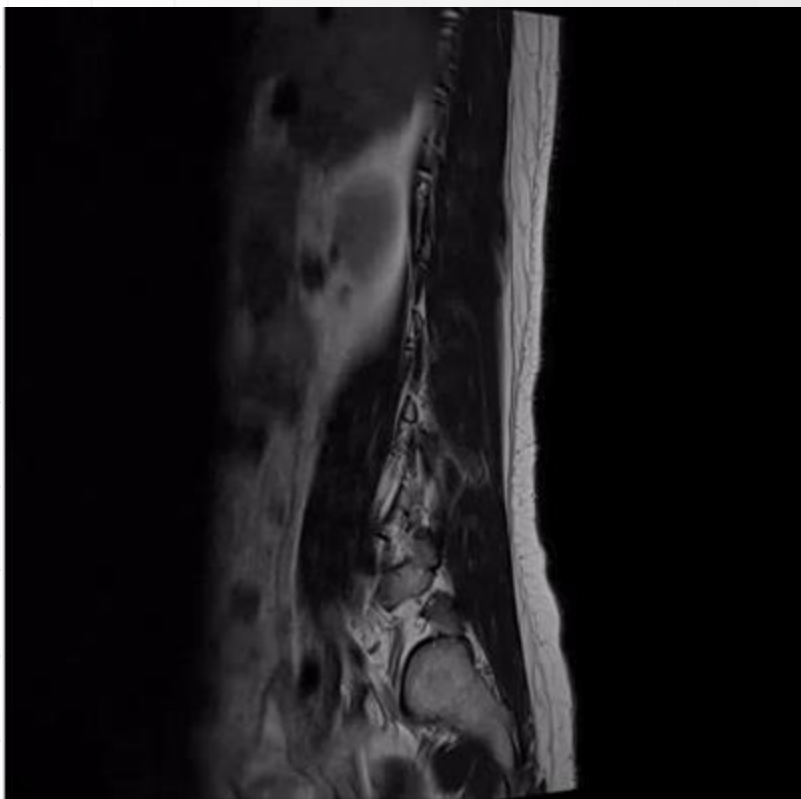
right_neural_foraminal_narrowing_I4_I5



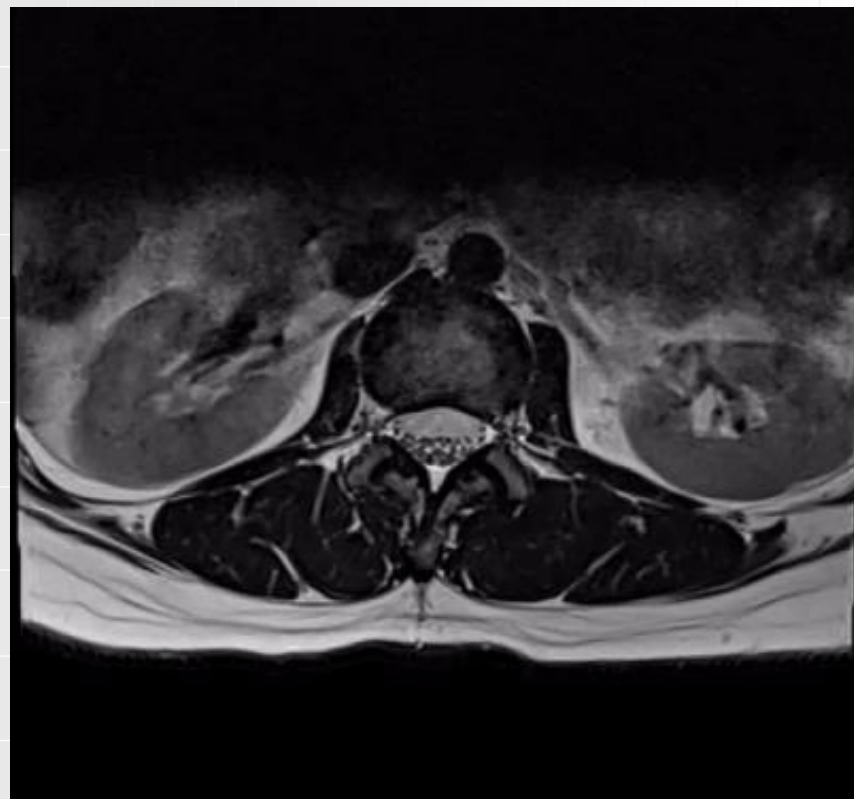
spinal_canal_stenosis_I4_I5



Przykładowe dane



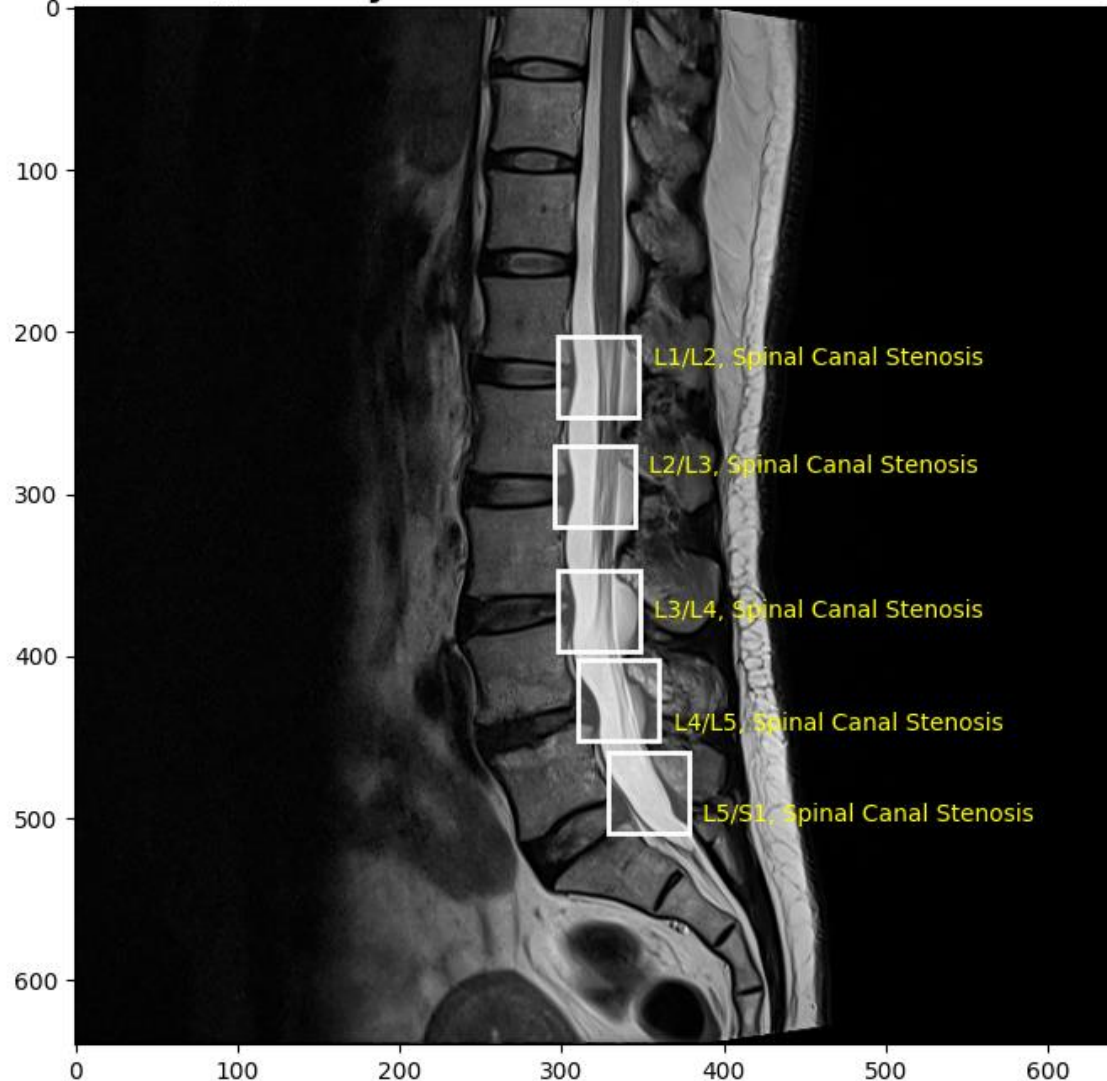
Sagittal T2/STIR



Axial T2

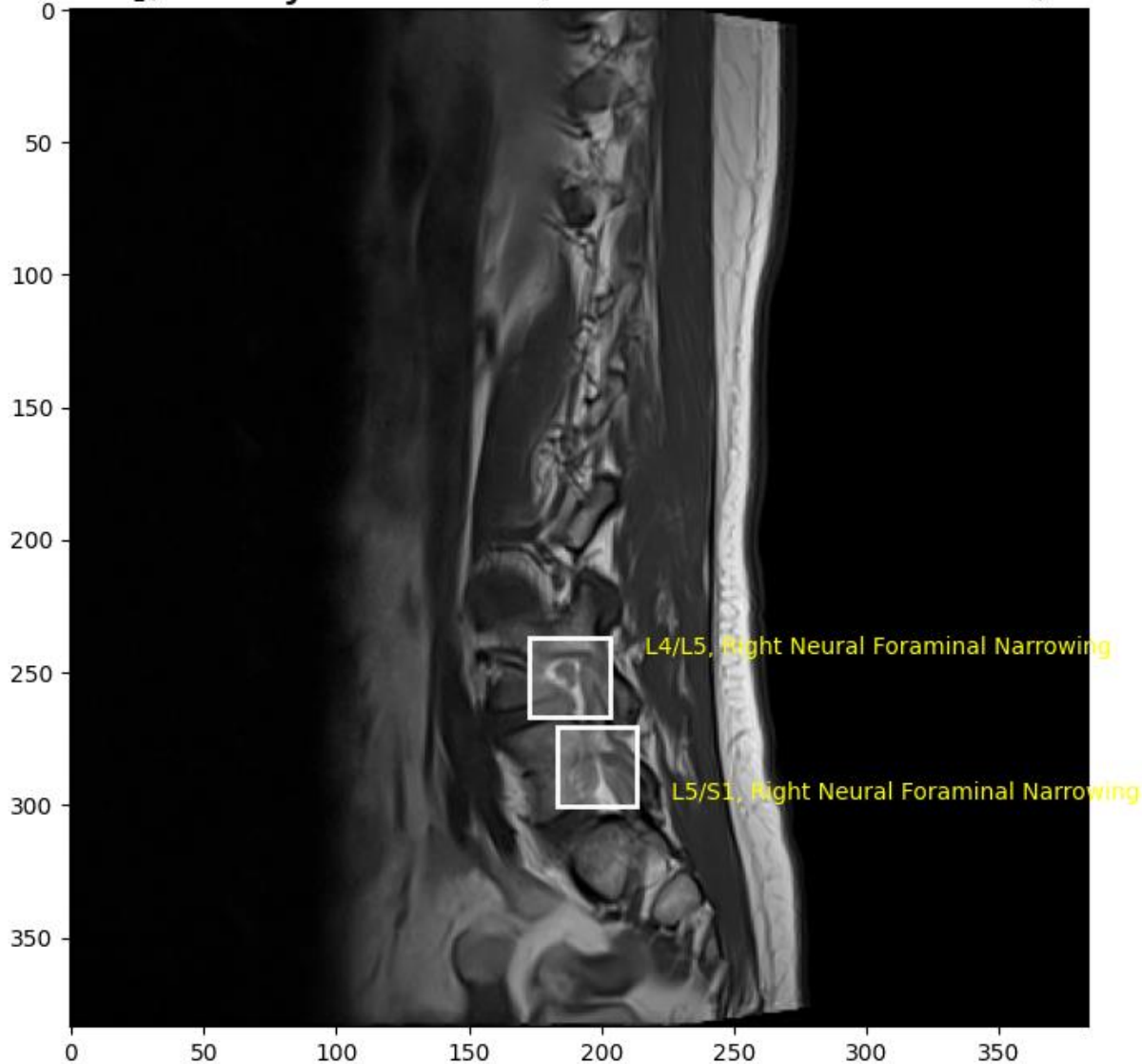
Przykładowe dane

['Sagittal T2/STIR'], Study: 4003253, Series: 702807833, Instance: 8



Przykładowe dane

['Sagittal T1'], Study: 4003253, Series: 1054713880, Instance: 4

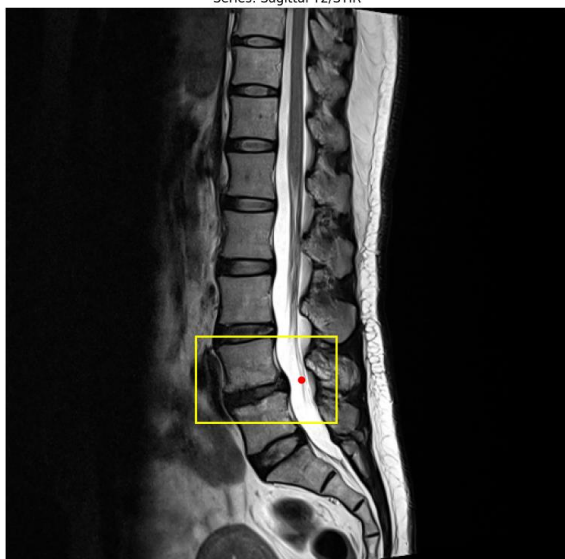


Wstępne przetwarzanie danych

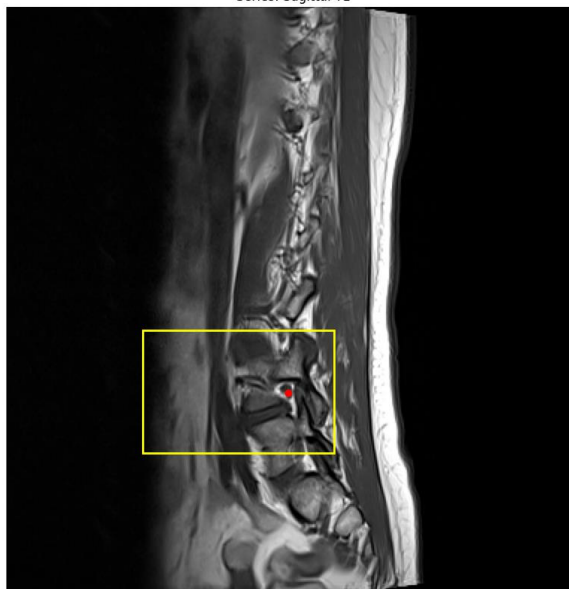
Condition: Spinal Canal Stenosis
 Series: Sagittal T2/STIR



Study ID: 4003253, Series ID: 702807833
 Condition: Spinal Canal Stenosis
 Series: Sagittal T2/STIR



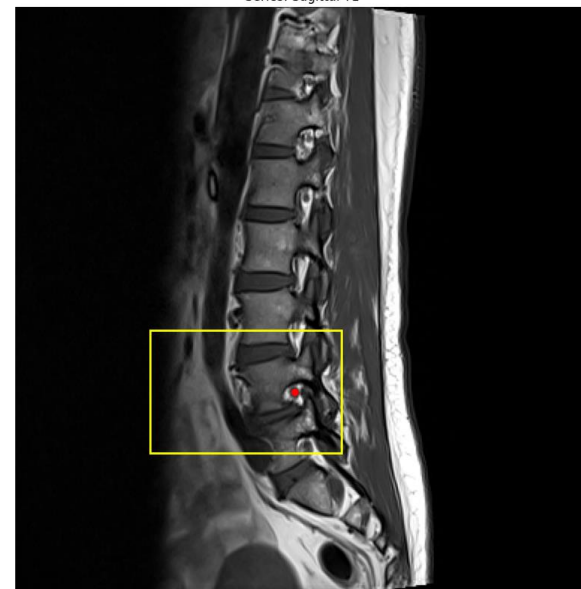
Condition: Right Neural Foraminal Narrowing
 Series: Sagittal T1



Study ID: 4003253, Series ID: 1054713880
 Condition: Right Neural Foraminal Narrowing
 Series: Sagittal T1



Condition: Left Neural Foraminal Narrowing
 Series: Sagittal T1

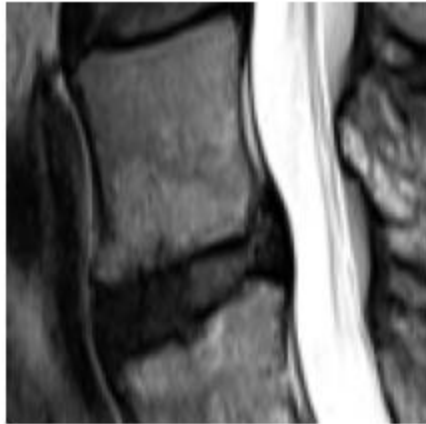


Study ID: 4003253, Series ID: 1054713880
 Condition: Left Neural Foraminal Narrowing
 Series: Sagittal T1

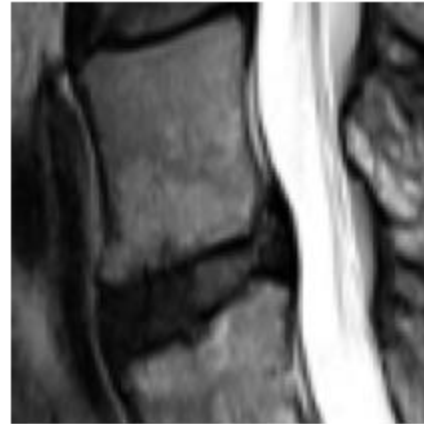


Wymiar tensorów

128x128

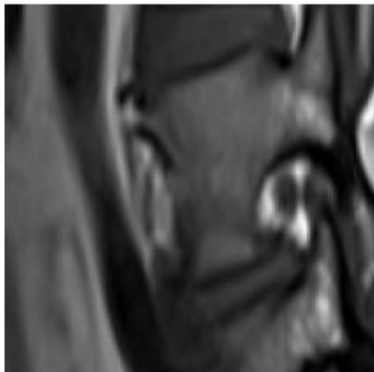


96x96

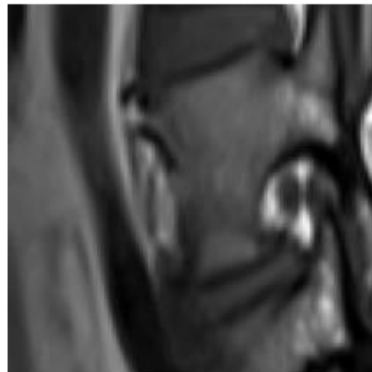


SCS

128x128



96x96



LNFN

128x128

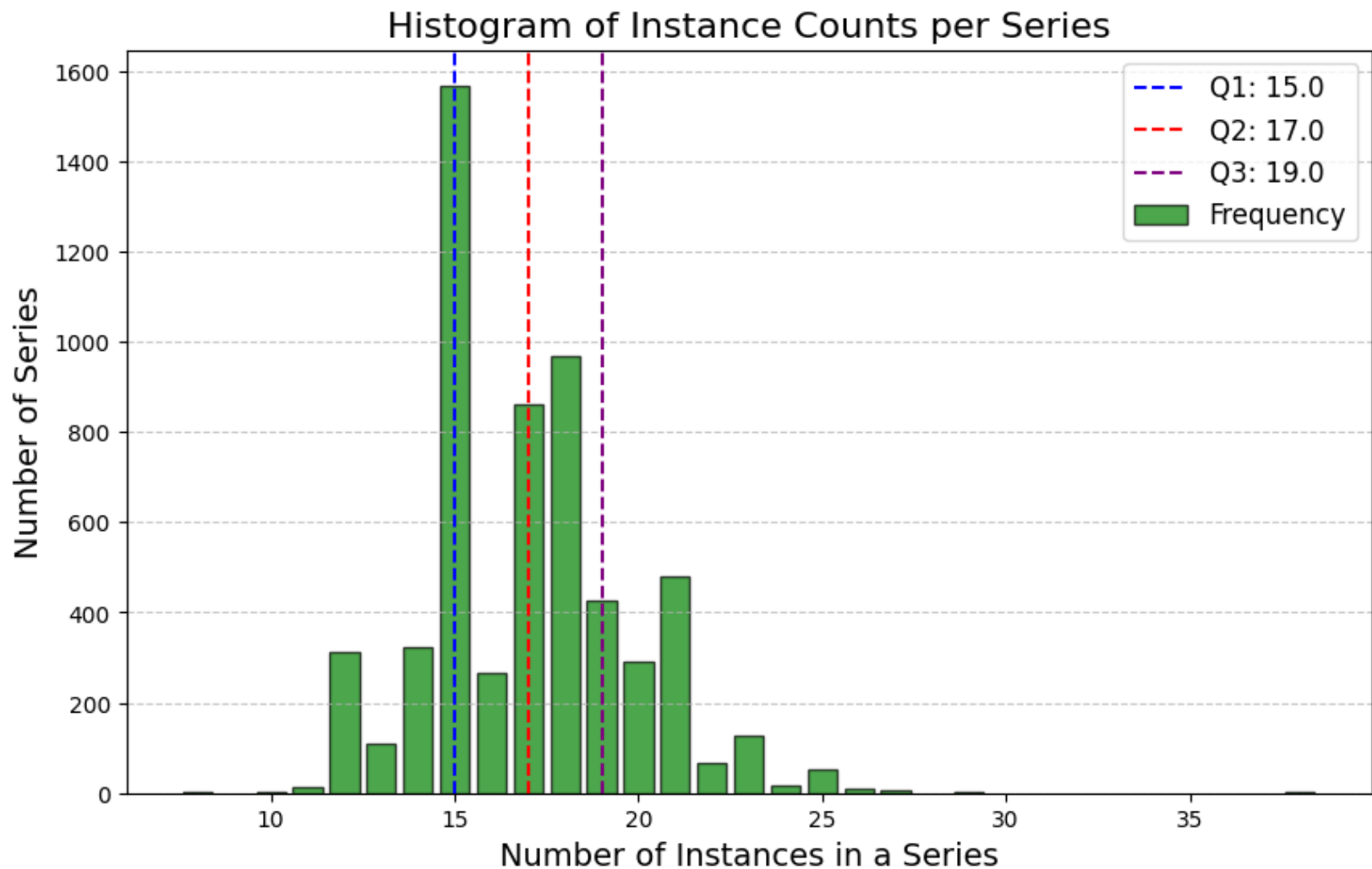


96x96



RNFN

Wymiar tensorów



Wymiar tensorów

```
git/ml_lumbar_mri/data/interim/  
├── full_series_128x128_15D  
├── full_series_128x128_17D  
├── full_series_128x128_19D  
├── full_series_128x128_1D  
├── full_series_128x128_25D  
├── full_series_96x96_15D  
├── full_series_96x96_17D  
├── full_series_96x96_19D  
├── full_series_96x96_1D  
├── full_series_96x96_25D  
├── target_window_128x128_3D_B1A1  
├── target_window_128x128_5D_B2A2  
├── target_window_96x96_3D_B1A1  
└── target_window_96x96_5D_B2A2
```

Wymiar tensorów

Slice 1



```
torch.Size([1, 128, 128])
```

Slice 1



Slice 2



Slice 3



Slice 4



Slice 5



```
torch.Size([5, 128, 128])
```

Slice 1



Slice 2



Slice 3



Slice 4



Slice 5



Slice 6



Slice 7



Slice 8



Slice 9



Slice 10



Slice 11



Slice 12



Slice 13



Slice 14

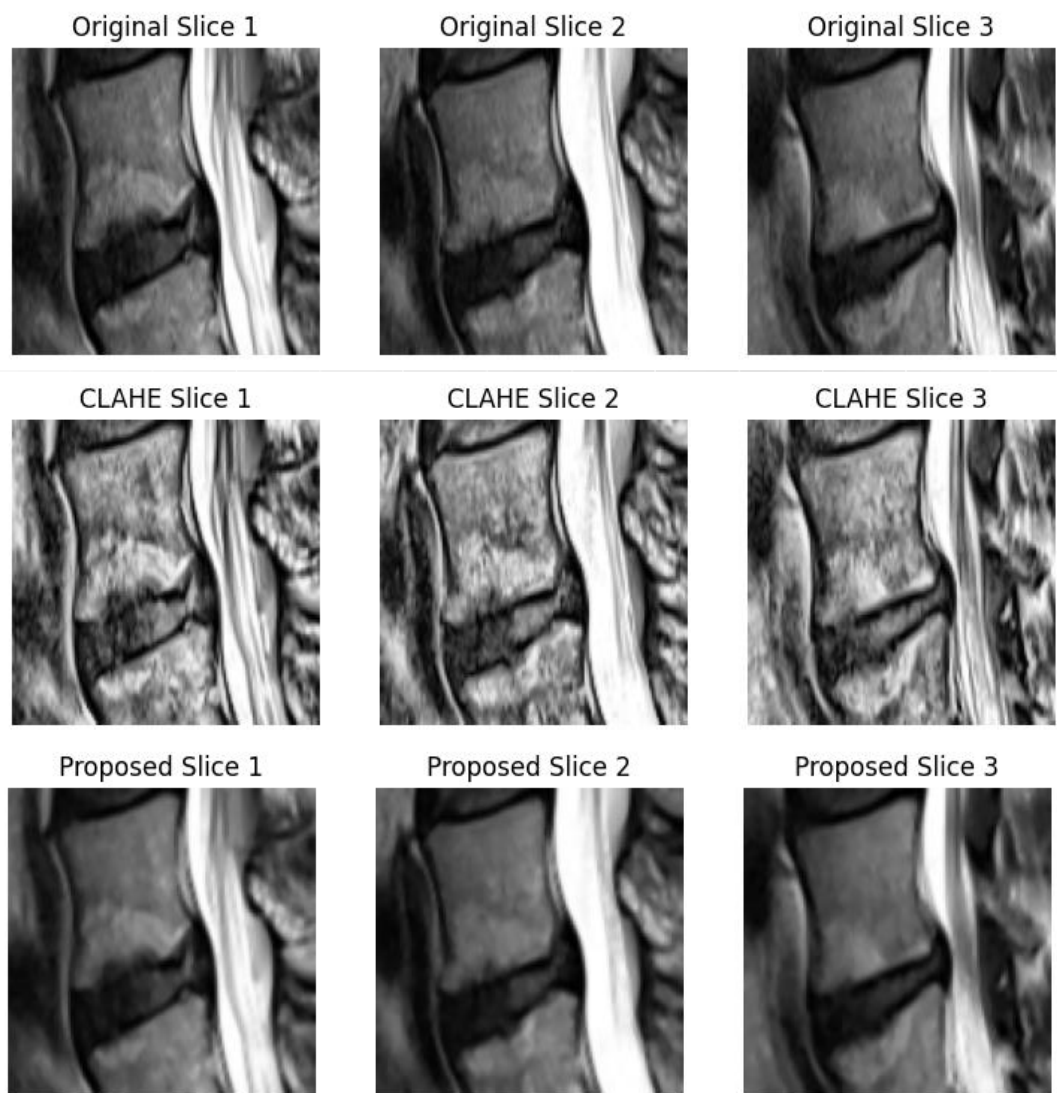


Slice 15



```
torch.Size([15, 128, 128])
```


Dalszy preprocessing



S. Suhas and C. R. Venugopal, "MRI image preprocessing and noise removal technique using linear and nonlinear filters," *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECOT)*, Mysuru, India, 2017, pp. 1-4, doi: 10.1109/ICEECOT.2017.8284595.

Implementacja modeli

Modele zostały zaimplementowane jako klasyfikatory wieloklasowe i wieloetykietowe, gdzie dla każdej patologii przewidywane są trzy poziomy nasilenia:

- **0 (Normal/Mild)** – stan normalny lub łagodny,
- **1 (Moderate)** – umiarkowane nasilenie,
- **2 (Severe)** – ciężkie nasilenie.

Dodatkowo, dla każdej patologii przeprowadzono klasyfikację binarną, w której uwzględniono dwa stany:

- **0 (Brak patologii lub łagodna)** – stan normalny lub łagodny,
- **1 (Umiarkowana lub ciężka)** – stan umiarkowany lub ciężki.

Cztery podejścia klasyfikacji

- **Single Multiclass** (single_multiclass)
 - Klasyfikacja tylko jednej choroby z trzema możliwymi poziomami zaawansowania (0 – Normal/Mild, 1 – Moderate, 2 – Severe).
- **Single Binary** (single_binary)
 - Klasyfikacja jednej choroby w trybie binarnym (brak zmian vs. obecne zmiany).
- **Multi Multiclass** (multi_multiclass)
 - Jednoczesna klasyfikacja trzech chorób (np. *Spinal Canal Stenosis, Left Neural Foraminal Narrowing, Right Neural Foraminal Narrowing*).
- **Multi Binary** (multi_binary)
 - Jednoczesna klasyfikacja tych samych trzech chorób, ale każda w trybie binarnym.

Przegląd architektur modeli

Zaimplementowane:

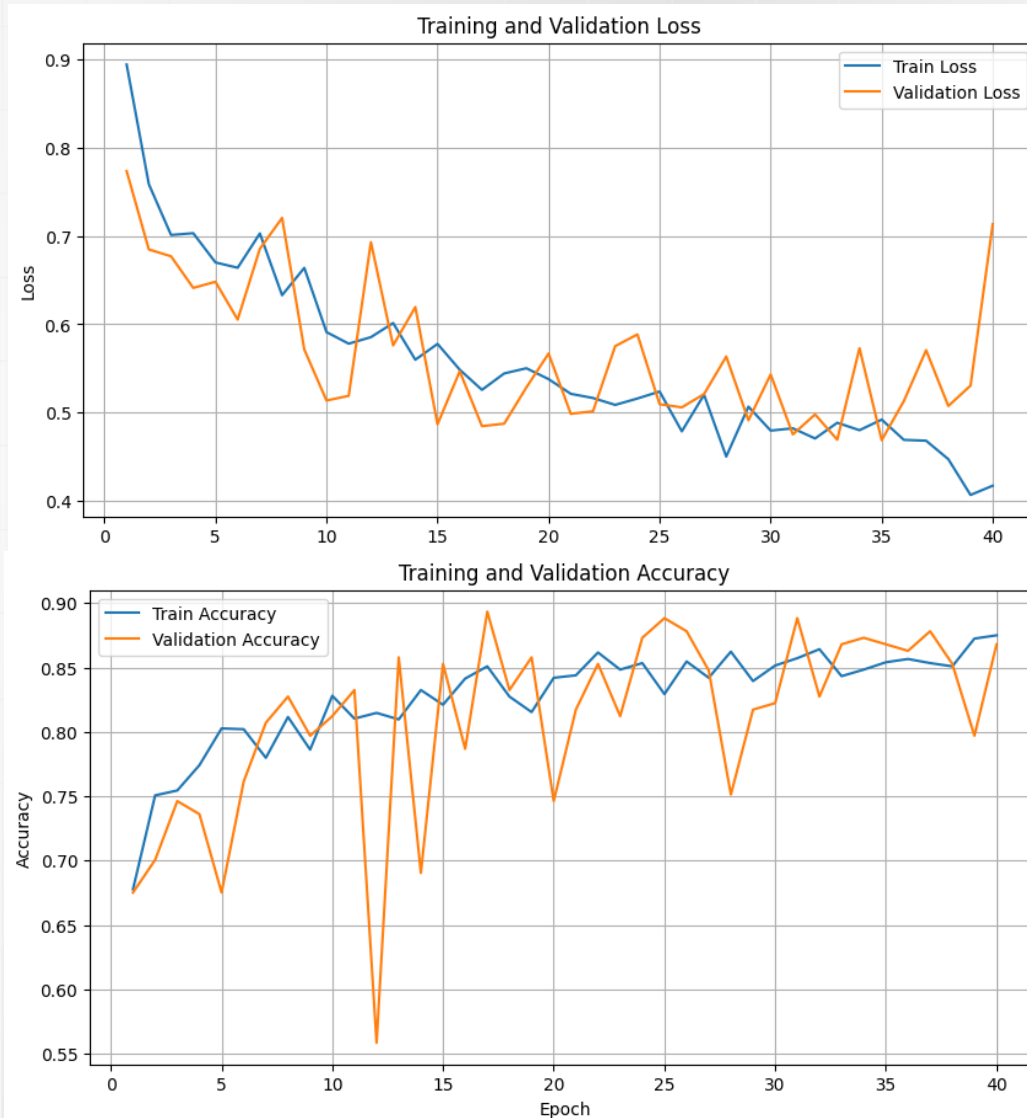
- 3D CNN
- 3D CNN + LSTM + Attention
- ResNet3D (wiele wariantów: 10, 18, 34, 50, 101)
- MedicalNet (ResNet3D)

Plany na przyszłość:

- ResNeXt
- DenseNet
- EfficientNet

Single Multiclass - SCS

3D CNN + LSTM + Attention



Single Multiclass - SCS

3D CNN + LSTM + Attention

Accuracy: 0.8737

Classification Report:

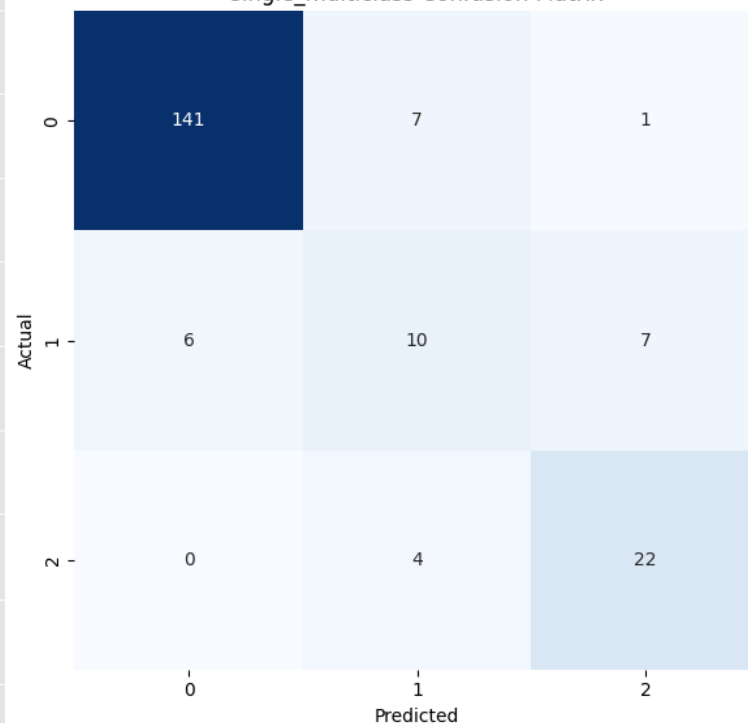
	precision	recall	f1-score	support
0	0.96	0.95	0.95	149
1	0.48	0.43	0.45	23
2	0.73	0.85	0.79	26

accuracy			0.87	198
macro avg	0.72	0.74	0.73	198
weighted avg	0.87	0.87	0.87	198

Confusion Matrix:

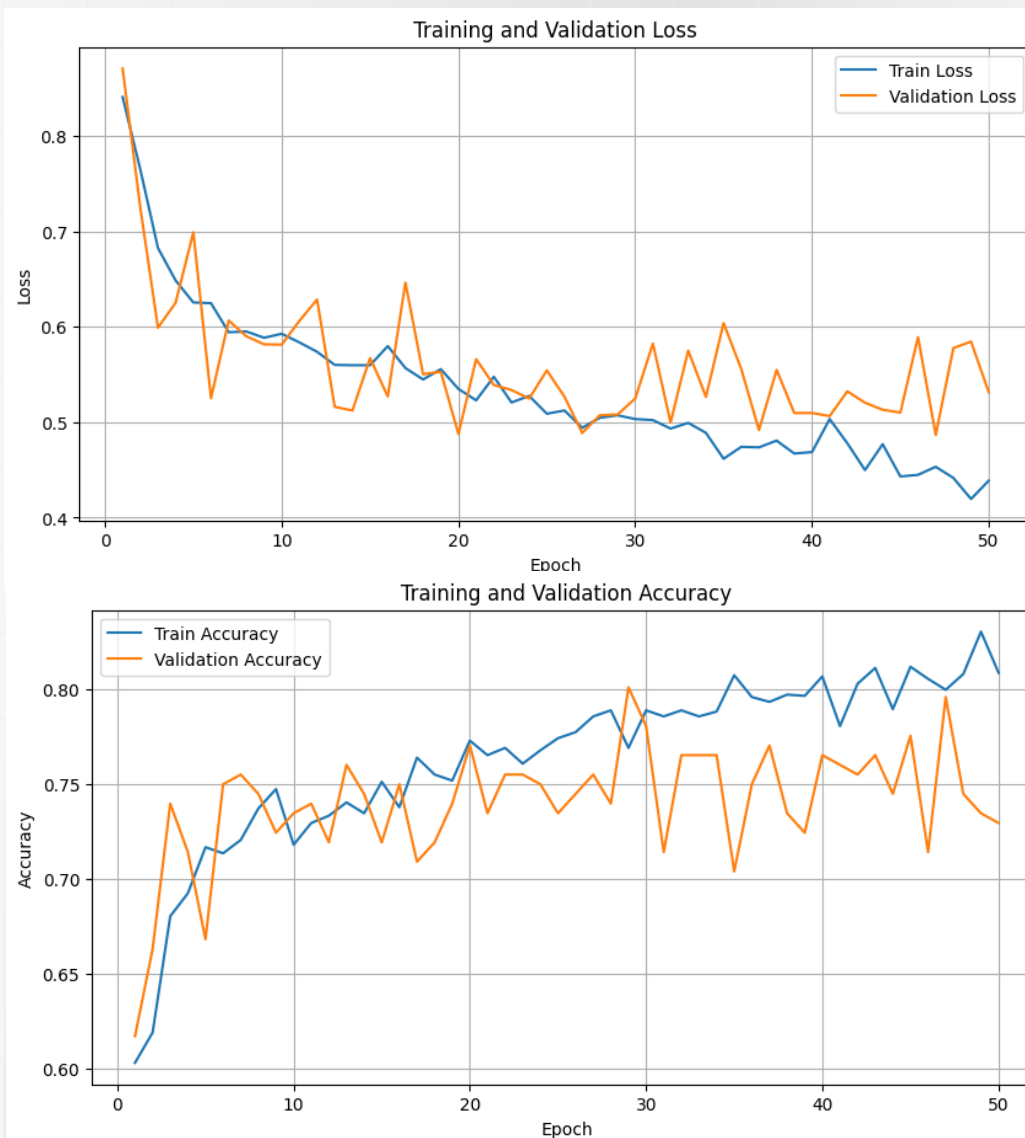
```
[[141  7  1]
 [ 6 10  7]
 [ 0  4 22]]
```

single_multiclass Confusion Matrix



Single Multiclass - RNFN

3D CNN + LSTM + Attention



Single Multiclass - RNFN

3D CNN + LSTM + Attention

Accuracy: 0.7677

Classification Report:

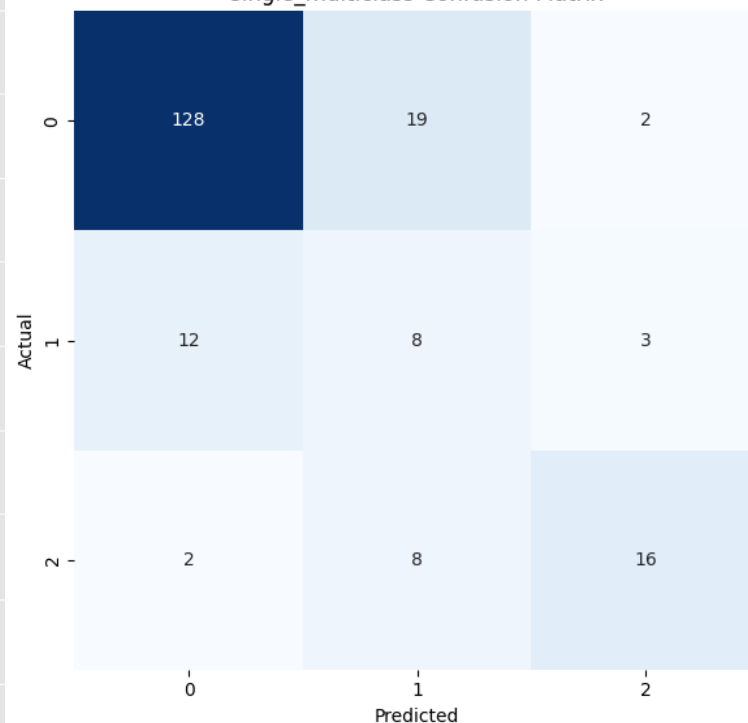
	precision	recall	f1-score	support
0	0.90	0.86	0.88	149
1	0.23	0.35	0.28	23
2	0.76	0.62	0.68	26

accuracy			0.77	198
macro avg	0.63	0.61	0.61	198
weighted avg	0.80	0.77	0.78	198

Confusion Matrix:

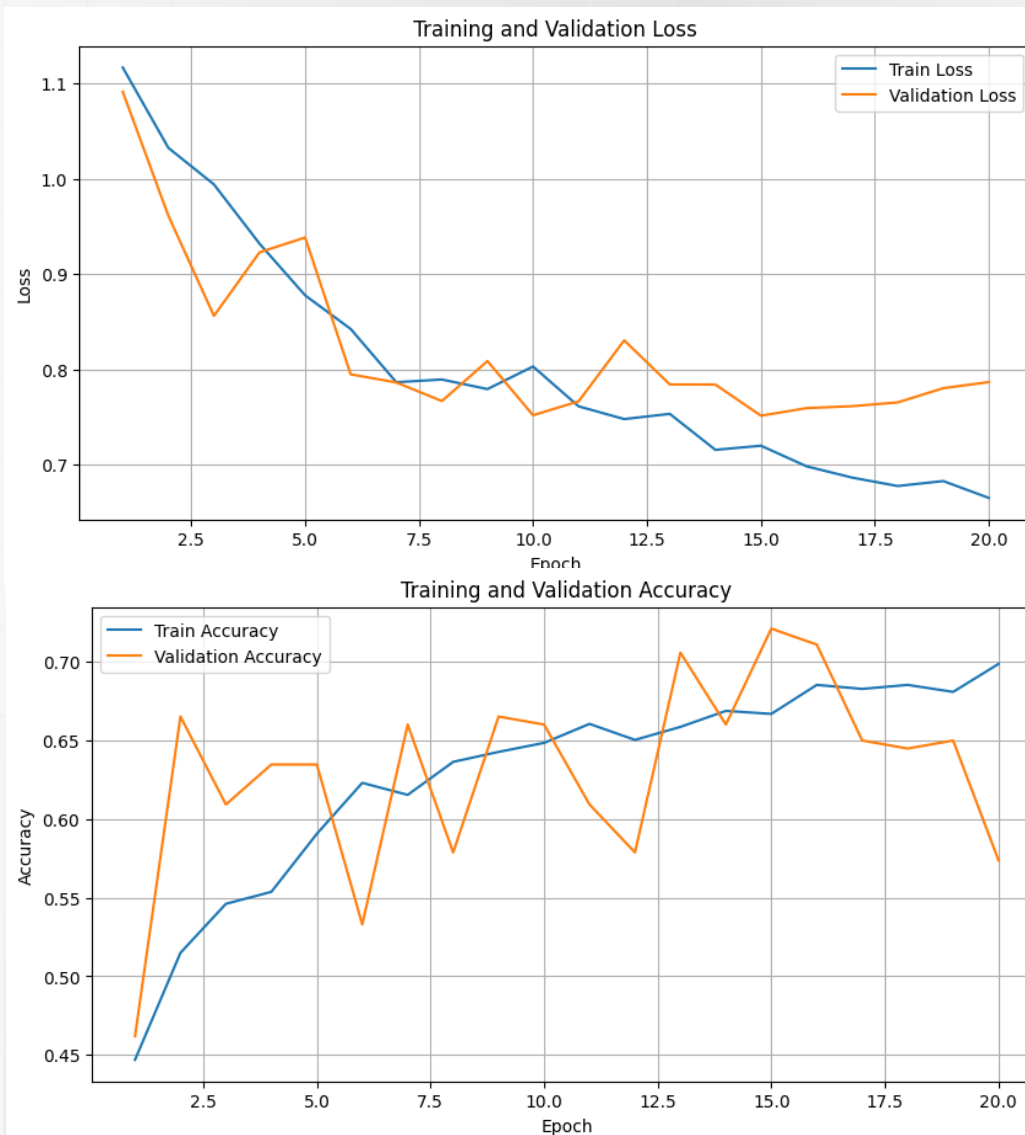
```
[[128 19  2]
 [ 12  8  3]
 [  2  8 16]]
```

single_multiclass Confusion Matrix



Single Multiclass - LNFN

3D CNN + LSTM + Attention



Single Multiclass - LNFN

3D CNN + LSTM + Attention

Accuracy: 0.7121

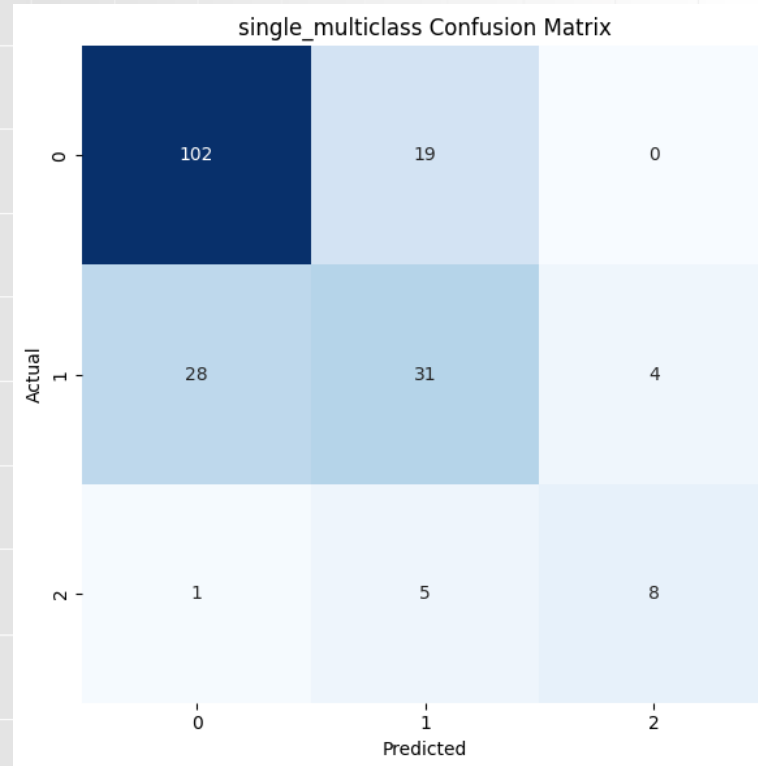
Classification Report:

	precision	recall	f1-score	support
0	0.78	0.84	0.81	121
1	0.56	0.49	0.53	63
2	0.67	0.57	0.62	14

accuracy			0.71	198
macro avg	0.67	0.64	0.65	198
weighted avg	0.70	0.71	0.71	198

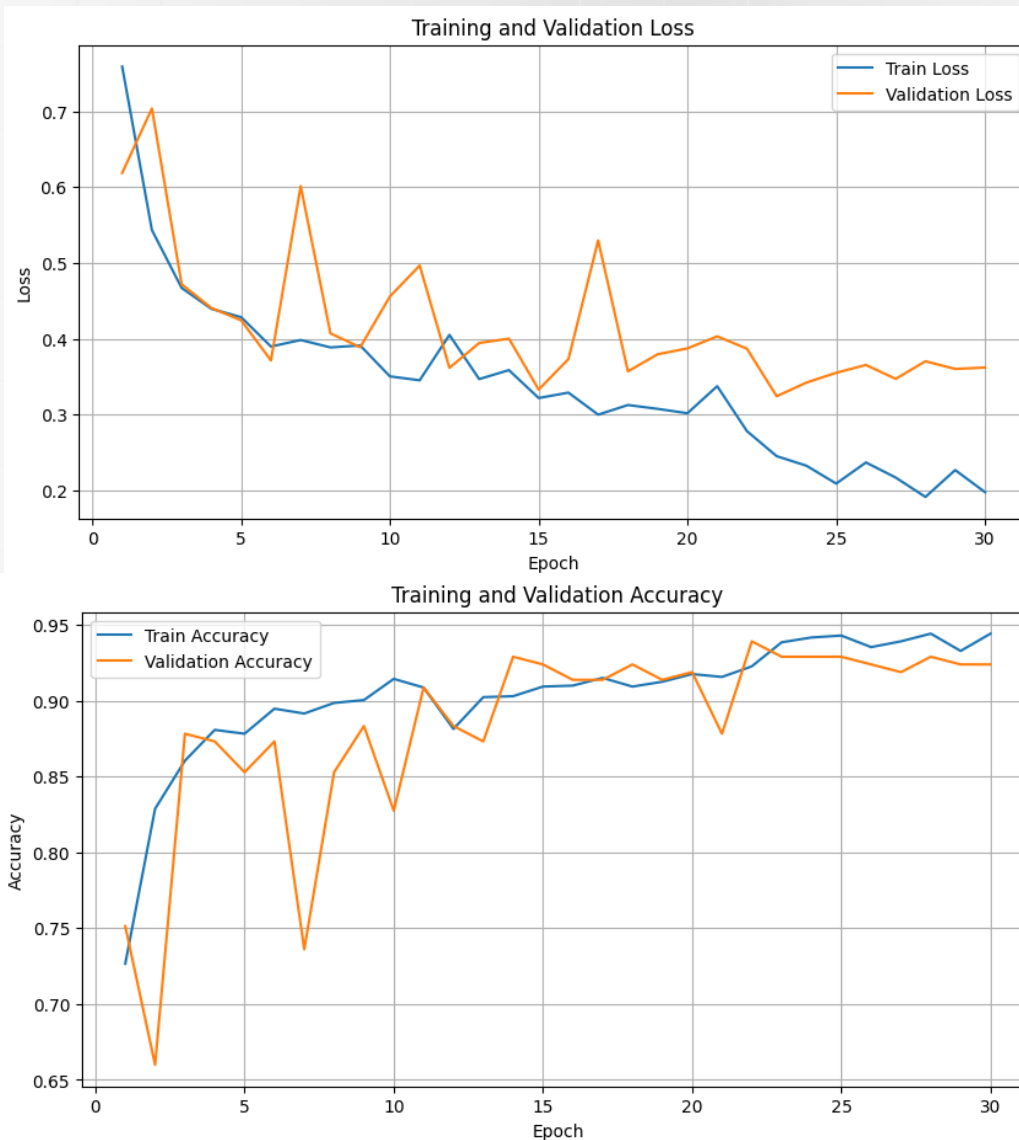
Confusion Matrix:

```
[[102  19   0]
 [ 28  31   4]
 [  1   5   8]]
```



Single Binary - SCS

3D CNN + LSTM + Attention



Single Binary - SCS

3D CNN + LSTM + Attention

Accuracy: 0.9444

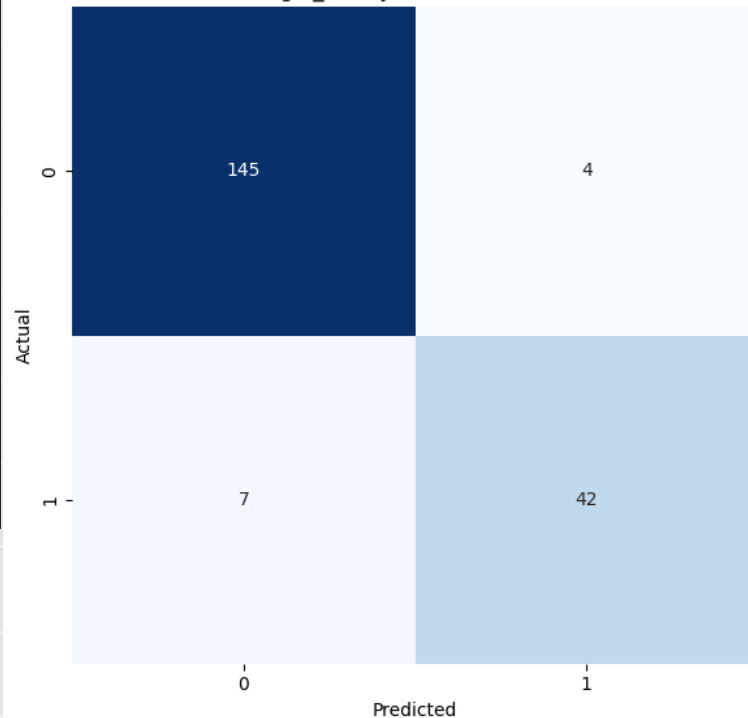
Classification Report:

	precision	recall	f1-score	support
0.0	0.95	0.97	0.96	149
1.0	0.91	0.86	0.88	49
accuracy			0.94	198
macro avg	0.93	0.92	0.92	198
weighted avg	0.94	0.94	0.94	198

Confusion Matrix:

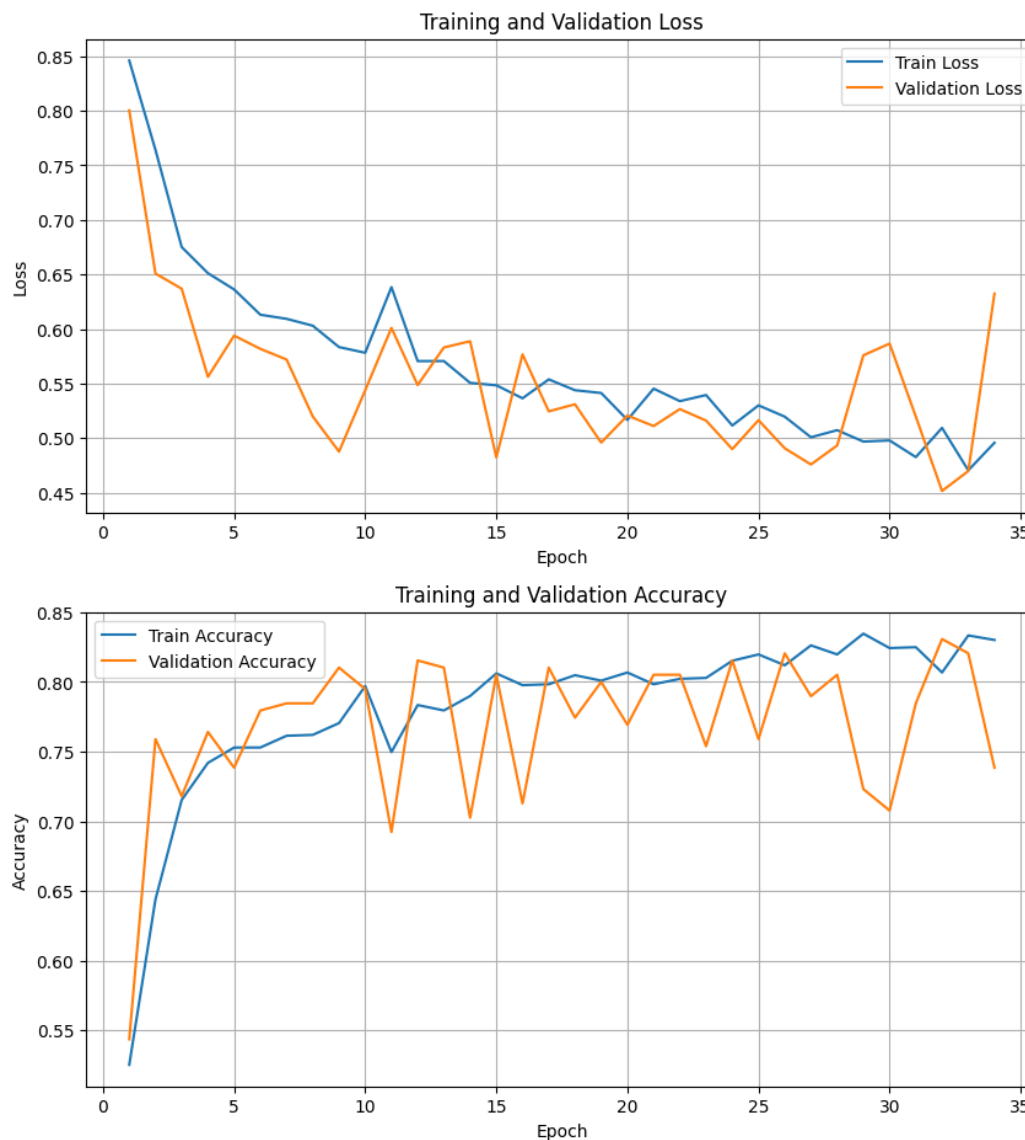
```
[[145  4]
 [ 7 42]]
```

single_binary Confusion Matrix



Single Binary - RNFN

3D CNN + LSTM + Attention



Single Binary - RNFN

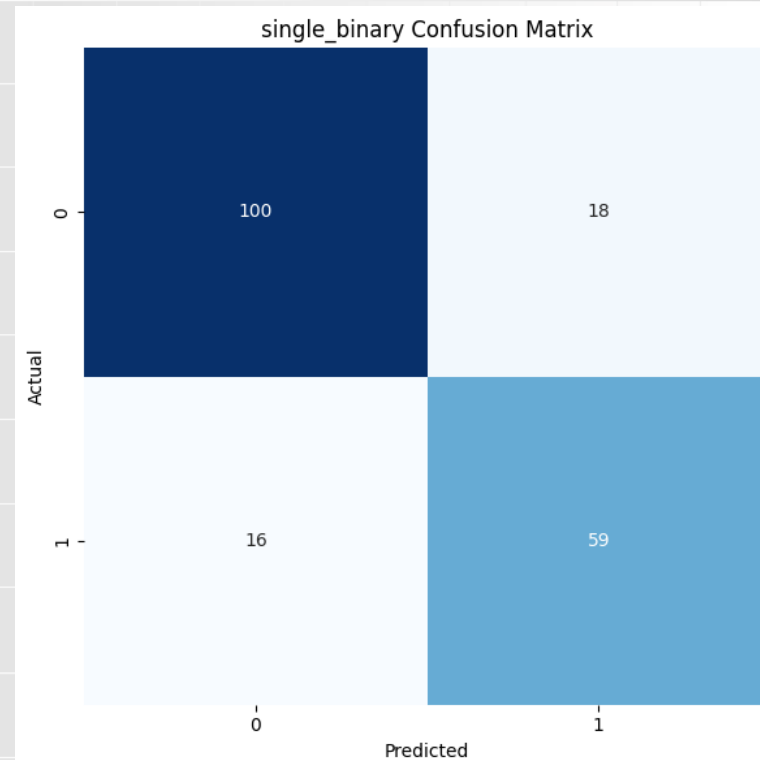
3D CNN + LSTM + Attention

```
Accuracy: 0.8238
Classification Report:
              precision    recall  f1-score   support

     0.0         0.86      0.85      0.85        118
     1.0         0.77      0.79      0.78         75

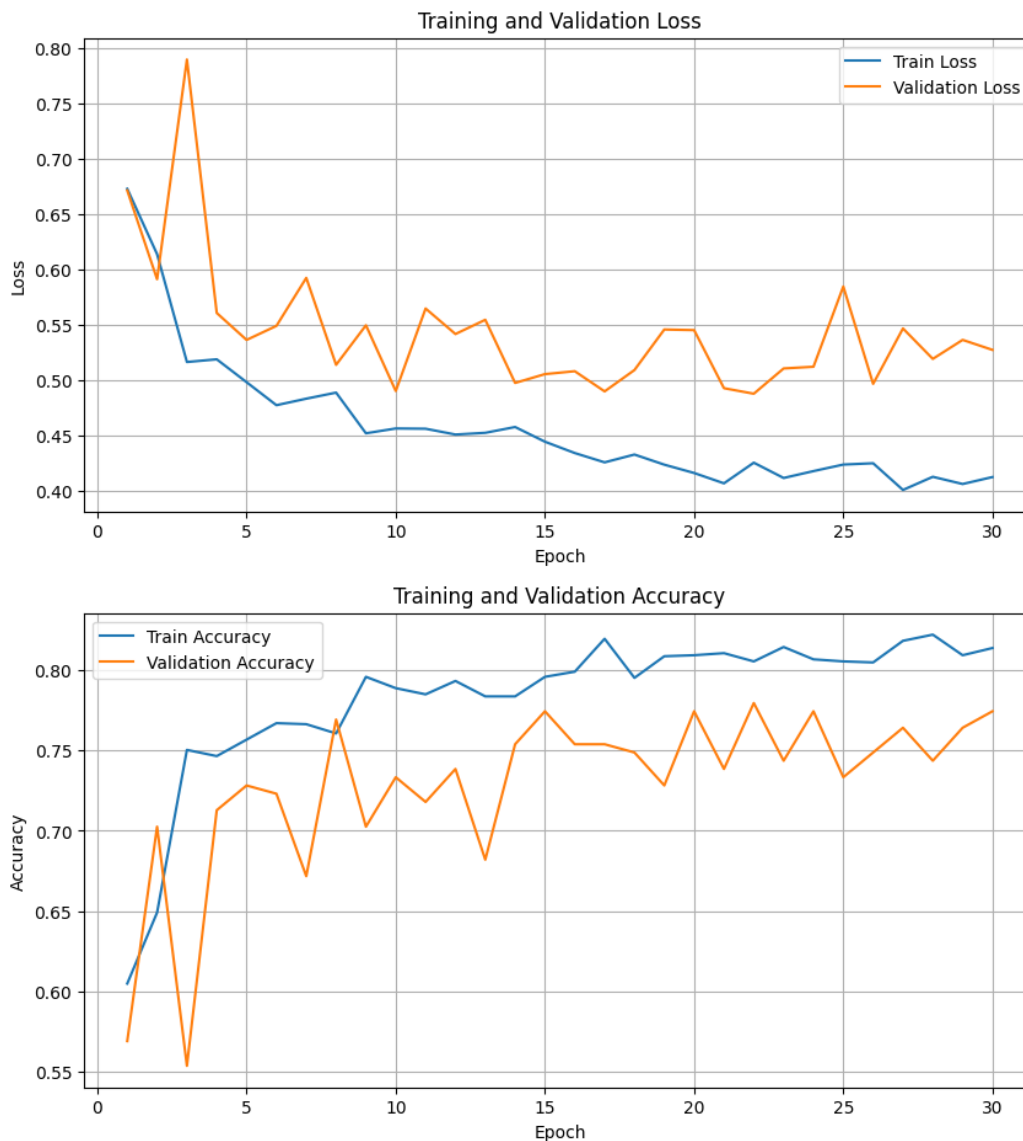
 accuracy          0.82          0.82          0.82        193
 macro avg         0.81          0.82          0.82        193
 weighted avg      0.82          0.82          0.82        193

Confusion Matrix:
[[100  18]
 [ 16  59]]
```



Single Binary - LNFN

3D CNN + LSTM + Attention



Single Binary - LNFN

3D CNN + LSTM + Attention

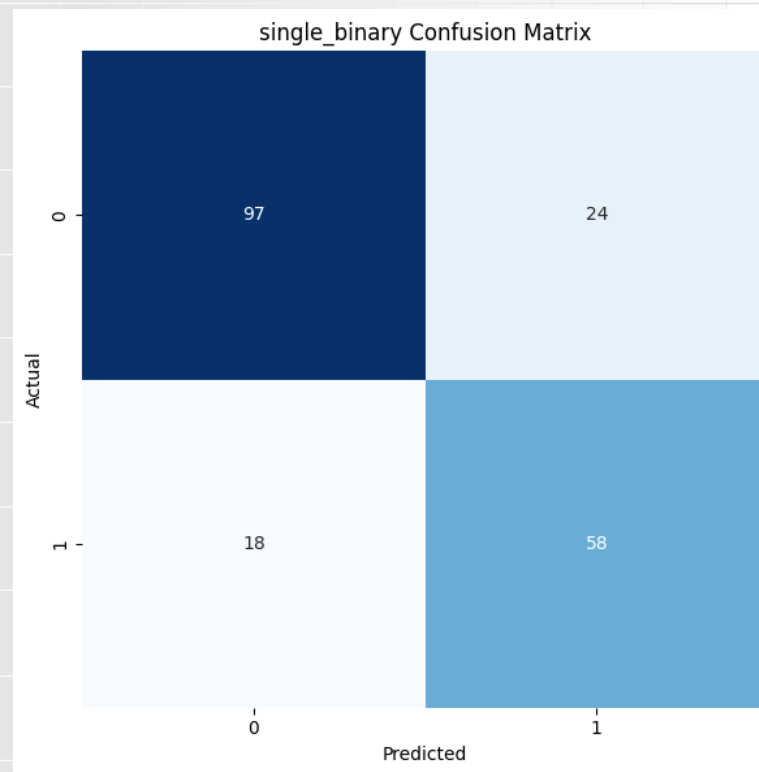
Accuracy: 0.7868

Classification Report:

	precision	recall	f1-score	support
0.0	0.84	0.80	0.82	121
1.0	0.71	0.76	0.73	76
accuracy			0.79	197
macro avg	0.78	0.78	0.78	197
weighted avg	0.79	0.79	0.79	197

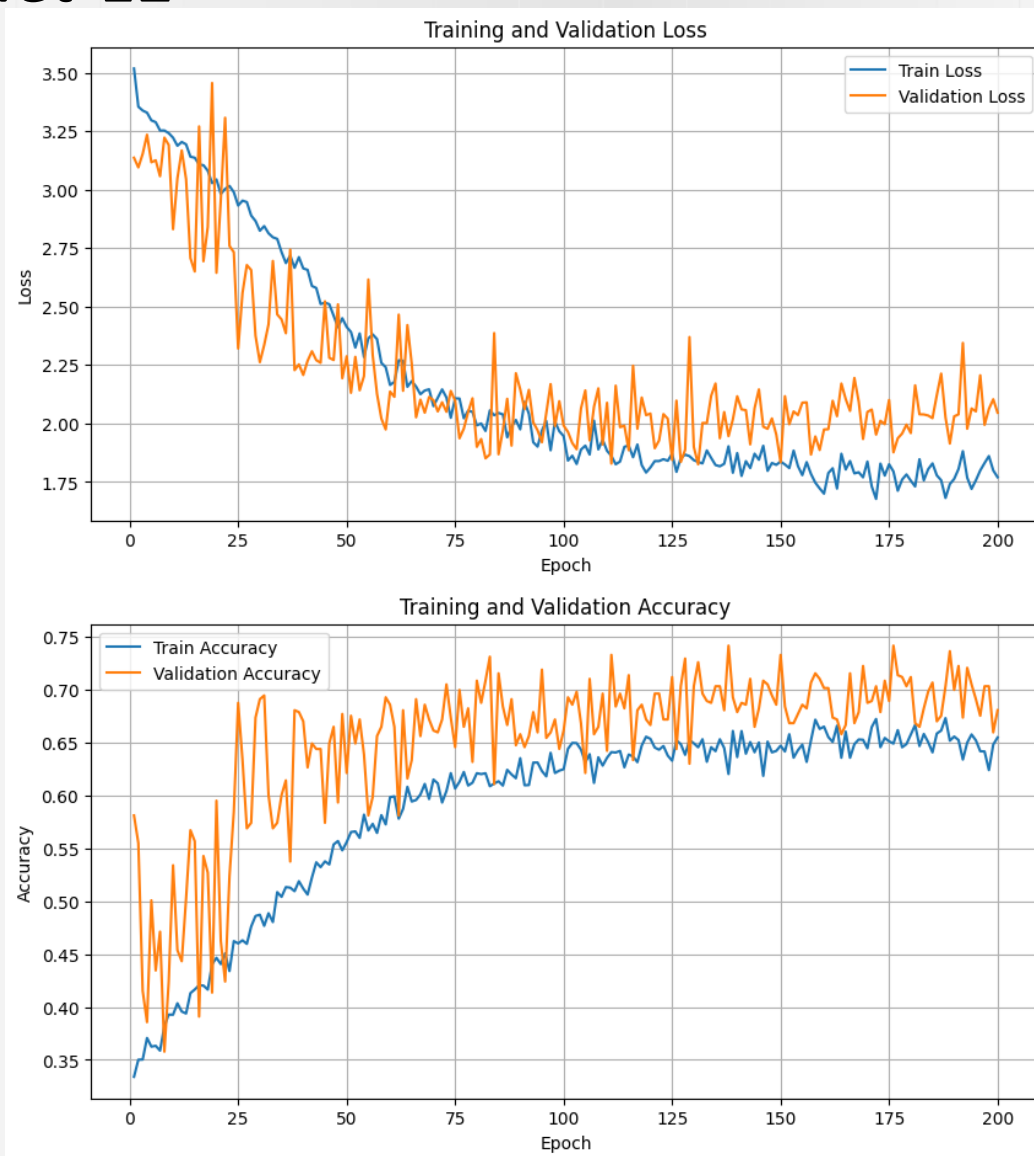
Confusion Matrix:

```
[[97 24]  
 [18 58]]
```



Multi Multiclass

Custom ResNet-12



Multi Multiclass

Custom ResNet-12

=== Metrics for SCS ===

Accuracy: 0.8624

Classification Report:

	precision	recall	f1-score	support
0	0.92	0.99	0.95	143
1	0.50	0.17	0.26	23
2	0.64	0.78	0.71	23
accuracy			0.86	189
macro avg	0.69	0.65	0.64	189
weighted avg	0.84	0.86	0.84	189

Confusion Matrix:

```
[[141  1  1]
 [ 10  4  9]
 [  2  3 18]]
```

=== Metrics for RNfN ===

Accuracy: 0.7302

Classification Report:

	precision	recall	f1-score	support
0	0.77	0.85	0.80	117
1	0.64	0.54	0.58	65
2	0.80	0.57	0.67	7
accuracy			0.73	189
macro avg	0.73	0.65	0.68	189
weighted avg	0.72	0.73	0.72	189

Confusion Matrix:

```
[[99 18  0]
 [29 35  1]
 [ 1  2  4]]
```

=== Metrics for LNfN ===

Accuracy: 0.6878

Classification Report:

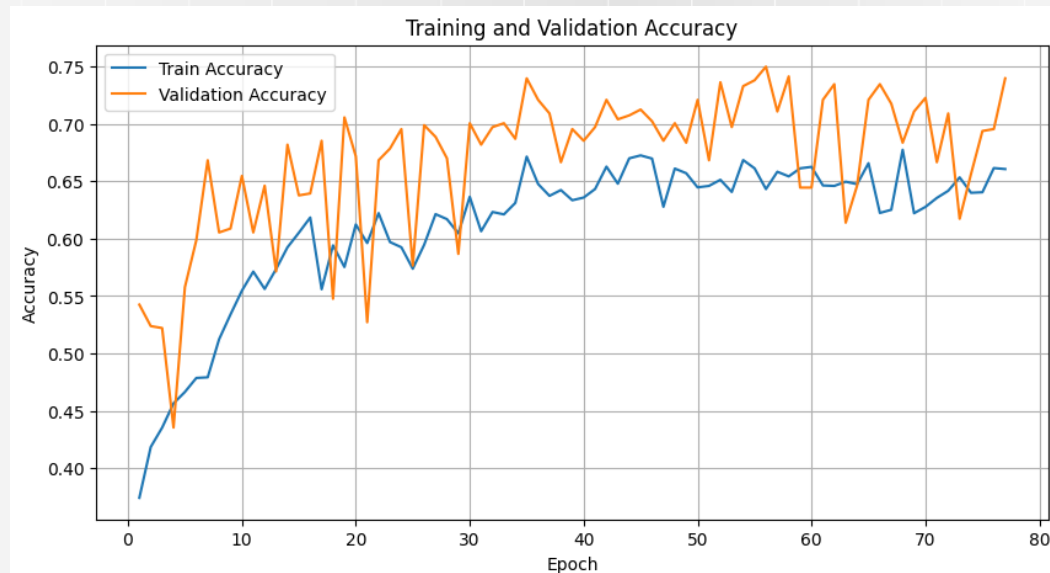
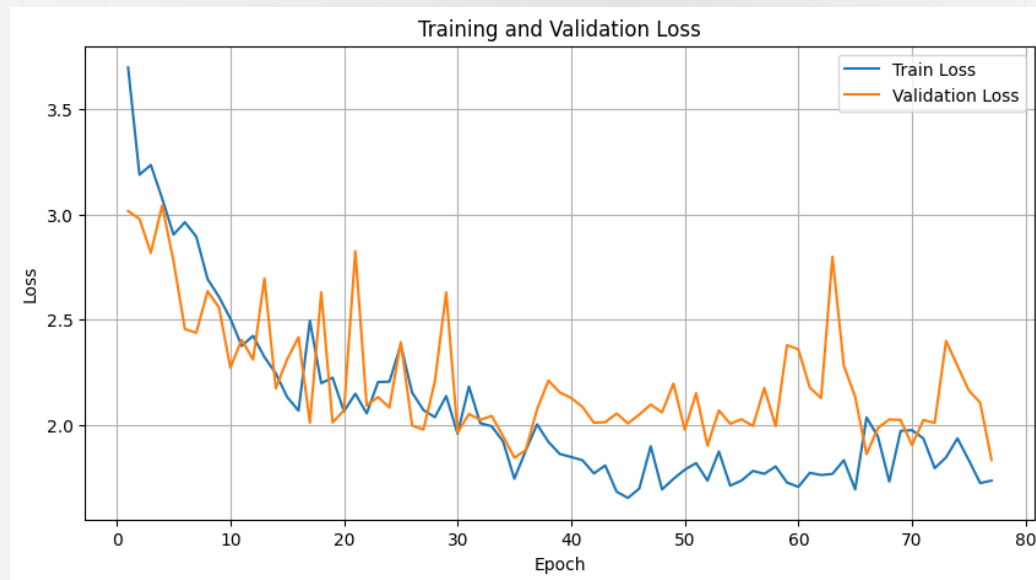
	precision	recall	f1-score	support
0	0.81	0.74	0.77	115
1	0.57	0.62	0.60	64
2	0.36	0.50	0.42	10
accuracy			0.69	189
macro avg	0.58	0.62	0.60	189
weighted avg	0.70	0.69	0.69	189

Confusion Matrix:

```
[[85 25  5]
 [20 40  4]
 [ 0  5  5]]
```


Multi Multiclass

ResNet-18



Multi Multiclass

ResNet-18

=== Metrics for SCS ===

Accuracy: 0.8579

Classification Report:

	precision	recall	f1-score	support
0	0.91	0.95	0.93	149
1	0.46	0.46	0.46	24
2	0.94	0.67	0.78	24
accuracy			0.86	197
macro avg	0.77	0.69	0.72	197
weighted avg	0.86	0.86	0.86	197

Confusion Matrix:

```
[[142  6  1]
 [ 13 11  0]
 [  1  7 16]]
```

=== Metrics for RNfN ===

Accuracy: 0.7157

Classification Report:

	precision	recall	f1-score	support
0	0.77	0.87	0.82	124
1	0.63	0.44	0.52	66
2	0.40	0.57	0.47	7
accuracy			0.72	197
macro avg	0.60	0.63	0.60	197
weighted avg	0.71	0.72	0.70	197

Confusion Matrix:

```
[[108 15  1]
 [ 32 29  5]
 [  1  2  4]]
```

=== Metrics for LNfN ===

Accuracy: 0.7208

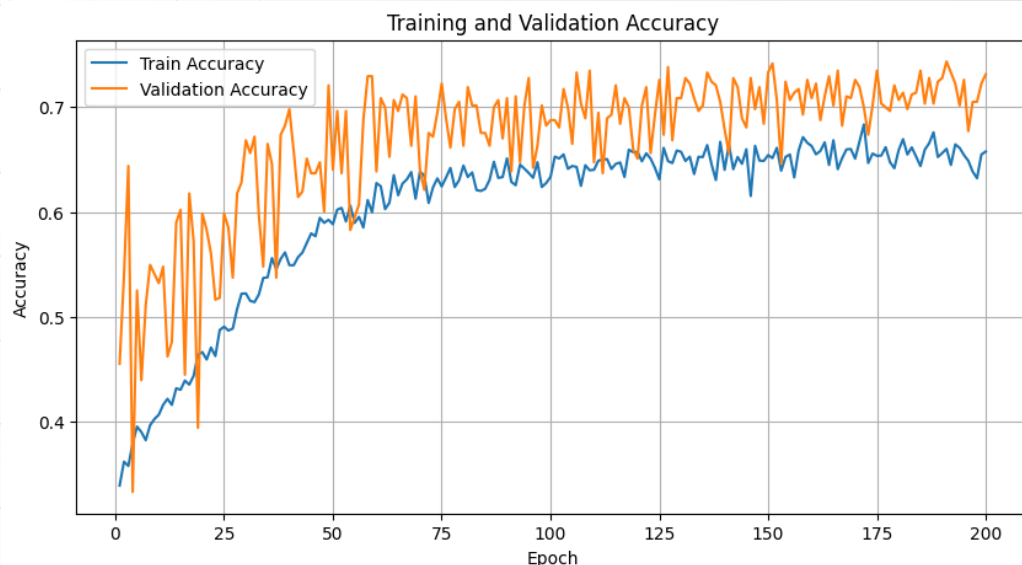
Classification Report:

	precision	recall	f1-score	support
0	0.79	0.87	0.82	122
1	0.60	0.48	0.53	65
2	0.50	0.50	0.50	10
accuracy			0.72	197
macro avg	0.63	0.62	0.62	197
weighted avg	0.71	0.72	0.71	197

Confusion Matrix:

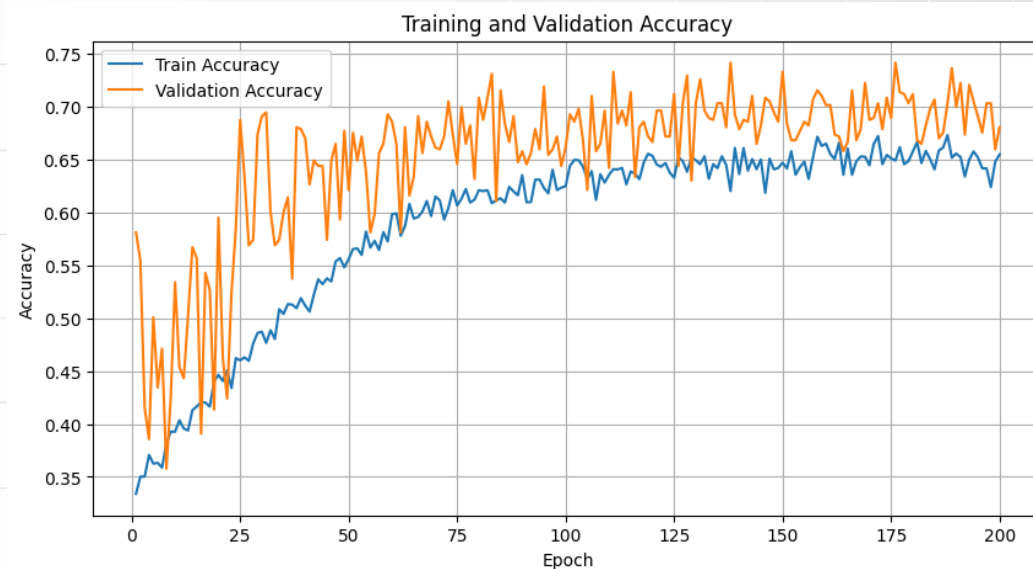
```
[[106 16  0]
 [ 29 31  5]
 [  0  5  5]]
```

Porównanie wymiarów tensorów



96x96x5

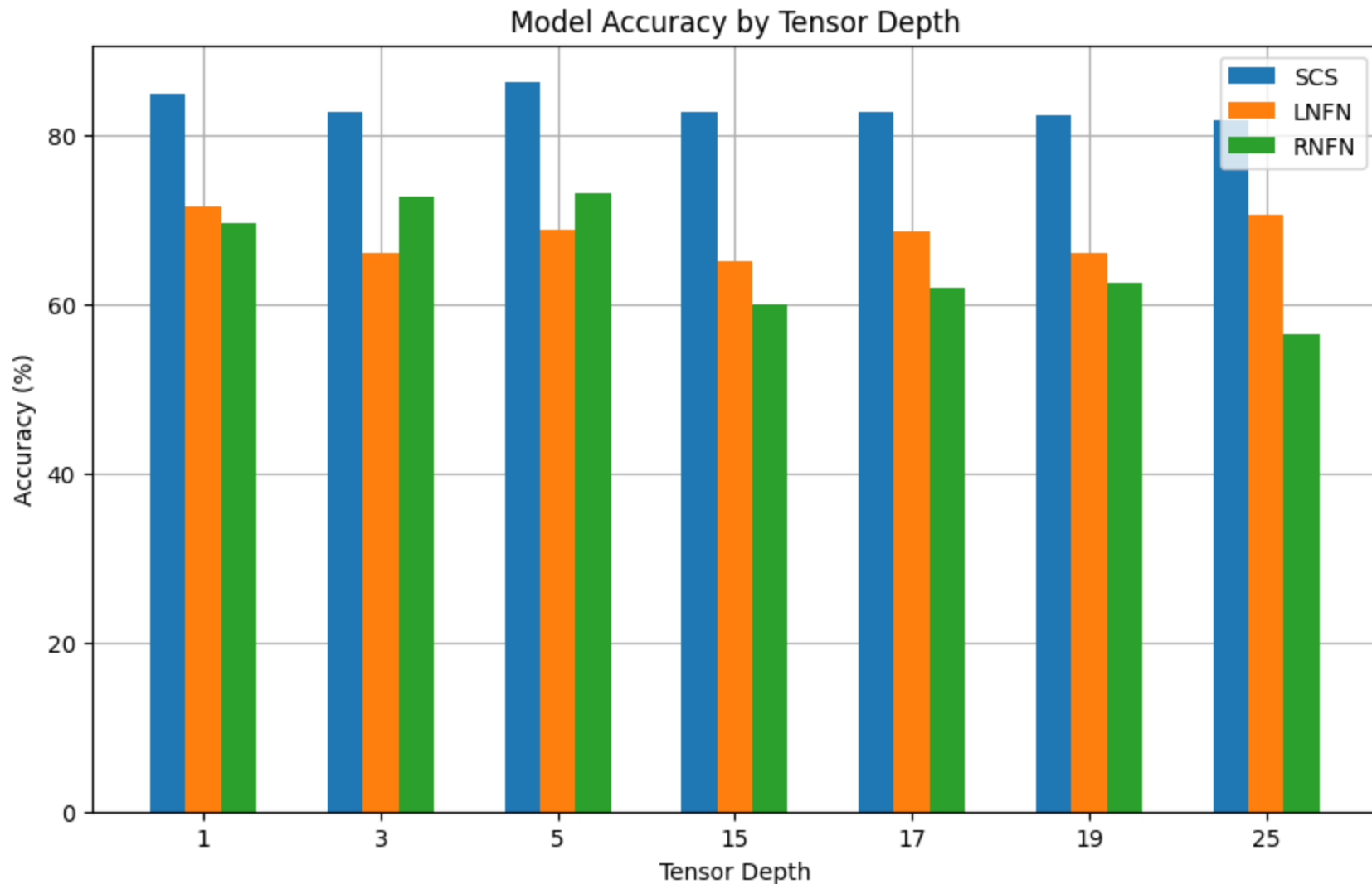
Choroba	Dokładność [%]
SCS	86
LNFN	69
RNFN	67



128x128x5

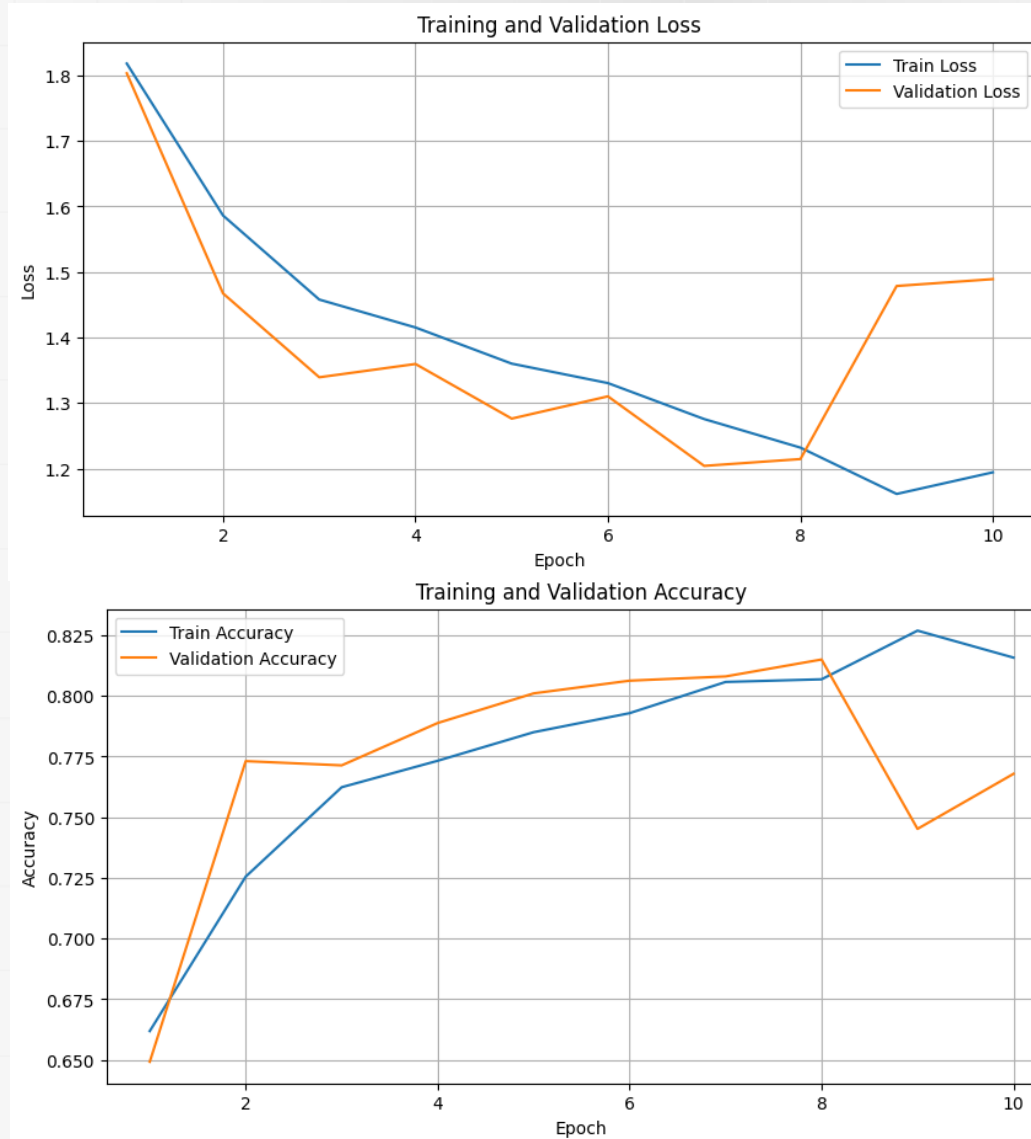
Choroba	Dokładność [%]
SCS	86
LNFN	69
RNFN	73

Porównanie wymiarów tensorów



Multi Binary

ResNet-18



Multi Binary

ResNet-18

=== Metrics for SCS ===

Accuracy: 0.9206

Classification Report:

	precision	recall	f1-score	support
0.0	0.96	0.94	0.95	143
1.0	0.82	0.87	0.84	46
accuracy			0.92	189
macro avg	0.89	0.90	0.89	189
weighted avg	0.92	0.92	0.92	189

Confusion Matrix:

```
[[134  9]
 [  6 40]]
```

=== Metrics for LNfN ===

Accuracy: 0.7884

Classification Report:

	precision	recall	f1-score	support
0.0	0.79	0.89	0.84	115
1.0	0.78	0.64	0.70	74
accuracy			0.79	189
macro avg	0.79	0.76	0.77	189
weighted avg	0.79	0.79	0.78	189

Confusion Matrix:

```
[[102 13]
 [ 27 47]]
```

=== Metrics for RNfN ===

Accuracy: 0.7884

Classification Report:

	precision	recall	f1-score	support
0.0	0.82	0.85	0.83	117
1.0	0.74	0.69	0.71	72
accuracy			0.79	189
macro avg	0.78	0.77	0.77	189
weighted avg	0.79	0.79	0.79	189

Confusion Matrix:

```
[[99 18]
 [22 50]]
```

Podsumowanie

Klasyfikacja ze stopniowaniem:

Choroba	Artykuł	Moje rozwiązanie
Spinal Canal Stenosis	78.6%	87/86%
Foraminal Stenosis	74.3%	73/71%

Klasyfikacja binarna:

Choroba	Artykuł	Moje rozwiązanie
Spinal Canal Stenosis	85.9%	94/92%
Foraminal Stenosis	89.0% (Axial + Sagittal)	81/79%

Podsumowanie

W ramach dalszych prac planowane są następujące kroki:

- **Dalszy preprocessing**
- **Augmentacja danych ?**
- **Analiza hiperparametrów**
- **Eksploracja metod głębokiego uczenia (np. ResNeXt, DenseNet, EfficientNet)**

Literatura

- Merali, Z., Wang, J. Z., Badhiwala, J. H., et al. A deep learning model for detection of cervical spinal cord compression in MRI scans. Scientific Reports, 2021.
- Abuhayi, B. M., Bezabh, Y. A., Ayalew, A. M. Lumbar Disease Classification Using an Involutional Neural Based VGG Nets. IEEE Access, 2024.
- Wang, T., Chen, R., Fan, N., et al. Machine Learning and Deep Learning for Diagnosis of Lumbar Spinal Stenosis. J Med Internet Res, 2024.
- Liawrungrueang, W., Kim, P., et al. Automatic Detection, Classification, and Grading of Lumbar Intervertebral Disc Degeneration. Diagnostics, 2023.



Politechnika
Wrocławska

Dziękuję za uwagę



HR EXCELLENCE IN RESEARCH