Determining the Impact of Motion Blur on the Segmentation of Tumors in Brain MR Images Using U-Net Architecture

Aidan Rogers, Ruochen Wang

Abstract:

Acquiring Magnetic Resonance Images of the brain is one of the most useful tools for identifying neural trauma and/or tumors. This accuracy comes at the expense of the patient who must have their heads uncomfortably fixed to a bed while subjected to a long, noisy scanning period. The head constraint is necessary to reduce motion artifacts, which makes it very difficult for experts to identify where trauma or tumors are in the MR images. By identifying which motion blur artifacts are most detrimental to the data, we can determine what sides of the head must be more constrained. After artificially blurring our images with a stride of 1x1 and a 3x3 Kernel, we used a Convolutional Neural Network (CNN) with U-Net architecture to optimize deblurring while tracking what motion blurs were most and least detrimental to tumor identification or segmentation. Our results yielded >98% accuracy when deblurring and segmenting MRI data. Using probabilistic optimization we also concluded that all motion blurs negatively impacted our results but that motion blurs to the left, right and down-left were the most detrimental across models and trials.

1. Introduction

Magnetic Resonance Imaging (MRI) is an extremely useful tool for creating 3D models of human anatomy, including the brain [1]. Image segmentation is effective for a multitude of tasks including quantifying and visualizing structures of the brain's anatomy, detecting changes in brain activity, delineation of pathological regions and giving insights to surgical planning or interventions based on images [2]. When used in combination, researchers can parse through and identify specific regions in MRI data for analysis. Segmentation specifically uses a technique to identify edges, which are defined by sharp changes in image intensity, to identify desired sections in the brain. However, generating detailed maps of complex structures like the brain comes with drawbacks, mainly background noise and image scanning time [2]. Edge detection algorithms may suffer from noise presented in images and require preprocessing steps [3,4]. This in combination with high costs and long image scanning times raises the necessity for head fixtures to ensure that patients cannot move their heads [5]. This poses a major issue since many patients struggle with staying completely still during imaging due to stress, loud noises and pain; in fact, 2,000,000 (2.3% of all MRI sessions) are canceled due to patients experiencing claustrophobia [6]. Because of this our team is aiming to determine which directions of motion blur are most detrimental to our MRI images so as to optimize head fixation and therefore increase patient comfort.

Our team is specifically using the Medical Image Computing and Computer Assisted Intervention (MICCAI) 2020 dataset of carcinogenic (tumorous) 3D MRI scans of 369 patients [6]. The dataset includes 3D voxel-wise maps using a multitude of MRI slices for each patient, as well as prelabeled data for training our CNN. Our team will focus on how motion will affect our CNN segmentation results. We would like to propose and train an end-to-end CNN that will segment brain tissues on these motion blurred images and correct for the blur. Our neural net

aims to replicate possible motion-blur from Brain MRI scans and identify which is the most detrimental to the data. Determining this could lead to better instruction for patients before MRI scans and for future designs for head constraints.

1.1 DataSet

The dataset used in this project is MICCAI Brain Tumor Segmentation (BraTS) 2020 [7]. The BraTS challenge started in 2012, and it is used for evaluating and comparing various algorithms for automatic segmentation of brain tumors. Multiple MRI scans are released every year and come with multiple modalities: T1-weighted, T1-post contrast, T2-weighted and FLAIR (fluid-attenuated inversion recovery). The gold standard of the brain tumor segmentation mask provided by the BraTS dataset was manually segmented and evaluated by neurologists [7]. The segmentation masks finally have 4 labels: the GD-enhancing tumor (ET — label 4), the peritumoral edema (ED — label 2), the necrotic and non-enhancing tumor core (NCR/NET — label 1), and the other normal tissues in the brain and background (label 0).

2. Related Works

Many previous works are done on the topic of image segmentation with deep neural networks. Among these works, U-Net is a CNN structure that was originally presented with the application of biomedical image segmentation and showed fast inference (less than a second for an image with one GPU) [8]. There are also generalized structures for 3D image segmentation with U-Net, which have been proved to be successful in image segmentation tasks [9].
Deep neural network approaches for removing the blur from images have also been proven to be efficient. Instead of estimating the blur based on priors, Gong et al. proved deep learning architectures can learn based on training so as to estimate the motion kernel [10]. According to these related works, we proposed using U-Net to conduct segmentation on artificially blurred images.

3. Methods

3.1 Preprocessing

To manage the computation complexity to a reasonable scale, the model was only trained and evaluated with FLAIR modality images. We chose FLAIR images because they tend to attenuate the view of the fluid in the brain and accentuates the tumor region [11]. Moreover, only the center horizontal slices were extracted for each subject.

The intensity of the FLAIR images vary for all subjects since these images were obtained from different individuals and multiple institutions [7]. The aim of the preprocessing step for the pipeline is to bring all brain scans to the same scale. Each slice is normalized based on the maximum intensity value in that slice before further processing in the model.

Even though multiple labels are present in the dataset, the labels could be further categorized into tumor and non-tumor. After transforming the label, there are 2 labels present in the dataset, with 0 as normal tissue and background, 1 as the tumor tissues. To evaluate and compare with the binary label segmentation, multi-label segmentation is performed by reserving all the 4 labels in the dataset.

3.2 Physical Layer

Since our work is focused on how motion-blur affects our images, our physical layer consists of eight 3x3 motion blur kernels corresponding to 8 different directions: up, down, left, right, up left, up right, down left and down right. The gradients of all the kernels are kept the same as control, so that the performance could be compared between these blur kernels. To simulate this physical layer, MRI scans are treated as ground truth and a convolution with a stride of one is applied onto an image. This was done 8 times total, once for each of the blur kernels, and then the images were fed into an encoding layer. We used probabilistic optimization to track which kernels were used by the network in the encoding layer.

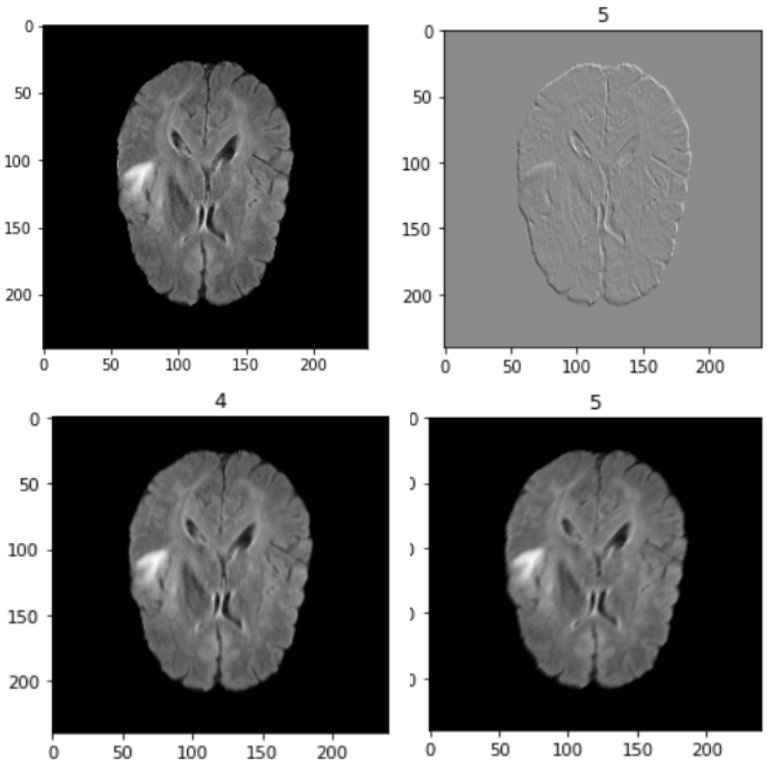


Figure 1: (top left) *Original Slice -* ***(top right)*** *subtraction of two blurred images -* ***(bottom left)*** *image blurred to the right -* ***(bottom right)*** *image blurred to the top left*

3.3 Probabilistic Optimization

Eight images were produced from the physical layer. Probabilistic optimization is a technique used to identify which image is passed into the network by sampling from a random distribution. By optimizing the probability distribution, the network will increase the probability of choosing one blurred image that favors segmentation accuracy out of eight artificially created blurred images. It selects the blurred image using a relaxed one hot categorical selection technique [12] (Figure 2).

3.4 U-Net Architecture

The encoding layer's output image undergoes a single batch normalization before being input to the U-Net since we use varying convolutions across epochs (Figure 2). The images are all of size 240x240 pixels and the U-Net starts with 8 channel filters. As we descend the contracting path of the U-Net, we use max pooling after every 2 convolutional layers (with increasing filter size) to reduce the image size down to 30x30 pixels as we increase the feature space to 64 channel filters. We then traverse the expanding path of the U-Net using upsampling (inverse of max pooling) after every 2 transpose convolutions eventually reaching an image of size 240x240 with 8 channels. This is then output to our softmax in our categorical cross-entropy model and sigmoid in our binary cross-entropy model (Figure 2). Additionally, we implemented skip connections for layers on either side of the U-Net (Figure 2), to make sure we retain key features detected before max pooling on the descending pathway.

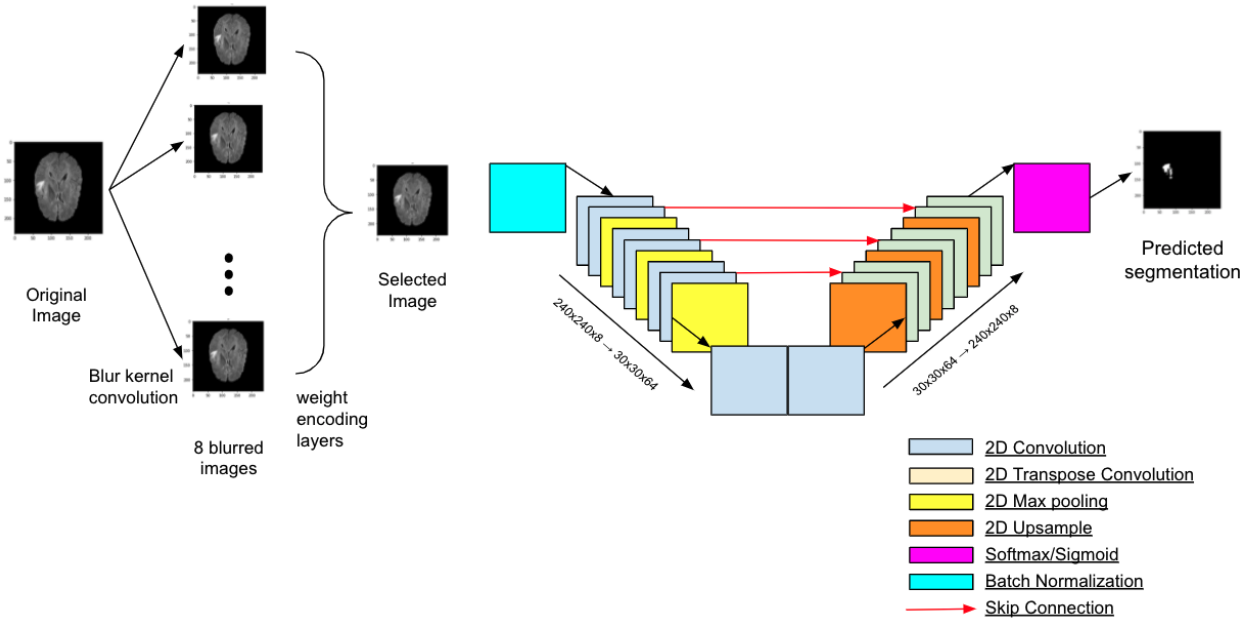


Figure 2: Full Pipeline

4. Results

The dataset we used has labels distinguishing where tumors are present in the MR images. We used our CNN to classify the images for both binary and categorical cross-entropy. 240x240 images were fed into the network and output as the same size. However, for the binary model, our labels were normalized between 0 and 1, whereas for the categorical they were between 0 and 3. Figures 3a and 3b depict our loss and accuracy for our Binary model and Figures 4a and 4b depict the same for our categorical model.

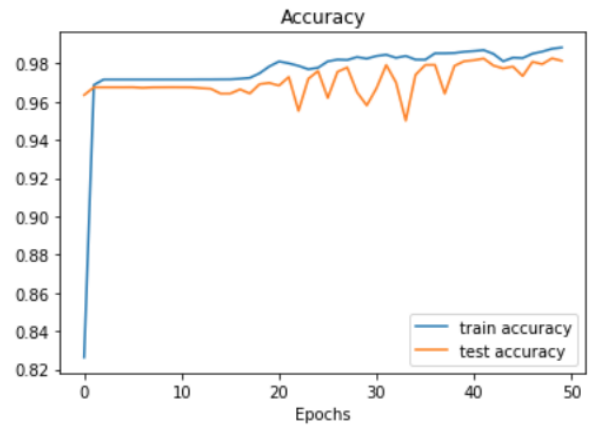


Figure 3a. Binary Model Accuracy

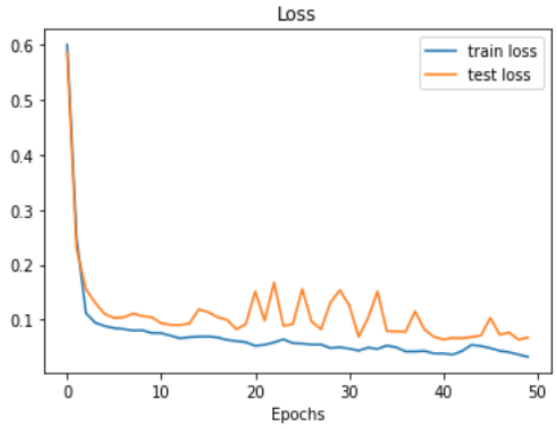


Figure 3b. Binary Model Loss

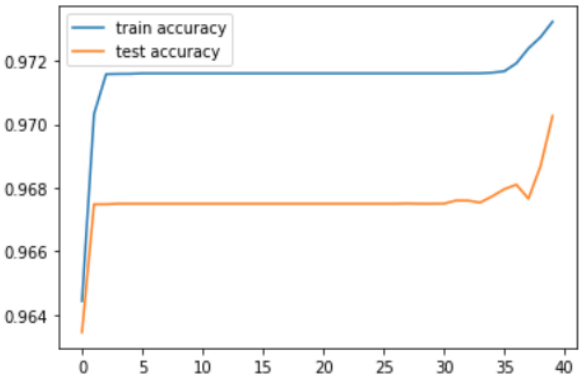


Figure 4a. Categorical Model Accuracy

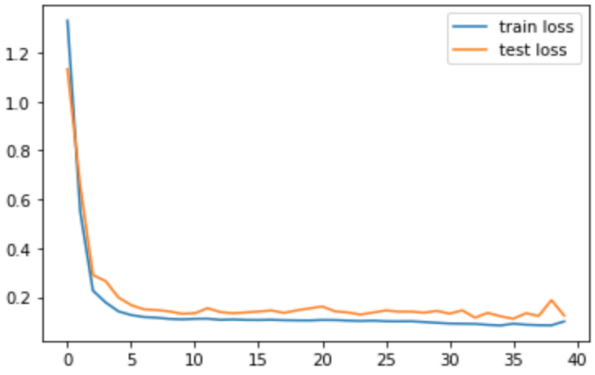


Figure 4b. Categorical Model Loss

Both models showed extremely impressive Figures for both testing and training accuracy or loss. In both models, we trained for over 30 epochs because they stagnated and after about 5 to 10 and then exceeded their previous local maxima of around 97.16% training accuracy. While we were analyzing the accuracy and loss of the model we also tracked the 8 weights attached to our blur kernels. The CNN optimizes itself, when choosing weights it would choose to utilize the weight associated with the blur kernel that was least detrimental to the CNN's accuracy/segmentation results. Those kernels that were most detrimental would have lower probabilities of being chosen. Our results are shown for two trials for each model in Figures 5 and 6.

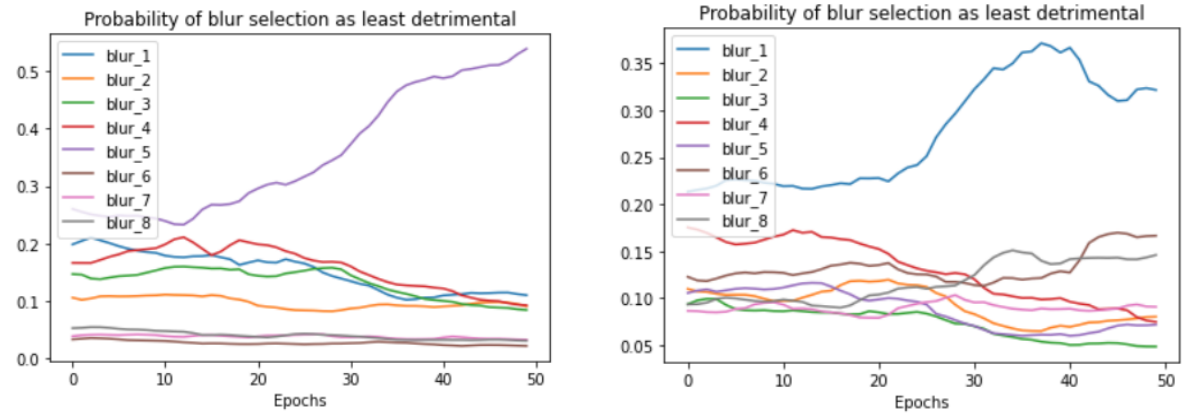


Figure 5 Binary Model Blur Kernel Probability Weights

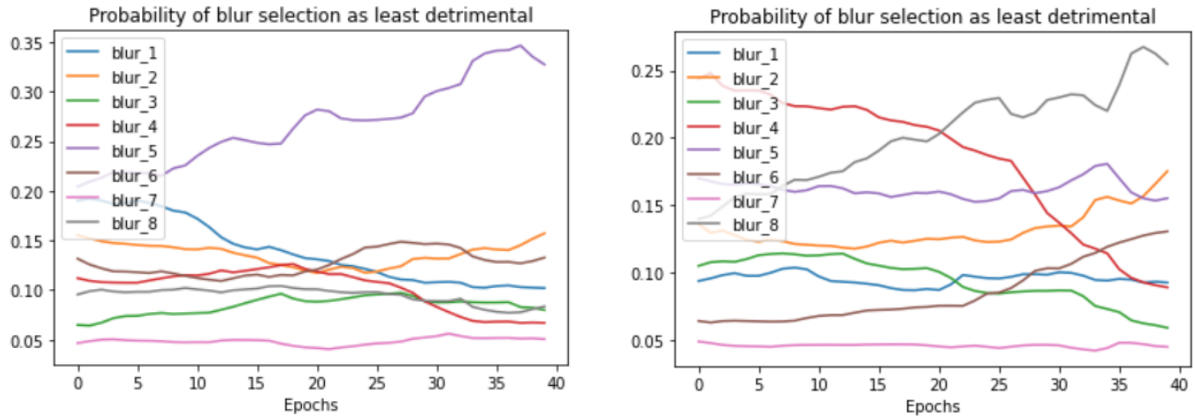
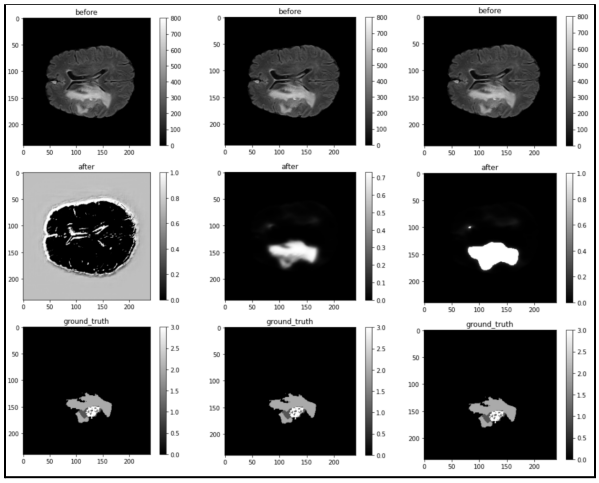


Figure 6 Categorical Model Blur Kernel Probability Weights

Once we saw our results were successful we had our model predict the segmentation labels before and after training for validation. We also thresholded our segmentation to see if that would improve our results (Figures 7 and 8).



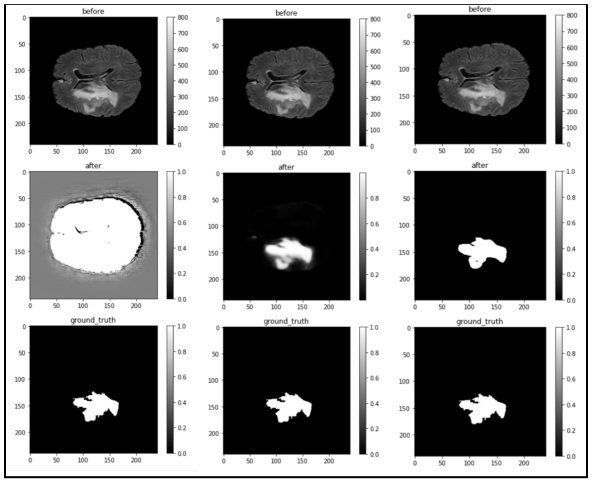


Figure 7 (left) Categorical: A untrained prediction, B: trained predicted, C: thresholded prediction
Figure 8 (right) Binary: A untrained prediction, B: trained predicted, C: thresholded prediction

5. Discussion

Our neural network produced very promising results. Both the categorical and binary cross-entropy models were able to segment with high accuracy and reproduce the labels provided by the MICCAI 2020 data set. Additionally, the probabilistic optimization of our physical layer yielded interesting results. Looking at Figures 5 and 6, it is clear that the majority of the blur kernels showed a similar probability to be selected for segmentation. Even though some kernels appeared as the least detrimental in specific runs, this was not consistent across trials or models. What was consistent was the low probability of blurs 7,3 and 4 (pink, green, red). These blur kernels corresponded with blurs down and to the left (pink), horizontally left (green)and horizontally right (red). The low probability of their selection deems them most likely to be most detrimental to the models ability to segment images with high accuracy. Kernels 3 and 4 specifically showed trends moving from high to low probabilities across epochs confirming that they were extremely detrimental to the network. Given our findings, we could take our research and attempt to design an updated head fixation cage. Given that our slices are horizontal blur kernels 3 and 4 deem that we should prioritize reducing head movement up and down the most.

Based on our promising findings it would make sense to continue working on this model, with which there are things we can still improve. The first would be to ensure that the blur kernels we were using were more realistic. We attempted this in our model by increasing the magnitude of the blurs, but this did not improve the efficacy all that greatly. To verify their efficacy we would use gaussian blurs or different size kernels. Another step we could take to improve the model would be to increase the training data size. The best way to make the data set more robust would be to modify our U-Net to be able to analyze 3D images instead of 2D central slices. We could do every single slice for every single subject, however by generating 3D images we would have a much more useful tool for segmentation. Additionally, we would use a 3x3x3 (3D) kernel, which would allow us to apply 54 (9x6) directional blurs rather than just 8.

References:

[1] T. Martín-Noguerol, F. Paulano-Godino, R. F. Riascos, J. Calabia-del-Campo, J. Márquez-Rivas, and A. Luna, "Hybrid computed tomography and magnetic resonance imaging 3D printed models for neurosurgery planning," Annals of Translational Medicine, vol. 7, no. 22, pp. 684–684, 2019.

[2] I. Despotović, B. Goossens, and W. Philips, "MRI Segmentation of the Human Brain: Challenges, Methods, and Applications," Computational and Mathematical Methods in Medicine, vol. 2015, pp. 1–23, 2015.

[3] J. Canny, "A computational approach to edge detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, no. 6, pp. 679–698, 1986.

[4] D. Marr and E. Hildreth, "Theory of edge detection," Proceedings of the Royal Society of London—Biological Sciences, vol. 207, no. 1167, pp. 187–217, 1980.

[5]Mandija, Stefano, et al. "Brain and Head-and-Neck MRI in Immobilization Mask: A Practical Solution for MR-Only Radiotherapy." Frontiers in Oncology

[6] J. Enders, E. Zimmermann, M. Rief, P. Martus, R. Klingebiel, P. Asbach, C. Klessen, G. Diederichs, T. Bengner, U. Teichgräber, B. Hamm, and M. Dewey, "Reduction of claustrophobia during magnetic resonance imaging: methods and design of the 'CLAUSTRO' randomized controlled trial," BMC Medical Imaging, vol. 11, no. 1, 2011.

[7] Menze B. H., Jakab A., Bauer S., Kalpathy-Cramer J., Farahani K., Kirby J., et al. . (2015a). The multimodal brain tumor image segmentation benchmark (brats). IEEE Trans. Med. Imaging 34, 1993–2024. 10.1109/TMI.2014.2377694

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Lecture Notes in Computer Science, pp. 234–241, 2015.

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

[10]D. Gong, J. Yang, L. Liu, Y. Zhang, I. Reid, C. Shen, A. Van Den Hengel, and Q. Shi, "From Motion Blur to Motion Flow: A Deep Learning Solution for Removing Heterogeneous Motion Blur," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[11]G. Vishnuvarthanan, M. P. Rajasekaran, N. A. Vishnuvarthanan, T. A. Prasath, and M. Kannan, "Tumor detection in T1, T2, FLAIR and MPR brain images using a combination of optimization and fuzzy clustering improved by seed-based region growing algorithm," International Journal of Imaging Systems and Technology, vol. 27, no. 1, pp. 33–45, 2017.

[12]E. Jang, S. Gu, and B. Poole, "Categorical Reparameterization with Gumbel-Softmax," ICLR 2017.