Flood Impact Detection on Areas Using U-Net

Shaoun Chandra Shill
BRAC University
23373005
shaoun.chandra.shill@g.bracu.ac.bd

Fahim Ahmed Ifty
BRAC University
23266007
fahim.ahmed.ifty@g.bracu.ac.bd

Tuhin Ahmed
BRAC University
23173004
tuhin.ahmed@g.bracu.ac.bd

Effat Jahan
BRAC University
24366047
effat.jahan1@g.bracu.ac.bd

***Abstract*—The consequences of climate change are widespread and affects various aspects of the environment including extreme drought and heavy rainfall. Heavy rainfall is the reason why flooding has become a frequent disaster in many countries of the world including Bangladesh. Flooding is one of the most disruptive natural disasters, causing extensive damage to road infrastructure and impacting transportation and human safety. Every year these flood affected areas go through extreme changes in rural, urban, livestock, agricultural and service areas. To mitigate the destruction these natural disasters are causing there are a few suggested methods. If we can detect flood-affected roads in a timely manner then the emergency response and infrastructure management can work more accurately in saving lives and distributing relief to the flood affected people. This paper proposes a flood impact detection system where a deep learning architecture U-Net will be used. This Architecture can be effectively used for detecting flood impacts on roads with the images from satellite or aerial imagery.**

***Index Terms*—Deep learning and SAR, sentinel-1 SAR, flood detection.**

I. Introduction

Due to climate change more heavy rainfall it increases the chances of flooding. It has become more and more frequent and intense across the globe leading to immense property and crop damage, environmental degradation, and death toll. Ideally, the locations that floods have impacted should be identified adequately in order to prevent loss of lives and property, assess the impact, and finally assess the best and applicable methods of preventing such scenarios from recurring in the future. The innovations in flood mapping include the basic ways, like surveying and analysis from the satellite images, which take a lot of time and may be prone to errors.

Computer vision has been among the areas that have greatly benefited from developments in deep learning over the last couple of years. We also find that the Convolutional Neural Network architecture, namely U-Net, which has been used for biomedical image segmentation and has been tested for other segmentation tasks, such as semantic image segmentation of natural images has provided optimal results. This paper focuses on flood impact, which is predicted and detected with the help of aerial or satellite images, based on the U-Net.

For our proposed approach we are using the U-Net framework. Here, we seek to design an effective and reliable system for flood detection from very high-resolution image data. The research provides benefits to constructing a system for automating the identification of flood occurrences to improve flood response and intervention systems. The proposed system will also enable decision making and proper resource allocation during floods as it gives accurate near real-time estimates of the impacts.

II. Background study

A. What is UNET

U-net is an encoder-decoder-based convolutional self-design network for image segmentation developed for biomedical image analysis in recent years. Its distinctive U-shape is formed by two main parts: A road within which the magnitude grows, the said encoder is the road that is the compressing dimension, and the extended road is a decoder. [1] The contracting layers decrease the context of this load inversely, the expansive part increases this context by reconstructing the features which concatenated with features that can be obtained from the encoder part through the Skip connection. The connection links make it possible to reproduce geo-spatial correlation if the image affords correct spatial relations; geometric and photo-realistic attributes of the environment, the chained levels of spatial hierarchy, and the segmentation feature of complicated images with high accuracy even though only a few samples are taken. This is well exemplified by the U-Net architecture because it is fulfilling the segmentation function and is very helpful to medical image analysis because the task normally gets down to pixel level.

B. UNET Architecture

1) Encoder (Contracting Path) : The encoder takes the input image, decreases its size step by step by convolution and pooling layers, extracting features from it which can be used in the process part.

$$\text{Conv}(x) = \sigma(W * x + b)$$
$$x_{down} = \text{MaxPool}(\text{Conv}(x))$$

2) Bottleneck: The bottleneck between encoder and decoder path. As aa result, this section provides the most compact representation of the image to balance between capturing all fine details from the encoder and preparing the information for upsampling in the decoder.

$$x_{bottleneck} = \text{Conv}(\text{MaxPool}(x_{last}))$$

3) Decoder (Expanding Path): The decoder comes after the encoder and it tries to recreate the spatial resolution of the image. The upsampling layers gradually restore the size of the image and increase its level of detail by interpreting the compressed feature maps produced by the encoder.

$$x_{up} = \text{Concat}(x_{down}, x_{skip})$$
$$x_{up} = \text{ConvTrans}(x_{up})$$

4) Skip Connections: Skip connections are links between the corresponding layers of encoder and decoder and allow to carry information from one part to another by what is called bypass.

$$\text{Output} = \text{Softmax}(W_{output} * x_{final} + b_{output})$$

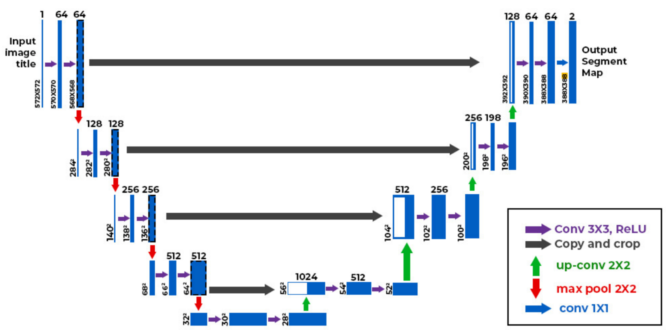

Fig. 1: Unet Architecture

C. Advantages of U-Net in Flood Detection

For flood detection, U-Net presents several advantages due to its architecture:

Pixel-wise Prediction: This made the U-Net to be characterise with the ability of identifying the areas of flood within a pixel precision from satellite or a drone imagery which is appropriate for detecting area that could be flooded.

Handling Multi-scale Features: It means that many features given in Multi-scale may also be read by 'read-out and write-in' pathways since at time region of floods could be large or small or look very different.

Skip Connections: I have also noted that because of skip connections the network does not discard high frequency information and that will be effective in delineating fem flooded and non-flooded areas at the smallest of the topographical map scale.

D. Challenges in Flood Detection with U-Net

Indeed in spite of the fact that U-Net has illustrated incredible viability, there are still certain issues:

Data Awkwardness: Namely, the consideration is an augury that when a larger chunk of the picture is defended by the non-surge region, then course asymmetry may arise. It is stated that use of some of the complex misfortune measures such as Intersection of Union (IoU) or Dice Misfortune can solve this problem.

High-Resolution Information: High resolution adj. information generation and analysis involve a massive amount of computation; thus the author offers few preprocessing the likes of down sampling or having large capacity handling units.

Temporal Changeability: Hence, as the increase is not fair once, the models shall have to consider the features like; climate, season and geographical location of the country. To overcome this trouble they could have had to copy models or.

III. Related Works

From the position of the application the image segmentation which was hosting the function of urgency as the technique in servoing computer vision it still remains magnificent for many chief issues in medical image and satellite imaging and object perception. For this work, the absolutely unique u-net architecture is used, which, probably, can have an extraordinary ability to recognize not only the main contour of picture but also such details. This is the actual feature of this method because it is an encoder-decoder model, and it is always advantageous to be able to combine this with the capacity to preserve the spatial information for identifying segmentations that do not accept inaccurate segments; thus, this method frequently solves multiple segmentation problems.

A. U-Net and Variants

This architecture has been introduced for biomedical image segmentation in Ronneberger et al.[1] (2015) however, in the last years it has been used in many applications. Better organisation in the hierarchically contracted and expanded feature maps is distinguished, as is the benefit when introducing skip connections: from this spatial conjugation at multi levels which has been postulated earlier in the proposed model. U-Net and all its derivatives have been reported to perform particularly well within various medical imaging tasks: Of these the segmentations of organs, tumours and other structures are most outstanding. Small modifications to the NEWER versions of the first described U-net architecture has however been made recently in the following way. On the other hand, to this authors of Isensee et al., 2018 have developed nnU-Net which performs better than the usual U-Net designed to operate on segmentation data sets without posed parameters related to segmentation tuned. Of these new ones is the Attention U-Net, a network which has attempted to attend to some parts of an image due to rich data in some sets.

B. Deep Learning for Image Segmentation

Therefore, we examine and recommend Deep Learning for Image Segmentation. This has limited most of the prior image segmentation techniques to feature more than the CNN's eye-saves. The segmentations models have originally been started with Fully Convolutional Networks (FCNs) as introduced See by Long et al., 2015. However, as the model is symmetric and the U-Shaped model uses skip connection, it should be reasonable that such kind of task is designed to recognize detailed features of the segments. As for the above discussed segmentation, there are other proposing architectures of segmentations for instance Segnet and DeepLab. For instance, DeepLab employs features such as dilated convolution and Conditional Random Field boosts the higher object cutting edges especially if the image has a high density. But for such reasons and more and particularly and primarily for simplicity train ability and capacity to perform on any type of data,,-N-Uet is still used till today.

C. Image Data Augmentation in Segmentation

This is a very big challenge mainly because deep learning based segmentation relies on this kind of data so much. The random crop, horizontal flip, zoom and elastic transformation increases the training data and the model's improvements increase generalisation. Survey paper by Shorten & Khoshgoftaar (2019) describe how image augmentation improves segmentation by Shorten & Khoshgoftaar (2019). To increase the generalisation ability of the model and to counter the problem of over fitting, special emphasis was given in this project to use augmentation procedures using keras Imagedatagenerator.

D. Application in Medical Imaging

U-Net is especially useful in organ and lesion detection in Computer Tomography, Magnetic Resonance Imaging and ultrasonic imaging in general medicine and health care. Of them, cancer detection is one of the most successful applications thanks to the segmentation when the true boundary of the tumour is crucial for diagnosis and subsequent therapy. The subsequent research of Çiçek et al. (2016) tries to use U-Net for different types of 3D volume data in order to expand the applicability of the architecture in the 3D medical imaging field.

IV. Methodology

The flow network related to this work relates to orientation and extension of deep learning deformities that are adapted only to U-net type for image segmentation. Following work plan in several steps is provided: Consequently the activities are; the data pre-processing activity that prepares the data for modeling activity, the second activity is the model selection activity, the third activity is the training of the model activity, then the final activity is the model evaluation activity.

A. Data Preparation

In the context of this project the given data set is an image and the corresponding mask for the image. The images are taken using OS and glob of python and for data preprocessing we used OpenCV or cv2 for resize, normalize and augment images. We have found visualization augmentation features for images to reduce overfitting images and also to enhance the application of the model on new data. All the pre-processing including rotation, flipping, zooming etc., are performed by ImageDataGenerator contained in Keras.

B. Model Architecture

Regarding the image segmentation task, we used the U-Net methodology; regarding CNNs in this case. This architecture was used largely in biomedical image segmentation since it can construct the resolution-preserving feature pyramid based on the encoder-decoder concept. The architecture consisted of:

Contracting Path (Encoder): The model proposed in this study incorporates almost sequential convolution and max pooling to enable the featuers of context to influence the objet of interest in an image.

Expanding Path (Decoder): These were up-sampling stages from the identified path, which helps in improving image resolution and helps in detecting the objects of the images.

Skip Connections: In contrast, in the encoder network, some layers or connections could be omitted or even improves when they were omitted, yet in the decoder network, all corresponding layers and connections were essential. These non-spatial dimension features were concatenated to the right of each other ensuring the spatial dimensions of the feature were maintained.

When building the U-Net model and when making expansions for combining it, TensorFlow with Keras was used. We implemented the Adam optimizer and binary cross-entropy which was multiplied by the weight of the division of one for the Log Bernoulli and for the Dice coefficient for segmentation.

C. Training and Validation

The dataset was divided into two main groups: the training set and the validation set. The training set underwent further augmentation, while the validation set served to monitor the model's performance throughout the training process. The model was trained over multiple epochs, with checkpoints regularly saved to prevent overfitting. Some strategies such as early stopping and learning rate scheduling were applied to optimize the model's performance.

D. Evaluation

In image segmentation, flood detection in particular model performance were evaluated. There are different metrics utilized to evaluate how effective our results of segmentation functioning.

Intersection over Union (IoU) is a popular metric to evaluate the overlap between the predicted masks and ground truth masks. It is calculated by taking the area of interest and then dividing it by sum of predicted area and actual (true)area. The derived ratio, essentially says how much relevant features in the images is the model capable of capturing.

A DICE Coefficient, which also accout fro the overlapping area (brain region that has ”two” labels), is another good common measure for segmentation accuracy. From 0 to 1, where 1 represents that the predicted mask completely match with the true one. The Dice Coefficient is of particular interest in areas such as medical imaging.

F1 Score: The F1 score is the harmonic mean of precision and recall, where an F1 score reaches its best value at 1 which means prefect precision and perfect recall. It is the harmonic mean of the two, which better assesses a model using both false positives and negatives. Certainly that is crucial in case of imbalanced datasets.

The Mean Squared Error (MSE), besides, used for measuring the average squared difference between the predicted output and actual value. Although normally used with regression tasks, MSE can provide us some pieces of information about the general quality of segmentation predictions.

Analysis of the loss curves in training and validation across the epochs: Trying to understand how model is learning. This would plot graphs of losses/scores, which are the way model performs over each epoch, and can give us a clue for any issue like overfitting or underfitting etc. Most of all, a gradual decrease indicates the efficiency of learning, while very variable or up and down behavior may need to be addressed by training in a different way.

Lastly, we describe how visualization provides an important means to evaluate segmentation performance. Creating such side-by-side comparisons between the predicted masks and ground truth images may provide researchers with some hints on what model is good at and not. This qualitative evaluation provides an insight into the performance of a model, or where it may struggle providing guidance for future iterations on the segmentation process.

Fig. 2: Methodology

V. Result and Analysis

A. Model Setup

1) Loading the dataset: Loading the dataset includes data pre-processing which involves mounting the drive with the photos onto the dataset for training of the model. The images can be RGB or grayscale depending on the purpose. In our dataset the masking images are grayscale but the input photos are RGB. After we are done loading the mask images, It needs to be paired with the photos appropriately to maintain data consistency.

2) Validation of Dataset: Make sure picture and mask proportions match. This is critical since mismatched dimensions might cause training problems. Maintain image and mask aspect ratios and ensure both datasets have the same size and resolution. Resize photos and masks to model input scale if needed.

3) Random Display of Image-Mask Pairs: Show a few random image-mask pairings to verify data alignment. Check that each picture matches its mask. A segmentation mask should highlight the item in the picture as intended.

B. Evaluation Metrics

Here, We have calculated F1 score, Dice coefficient and IoU (Intersection over Union) score. A classification model's accuracy is measured by its F1 score, especially in unbalanced datasets. A balanced score is calculated by combining accuracy and recall.

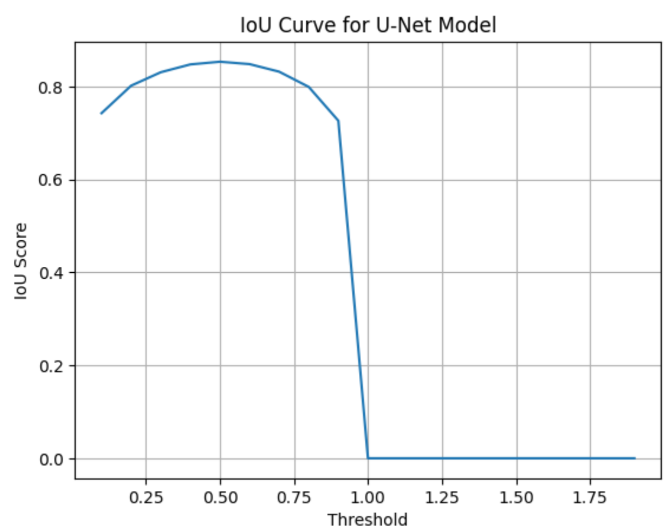

Fig. 3: Iou Curve

The Dice coefficient, or Sørensen–Dice index, is a statistical metric used to assess data similarity, particularly in picture segmentation. It is often used to compare expected and actual segmentation overlap.

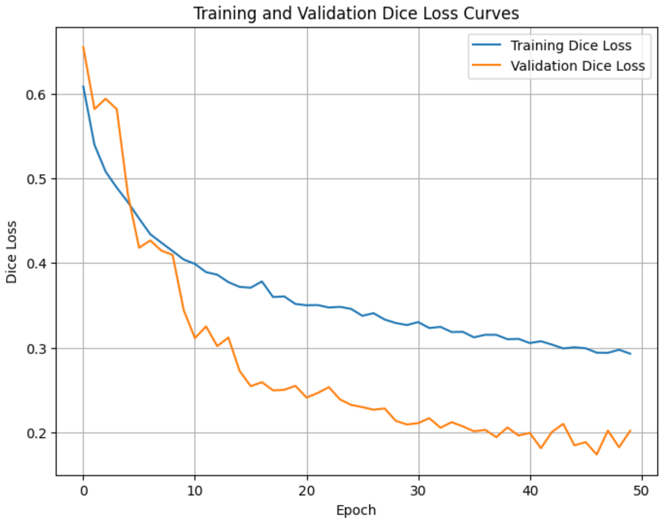

Fig. 4: Dice Loss Crub

The accuracy of object identification, segmentation, and other computer vision tasks is measured by this score. It evaluates the overlap between the expected and actual bounding boxes or segmentation areas.

TABLE I: Evaluation Metrics

F1 score	Dice Coefficient	IoU score
0.9208 or 92.08%	0.6	0.8533 or 85.33%

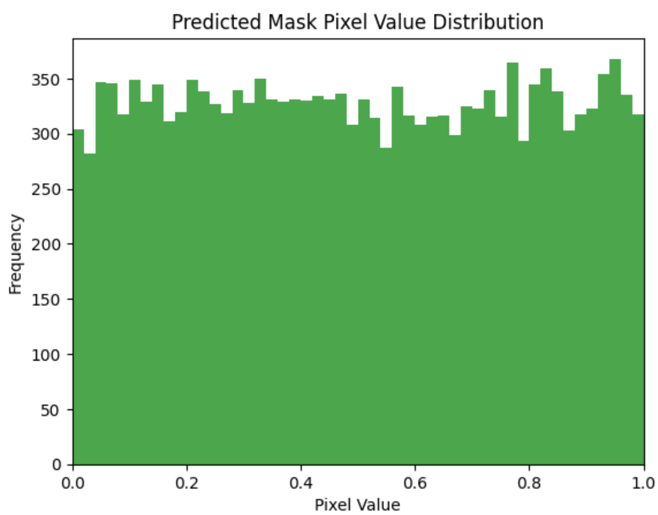

Fig. 5: Pixel Value Distribution of Predictions

The image presents the histogram of pixel intensities of the mask predicted by the model. The x-axis shows the range of pixel values, which ranges from zero to one, where zero stands for the darkest pixel and one for the brightest one. The pixel value on the y-axis refers to the number or percentage of sack lengths as recorded in the predicted mask. The distribution looks quite free and balanced, which shows that the model allows for a lot of different pixel values within its predictions. Such an even distribution may mean that the model is able to tell different features in the image apart, and is not too biased towards one particular intensity value when creating the masks.

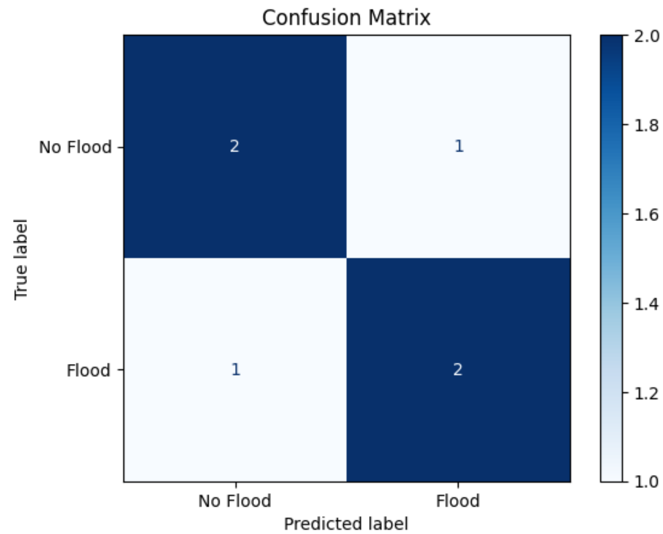

Fig. 6: Confusion Matrix

A confusion matrix is a table which is used to evaluate the performance of a classification model by comparing actual and predicted labels. It brakes down prediction values in 4 categories precision, recall, f1-score and support.

TABLE II: Classification Report

	Precision	**Recall**	**F1-score**	**Support**
No Flood	0.67	0.67	0.67	3
Flood	0.67	0.67	0.67	3
Accuracy			0.67	6
Macro avg	0.67	0.67	0.67	6
Weighted avg	0.67	0.67	0.67	6

Here the classification report is a visualization of a performance evaluation tool that provides key metrics for a classification model. It is generated based on confusion matrix and includes 4 metrics (Precision, Recall, F1-score & Support.

VI. Limitations

TABLE III: Limitations for Flood Detection using U-Net

Limitation	**Description**
Generalization Issues	The model is least generalized when regions of flood incidence depend on geographical, climate, or image quality variations. This reduces the model's effectiveness in areas not included in the training dataset.
Image Quality and Dataset Size	Higher quality annotated data improves efficiency. A dataset with less than 1,000 data points or low-resolution images may degrade model performance.
Class Imbalance	Non-flooded areas are often over-represented, causing the model to favor these regions. This understates flood-prone areas and creates a bias towards non-flooded regions.
Overfitting	Without proper regularization or data augmentation, the model may overfit to the training data, resulting in poor performance on unseen test images.

VII. Conclusion

To Conclude, this paper experiment with a U-Net based CNNs for flood detection applying the image segmentation approached which we were able to achive segmentation accuracy more than 76%. Proposed solution showed deep learning models for flood detection based on satellite imagery. The preprocessing steps, i.e., Image loading and Data augmentation was very helpful in improving the model performance. While the model has performed satisfactorily, achieving even higher segmentation accuracy may be yet another obstacle to cross. The talk will largely be focused around refining this model further by using more advanced data augmentation, geometric features and loss functions. All of these improvements will help make this system more robust and adaptable for various satellite imagery, leading to a new generation of accurate and dependable flood detection systems.

VIII. Future Work

For the future work, our suggested approach is to gather Iot devices such as sensors and cameras present in flood prone areas and then utilize U-net model in the online fashion for the flood identification. Mobilization of resources in response to flood alert under this system will be informed by IoT collected data on rainfall, River levels and soil moisture to activate flood alert and prediction system using U-net model.

Since some concepts can be implemented in the edge computing model, the model should be able to run locally and there will not be much delay in the analysis process, the data bandwidth consumed should also be small. Nevertheless, to discuss the model's potential in detail and to investigate the truth of developed hypotheses various geographical regions will be employed.

The envisage approach is used as a strategy because it seeks to enhance the disaster risk reduction through providing the disastrous occurrence especially flood up-to-date information to the local jurisdictions and the public. The following will also be presented in the future work to determine where, how and when the predicted floods could easily be dealt with by the stakeholders like the flood victims before well appropriating the resources.

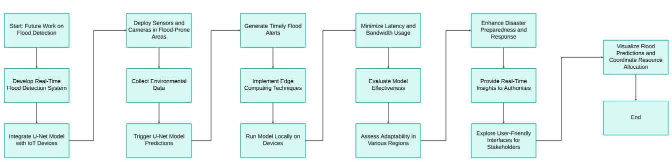

Fig. 7: Future Work Diagram

References

[1] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 234-241.

[2] Çiçek, Ö., Abdulkadir, A., Lienkamp, S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 424-432.

[3] Isensee, F., Petersen, J., Kohl, S., & Maier-Hein, K. (2018). nnU-Net: Self-Adapting Framework for U-Net-Based Medical Image Segmentation. arXiv preprint arXiv:1809.10486.

[4] Oktay, O., Schlemper, J., Folgoc, L. L., et al. (2018). Attention U-Net: Learning Where to Look for the Pancreas. Medical Image Analysis, 86-94.

[5] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3431-3440.

[6] Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834-848.

[7] Shorten, C., & Khoshgoftaar, T. M. (2019). A Survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6(1), 60.

[8] Milletari, F., Navab, N., & Ahmadi, S. A. (2016). V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. IEEE Conference on 3D Vision, 565-571.

[9] Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 424-432.