Attention Semi-Siamese U-Net for quantification of biomarkers in tauopathies

Luis A. Campero-Garcia
Tecnologico de Monterrey
Monterrey, 64849, Mexico
a01019685@tec.mx

Jose A. Cantoral-Ceballos
Tecnologico de Monterrey
Monterrey, 64849, Mexico
joseantonio.cantoral@tec.mx

Miguel A. Ontiveros-Torres
Tecnologico de Monterrey
Monterrey, 64849, Mexico
miguelontiveros@tec.mx

Andres E. Gutierrez-Rodriguez
MAHLE Shared Services
Monterrey, 64650, Mexico
andres.eduardo.gutierrez@mahle.com

Abstract

Efforts to diagnose and predict neurodegenerative diseases using imaging techniques have increased, particularly for tauopathies, which contribute to rising morbidity and mortality. Current approaches, however, do not focus on the molecular mechanisms causing tau protein's abnormal behavior, leading to neurofibrillary tangles in hippocampal and cortical regions. This work aims to develop a novel quantification protocol for tauopathy biomarkers based on pathological tau modifications. Utilizing a U-Net-based neural network with attention modules, we analyze segmented masks to quantify combined fluorescent signals from molecular changes in neurofibrillary tangles. By studying these tangles individually, we can detect unconventional interactions between biomarkers, providing valuable information for understanding dementia pathogenesis and informing future computational analysis approaches.

1. Introduction

Alois Alzheimer first described the formation of neurofibrillary tangles (NFTs), tau protein deposits in neurons, and cholinergic neuron destruction in Alzheimer's disease in 1906 [5]. The term "tauopathies" was later coined in the 1980s by Bernardino Ghetti and Michel Goedert, referring to neurodegenerative disorders characterized by tau polypeptide deposits in neurons, glial cells, and extracellular space [4]. In tauopathies, tau polypeptides undergo abnormal post-translational modifications (PTMs), such as hyperphosphorylation, proteolysis, and conformational changes. These alterations cause tau proteins to dissociate from microtubules and polymerize into insoluble aggregates called paired helical filaments (PHFs), which form NFTs [1]. NFTs, primarily composed of tau, are hallmarks of numerous diseases, prompting efforts to study their characteristics for insights into tauopathy pathogenesis. As tau deposits form 10-15 years before clinical symptoms appear, advanced visualization methodologies are continuously developed and refined for potential diagnostic and monitoring applications [10].

We propose a novel protocol for quantitative tauopathy analysis in previously unused post-mortem brain tissue immunofluorescence images, obtained through a collaboration between the National Dementia Biobank and Tecnologico de Monterrey's School of Engineering and Sciences. The protocol incorporates a Semi-siamese U-Net-based neural network model [8] with added attention modules for deep learning (DL) image segmentation, enabling quantification of tau protein expression in specific immunofluorescence image regions.

2. Related work

The U-Net's success in biomedical image segmentation is demonstrated by the numerous models based on it, which is further supported by the modularity and flexibility of neural networks and their training methods [15]. In this section we present some of the models in the state of the art.

2.1. Semi-siamese U-Net

The Semi-siamese U-Net [8] was designed for processing lung electrical tomographies, specifically for separating the central chest region containing the heart from the surrounding lung area. This architecture features two distinct expanding paths following the contracting path, with one path segmenting the lung region and the other targeting the heart region. Unlike typical segmentation models, the Semi-siamese U-Net outputs a conductivity distribution map for lung electrical tomographies instead of conventional segmentation maps.

2.2. MultiResUnet

The MultiResUNet builds upon the original U-Net model, replacing skip connections with custom Res Paths [6]. The Res Path addition addresses medical image scale variation, inspired by the inception-like architecture explored in [17]. Additionally, the authors identify a semantic gap between the contracting and expanding paths' same levels. They argue that directly joining lower-level features from the contracting path with expanding path features could negatively impact the network's performance. Instead, they propose adding convolutional layers at each level rather than concatenating directly via skip connections. The encoder's feature map is fed to the Res Path before the pooling layer, and the Res Path's outputs are concatenated to the upsampled feature map from the level below. This addition enables the MultiResUNet to outperform the traditional U-Net in five test datasets, with four having over a 1% IOU metric improvement and one exceeding a 10% increase. Moreover, the model demonstrates better border delineation, enhanced outlier reliability, and superior performance in fewer epochs [6].

2.3. RCU-Net

Yu et al. build upon the MultiResUNet by adding a custom block to the architecture for tumor segmentation in breast ultrasound images. This new architecture, called RCU-Net, incorporates a custom dense block at the U-Net's top level while also replacing direct skip connections with Res Path blocks [19]. The dense block connects the network input to the last upsampled features before the final convolution set for segmentation. Additionally, the RCU-Net's Res Paths are slightly different from the MultiResUNet's, as they consist of fewer convolutions at each decreasing network level. The dense block addition enables the RCU-Net to outperform both the traditional U-Net and the U-Net with Res Paths in tumor segmentation.

3. The attention semi-siamese U-Net

Our proposed architecture is based on the U-Net, popular for biomedical image segmentation due to its performance with few labeled images [14]. The U-Net comprises an encoder path and a decoder path, with skip connections transferring feature maps between them. This design suits our study, where effective segregation of affected regions with limited training samples is the goal. The Semi-siamese U-Net is particularly relevant, as it separates decoders for specialized tasks. Building on this, our study employs the Semi-siamese U-Net to simultaneously segment four immunofluorescence channel combinations in dementia images, with each combination managed by an individual decoder path. Additionally, we incorporate attention modules in every upsampling step across all decoders to enhance performance, we show improved performance from previous efforts [2]. The proposed model is shown in Fig. 1.

Our network architecture features one encoder path and four decoder paths for multi-task image segmentation. All segmentation tasks share parameters in the contracting path. The model takes a three-channel 256×256 pixel image as input, outputting four one-channel 256×256 pixel masks. Each encoder level has two convolutional layers, with filter counts of 16, 32, 64, 128, and 256 at each successive level. Filters have a size of 5×5, with "same" padding and a stride of 1. A 0.2-rate dropout layer follows each max-pooling to improve generalization. The decoder comprises upsampling steps with attention blocks. Upsampling is performed in two stages. First, the previous level's last feature map is fed into the attention block as gating signal g, with the skip connection from that level input as x^l. Then, x^l goes through a 2-stride $1 \times 1 \times 1$ convolutional layer, matching the gating signal's dimensions. Meanwhile, the gating signal enters its corresponding $1 \times 1 \times 1$ convolutional layer. The second upsampling stage follows the attention block output. The lower-level feature map is upsampled using nearest-type interpolation, and the attention block output is concatenated. Two convolutional layers are applied, with dropout layers (0.2 rate) following. Convolutional layers use ReLU activation, except for the final output layer, which uses sigmoid activation for segmentation with a 1×1 convolution. The attention blocks improve segmentation accuracy for structures of varying sizes [9], crucial for capturing immunofluorescence images with different optical zoom levels.

4. Experimental Framework and Results

In order to compare the performance of our model against four state of the art models, we evaluated the experimental results using four metrics: true positive (TP), false positive (FP), the Dice coefficient (DC) [3], and the intersection over union (IOU, also known as the Jaccard index) [7]. These metrics are defined based on the ground truth (GT) and the segmentation results (SR) obtained from the forward pass of the evaluated method. We used 5-fold cross-validation for our experiments. The 97 images were randomly divided into 5 groups. In each group, 78 images (80%) were used for training, while the remaining 19 images (20%) were used for validation. The training set in each fold underwent the same transformations to further augment the data. First, each image was randomly rotated by 90, 180, or 270 degrees. Next, random elastic deformations [16] were applied to both the original and rotated images, resulting in a total of 312 training images after augmentation. To enhance training and inference times, the images and masks were resized to 256×256 pixels using the OpenCV Python library.

For each model, five sets of numerical weights were gen-

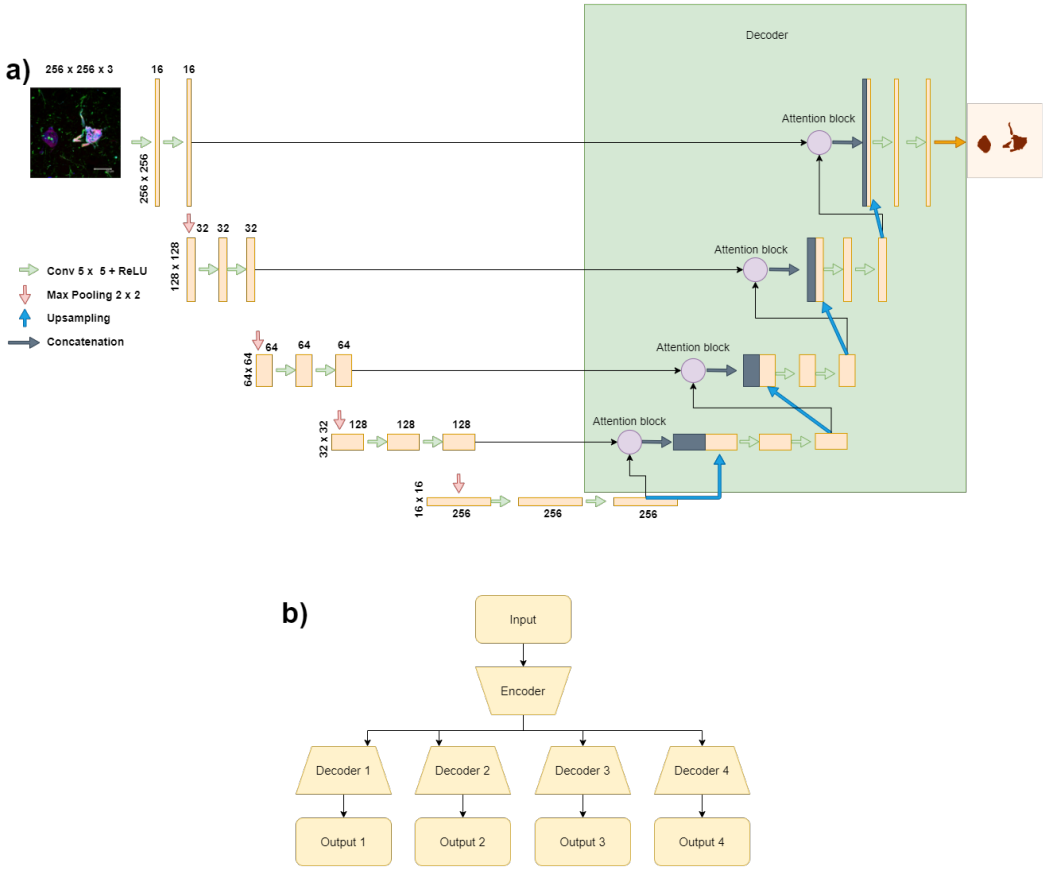

Figure 1. **Proposed model of the extended semi-Siamese U-Net neural network.** (**a**) The proposed model showing only one decoder. Each box corresponds to a multi-channel feature map. The number of channels is denoted at the top of each box; the height and width of each feature map are provided at the edge of each box and are the same in the encoder and the decoder. The encoder path on the left consists of five levels (including the bottleneck at the bottom), each one consisting of two convolutional layers. The outputs of the encoder are fed into attention blocks and its outputs are concatenated with the upsampled feature maps in the decoder before being passed through two additional convolutional layers. Each of the four decoding paths gets its own copy of the encoder feature map corresponding to its level. Attention blocks help the model to detect features of smaller size in the images. (**b**) The final model. Four parallel decoders are in charge of segmenting different biomarker interactions.

erated. The testing images were input into each model, and for each one, four segmentation masks were obtained. The output images were used to calculate the metrics for the four outputs. These measurements were then averaged to obtain the final results per model using each set of weights. Tab. 1 summarizes the mean and standard deviation of the four metrics, along with the four outputs of the models.

Figure 2 displays the boxplots representing the performance of each model across the four outputs. The proposed model demonstrates a better average performance in all metrics. Compared to the traditional U-Net, the DC, FP, IOU, and TP of the proposed model improved by 3.33%, 3.86%, 5.47%, and 3%, respectively. Additionally, the proposed model appears to be slightly more stable (i.e., has a lower standard deviation) than the models incorporating a residual path, which is often utilized to enhance stability [15]. In Figure 3, we randomly selected a representative AD image from the test set, with different regions highlighted across each color channel combination, to demonstrate the segmentation capabilities of each model on the desired regions. Although the performance metrics across models were similar, Figure 3 shows that the proposed model performs better in including all the desired areas across all four outputs in the segmentation while avoiding the undesired ones, as specified by the ground truth labels.

5. Discussion

The detailed interpretation of immunofluorescence image quantification is beyond the scope of this research. However, our result show higher association of Red-Blue channels (thiazine red and pS396 [13]). This indicates

Table 1. The mean and the standard deviation of the four metrics across the 4 outputs and the 5-fold cross-validation. The numbers after the $\pm$ represent the standard deviation.

Model	DC	FP	IOU	TP
Attention Semi-Siamese U-Net	**0.9227 $\pm$ 0.0206**	**0.0363 $\pm$ 0.0222**	**0.8572 $\pm$ 0.0349**	**0.8918 $\pm$ 0.0283**
Semi-Siamese U-Net	0.9064 $\pm$ 0.0223	0.0437 $\pm$ 0.0201	0.8268 $\pm$ 0.0402	0.8690 $\pm$ 0.0424
Inception U-Net	0.8938 $\pm$ 0.028	0.0638 $\pm$ 0.0252	0.8092 $\pm$ 0.0459	0.8593 $\pm$ 0.0408
RCU-Net	0.8984 $\pm$ 0.0233	0.0519 $\pm$ 0.0202	0.8164 $\pm$ 0.0385	0.8536 $\pm$ 0.0424
Res U-Net	0.893 $\pm$ 0.0335	0.0699 $\pm$ 0.0561	0.8083 $\pm$ 0.0535	0.8589 $\pm$ 0.0439
U-Net	0.8894 $\pm$ 0.0333	0.0749 $\pm$ 0.0443	0.8025 $\pm$ 0.0539	0.8618 $\pm$ 0.0396

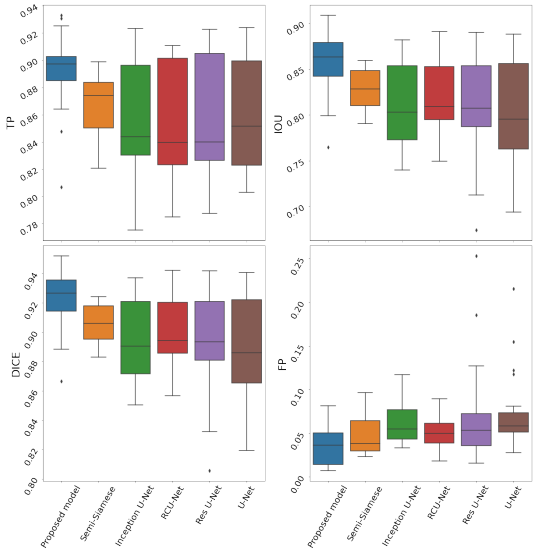

Figure 2. **Performance of the models in the four metrics**. TP, IOU, DC, and FP averaged across the four outputs using 5-fold cross-validation. The x-axis represents the model name and the y-axis the values of the metric. It can be seen that the proposed model has a better average performance in all metrics except the TP.

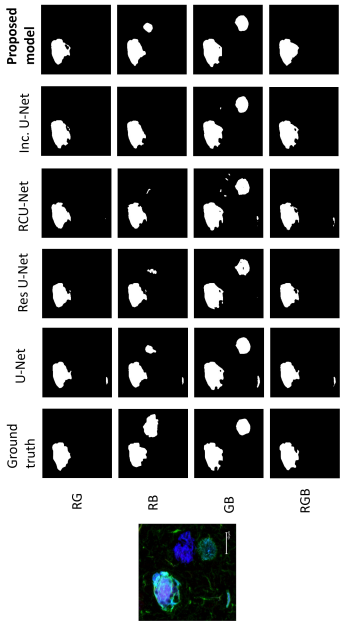

Figure 3. **Experimental segmentation results obtained with the different neural network models**. Each column represents the outputs of the corresponding model and each row one combination of channels. R: red; G: green; B: blue.

that phosphorylation at serine 396 matures into the polymeric fibrillar form with a value of 6.54%, signifying a more advanced stage in fibrillar filament formation than the biomarker for dual phosphorylation at serine 202 and threonine 305 (AT8 antibody), which has a lower value of 3.96% [11]. These results align with experimental studies demonstrating that tau polypeptide is highly phosphorylated at serine 396 in advanced polymeric stages [12, 18]. This emphasizes the significance of the quantification protocol for understanding neurofibrillary tau polypeptide changes in various tauopathies and brain regions. Quantification would be based on patients' unique fibrillar lesion structures and include analysis of other biomarkers controlled by fluorescence. Our proposed method can localize different types of PTMs in NFT bodies and distinguish them from other pathological events in the quadrant with an IOU of 0.8572. Prior to this, no automatic method existed for obtaining quantitative information on fluorescence images that focused on specific biological events and biomarkers based on unique fibrillar lesion structures. Manual segmentation of four signals in a three-channel image typically takes 15 to 30 minutes, depending on the complexity of the biological disease events in the image. In contrast, our method takes just a few seconds to perform segmentation on a standard CPU. Furthermore, the segmentation obtained using our deep learning model is not only faster and simpler but

also more reliable, as it is less susceptible to variability due to differences in human raters' criteria, subjectivity, or image quality. The improved performance of our method is mainly attributed to each upsampling path focusing on a single task, rather than sharing weights during the upsampling with other tasks, as seen in U-Net, Inception U-Net, Res U-Net, and RCU-Net.

References

[1] Diana M. Acosta, Chiara Mancinelli, Clay Bracken, and David Eliezer. Post-translational modifications within tau paired helical filament nucleating motifs perturb microtubule interactions and oligomer formation. *Journal of Biological Chemistry*, 298(1), Jan. 2022. Publisher: Elsevier. 1

[2] Luis A. Campero-Garcia, Jose A. Cantoral-Ceballos, Alejandra Martinez-Maldonado, Jose Luna-Muñoz, Miguel A. Ontiveros-Torres, and Andres E. Gutierrez-Rodriguez. A novel automatic quantification protocol for biomarkers of tauopathies in the hippocampus and entorhinal cortex of post-mortem samples using an extended semi-siamese u-net. *Biology*, 11(8), 2022. 2

[3] Lee R. Dice. Measures of the Amount of Ecologic Association Between Species. *Ecology*, 26(3):297–302, 1945. Publisher: Ecological Society of America. 2

[4] Jürgen Götz, Glenda Halliday, and Rebecca M Nisbet. Molecular pathogenesis of the tauopathies. *Annual Review of Pathology: Mechanisms of Disease*, 14:239–261, 2019. 1

[5] Hanns Hippius and Gabriele Neundörfer. The discovery of alzheimer's disease. *Dialogues in clinical neuroscience*, 2022. 1

[6] Nabil Ibtehaz and M Sohel Rahman. Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation. *Neural networks*, 121:74–87, 2020. 2

[7] Paul Jaccard. The Distribution of the Flora in the Alpine Zone. *The New Phytologist*, 11(2):37–50, 1912. Publisher: [Wiley, New Phytologist Trust]. 2

[8] Yen-Fen Ko and Kuo-Sheng Cheng. Semi-Siamese U-Net for separation of lung and heart bioimpedance images: A simulation study of thorax EIT. *PLOS ONE*, 16(2), Feb. 2021. Publisher: Public Library of Science. 1

[9] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, et al. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018. 2

[10] Rocco Pizzarelli, Natalia Pediconi, and Silvia Di Angelantonio. Molecular imaging of tau protein: New insights and future directions. *Frontiers in Molecular Neuroscience*, 13:586169, 2020. 1

[11] Marta Querol-Vilaseca, Sònia Sirisi, Laura Molina-Porcel, Beatriu Molina, Jordi Pegueroles, Paula Ferrer-Raventós, Raúl Nuñez-Llaves, Oriol Dols-Icardo, Mircea Balasa, Maria Florencia Iulita, et al. Neuropathology of a patient with alzheimer disease treated with low doses of verubecestat. *Neuropathology and Applied Neurobiology*, 48(3):e12781, 2022. 4

[12] Lata Rani and Sairam S Mallajosyula. Phosphorylation-induced structural reorganization in tau-paired helical filaments. *ACS Chemical Neuroscience*, 12(9):1621–1631, 2021. 4

[13] Mamen Regalado-Reyes, Diana Furcila, Félix Hernández, Jesús Ávila, Javier DeFelipe, and Gonzalo León-Espinosa. Phospho-tau changes in the human ca1 during alzheimer's disease progression. *Journal of Alzheimer's Disease*, 69(1):277–288, 2019. 3

[14] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv:1505.04597 [cs]*, May 2015. arXiv: 1505.04597. 2

[15] Nahian Siddique, Sidike Paheding, Colin P. Elkin, and Vijay Devabhaktuni. U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. *IEEE Access*, 9:82031–82057, 2021. 1, 3

[16] Patrice Y Simard, David Steinkraus, John C Platt, et al. Best practices for convolutional neural networks applied to visual document analysis. In *Icdar*, volume 3, 2003. 2

[17] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016. 2

[18] Angie K Torres, Bastián I Rivera, Catalina M Polanco, Claudia Jara, Cheril Tapia-Rojas, et al. Phosphorylated tau as a toxic agent in synaptic mitochondria: implications in aging and alzheimer's disease. *Neural regeneration research*, 17(8):1645, 2022. 4

[19] Kailuo Yu, Sheng Chen, and Yanghuai Chen. Tumor segmentation in breast ultrasound image by means of res path combined with dense connection neural network. *Diagnostics*, 11(9):1565, 2021. 2