FAKULTÄT FÜR
INFORMATIK

DIGITAL ENGINEERING PROJECT

Automated Detection of MR Imaging Biomarkers of Cerebral Small Vessel Disease

Date:25-Aug-2020

Abhivanth Murali
Balakrishnan Ramakrishnan
Harish Kumar Harihara Subramanian
Logesh Babu Radhakrishnan

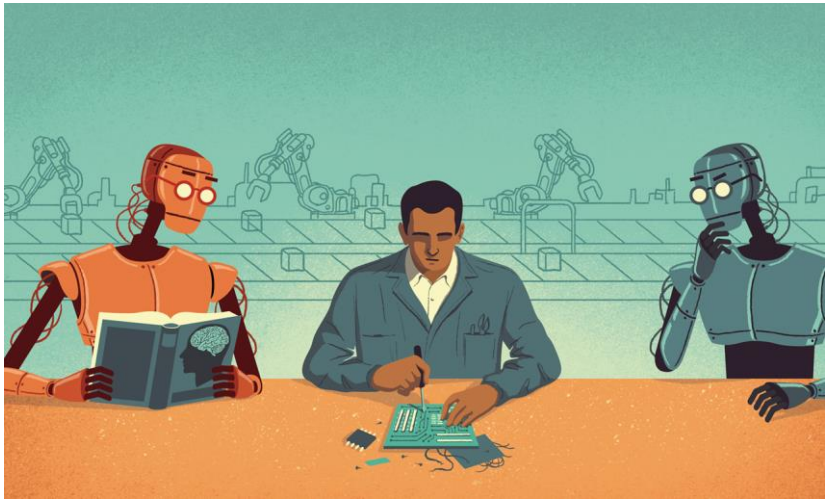
Supervisors:
Alessandro Sciarra
Max Dünnwald

Table of contents

Motivation

⇦ Problem statement

Solution ⇨



Introduction

- Cerebral Small Vessel Disease (CSVD) is an umbrella term for disorders related to small blood vessels in brain.

- Its presence is indicated by biomarkers in MRI scans.

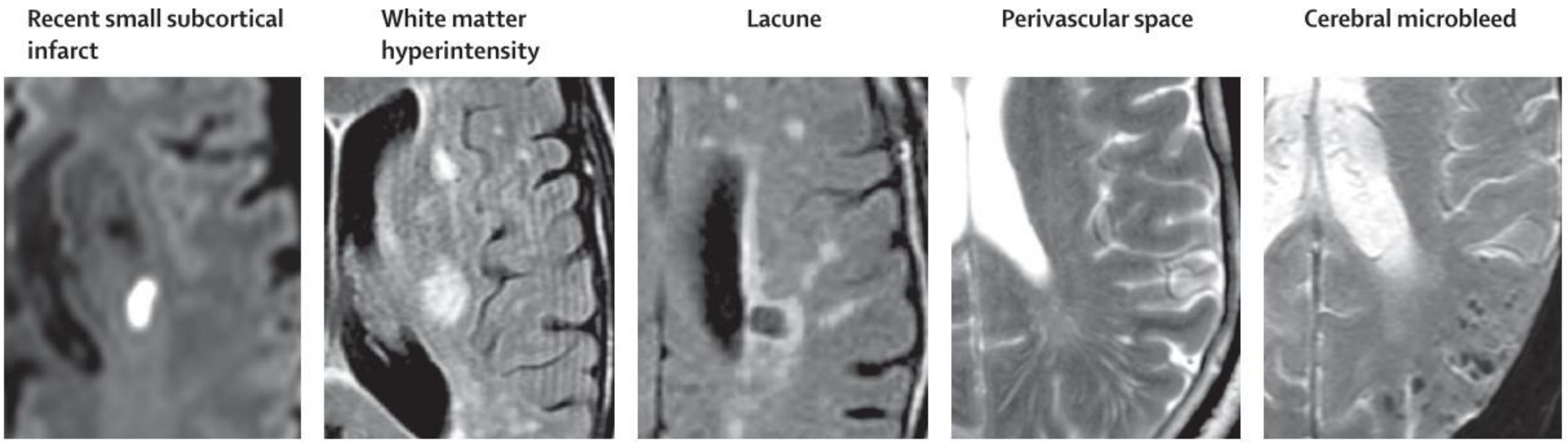


Fig 1: Cerebral Small Vessel Disease Biomarkers

The below biomarkers are selected for this project

White Matter Hyperintensity

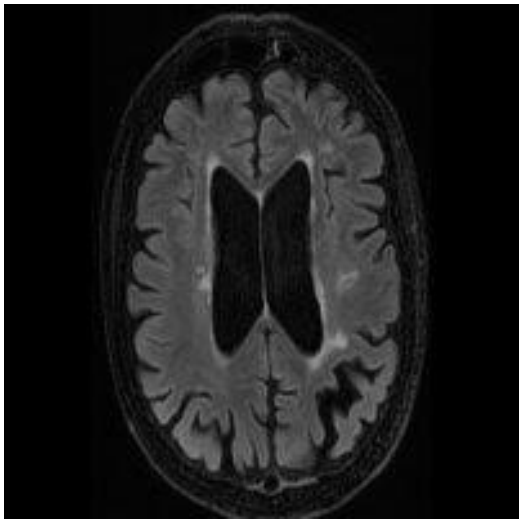


Fig 2: WMH

Cerebral Microbleeds

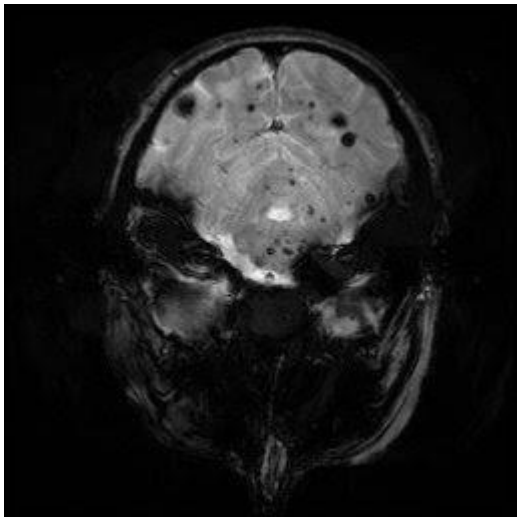


Fig 3: CMB

Literature research - WMH

The implementation of WMH biomarker is based on the paper published by **Hongwei Li et al [1]**, because

- This algorithm is evaluated and ranked 1st in the MICCAI 2017 challenge
- It performed well on a hidden set of 110 cases from 5 different scanners – *Good accuracy across different scanners*
- It is evaluated against 4 other methods and achieved the highest results in 3 out of 5 evaluation metrics
- High scores - **Averaged dice score - 80%, Precision - 84% and robust Hausdorff distance – 6.30mm**
- Low Scores – **Average volume difference - 21.88 , F1-Score - 78%**
- So this method is recognised as the state-of-the-art method

[1] H. Li et al., "Fully convolutional network ensembles for white matter hyperintensities segmentation in MR images", *Neuro Imag.*, vol. 183, pp. 650-665, Dec. 2018

Literature research – CMB

The implementation of CMB biomarker is based on the paper published by **Qi Dou et al [2]**, because

- This method achieves much better detection accuracy, as it takes full advantage of spatial contextual information in MR volumes to extract more representative high-level features for CMB's
- This model is designed based on the cascaded framework which is a unique technique to improve the detection performance while reducing the computational cost
- So, compared to the traditional sliding window strategy, this method removes the redundant computations and dramatically speed up the detection process

[2] Dou, Q. , Chen, H. , Yu, L. , Zhao, L. , Qin, J. , Wang, D. , Mok, V.C. , Shi, L. , Heng, P.-A. , 2016b. Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. IEEE Trans. Med. Imaging 35, 1182–1195 .

Dataset

White matter hyperintensities (WMH):

- Dataset acquired from the event “MICCAI WMH Segmentation Challenge”
- https://wmh.isi.uu.nl/data/.
- 60 subjects from 3 different hospital

Cerebral microbleeds (CMB):

- Dataset acquired from Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset
- 3,605 subjects in total with ground truth coordinates of CMB in meta file

White matter hyperintensities - Pre-processing

- Uniform size
- Normalizing intensity – Gaussian normalization & morphological fill operations
- Extract axial slices

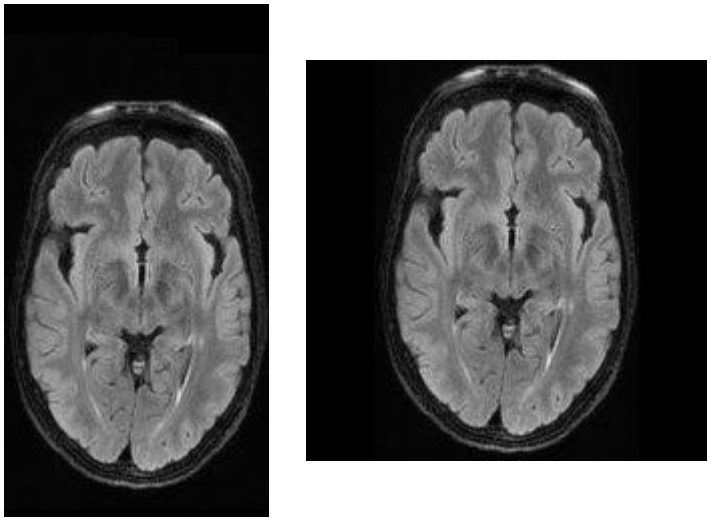


Fig 4: Actual and pre-processed FLAIR slice

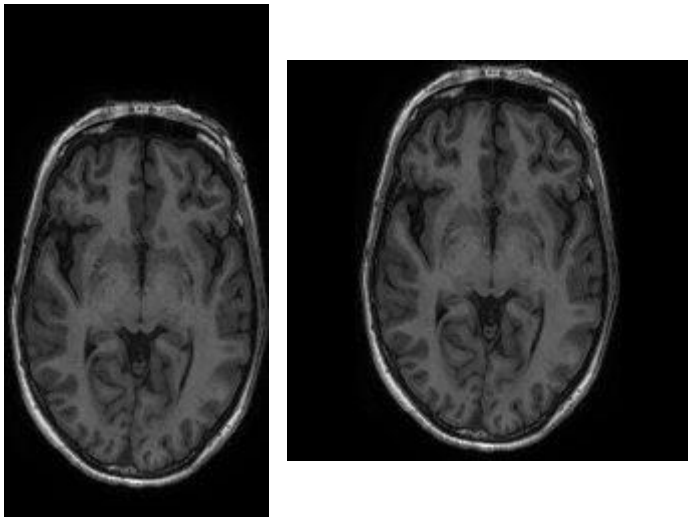


Fig 5: Actual and pre-processed FLAIR slice

White Matter Hyperintensities - Architecture

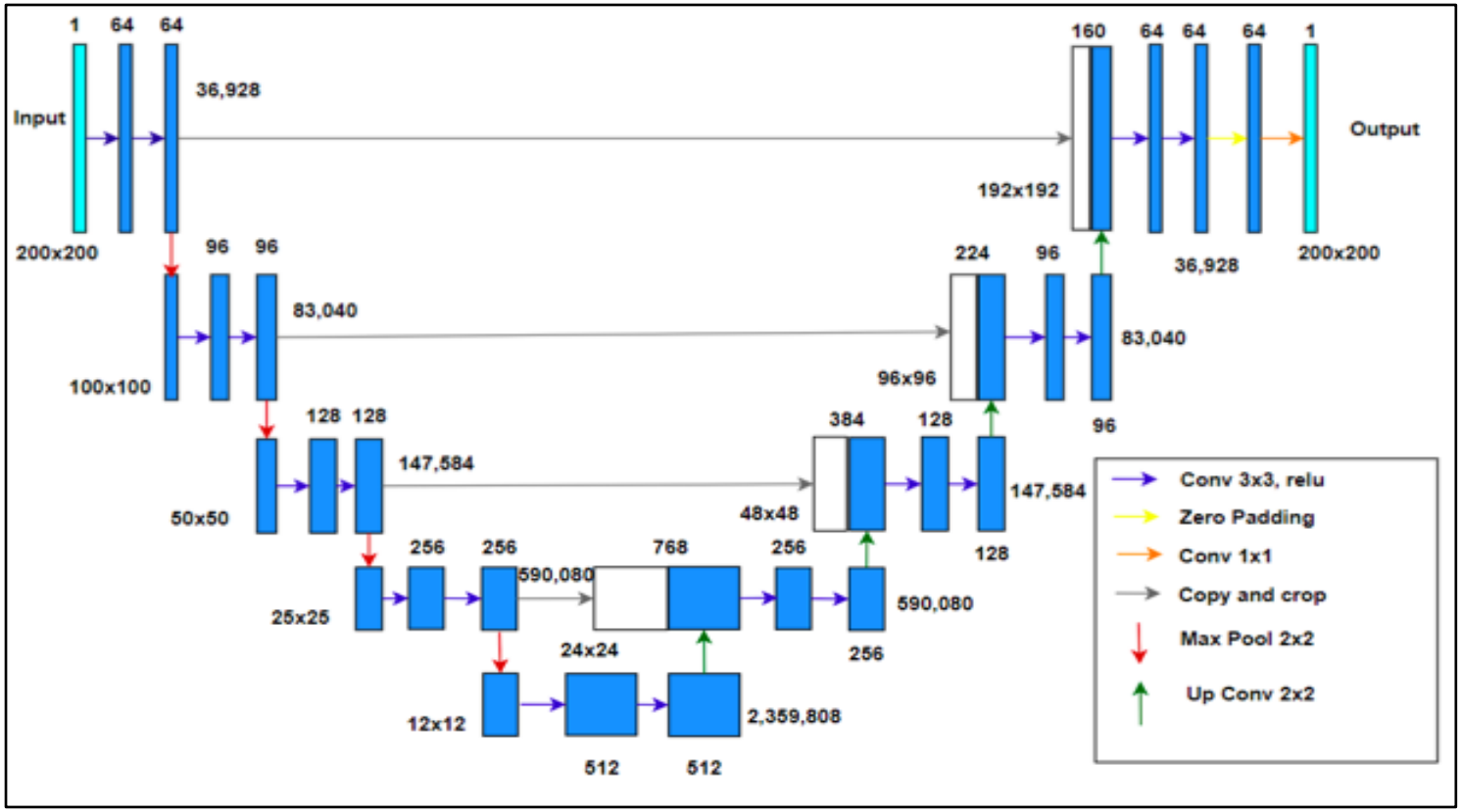


Fig 6: U-Net Architecture [3]

[3] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 234–241, 2015..

White matter hyperintensities – Held Out Evaluation

- In this method, the dataset is split into 'Train' and 'Test' set
- Using the training set the model will be trained and the test set is used to see the model performance on the unseen data.
- The common split will be 80% of data for training and 20% for testing
- So, we used 85% of data for training and the remaining 15% for testing

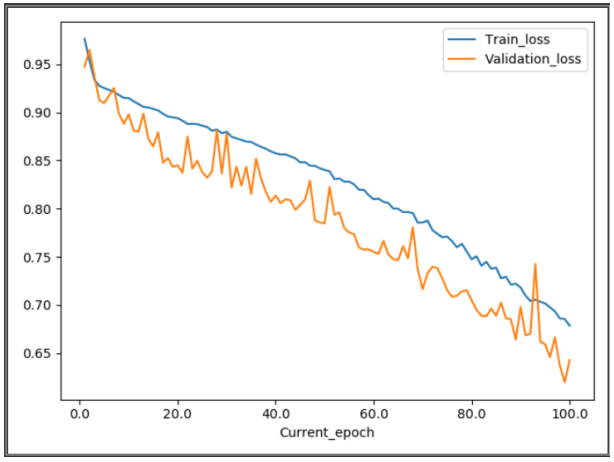


Fig 7.1: Held out with out augmentation (0 - 200 epochs)

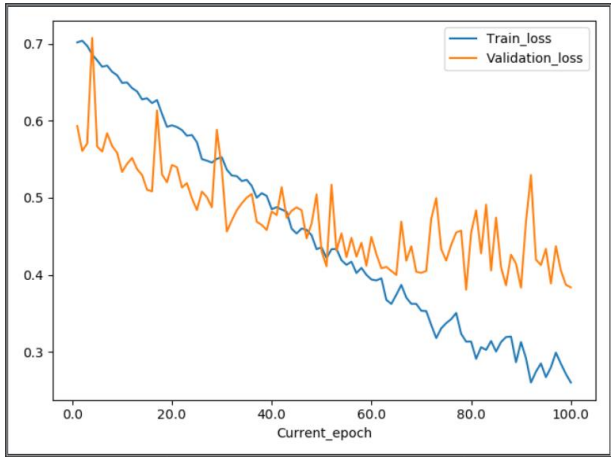


Fig 7.2: Held out with out augmentation (100 - 200 epochs)

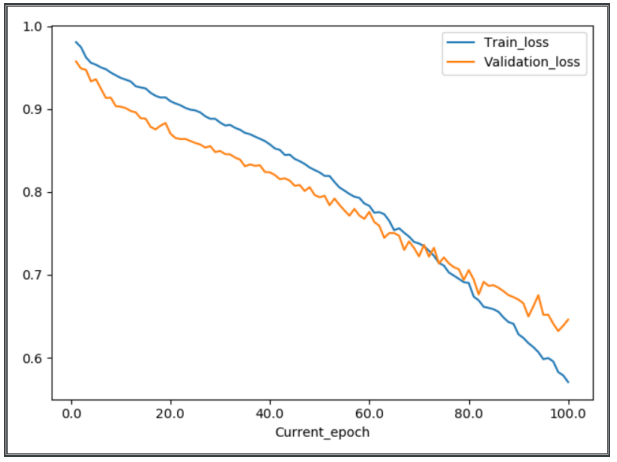


Fig 7.3: Held out with augmentation (0 - 200 epochs)

Hyper Parameters- Optimizer : Adam, Learning rate : 2e-4, Batch size : 16, Epochs : 100

White matter hyperintensities – Cross Scanner Evaluation

- In this method, the dataset is split with respect to the scanners
- The common split will be 1 scanner will be used to test and the remaining for training
- As we got 3 datasets from 3 different scanners, we used 2 datasets to train and 1 for testing

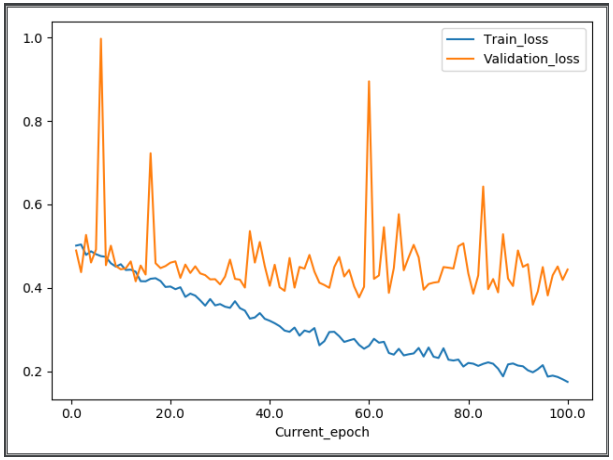


Fig 8.1: Cross scanner - Singapore dataset (100 – 200 epochs)

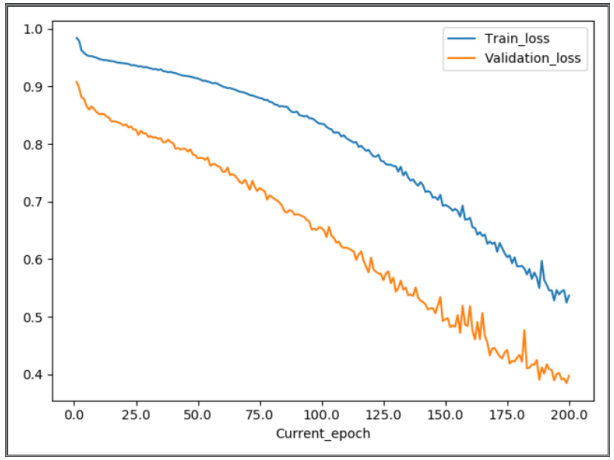


Fig 8.2: Cross scanner - Utrecht dataset (0 – 200 epochs)

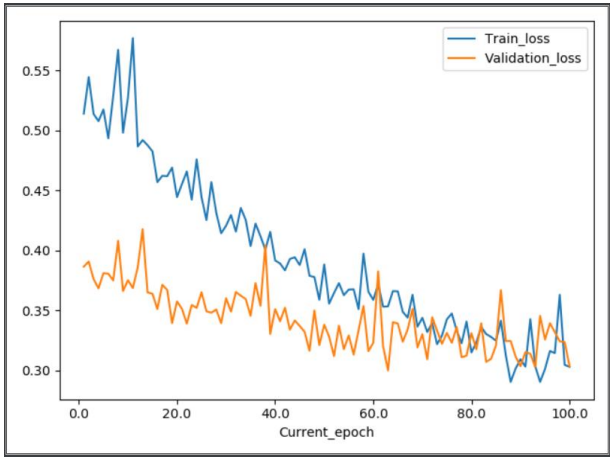


Fig 8.3: Plot loss Cross scanner - Utrecht dataset (200 – 300 epochs)

Hyper Parameters- Optimizer : Adam, Learning rate : 2e-4, Batch size : 16, Epochs : 100

White matter hyperintensities – Cross Scanner Evaluation

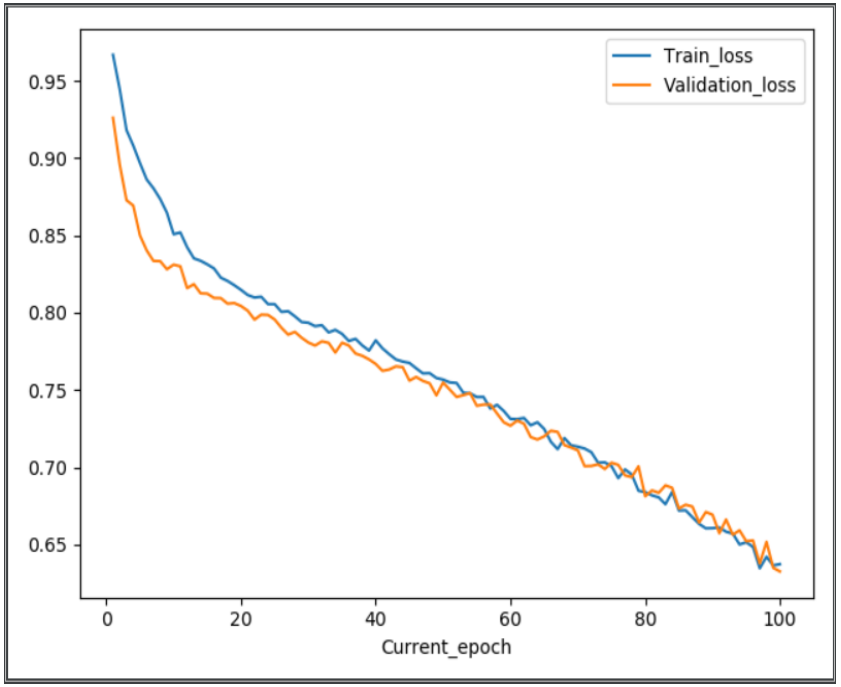


Fig 8.4: Cross scanner – Ge3t dataset (0 – 100 epochs)

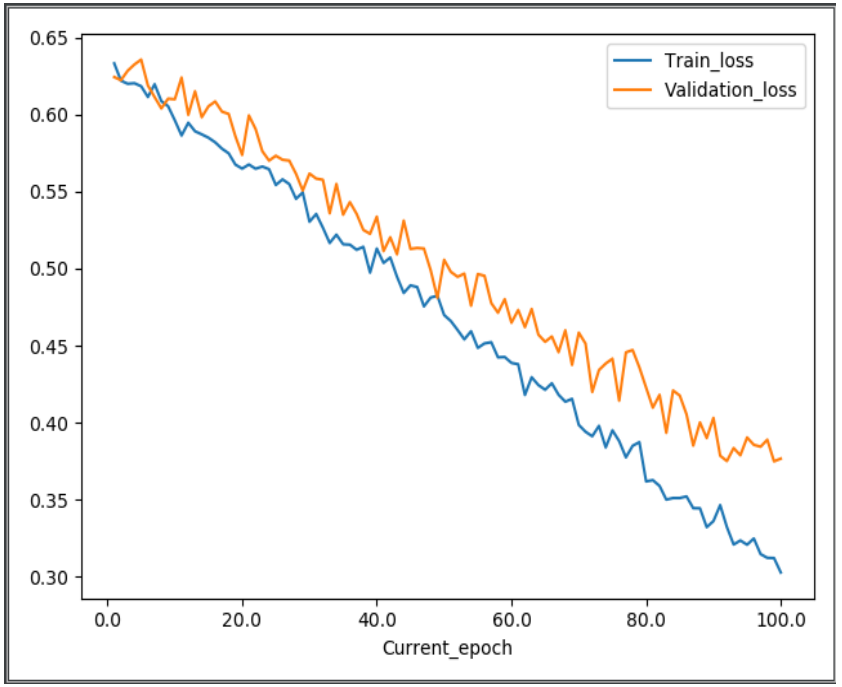


Fig 8.5: Cross scanner – Ge3t dataset (100 – 200 epochs)

White matter hyperintensities - Results

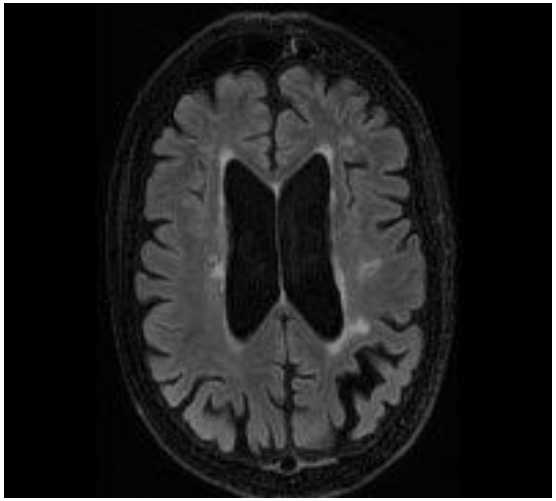


Fig 9: Pre-processed Slice

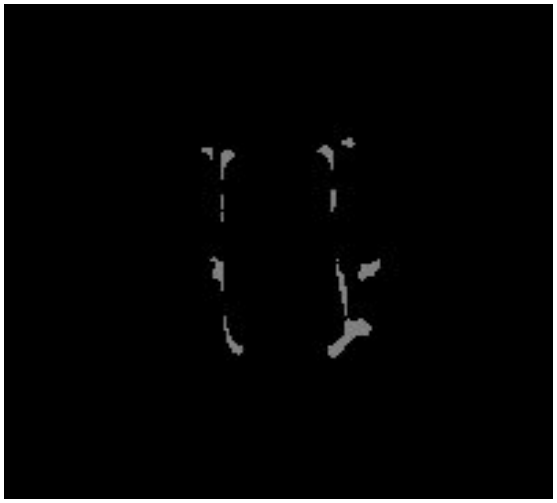


Fig 10: Ground truth Slice

Fig 11: Predicted result (Held-Out)

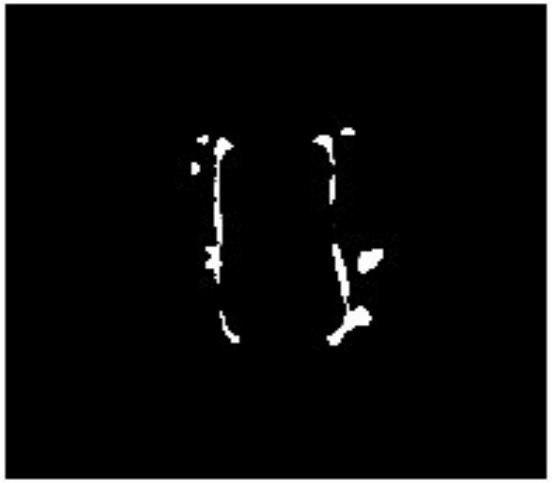


Fig 12: Predicted result (Cross-Scanner)

White matter hyperintensities - Visualization

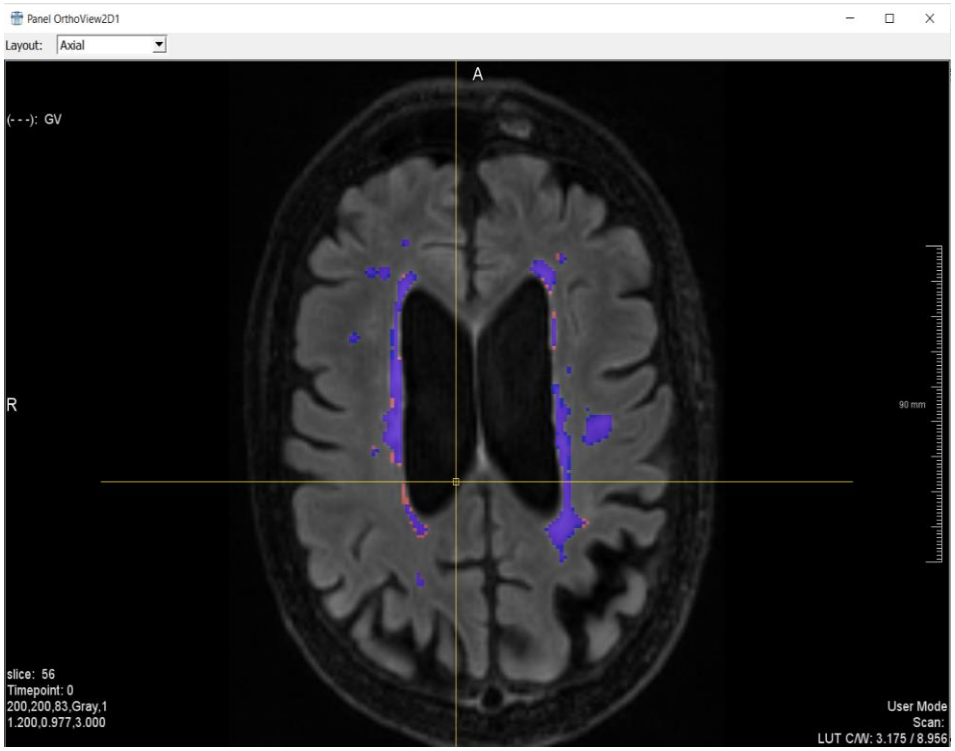


Fig 13: Overlaid view: Held-Out

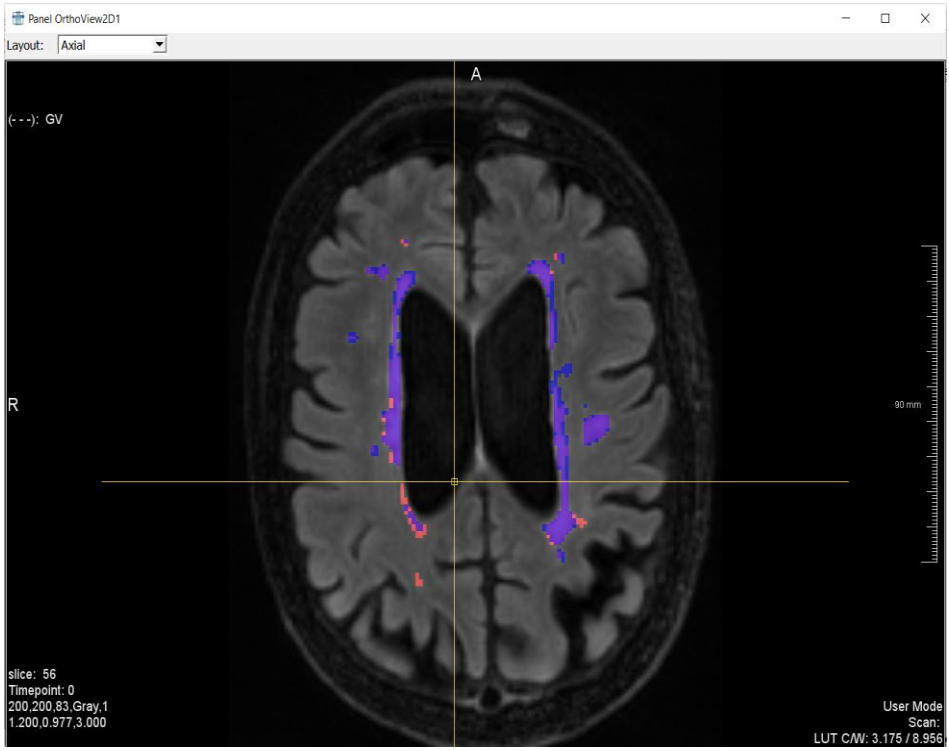


Fig 14: Overlaid view: Cross-Scanner

Actual Ground Truth

Predicted Ground Truth

White matter hyperintensities - Metrics

- The held-out metrics in the research paper are higher compared to our implementation as the results were obtained from the 110 unseen subjects (5 different scanners)

- The cross-scanner metrics in the research paper is close to our results from Singapore dataset. They have achieved the results from 2 unseen scanners

- Unable to fetch the results for Cross-Scanner evaluation for Utrecht dataset due to overfitting issue.

Predicted	Average DSC*	Average AVD*	Average Recall	Average F1 score
Held-Out (*with out augmentation*)	0.59	62.12	0.65	0.47
Held-Out (*with augmentation*)	0.56	67.11	0.60	0.48
Cross-Scanner (*Singapore*)	0.65	44.09	0.74	0.66
Cross-Scanner (*Ge3t*)	0.63	48.30	0.66	0.68

Table 1: Predicted results

Actual Results from paper	Average DSC*	Average AVD*	Average Recall	Average F1 score
Held-Out	0.80	21.88	0.84	0.76
Cross-Scanner	0.745	26.2	0.87	0.725

Table 2: Actual results

***DSC = Dice Similarity Coefficient** ***AVD = Average Volume Difference** ***H95 - Hausdorff distance**

Cerebral Microbleeds - Pre-processing

- Convert DICOM to NIfTI images
- Generate ground truth
- Normalize input images
- Split the subjects based on number of CMB
- Generate 3D patches of size “16x16x10”

	Subjects	Positive patches
Train	325	2833
Validation	37	571
Test	139	738
Total	501	4142

Table 3: CMB Dataset split

Cerebral Microbleeds - Cascaded Network

Layer	Kernel Size	Stride	Output Size	Feature Volumes
Input	-	-	16 x 16 x 10	1
Convolution-1	5 x 5 x 3	1	12 x 12 x 8	64
Max pooling-1	2 x 2 x 2	2	6 x 6 x 4	64
Convolution-2	3 x 3 x 3	1	4 x 4 x 2	64
Convolution-3	3 x 3 x 1	1	2 x 2 x 2	64
Fully Convolution-1	2 x 2 x 2	1	1 x 1 x 1	150
Fully Convolution-2	1 x 1 x 1	1	1 x 1 x 1	2

Table 4: Screening Stage Network Architecture

Cerebral Microbleeds - Cascaded Network

Layer	Kernel Size	Stride	Output Size	Feature Volumes
Input	-	-	20 x 20 x 16	1
Convolution-1	7 x 7 x 5	1	14 x 14 x 12	32
Max pooling-1	2 x 2 x 2	2	7 x 7 x 6	32
Convolution-2	5 x 5 x 3	1	3 x 3 x 4	64
Fully Connecetd-1	-	-	1 x 1 x 1	500
Fully Connected-2	-	-	1 x 1 x 1	100
Fully Connected-3	-	-	1 x 1 x 1	2

Table 5: Discrimination Stage Network Architecture

Screening Stage

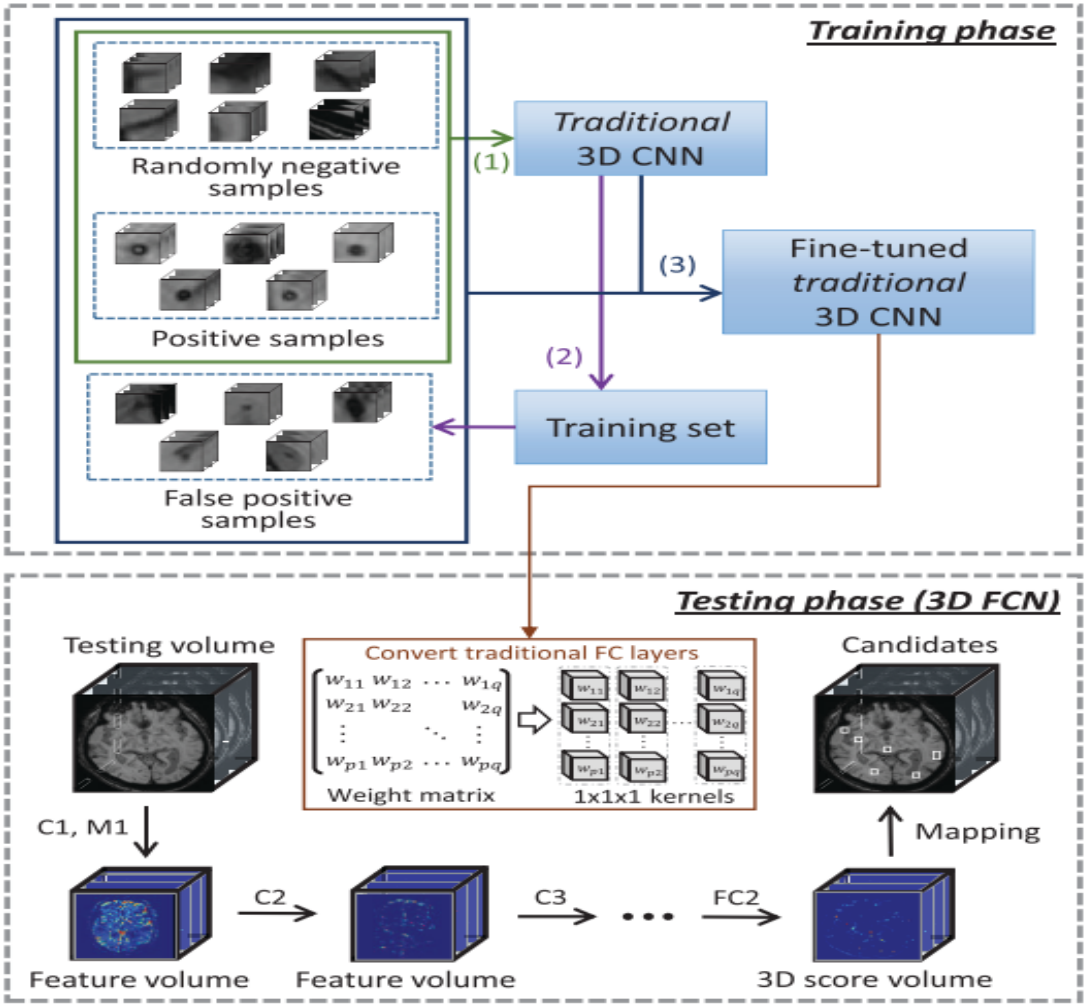


Fig 15: Screening stage workflow

Cerebral Microbleeds - Results

- Screening stage is divided into 3 sub-steps during training the network.
  - *Step 1 – The network is trained with balanced dataset (Equal number of positive and negative patches) and the Sensitivity is calculated as 95%*
  - *Step 2 – The false positive patches(mimcs) are extracted and the dataset is enhanced with true positives (23.63%), false positives (28.80%) and true negatives (47.5%)*
  - *Step 3 – The network is further trained with the enhanced dataset and the Sensitivity is calculated as 41%*
- The network did not perform well and the metrics calculated are very less compared to the results from the research paper.
- During testing, the whole volume input is provided to the network. The non max suppression technique is implemented and the score map is calculated to trace to the original coordinates.
- Due to the poor performance of the network, we are not able to trace to the original coordinates
- Unable to perform the Discrimination Stage, as the score map does not trace to the original coordinates.

Web API – Home Page

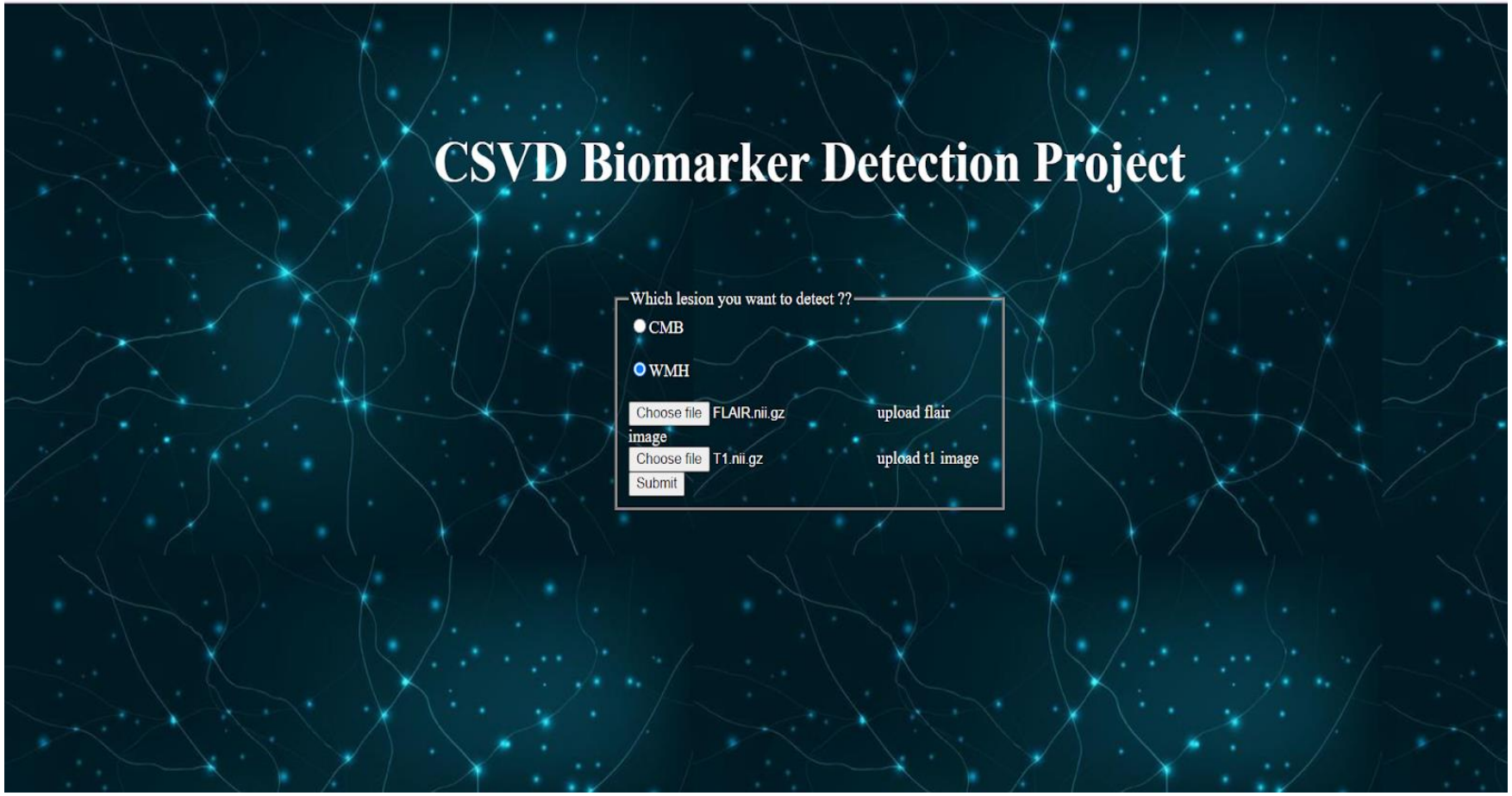


Fig 16: UI Home Page

Web API – Output Viewer Page

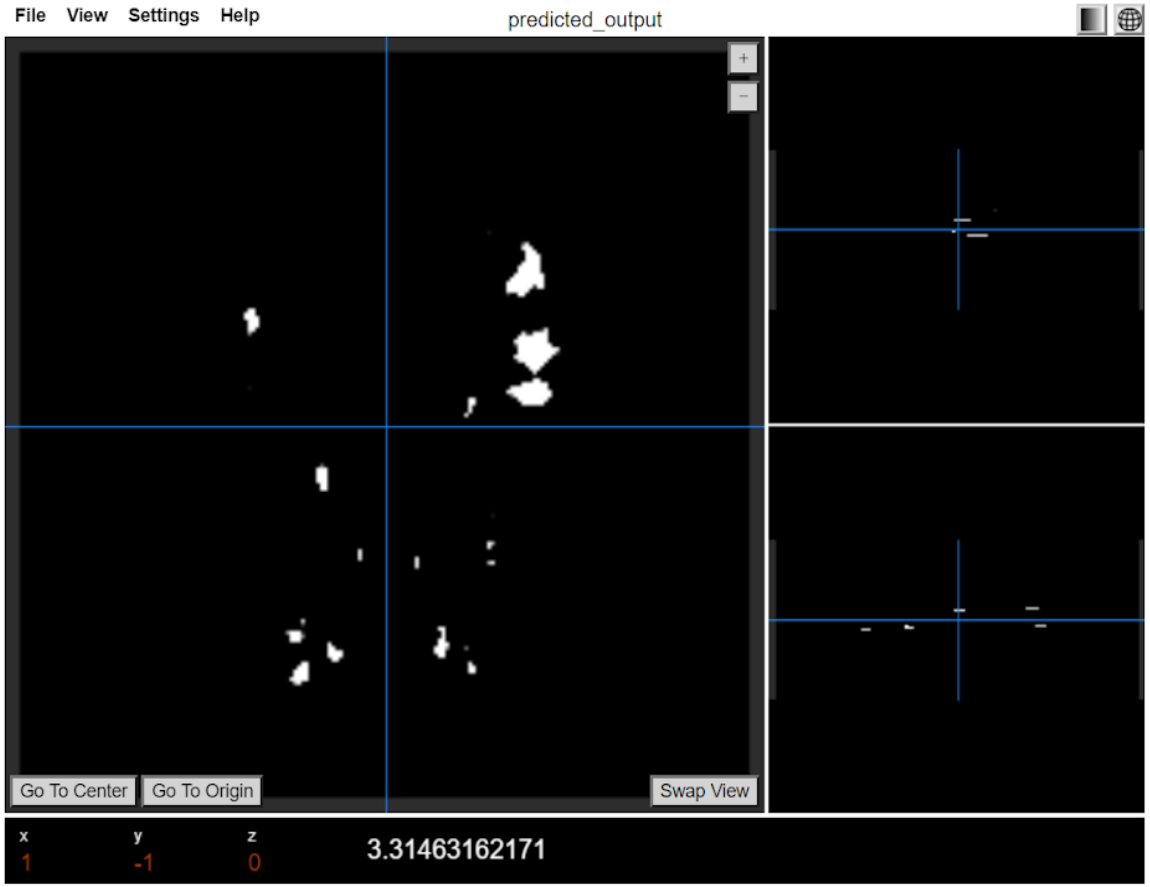
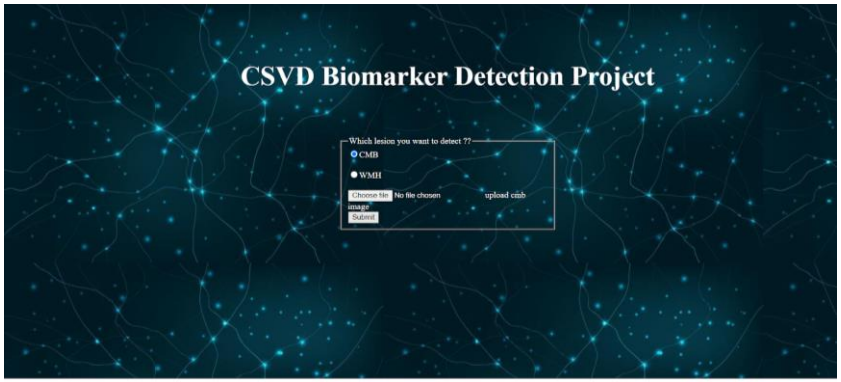
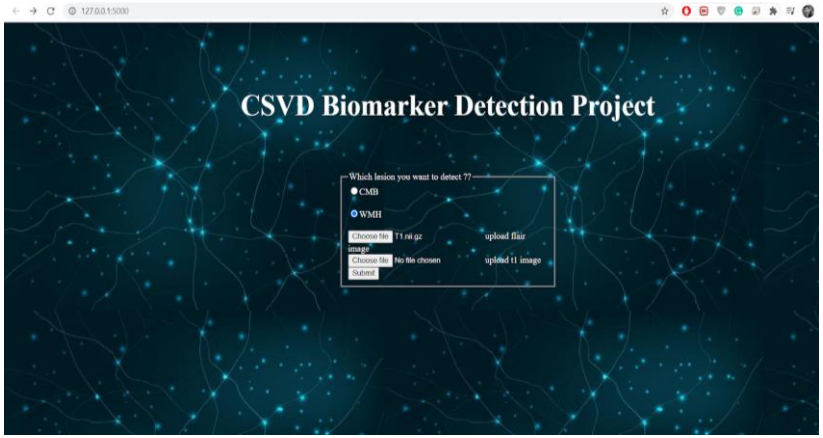
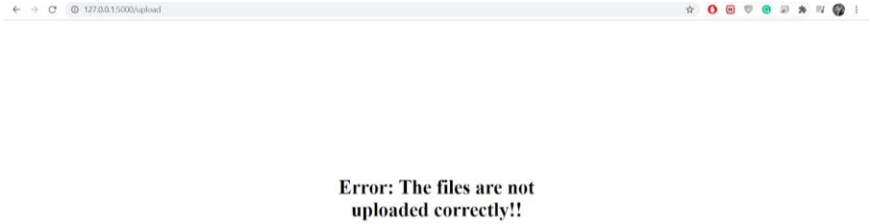


Fig 17: Output displayed in papaya viewer plugin

User interface – Error Handling







Q & A

THANK YOU