Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

Left Atrial Segmentation in Cardiac MRI using Deep Learning Approach

Rajendra Kumar Pandey
Department of Computer Science and Engineering
Shri Shankaracharya Institute of Professional Management and Technology, Raipur, India

Dr. Yogesh Kumar Rathore
Department of Computer Science and Engineering
Shri Shankaracharya Institute of Professional Management and Technology, Raipur, India

***Abstract*— Accurate segmentation of the Left Atrial (LA) in cardiac MRI is particularly crucial for proper diagnosis and treatment of a number of heart disorders. In this study, the authors offer a deep learning method employing the SegNet model for automatic LA segmentation from cardiac MRI data. The suggested methodology comprises extensive procedures including preprocessing in terms of scaling and normalization, followed by training and validation using a publically available dataset. Model performance is tested against standard segmentation metrics: Precision, Recall, and the F1 Score. In this situation, validation accuracy reached by the SegNet model is quite high, at 99.56%, while a DiceCoefficient of 0.378 and a Jaccard Index of 0.236 show excellent segmentation but also bring out areas where there may be space for additional progress in this respect. Precisions and recalls are both 0.739 and 0.039, respectively; consequently, it performs exceptionally well in two separate elements of segmentation. The training took roughly 2720.62 seconds, which comes at an acceptable memory consumption. These results illustrate the great potential of SegNet for enhancing LA segmentation accuracy from cardiac MRI and give a baseline for further improvement and application in clinics.**

***Keywords*— SegNet, Cardiac MRI, Left Atrium Segmentation, Deep Learning, Medical Imaging**

I. INTRODUCTION

Cardiac magnetic resonance imaging is a significant modality for the delineation of cardiac anatomy and pathology and gives highly essential insights into diagnosis and therapy methods related to cardiovascular disorders [1], [2]. In cardiac MRI, correct LA segmentation can be particularly valuable in evaluating atrial size, function, and related heart problems. Traditionally, such was already done using manual segmentation methods; however, these techniques are not only very time-consuming but also prone to variable due to observer competence[3].

Given these challenges, a considerable trend has lately been noted toward automatic segmentation systems. Deep learning-based techniques among them have showed substantial promise primarily because they are capable of managing complicated image data and delivering outcomes that are more consistent [4]. In this study, we will demonstrate a deep learning methodology employing the SegNet model to autonomously separate the left atrium from cardiac MRI data[5]. In this work, the SegNet architecture is particularly efficient for semantic image segmentation and is being harnessed and applied toward higher accuracy in anatomical delineations.

Its encoder-decoder architecture allows it to catch complexities of anatomy for optimal output with correct segmentation masks. The strategy increases the accuracy and efficiency of LA segmentation when compared to standard methods and gives eventual assistance for better diagnostic and treatment planning in cardiology [6], [7].

It shall study the performance of the SegNet model under these circumstances, compare its effectiveness against existing approaches, and explain the potential implications for clinical practice.

II. RELATED WORK

Deep learning has given birth to new issues in automated segmentation of cardiac structures. More classical methods were earlier and connected to standard image processing with statistical models, having limited relevance while dealing with cardiac anatomy complexity [4], [8].

Recent developments underscore the performance coming from deep learning models, most obviously the architecture of U-Net proposed by Ronneberger et al. in 2015 [9]. This structure of the encoder-decoder design, together with skip connections, makes U-Net one of the efficient models for biomedical image segmentation tasks. Its success contributed to the creation of 3D variations for better handling of volumetric data[10], [11].

Deep learning-based approaches have been investigated in the area of left atrial segmentation. Left Atrial Segmentation Challenge presented benchmarks, which determined that statistical models using region-growing approaches can be quite effective[3], [4]. Recent research uses higher order models such as DeepLabV3 and V-Net, which seem to ensure more reliable results in segmentation [12].

Despite these advances, anatomical heterogeneity and generalizability to unexplored datasets and clinical contexts remain substantial hurdles [13]. Proposed by Badrinarayanan et al. 2017, SegNet shares an encoder-decoder structure in which symmetric skip routes enable the better retention of spatial information and consequently increase segmentation accuracy[7]. Cardiac MRI is yet another significant diagnostic technique for cardiovascular diseases. Accurate left atrial segmentation is a critical preparatory step in the estimate of atrial size and function. Traditional manual methods are time-consuming and consequently variability-prone; hence, they

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

need an automated solution [14], [15]. In this presentation, authors employed SegNet to produce the automatic segmentation of the left atrium, with additional preprocessing procedures applied to reach higher performance: scaling, normalization, and lastly binarization. The model is assessed in a thorough approach to better its effectiveness, accuracy, and consistency across clinical situations.

III. METHEDOLOGY

A. Data

In this work, the authors have used a dataset accessible publically under the term Heart MRI Image Dataset for Left Atrial Segmentation. As how it is expressly indicated by its name, the described dataset contains cardiac MRI images and hand segmentations of the left atrium correspondingly. For training and evaluation of the model, the dataset would be separated into two: a training subset and a validation subset to eliminate biases during model performance assessment [3], [16]. Fig 1 shows the sample images form the dataset below.

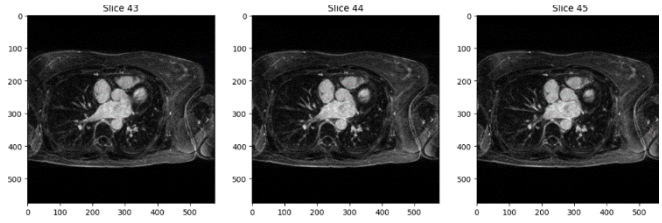

Fig. 1. Sample Images from the Cardiac MRI Dataset - This figure displays a selection of MRI slices from the cardiac dataset.

B. Preprocessing

Several preprocessing procedures are employed to guarantee the consistency and quality of the supplied data[17].
Reading and Resizing: It reads the NIfTI MRI images and resizes them to a fixed spatial resolution of 256 by 256 pixels using interpolation algorithms suitable in each modality that do not lead to loss of information from images.

$$I_{\text{resized}}(x', y') = I\left(\left\lfloor x \cdot \frac{W'}{W} \right\rfloor, \left\lfloor y \cdot \frac{H'}{H} \right\rfloor\right) \qquad (1)$$

The pixel values are normalized to fall within the interval [0, 1], generated by dividing the pixel values by the greatest intensity recorded in the dataset.

$$I_{\text{normalized}}(x, y) = \frac{I(x, y)}{255} \qquad (2)$$

The segmentation label is binarized, which is advantageous for clearly identifying the left atrium from other portions.

$$I_{\text{binary}}(x, y) = \begin{cases} 1 \ \ if \ \ I(x,y) > t \\ 0 \ \ otherwise \end{cases} \qquad (3)$$

Gaussian smoothing is done to minimize noise by convolving the image using a Gaussian kernel.

$$I_{\text{smoothed}}(x, y) = \sum_{u=-k}^{k} \sum_{v=-k}^{v} I(x-u, y-v) \cdot G(u, v) \qquad (4)$$

C. Model Architrecture

The model utilized for semantic segmentation in this work is SegNet. consists of an encoder-decoder system where the encoder collects hierarchical features through convolutional and max-pooling layers, while the decoder reconstructs high-resolution segmentation maps using upsampling and convolutional layers [6], [7]. Skip links between corresponding encoder and decoder layers to retain critical spatial information to assure accuracy of a higher order. The input layer can process images of size 128×128×128×1. The encoder progressively captures features with filters of 64, 128, 256, and 512, while the bottleneck employs 1024 filters with dropout to avoid overfitting. On the reconstruction side, the decoder finally outputs a feature map with filters of 512, 256, 128, and 64. The output layer will generate a binary segmentation mask through the sigmoid activation function [4], [18]. ***Fig 2*** displays the model architecture. This diagram demonstrates movement from the encoder component to the decoder portion, elaborating on hierarchical feature extraction and retaining spatial information through skip links in the SegNet.

D. Metrics

The binary cross-entropy loss function is quite suited for the task and will be used to optimize the model [7], [19]. In essence, it assesses how well the projected probabilities correspond with the actual labels, described below in Equation **5**.

$$\text{Loss} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \qquad (5)$$

The model's performance will be quantified using numerous indicators. Similar to Equation, the Dice coefficient is defined as a measure of the overlap between anticipated and ground-truth segmentation masks, where a value of 1 corresponds to complete agreement [19], [20]. The Jaccard index is another set similarity metric that gives the similarity between the predicted and target sets, as described in Equation **6** and **7**, and greater values are better.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

Layer Type	Description
Input	Input Image
Conv1_1	Convolution + Filters: 64
BatchNorm1_1	Batch Normalization
ReLU1_1	ReLU Activation
Conv1_2	Convolution + Filters: 64
BatchNorm1_2	Batch Normalization
ReLU1_2	ReLU Activation
MaxPooling1	MaxPooling + Indices
Conv2_1	Convolution + Filters: 128
BatchNorm2_1	Batch Normalization
ReLU2_1	ReLU Activation
Conv2_2	Convolution + Filters: 128
BatchNorm2_2	Batch Normalization
ReLU2_2	ReLU Activation
MaxPooling2	MaxPooling + Indices
Conv3_1	Convolution + Filters: 256
BatchNorm3_1	Batch Normalization
ReLU3_1	ReLU Activation
Conv3_2	Convolution + Filters: 256
BatchNorm3_2	Batch Normalization
ReLU3_2	ReLU Activation
Conv3_3	Convolution + Filters: 256
BatchNorm3_3	Batch Normalization
ReLU3_3	ReLU Activation
MaxPooling3	MaxPooling + Indices
Conv4_1	Convolution + Filters: 512
BatchNorm4_1	Batch Normalization
ReLU4_1	ReLU Activation
Conv4_2	Convolution + Filters: 512
BatchNorm4_2	Batch Normalization
ReLU4_2	ReLU Activation
Conv4_3	Convolution + Filters: 512
BatchNorm4_3	Batch Normalization
ReLU4_3	ReLU Activation
MaxPooling4	MaxPooling + Indices
Conv5_1	Convolution + Filters: 512
BatchNorm5_1	Batch Normalization
ReLU5_1	ReLU Activation
Conv5_2	Convolution + Filters: 512
BatchNorm5_2	Batch Normalization
ReLU5_2	ReLU Activation
Conv5_3	Convolution + Filters: 512
BatchNorm5_3	Batch Normalization
ReLU5_3	ReLU Activation
MaxPooling5	MaxPooling + Indices
Upsampling5	Upsampling + Indices
DeConv5_3	Deconvolution + Filters: 512
BatchNorm5_3	Batch Normalization
ReLU5_3	ReLU Activation
DeConv5_2	Deconvolution + Filters: 512
BatchNorm5_2	Batch Normalization
ReLU5_2	ReLU Activation
DeConv5_1	Deconvolution + Filters: 512
BatchNorm5_1	Batch Normalization
ReLU5_1	ReLU Activation
Upsampling4	Upsampling + Indices
DeConv4_3	Deconvolution + Filters: 512
BatchNorm4_3	Batch Normalization
ReLU4_3	ReLU Activation
DeConv4_2	Deconvolution + Filters: 512
BatchNorm4_2	Batch Normalization
ReLU4_2	ReLU Activation
DeConv4_1	Deconvolution + Filters: 512
BatchNorm4_1	Batch Normalization
ReLU4_1	ReLU Activation
Upsampling3	Upsampling + Indices
DeConv3_3	Deconvolution + Filters: 256
BatchNorm3_3	Batch Normalization
ReLU3_3	ReLU Activation
DeConv3_2	Deconvolution + Filters: 256
BatchNorm3_2	Batch Normalization
ReLU3_2	ReLU Activation
DeConv3_1	Deconvolution + Filters: 256
BatchNorm3_1	Batch Normalization
ReLU3_1	ReLU Activation
Upsampling2	Upsampling + Indices
DeConv2_2	Deconvolution + Filters: 128
BatchNorm2_2	Batch Normalization
ReLU2_2	ReLU Activation
DeConv2_1	Deconvolution + Filters: 128
BatchNorm2_1	Batch Normalization
ReLU2_1	ReLU Activation
Upsampling1	Upsampling + Indices
DeConv1_2	Deconvolution + Filters: 64
BatchNorm1_2	Batch Normalization
ReLU1_2	ReLU Activation
DeConv1_1	Deconvolution + Filters: 64
BatchNorm1_1	Batch Normalization
ReLU1_1	ReLU Activation
Output	Pixel-wise Classification

Fig. 2. The Illustration depicts the encoder-decoder system, showcasing the layers involved in feature extraction and reconstruction

$$\text{Dice} = \frac{2|A \cap B|}{|A| + |B|} \quad (6)$$

$$\text{Jaccard} = \frac{|A \cap B|}{|A \cup B|} \quad (7)$$

Accuracy represents the proportion of voxels correctly classified out of the total voxel count. Precision is the proportion that predicts genuine positives among all positive predictions, while recall is the proportion of true positives out of the total actual positives. The F1 score is a composite measure of precision and recall, integrating all dimensions of performance into one value.

E. Training and Evaluation

During training, the model's performance will be overseen using validation loss and accuracy measures. Optimization is done using the Adam optimizer with an initial learning rate of $\mathbf{1\times10^{-4}}$. The number of epochs for training will be a total of **50**; early endings avoid overfitting. This evaluation is done on a separate validation set to see how well it generalizes to unseen data. The preceding measures fully quantify the quality of segmentation.

IV. RESULT AND DISCUSION

The SegNet model for the segmentation of the left atrium from cardiac MRI images was tested using a number of performance criteria on the validation set. It gave a validation loss of **0.0104**, so exhibiting very minimal disagreement between the predicted and real segmentation masks, thereby keeping error in predictions very efficiently at bay. Validation accuracy reached as high as **99.56%**, ergo most of the voxels were correctly identified. It reported a Dice Coefficient of **0.378**, showing reasonable agreement but also room for possible increases in precision with the amount of overlap between the anticipated segmentation mask and the ground truth. The Jaccard Index, which expresses how similar the anticipated and real segmentations are, was **0.236**. This will be the segmentation accuracy for this model. The precision was found to be **0.739**, implying that a substantial proportion of voxels accurately recognized as belonging to the left atrium are classified as such. However, the recall was quite low at **0.039**, which indicates it missed a lot of the real positive voxels, and thus there could be some space for improvement to capture all relevant areas. Next, the balanced value between accuracy and recall is the F1 score, also known as the harmonic mean of precision and recall, again showing an overall balanced performance of this model. **Fig 3** show the model's accuracy and loss plots during training, providing a visual picture of the model's performance throughout epochs. ***Table 1*** shows the detailed results of the validation metrics used to evaluate the performance of the cardiac image segmentation model.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

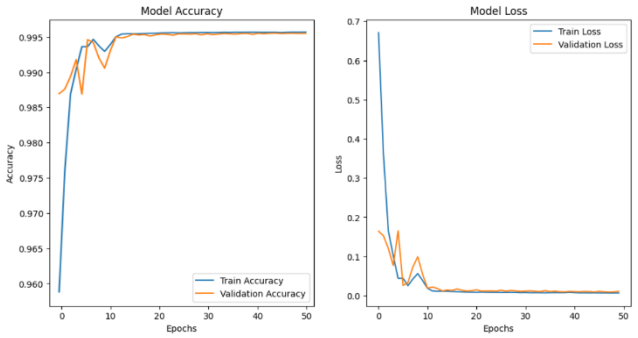

Fig. 3. This graph depicts the accuracy and loss curves for both training and validation sets, highlighting the model's performance and convergence over epochs.

It has to train a SegNet model for **50** epochs using an Adam optimizer while having a learning rate of $\mathbf{1\times10^{-4}}$. This equals approximately **2720.62** seconds to prove good convergence and learning in the available computational resources. Early stopping criterion utilized to avoid overfitting of the model and increase generalization on Validation data.

TABLE I. STYLESVALIDATION METRICS FOR CARDIAC IMAGE SEGMENTATION MODEL

Validation Metric	Value
Loss	0.0103788310661911
Accuracy	0.9956213235855103
Dice Coefficient	0.3783719539642334
Jaccard Index	0.2356243431568145
Precision	0.7390219569206238
Recall	0.0392952859401702
F1 Score	0.0746227213864206

A qualitative study of the segmentation data indicates that, in most cases, the SegNet model does highlight the left atrium, but not without some misgivings. Specifically, less accurate portions of the segmentation mask often correspond to locations with complicated anatomical structures or varying intensity levels in MRI images. ***Fig 4*** shows both the image labels and the predicted labels, displaying the model's performance and indicating the places where improvements are needed.

Results suggest that the SegNet model, having extremely good performance for accuracy and precision, does demonstrate clearly regions where improvement is necessary at recall and in the total Dice Coefficient. A lower recall might be interpreted as indicating that the model may omit a section of the left atrium; this may bring about clinical utility regarding the segmentation in comprehensive cardiac assessment. Future work will be dedicated toward further improvement of the model for improved recall and Dice Coefficient through more advanced techniques, including data augmentation, hyperparameter tweaking, and other hybrid models for better performance.

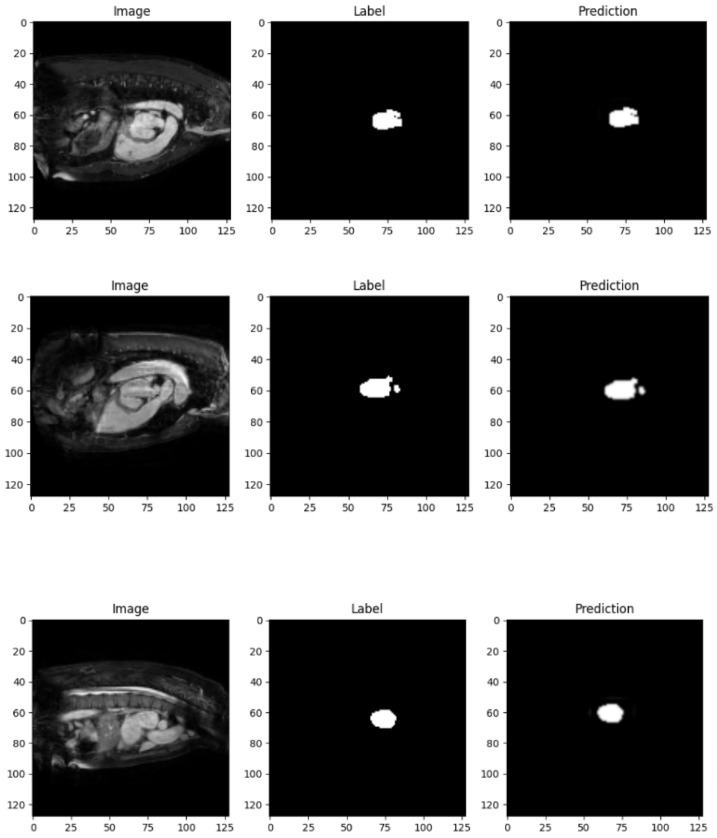

Fig. 4. This figure displays three sample images from the dataset, with their corresponding ground truth labels on the left and the predicted labels on the right, highlighting the model's segmentation performance.

V. CONCLUSION

In this research, an enabler of deep learning in the form of a SegNet model is proposed for autonomous left atrial segmentation in cardiac MRI. This yielded an accuracy of **99.56%** in validation, which proved to be highly useful in evaluating medical photos. However, improvements are needed in recall and segmentation completeness.

Future studies will focus on strengthening model generalizability and exploring more advanced methodologies, including attention mechanisms and multi-task learning.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

REFERENCES

[1] J. Selvanayagam, G. Nucifora, and S. Neubauer, "Cardiovascular Magnetic Resonance," Cardiac CT, PET and MR, pp. 38–90, Jan. 2019, doi: 10.1002/9781118754467.CH2.

[2] N. L. Muller, "Computed tomography and magnetic resonance imaging: past, present and future," European Respiratory Journal, vol. 19, no. Supplement 35, pp. 3S – 12s, Feb. 2002, doi: 10.1183/09031936.02.00248202.

[3] "catactg/lasc: Left Atrial Segmentation Challenge 2013." Accessed: Feb. 20, 2024. [Online]. Available: https://github.com/catactg/lasc

[4] C. Chen et al., "Deep Learning for Cardiac Image Segmentation: A Review," Front Cardiovasc Med, vol. 7, Mar. 2020, doi: 10.3389/fcvm.2020.00025.

[5] C. Liguori, S. Tamburrini, G. Ferrandino, S. Leboffe, N. Rosano, and I. Marano, "Role of CT and MRI in Cardiac Emergencies," Tomography, vol. 8, no. 3, pp. 1386–1400, May 2022, doi: 10.3390/tomography8030112.

[6] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," 2018. Accessed: Mar. 22, 2024. [Online]. Available: https://github.com/tensorflow/models/tree/master/

[7] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," Nov. 2015, [Online]. Available: http://arxiv.org/abs/1511.00561

[8] Y. Xie, B. Yang, Q. Guan, J. Zhang, Q. Wu, and Y. Xia, "Attention Mechanisms in Medical Image Segmentation: A Survey," May 2023, [Online]. Available: http://arxiv.org/abs/2305.17937

[9] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[10] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," 2016, pp. 424–432. doi: 10.1007/978-3-319-46723-8_49.

[11] H. Zheng et al., "A New Ensemble Learning Framework for 3D Biomedical Image Segmentation." [Online]. Available: www.aaai.org

[12] F. Milletari, N. Navab, and S. A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," Proceedings - 2016 4th International Conference on 3D Vision, 3DV 2016, pp. 565–571, Jun. 2016, doi: 10.1109/3DV.2016.79.

[13] T. P. Szczykutowicz, G. V. Toia, A. Dhanantwari, and B. Nett, "A Review of Deep Learning CT Reconstruction: Concepts, Limitations, and Promise in Clinical Practice," Curr Radiol Rep, vol. 10, no. 9, pp. 101–115, Sep. 2022, doi: 10.1007/s40134-022-00399-5.

[14] W. Bai et al., "Automated cardiovascular magnetic resonance image analysis with fully convolutional networks," Journal of Cardiovascular Magnetic Resonance, vol. 20, no. 1, p. 65, Dec. 2018, doi: 10.1186/s12968-018-0471-x.

[15] J. Sander, B. D. de Vos, and I. Išgum, "Automatic segmentation with detection of local segmentation failures in cardiac MRI," Sci Rep, vol. 10, no. 1, p. 21769, Dec. 2020, doi: 10.1038/s41598-020-77733-4.

[16] L. Li, V. A. Zimmer, J. A. Schnabel, and X. Zhuang, "AtrialGeneral: Domain Generalization for Left Atrial Segmentation of Multi-center LGE MRIs," 2021, pp. 557–566. doi: 10.1007/978-3-030-87231-1_54.

[17] C. M. Scannell et al., "Deep-Learning-Based Preprocessing for Quantitative Myocardial Perfusion MRI," Journal of Magnetic Resonance Imaging, vol. 51, no. 6, pp. 1689–1696, Jun. 2020, doi: 10.1002/jmri.26983.

[18] R. M. Wehbe et al., "Deep Learning for Cardiovascular Imaging," JAMA Cardiol, vol. 8, no. 11, p. 1089, Nov. 2023, doi: 10.1001/jamacardio.2023.3142.

[19] J. Alroy, "A new twist on a very old binary similarity coefficient," Ecology, vol. 96, no. 2, pp. 575–586, Feb. 2015, doi: 10.1890/14-0471.1.

[20] D. Brunet, E. R. Vrscay, and Zhou Wang, "On the Mathematical Properties of the Structural Similarity Index," IEEE Transactions on Image Processing, vol. 21, no. 4, pp. 1488–1499, Apr. 2012, doi: 10.1109/TIP.2011.2173206.