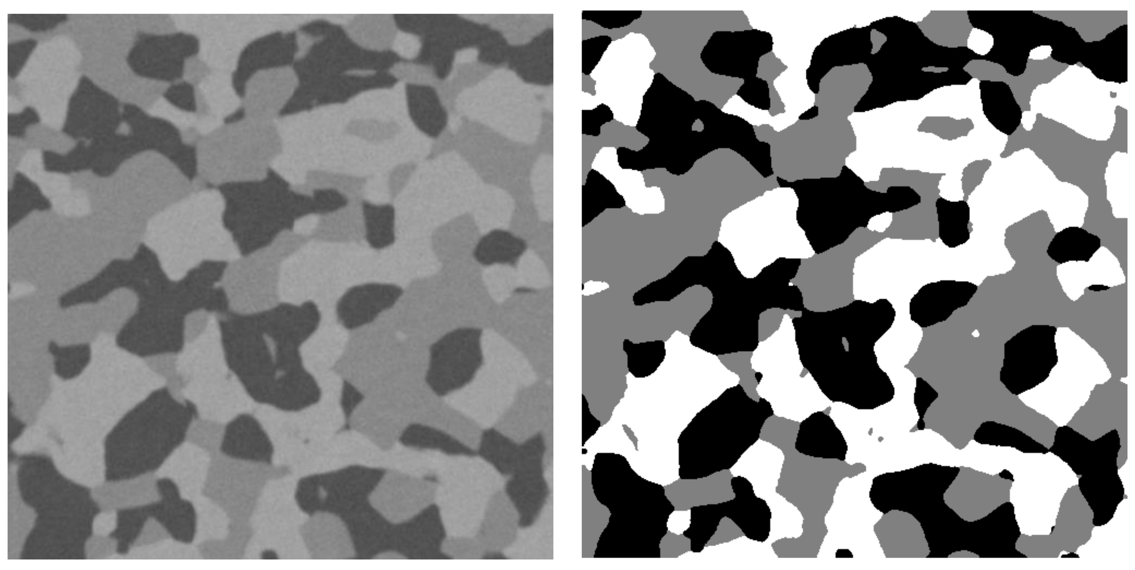
# Automating the Segmentation of X-ray Images with Deep Neural Networks

Project: 28
Supervisor: Salvatore De Angelis & Peter Stanley Jørgensen

Cassandra Desvig-Forss, Ludvík Petersen & Ulrika Woulhøj Jakobsen
s194115, s194613, s194124

## 1. Introduction

Analysis of tomographic X-ray datasets has become very important in different aspects of applications. The raw data is cumbersome to segment manually. The goal here is to automate this segmentation.
The dataset consists of 500 images and 500 labels of size 501x501. There are three classes: black, grey, and white. Raw data is greyscale (16-bit). The images are cropped to sizes of 128x128. The dataset is normalized, and three masks are defined from the label images.



*Left: Raw 501x501 16-bit image.*

*Right: True label with the three classes.*

## 2. Baseline

U-Nets are good for semantic segmentation. The model consists of a contracting path and an expanding path, with an encoder and decoder respectively. Skip connections are used to transfer spatial information from the contracting path to the expanding path. At the bottom of the paths, there is a bottleneck [1]. For the encoder, we had a 3x3 convolutional layer, a batch norm, and a ReLU activation function [2]. The same was used for the decoder. No weight decay was used. Cross entropy loss was used, as well as the Adam optimizer with a learning rate of 1e-3. With this baseline model, the resulting training accuracy was 98.9%

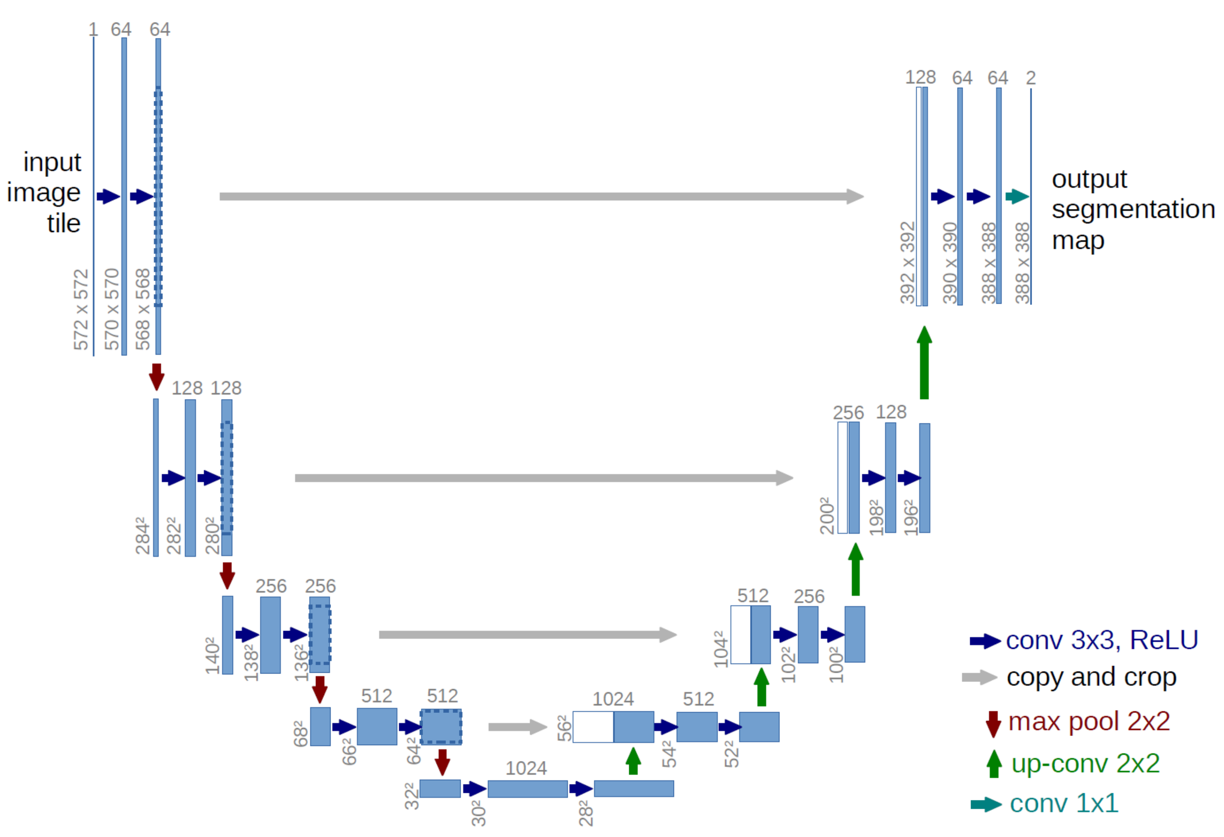
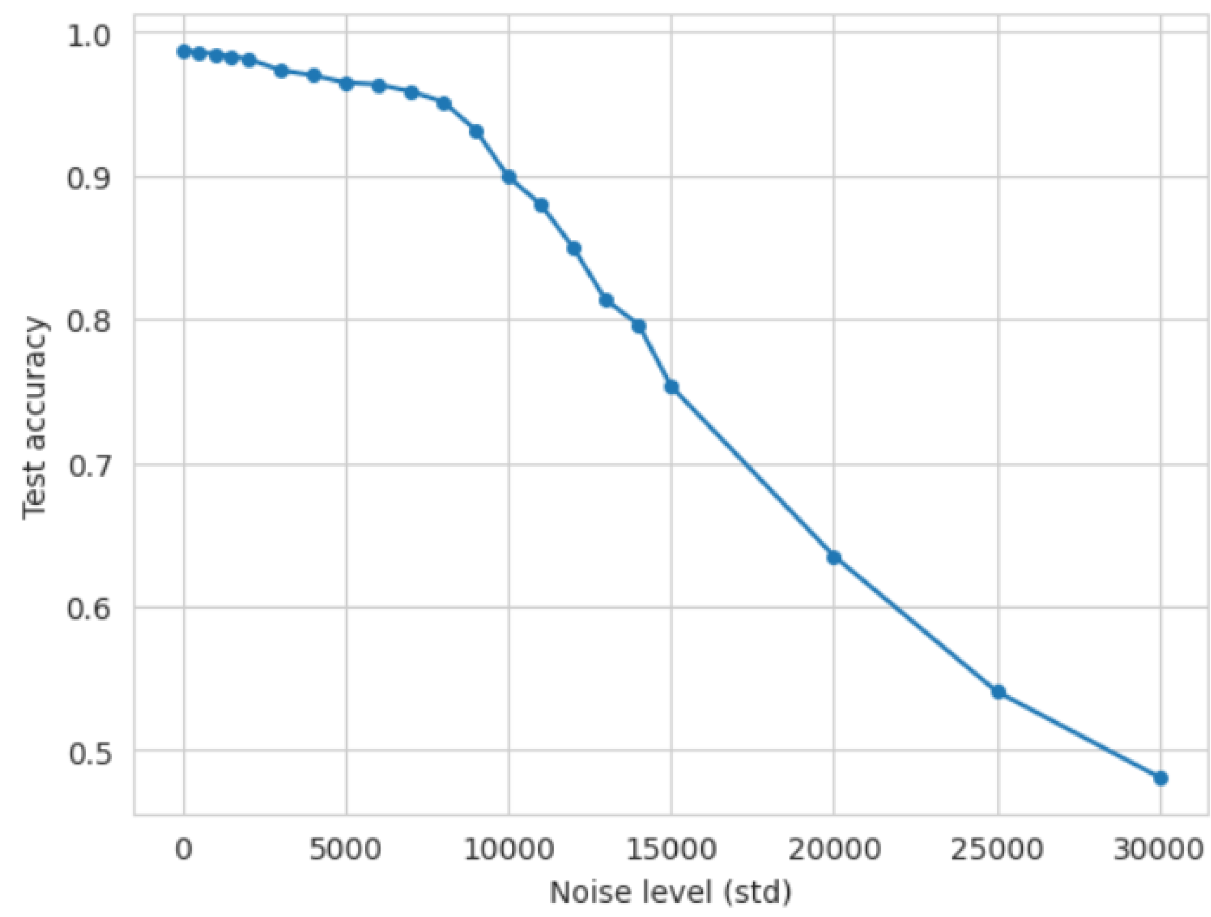



*Figure showing the U-Net architecture with other image sizes. The figure is from [2].*

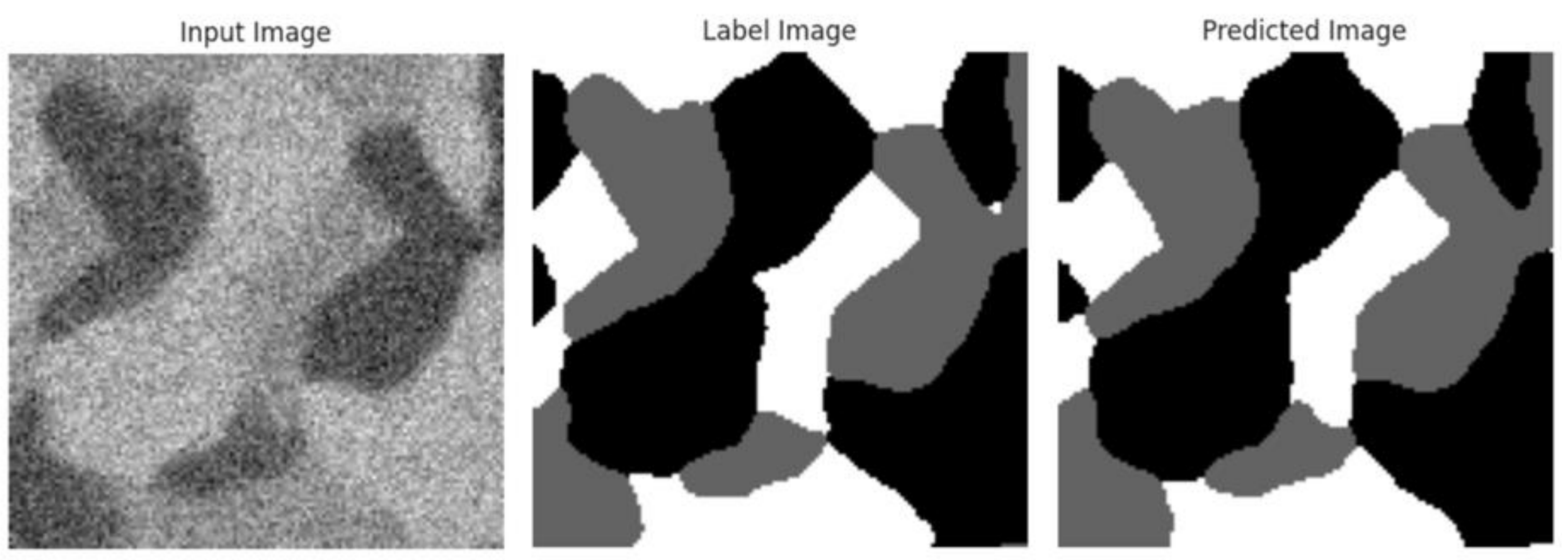
## 3. Casework: Noise and larger dataset

To challenge our baseline noise was implemented on specific parts of the dataset. The noise is simply Gaussian noise with a mean of zero and a standard deviation which can be modified such that the noise level is different. All test data have noise added, but only a percentage of the training and validation data have noise added. It was found that adding noise affected the performance of the network negatively.

To counter that more data was introduced by cropping the original images into 9 smaller images (128x128), which resulted in 4500 images in total.




*Accuracy for the model as a function of the noise level. The model is trained on data with varying noise levels on 50% of the data.*




*Left: The input image with added noise (std = 5000). Middle: The label image Right: The predicted image. As can be seen, there is a good overlap between the label image and predicted image.*

| | Baseline | 0% | 25% | 50% | 75% | 100% |
|---|---|---|---|---|---|---|
| Testing (small) | 0.968 | 0.844 | 0.863 | 0.841 | 0.866 | 0.869 |
| Valid (small) | 0.990 | 0.987 | 0.977 | 0.972 | 0.970 | 0.967 |
| Testing (large) | 0.990 | 0.545 | 0.936 | 0.954 | 0.957 | 0.960 |
| Valid (large) | 0.991 | 0.990 | 0.982 | 0.981 | 0.977 | 0.977 |

*The obtained accuracies in the various evaluations/tests. The percentages represent the amount of training data that have noise added. Small and large represents the size of the dataset.*

## 4. Conclusion

A U-Net model has been implemented. The baseline was found to perform well. After this, the model was challenged on its ability to generalize, by adding noise to a certain part of the images as well as creating a larger data set. It was found that the best performance was with a large dataset with noise on over 50% of the training data.

## 5. References

[1] Aditya_taparia, U-Net Architecture Explained. URL: https://www.geeksforgeeks.org/u-net-architecture-explained/
[2] Aladdin Persson / machine learning repository. URL: https://github.com/aladdinpersson/Machine-Learning-Collection/tree/master/ML/Pytorch/image_segmentation/semantic_segmentation_unet