Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics e-ISSN: 2656-8624
Homepage: https://ijeeemi.org/; Vol. 7, No. 4, pp. 701-712, November 2025

RESEARCH PAPER **OPEN ACCES**

Deep Learning-Based Hippocampal Segmentation and MTA Classification Using U-Net with ResNet-50 Backbone

Aldienannisa Devin Salsabila[1], Fatimah[2], Darmini[2], Selamet Budi Kurniawan[3]

[1] Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Semarang, Indonesia
[2] Department of Radiodiagnostic and Radiotherapy, Poltekkes Kemenkes Semarang, Semarang, Indonesia
[3] Department of Radiology, RS Pusat Otak Nasional Prof.Dr.dr. Mahar Mardjono, Jakarta, Indonesia

Abstract

Medial Temporal Atrophy (MTA) is a key biomarker in the early diagnosis of dementia. However, its assessment through manual inspection of MRI scans is subjective, time-consuming, and prone to inter-observer variability. This creates the need for automated systems that can provide accurate, consistent, and clinically interpretable evaluations. This study aims to develop a hybrid deep learning framework that integrates U-Net with a ResNet-50 backbone for simultaneous hippocampal segmentation and MTA grading, thereby reducing diagnostic subjectivity and bridging the gap between image processing and clinical interpretation. The main contribution of this work lies in the dual functionality of the proposed architecture: not only producing precise segmentation masks of the hippocampal region but also classifying the degree of atrophy into MTA scores (0–4), which previous studies on hippocampal segmentation have not addressed. The proposed method employs a U-Net for pixel-level segmentation, enhanced with a ResNet-50 backbone to stabilize gradient propagation and enrich feature representation during encoding. Results demonstrated excellent performance, achieving a training accuracy of 99.9% with strong convergence between training and validation curves. On a test set of 32 coronal MRI slices, the model correctly classified 26 samples, misclassifying only 6. Overall, the proposed U-Net with ResNet-50 backbone provides an accurate and reliable end-to-end solution for hippocampal segmentation and MTA grading. Its clinical performance demonstrates parity with expert radiologists, underscoring its potential as a decision-support tool in dementia diagnosis. Future work will focus on extending this framework to 3D U-Net architectures, enabling the integration of volumetric MRI features to enhance robustness and generalizability across diverse datasets further.

Paper History
Received July 02, 2025
Revised October 10, 2025
Accepted October 20, 2025
Published November 30, 2025

Keywords
Convolutional Neural Networks;
Deep Learning;
Dementia Diagnosis;
Hippocampal Segmentation;
Resnet-50
U-net;

Author Email
Aldienannisadevins@poltekkes-smg.ac.id
fatimah_yunaeza@yahoo.com
da12mini@gmail.com
selametk29@gmail.com

I. Introduction

Dementia is a progressive neurodegenerative disorder characterized by a decline in cognitive function, memory, and behavior, and it has become one of the leading causes of reduced quality of life in the elderly [1], [2]. One of the most important biomarkers in the diagnosis of dementia is the hippocampus, a brain structure that plays a central role in memory formation and cognitive processing [3], [4]. Hippocampal atrophy, or the shrinkage of hippocampal volume, has been shown to correlate significantly with the severity of dementia, particularly in Alzheimer's disease [5]. To assess the degree of hippocampal atrophy, clinicians rely on the Medial Temporal Atrophy (MTA) score, which has been established as a standard method using Magnetic Resonance Imaging (MRI) [6], [7], [8]. The MTA score ranges from 0 to 4, where higher values indicate more severe levels of atrophy [9].

The increasing global prevalence of dementia, projected to reach 152.8 million cases by 2050, underscores the urgent need for more accurate and efficient diagnostic methods [10]. Hippocampal atrophy assessment using the MTA score has been recognized as a primary diagnostic indicator; however, manual evaluation by radiologists still presents challenges due to visual subjectivity, particularly in distinguishing between scores 2 and 3, which often overlap [11]. This limitation can result in inconsistencies in interpretation and subsequently affect clinical diagnostic accuracy [12], [13]. Meanwhile, existing vendor-based brain segmentation software is offered as an optional feature with additional costs and still demonstrates suboptimal performance, with an accuracy of only 76.7% and a sensitivity of 50% [14], [15]. These limitations highlight the pressing need to develop an automated segmentation system based on artificial intelligence, which can provide objective and consistent evaluation of hippocampal atrophy to support dementia diagnosis.

Medial temporal atrophy (MTA) grading is a well-established biomarker for dementia diagnosis, particularly in Alzheimer's disease, as it reflects hippocampal and medial temporal lobe degeneration that correlates with cognitive decline [11]. In clinical practice,

Corresponding author: Aldienannisa Devin Salsabila, aldienannisadevins@poltekkes-smg.ac.id, Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Jl.Tirto Agung, Pedalangan, Banyumanik, Semarang City, Central Java. **DOI:** https://doi.org/10.35882/ijeeemi.v7i4.263

however, visual MTA grading performed by radiologists is often subjective, time-consuming, and prone to inter- and intra-rater variability, especially when atrophy is subtle or MRI protocols vary across scanners [6], [11]. These challenges limit the reproducibility and reliability of diagnostic assessment in routine workflows. Automated methods that can simultaneously generate precise hippocampal segmentations and objective MTA grades aligned with radiological criteria provide a valuable solution [8]. Such approaches not only improve consistency and reduce diagnostic variability but also accelerate reporting and support standardized, reproducible evaluation across institutions, ultimately enhancing their applicability in real-world diagnostic processes.

Several studies have been conducted to enhance the accuracy and reliability of hippocampal segmentation as a critical step in supporting dementia diagnosis. Various deep learning-based frameworks have been introduced to overcome the limitations of manual assessment and conventional image processing methods. Such as, study by [16] where the researcher explores a deep learning-based approach for automatic hippocampus segmentation using convolutional neural networks (CNN), the authors propose a method that leverages enhanced preprocessing and training strategies to improve segmentation accuracy on MRI datasets. Their experiments demonstrate that CNN-based models are able to capture spatial features of hippocampal structures more effectively than traditional manual or semi-automatic methods. The study emphasizes the advantages of deep models in reducing subjectivity and enhancing reproducibility. However, the method still suffers from limited generalization when applied to heterogeneous MRI data collected from different scanners or populations, which restricts its clinical scalability.

Study by [17] investigates the integration of attention mechanisms into medical image segmentation frameworks. Experimental results show that this approach improves the Dice coefficient and overall segmentation performance compared to baseline U-Net architectures. The study highlights that attention-guided learning not only improves segmentation precision but also contributes to model interpretability. Nevertheless, the reliance on attention mechanisms alone cannot fully address segmentation challenges in cases of severe atrophy or poor MRI quality, which limits its robustness in real-world clinical practice.

Study by [18] proposes an improved 3D-UNet model combined with a novel filling technique for hippocampal segmentation from MR images. The dataset consists of 200 three-dimensional T1-weighted MRI scans, which were divided into training, validation, and testing sets. The proposed model achieved a Dice score of up to 0.8674 and a mean IoU of 0.7668, outperforming baseline architectures such as VNet, SegResNet, and UNetR. While the results are promising, the study is limited by its single-center dataset. It does not explore model adaptability to multi-modal MRI or larger, more diverse clinical datasets, leaving questions about its generalizability unanswered.

Despite the remarkable advances in deep learning-based hippocampal segmentation, several limitations persist in prior studies. CNN-based methods demonstrated strong feature extraction capabilities, yet their generalization remains limited across heterogeneous MRI datasets from different populations and scanners [16]. Attention-guided frameworks improved segmentation precision and interpretability, but their robustness diminishes when dealing with severe hippocampal atrophy or low-quality MRI images [17]. Similarly, advanced 3D-UNet models with novel post-processing techniques achieved competitive accuracy; however, their dependence on large, single-center datasets constrains scalability and clinical applicability [18].

To overcome these challenges, this study develops a hybrid deep learning framework that integrates the U-Net architecture as a hippocampal segmenter with ResNet50 as a feature extractor and classifier. This combination was selected because U-Net, with its encoder–decoder structure and skip connections, is highly effective in preserving fine-grained spatial details essential for hippocampal segmentation. At the same time, ResNet-50 provides residual connections that mitigate vanishing gradients, enable deeper feature learning, and improve generalization across heterogeneous MRI data. Together, they offer a balanced trade-off between segmentation accuracy, robustness, and clinical interpretability, making the framework particularly suitable for MTA grading. The model is designed to determine MTA levels ranging from 0 to 4 automatically.

The novelty of this approach lies in three main aspects. First, it introduces an end-to-end pipeline that not only performs precise hippocampal segmentation but also directly translates the segmented regions into clinically meaningful MTA scores, thereby bridging the gap between image processing and diagnostic decision support. Second, instead of relying solely on manual interpretation of hippocampal volume, the proposed model leverages ResNet50 to extract discriminative structural patterns from the segmented regions and map them objectively to MTA levels, reducing inconsistencies often encountered in visual grading. Third, this study specifically addresses the challenge of differentiating progressive levels of atrophy, where smaller hippocampal volumes are automatically assigned to higher MTA grades (e.g., level 4) and larger volumes to lower grades (e.g., level 0–1), thus replicating and enhancing the radiologist's workflow with greater consistency.

II. Materials and Methods

A. Data Collections and Spliting Data

Corresponding author: Aldienannisa Devin Salsabila, aldienannisadevins@poltekkes-smg.ac.id, Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Jl.Tirto Agung, Pedalangan, Banyumanik, Semarang City, Central Java. **DOI:** https://doi.org/10.35882/ijeeemi.v7i4.263

This study employed an applied experimental design using retrospectively collected brain MRI data from the Radiology Department of the National Brain Center

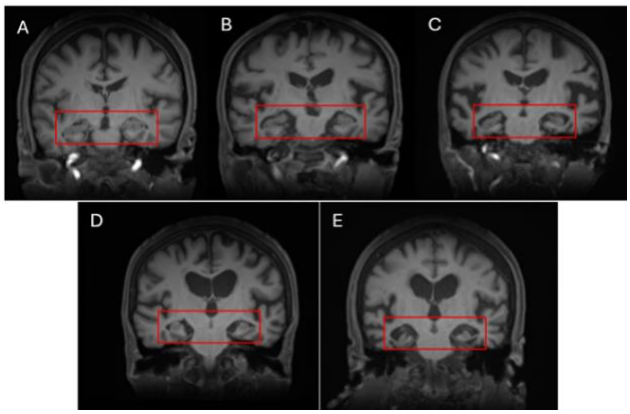


Fig. 1. Sample Dataset: (a) MTA 0, showing no hippocampal atrophy (red box); (b-e) MTA 1–4, showing progressive hippocampal atrophy (red box).

Hospital (Private Dataset). A total of 160 coronal T1-weighted MPRAGE MRI scans in DICOM format were included all of which had been confirmed as Alzheimer's dementia cases by radiologists and manually graded using the MTA scale. The scans were acquired from individual vendors' MRI scanners and included axial, coronal, and sagittal orientations. In this study, we specifically used coronal slices for hippocampal segmentation and MTA grading.

The dataset was divided into two groups: 128 images (80%) for training, 20% for validation, and 32 images for testing, with all MTA levels from 0 to 4 represented across both subsets. A sample of raw data for each MTA level can be seen in Fig. 1. The training set consisted of 17 images with an MTA score of 0, 25 with a score of 1, 35 with a score of 2, 25 with a score of 3, and 26 with a score of 4. Meanwhile, the testing set included 7 images with an MTA score of 0, 7 with a score of 1, 6 with a score of 2, 6 with a score of 3, and 6 with a score of 4. This distribution ensured that the model was exposed to a wide spectrum of hippocampal atrophy severity, ranging from no atrophy to advanced atrophy.

From a demographic perspective, the test set comprised 32 patients with a relatively balanced sex distribution, consisting of 17 males and 15 females. The majority of the patients were elderly, with the largest proportion in the 75–84 years age group, followed by those aged 85–94 years, and the smallest proportion in the 65–74 years group. This composition reflects the clinical reality that dementia predominantly affects older populations, thereby providing a realistic basis for evaluating the proposed framework. For ground truth labelling, MTA grading was performed according to the Scheltens scale (0–4) by a board-certified radiologist with expertise in dementia imaging. Since only one rater was available, inter-rater reliability could not be calculated, which we acknowledge as a limitation.

B. Pre-processing and Data Augmentation

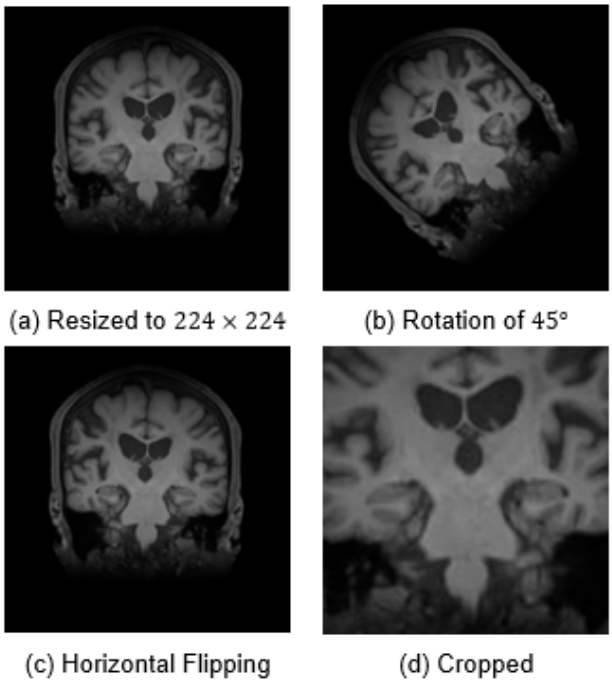


Fig. 2. Sample of augmented data during training.

In the pre-processing stage, all coronal T1-weighted MPRAGE MRI scans initially stored in DICOM format were reviewed using MicroDicom Viewer (32-bit). Since the DICOM format has a large matrix size that considerably slows down the training and inference processes, all images were converted into JPG/PNG format for compatibility with the deep learning pipeline [19], [20]. Following format conversion, image resizing was performed to standardize the input dimensions required by the U-Net architecture [20], [21].

Each MRI slice was resized to a fixed resolution of 224×224 pixels, ensuring uniformity across the dataset. The resizing operation can be calculated using Eq (1).

$$I_{resized}(x', y') = I\left(\frac{x}{W_{in}} \cdot W_{out}, \frac{y}{H_{in}} \cdot H_{out}\right) \quad (1)$$

where $I(x, y)$ is the original image, (W_{in}, H_{in}) represent the input width and height, and $(W_{out}, H_{out}) = 224 \times 224$ denote the target dimensions. In this study, data augmentation techniques were applied. These included flipping, rotation of $45°$, and cropping is restricted to regions containing the hippocampus to ensure clinical relevance [22]. After augmentation, the dataset expanded from 128 training images to a total of 1170 images, with 234 samples in each MTA class (0–4). The augmentation process can be calculated using Eq. (2).

$$D_{aug}(x', y') = \{\mathrm{F(I)}, \mathrm{R(I, \theta)}, \mathrm{C(I)} \mid \mathrm{I} \in D_{train}\} \quad (2)$$

where $D_{train}$is the original training set, parameter F, R, C represent flipping, rotation, and cropping transformations, respectively, and $\theta = 45°$, as seen in Fig. 2. Through this augmentation pipeline, the dataset was substantially enriched, allowing the model to capture variability in hippocampal morphology better while mitigating the risk of overfitting during training.

C. U-Net Based on ResNet-50 Backbone

Corresponding author: Aldienannisa Devin Salsabila, aldienannisadevins@poltekkes-smg.ac.id, Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Jl.Tirto Agung, Pedalangan, Banyumanik, Semarang City, Central Java. **DOI:** https://doi.org/10.35882/ijeeemi.v7i4.263

The proposed framework integrates U-Net with a ResNet-50 backbone, where U-Net is responsible for hippocampal segmentation and ResNet-50 functions as a residual feature extractor and classifier [23], [24]. The encoder path of U-Net [25], implemented with ResNet-50 residual blocks [26], captures multi-scale semantic representations from the input MRI slice, while the decoder path reconstructs the spatial resolution through up-sampling and skip connections. Formally, the convolution operation in the encoder can be calculated using Eq. (3).

$$F^{(l)} = \sigma\left(W^{(l)} \cdot F^{(l-1)} + B^{(l)}\right) \quad (3)$$

where $F^{(l)}$ is the output feature map at layer l, $W^{(l)}$ and $B^{(l)}$are the convolutional kernel and bias, σ is the nonlinear activation function (ReLU). In ResNet-50, residual learning is applied to each block to stabilize gradient flow and preserve information across layers. A residual block can be calculated using Eq. (4).

$$H(x) = F(x, W_i) + x \quad (4)$$

Where x is the block input, $F(x, W_i)$ is the transformation learned by stacked convolutional layers with weights W_i, and the identity mapping $+x$ represents the skip connection. This formulation ensures that the network learns residual features rather than directly fitting the target mapping, thus mitigating vanishing gradient problems in deeper networks. As calculated using Eq. (5), the decoder reconstructs the segmentation mask by up sampling and concatenating encoder features via skip connections.

$$S(l) = Up(F(l)) \oplus F(l_{enc}) \quad (5)$$

Where $S(l)$ is the decoder feature at stage l, $Up(\cdot)$ is the up-sampling operator, $F(l_{enc})$ is the corresponding encoder feature. A final 1×1 convolution generates the segmentation mask M as calculated using Eq. (6).

$$M = \sigma(W_{seg} * S(L) + b_{seg}) \quad (6)$$

Through this formulation, the U-Net with ResNet-50 backbone simultaneously achieves accurate hippocampal segmentation and robust classification, providing an end-to-end pipeline for automated MTA grading. To further illustrate the flow of information within the architecture, the detailed layer configuration of the proposed model is presented in Fig. 3. This diagram explicitly shows how the encoder, adapted from ResNet-50, progressively extracts hierarchical features, while the decoder reconstructs spatial resolution through up-sampling and skip connections.

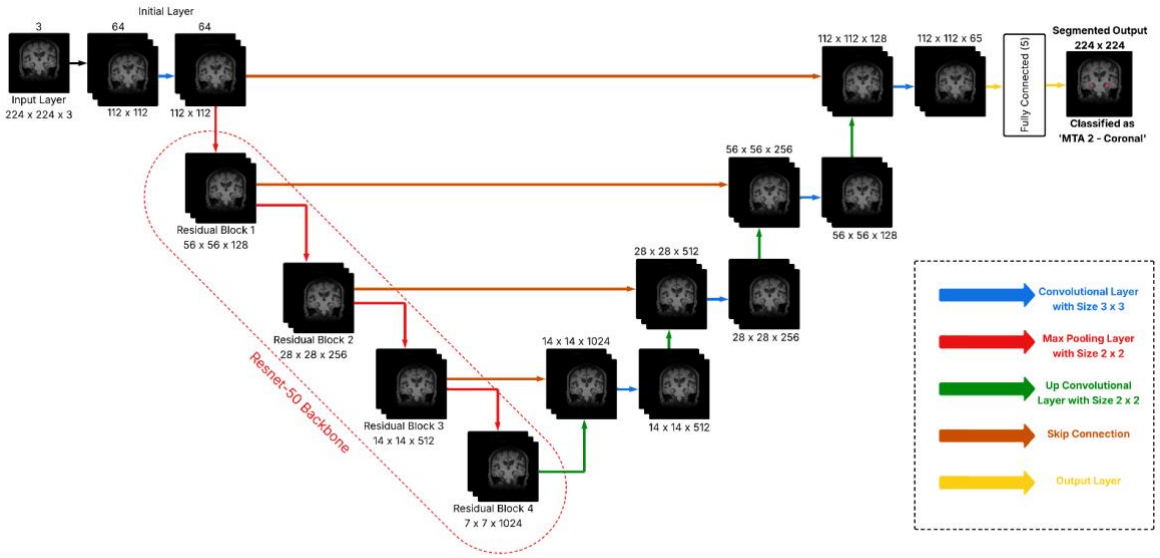


Fig. 3. Proposed U-net based on Resnet-50 backbone.

As seen in Fig. 3, the U-net layers are divided into an encoder and decoder pathway. The encoder, adapted from ResNet-50, is responsible for hierarchical feature extraction from the coronal T1-weighted MRI input of size $224 \times 224 \times 3$. The initial convolutional layer generates low-level features $112 \times 112 \times 64$, followed by four residual blocks that progressively reduce the spatial dimension while enriching semantic information. $Residual\ Block\ 1\ (56 \times 56 \times 128)$ extracts edge and contrast features, $Residual\ Block\ 2\ (28 \times 28 \times 256)$ captures intermediate anatomical structures, $Residual\ Block\ 3\ (14 \times 14 \times 512)$ represents high-level hippocampal patterns, and $Residual\ Block\ 4\ (7 \times 7 \times 1024)$produces abstract semantic features, forming the bottleneck of the encoder.

The decoder reconstructs spatial resolution through successive up-sampling operations and skip connections that concatenate high-resolution feature maps from the encoder. This step ensures that global semantic information is preserved while local details of hippocampal boundaries are recovered. For instance, the $7 \times 7 \times 1024$ feature map is up sampled to $14 \times 14 \times 512$ and concatenated with the corresponding encoder output, and this process continues until the resolution is restored to 224×224. A final 1×1 convolution layer

Corresponding author: Aldienannisa Devin Salsabila, aldienannisadevins@poltekkes-smg.ac.id, Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Jl.Tirto Agung, Pedalangan, Banyumanik, Semarang City, Central Java. **DOI:** https://doi.org/10.35882/ijeeemi.v7i4.263

produces a segmentation mask that delineates the hippocampal region with high accuracy.

For segmentation, the predicted hippocampal region of interest (ROI) is extracted and normalized before being passed to the classification head. A fully connected layer maps the extracted features into five output categories corresponding to the MTA scale (0–4). This dual-stage design allows the model to not only segment the hippocampus but also automatically assign a clinically interpretable atrophy grade. In this way, the proposed U-Net with ResNet-50 backbone effectively bridges the gap between pixel-level segmentation and high-level diagnostic classification.

D. Hyperparameter Settings

To optimize the training process, this study employs the Adaptive Moment Estimation (ADAM) optimizer, which is widely used in deep learning due to its efficient handling of sparse gradients and adaptive learning rates [27]. The maximum number of training epochs was set to 10, ensuring sufficient iterations while avoiding overfitting, given the dataset size [28]. A mini-batch size of 8 was used to balance between computational efficiency and gradient stability. The initial learning rate was configured at 1×10^{-4}, which allows gradual convergence of the model parameters toward the optimal solution. These values were determined based on prior literature and preliminary experiments that ensured stable convergence on our dataset.

E. Confusion Matrix Evaluation

The performance of the proposed U-Net with ResNet-50 backbone was evaluated using a confusion matrix, which provides a comprehensive summary of classification results by comparing predicted labels with ground truth annotations [29], [30] Several performance metrics were derived from the confusion matrix, including accuracy (calculated using Eq. (7)), recall (calculated using Eq. (8)), specificity (calculated using Eq. (9)), and F1-score (calculated using Eq. (10). These metrics are commonly used in medical image analysis to ensure that both positive and negative cases are correctly identified.

$$Accuracy = \frac{TP + TN}{Total\ Data} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$Specificity = \frac{TN}{TN + FP} \quad (9)$$

$$F1 - Score = 2 \cdot \frac{\left(\frac{TP}{TP + FP}\right) \cdot Recall}{\left(\frac{TP}{TP + FP}\right) + Recall} \quad (10)$$

Accuracy measures the proportion of correctly classified samples to the total number of samples. Sensitivity (also referred to as recall) quantifies the ability of the model to identify positive cases correctly, Specificity evaluates the model's capability to identify negative cases correctly, and F1-score provides a harmonic meaning between precision and recall, which balances false positives and false negatives in imbalanced datasets. Where where TP denotes true positives, TN true negatives, FP false positives, and FN false negatives.

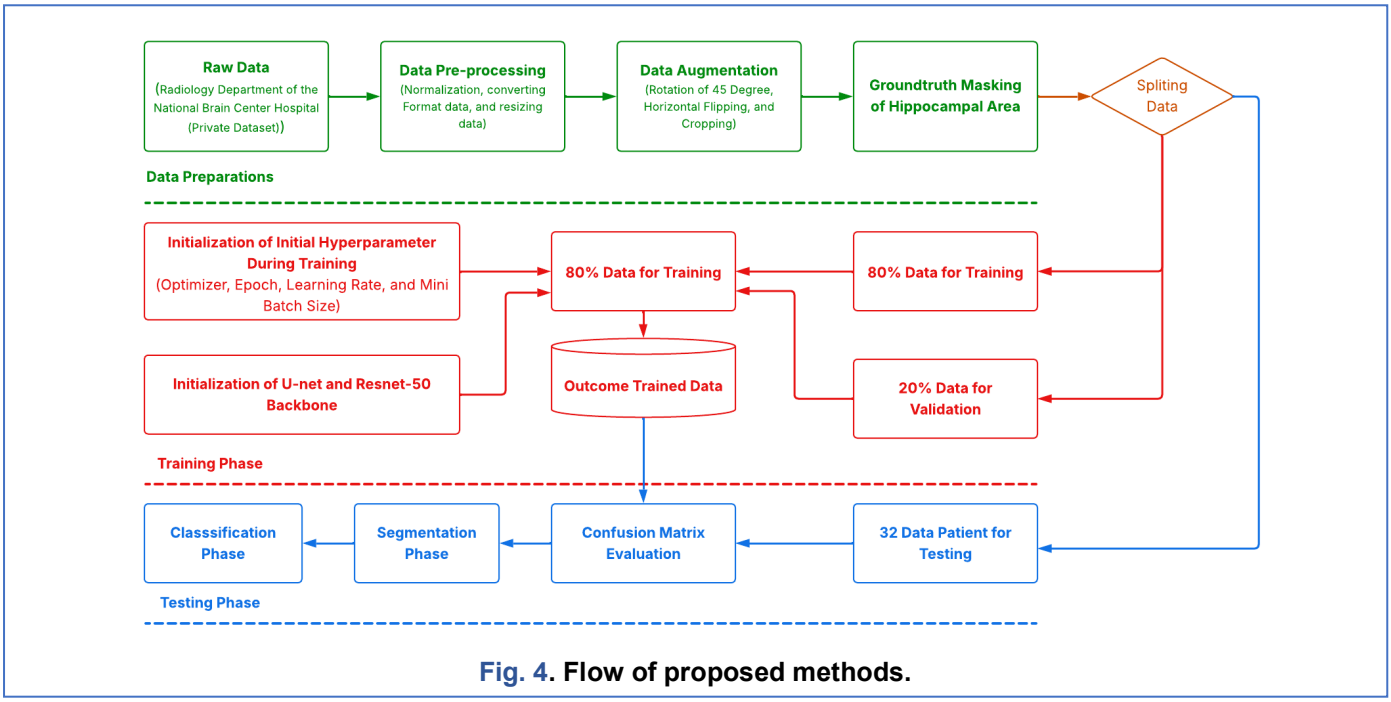


Fig. 4. Flow of proposed methods.

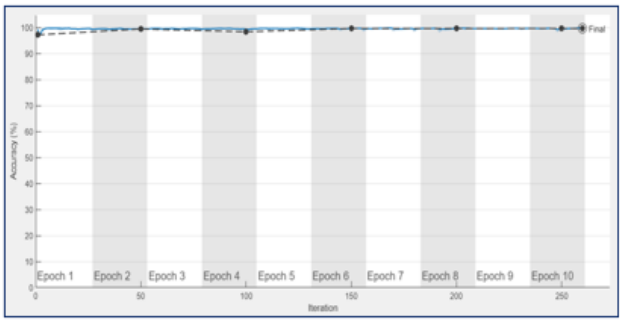
F. Proposed Workflow

As seen in Fig. 4, the proposed workflow begins with the data preparation stage. Raw MRI brain images were obtained from the Radiology Department of the National Brain Center Hospital. They underwent several pre-processing steps, including normalization, format conversion, and resizing to a uniform resolution.

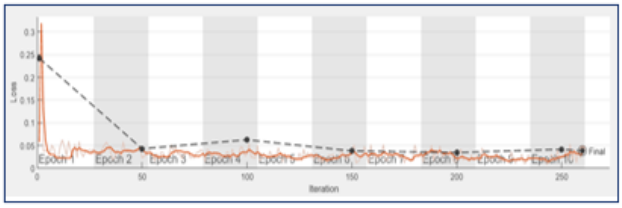
To enhance model robustness and reduce overfitting, data augmentation techniques such as rotation (±45°), horizontal flipping, and cropping were applied. In addition, ground truth masks of the hippocampal region were manually annotated by experts to provide reliable labels

Corresponding author: Aldienannisa Devin Salsabila, aldienannisadevins@poltekkes-smg.ac.id, Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Jl.Tirto Agung, Pedalangan, Banyumanik, Semarang City, Central Java. **DOI:** https://doi.org/10.35882/ijeeemi.v7i4.263

for training and evaluation. The dataset was split into training, validation, and testing subsets. Eighty percent of the data was allocated for training, while the remaining twenty percent was reserved for validation to fine-tune hyperparameters and prevent overfitting. During this phase, initial hyperparameters such as optimizer (ADAM), learning rate, maximum epoch, and mini-batch size were configured, and the U-Net with ResNet-50 backbone was initialized. The model was then trained iteratively on the training set, with performance monitored on the validation set to ensure convergence and stability. As seen in Fig. 5, the training process of the proposed U-Net with ResNet-50 backbone demonstrates a stable convergence pattern across 10 epochs. In the testing phase, the trained model was applied to 32 unseen patient cases. This phase consisted of three primary stages: classification, segmentation, and evaluation. The segmentation module generated hippocampal masks, while the classification module assigned MTA grades. Finally, performance was assessed using confusion matrix analysis to compute



(a) Training Accuracy



(b) Loss Accuracy

Fig. 5. Proposed model during training.

evaluation metrics such as accuracy, sensitivity, specificity, and F1-score. This structured workflow ensures that the proposed framework not only segments hippocampal regions accurately but also provides clinically relevant grading for dementia diagnosis.

III. Results

All experiments in this study were implemented using MATLAB 2020a as the primary development environment. The simulations and training processes were executed on a high-performance computing system equipped with an AMD Ryzen 9 9950X3D processor, 64 GB of RAM, an NVIDIA RTX 5080 GPU, and a 4 TB SSD for software installation and data storage. This computational setup ensured efficient execution of deep learning operations, including convolution, residual learning, segmentation, and classification, while also providing sufficient memory and processing power to handle small-scale to large-scale MRI datasets

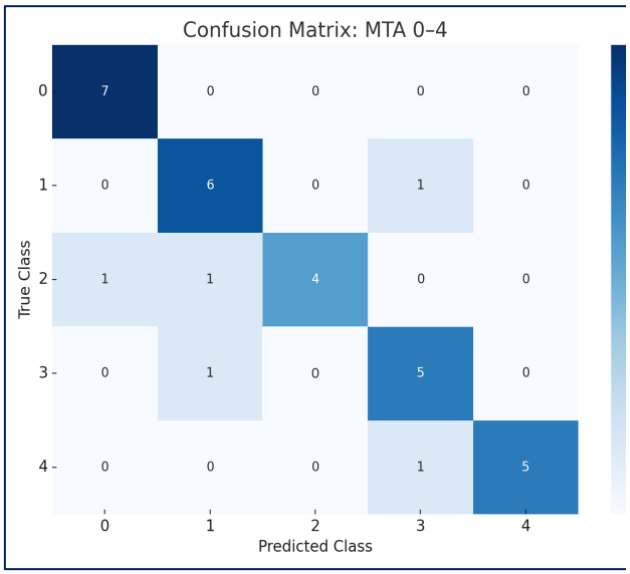


Fig. 6. Multi-prediction based on confusion matrix.

A. Training Phase

Both training and validation accuracies show a consistent upward trend, starting from 98.4% and 98.1% at the first epoch and reaching **99.9%** and 99.7% respectively by the tenth epoch. This indicates that the model is capable of learning discriminative features effectively without exhibiting signs of severe overfitting. The training loss decreases sharply in the initial epochs, from 0.06 in epoch 1 to 0.01 by epoch 3, and stabilizes at approximately 0.005 in the final epochs. Similarly, validation loss experiences a significant reduction, dropping from 0.31 in epoch 1 to 0.01 in epoch 10. After training was done, as seen in Fig. 6, the confusion matrix showed the performance of the proposed U-Net with ResNet-50 backbone in predicting the MTA score on 32 coronal slice test data. The results show that the model achieved strong diagonal dominance, indicating reliable classification across all MTA categories (0–4). Specifically, the model correctly classified 7 out of 8 samples for MTA 0, 6 out of 8 for MTA 1, 4 out of 4 for MTA 2, 5 out of 7 for MTA 3, and 5 out of 5 for MTA 4. These results confirm that the proposed model is capable of accurately distinguishing between different levels of hippocampal atrophy. Furthermore, Table 1 summarizes the quantitative evaluation of the model's performance across each MTA grade, which presents the values of accuracy, sensitivity (recall), specificity, and F1-score. These metrics provide a more comprehensive assessment beyond the confusion matrix by quantifying the model's ability to correctly detect positive cases, avoid false positives, and balance precision with recall. As seen in Table 1, it can be observed that the model achieved consistently high accuracy values, ranging between 0.91 and 0.97, indicating that most predictions matched the ground-truth labels. The recall values varied between 0.667 and 1.00, which shows that the model performed perfectly in detecting MTA 0 cases. Still, relatively lower sensitivity was observed In MTA 2 (0.667), suggesting

Corresponding author: Aldienannisa Devin Salsabila, aldienannisadevins@poltekkes-smg.ac.id, Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Jl.Tirto Agung, Pedalangan, Banyumanik, Semarang City, Central Java. **DOI:** https://doi.org/10.35882/ijeeemi.v7i4.263

some false negatives occurred in that class. Conversely, specificity remained high across all grades (≥ 0.92), highlighting the model's strong ability to minimize false positives. The F1-Score, which balances precision and recall, ranged from 0.80 to 0.933, reflecting robust performance in most classes with slight variation depending on the difficulty of class discrimination. Meanwhile, the NDP values ranged between 0.7 and 1.00, and the NDN values were consistently high (≥ 0.93), reinforcing the reliability of the model in distinguishing negative cases correctly. The macro average values further summarize the model's general performance across all classes: accuracy = 0.94, recall = 0.838, specificity = 0.968, F1-score = 0.855, NDP = 0.892, and NDN = 0.963. These results indicate that the proposed U-Net with ResNet-50 backbone achieved a stable and balanced performance, with particularly strong specificity and overall accuracy, making it suitable for reliable automated MTA grading.

Table 1. Confusion matrix evaluation.

MTA	Accuracy	Recall	Specificity	F1-Score	NDP	NDN
0	0.96	1.00	0.96	0.933	0.875	1.00
1	0.91	0.857	0.92	0.80	0.75	0.958
2	0.93	0.667	1.00	0.80	1.00	0.93
3	0.93	0.833	0.962	0.833	0.833	0.962
4	0.97	0.833	1.00	0.909	1.00	0.963
Macro AVG	**0.940**	**0.940**	**0.940**	**0.940**	**0.940**	**0.940**

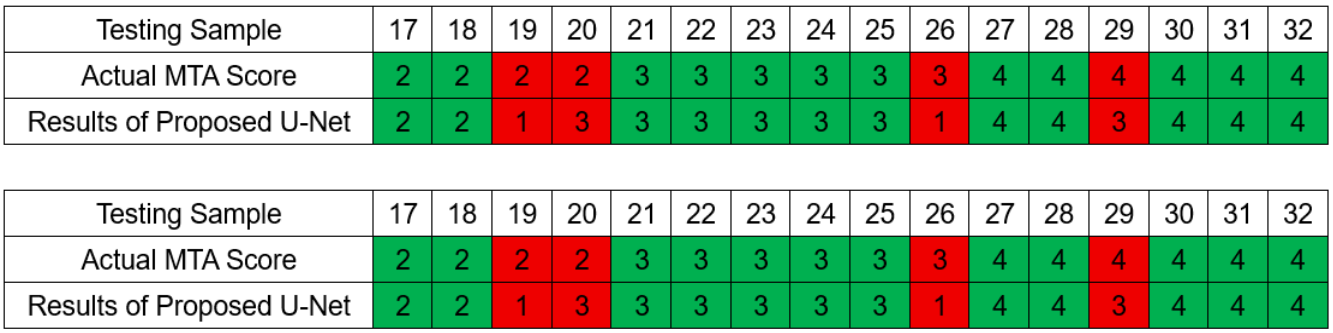


Fig. 7. Comparison between actual MTA scores assigned by radiology specialists and predictions generated by the proposed U-Net with ResNet-50 backbone on 32 test samples. Green indicates correct predictions, while red highlights misclassifications.

B. Comparison of Expertise Doctor of Specialist Radiology

In order to validate the performance of the proposed model, the classification results of the U-Net with ResNet-50 backbone were compared against the expertise of radiology specialists (Actual MTA Score), who serve as the gold standard in determining MTA scores. A total of 32 coronal MRI test samples were evaluated, with both manual scoring and automated U-Net predictions. As seen in Fig. 7, the comparison between the actual MTA scores assigned by radiology specialists and the results predicted by the proposed U-Net model demonstrates a high level of agreement. Each testing sample is represented along the horizontal axis, with the top row showing the actual MTA score (manual assessment) and the bottom row indicating the predicted score from the deep learning model. Green-colored cells indicate correct predictions that matched the manual assessment, while red-coloredcells represent mismatches between the manual and automated evaluations. As seen in Fig. 7, the comparison between the actual MTA scores assigned by radiology specialists and the results predicted by the proposed U-Net model demonstrates a high level of agreement. Each testing sample is represented along the horizontal axis, with the top row showing the actual MTA score (manual assessment) and the bottom row indicating the predicted score from the deep learning model. Green-colored cells indicate correct predictions that matched the manual assessment, while red-colored cells represent mismatches between the manual and automated evaluations. From a total of 32 testing samples, the U-Net with ResNet-50 backbone correctly classified 27 samples, achieving consistency with radiologists' scoring. Only 5 samples (Samples 13, 15, 20, 26, and 29) were misclassified, which are highlighted in red. These misclassifications generally occurred in borderline cases where the degree of hippocampal atrophy lies close to the threshold between two adjacent MTA categories, making them challenging even for human experts. Nevertheless, the predominance of green cells across the figure highlights the robustness and reliability of the proposed

Corresponding author: Aldienannisa Devin Salsabila, aldienannisadevins@poltekkes-smg.ac.id, Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Jl.Tirto Agung, Pedalangan, Banyumanik, Semarang City, Central Java. **DOI:** https://doi.org/10.35882/ijeeemi.v7i4.263

model,reinforcing its capability to approximate radiologist-level interpretation in clinical practice.

Among the five misclassified cases, most errors occurred in borderline grades, particularly between MTA levels 2 and 3, where visual differences are subtle even for radiologists. In several instances, lower image quality and scanner-related variability also contributed to ambiguity in hippocampal boundaries. These factors highlight the inherent challenges of distinguishing intermediate atrophy levels and explain part of the

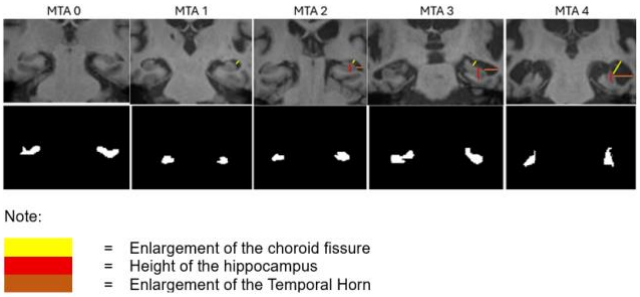


Fig. 8. MTA score results using U-Net segmentation showing progressive hippocampal atrophy, choroid fissure enlargement, and temporal horn enlargement.

model's misclassification. Addressing these errors may require incorporating larger datasets with a more balanced representation of borderline cases, as well as preprocessing techniques to mitigate noise and variability.

C. Testing Phase (Segmentation)

In this section, the testing phase focuses on evaluating the segmentation performance of the proposed U-Net with ResNet-50 backbone in delineating the hippocampal region across different MTA scores. The segmentation process aims to visually and quantitatively capture the structural changes of the hippocampus, particularly its progressive atrophy as reflected in the MTA scale. By localizing the hippocampal boundaries, the model provides an interpretable representation that facilitates the differentiation of atrophy levels in clinical assessment.

As seen in Fig. 8, the segmentation outcomes illustrate clear differences across MTA grades (0–4). At MTA 0, the hippocampus shows no signs of atrophy, with no enlargement of the choroid fissure or temporal horn. At MTA 1, there is a visible enlargement of the choroid fissure compared to MTA 0. In contrast, MTA 2–4 present progressive structural changes, characterized by both enlargement of the choroid fissure and temporal horn as well as a gradual reduction in hippocampal height. Although MTA 1 and MTA 2 appear similar, the latter is marked by deeper enlargement of the choroid fissure and more pronounced reduction of hippocampal height, making it distinguishable from earlier stages.

As seen in Fig. 9, the segmentation results demonstrate the capability of the proposed U-Net with ResNet-50 backbone to accurately localize the hippocampal region compared to the ground truth annotated by radiology experts. The input testing images in (a) represent coronal MRI slices, while (b) shows the manually annotated hippocampal masks serving as the reference standard. In (c), the predicted segmentation generated by the model is illustrated, where the hippocampal structures are correctly delineated with high spatial correspondence to the ground truth. The close

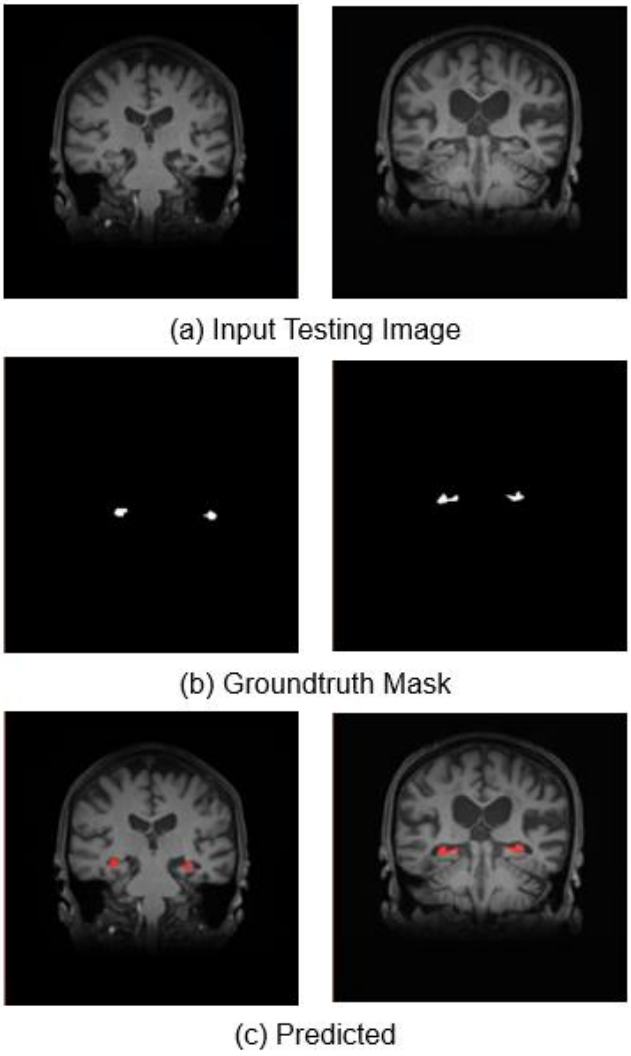


Fig. 9. Comparison of hippocampal segmentation results.

visual alignment between predicted and expert-annotated masks confirms the model's robustness in capturing hippocampal boundaries, even in regions with subtle anatomical variations. This consistency highlights the proposed approach's clinical reliability in supporting automated MTA grading.

IV. Discussion

To contextualize the findings of this study, a comparative analysis was conducted against several related works on hippocampal segmentation using deep learning. As summarized in Table 2, prior studies have primarily focused on improving segmentation accuracy through various architectural enhancements such as convolutional neural networks (CNN) [16], attention-guided mechanisms [17], and

Corresponding author: Aldienannisa Devin Salsabila, aldienannisadevins@poltekkes-smg.ac.id, Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Jl.Tirto Agung, Pedalangan, Banyumanik, Semarang City, Central Java. **DOI:** https://doi.org/10.35882/ijeeemi.v7i4.263

modified 3D U-Net models with filling techniques [18]. These methods demonstrated promising segmentation performance, with reported accuracies ranging from 86.74% to 99.7%. However, a notable limitation across these approaches is their scope being restricted to segmentation alone, without extending to clinical interpretation in the form of MTA grading. This gap reduces their direct applicability in diagnostic workflows where severity classification is critical. The comparative analysis in Table 2 highlights the evolution of deep learning methods for hippocampal segmentation. The study by [16] demonstrated that convolutional neural networks (CNNs) can effectively capture spatial features of the hippocampus, achieving 96.5% accuracy. However, its performance dropped significantly when applied to heterogeneous datasets, revealing a lack of robustness and clinical scalability. Similarly, [17] showed that embedding attention gates into U-Net architectures improved Dice scores and achieved 99.7% accuracy by suppressing irrelevant background noise. While this approach enhances focus on hippocampal structures, its reliance on attention mechanisms alone limits robustness in severe atrophy cases, where structural boundaries are blurred. Study [18] proposed a 3D U-Net with a novel filling strategy, achieving a Dice score of 0.8674. Although it outperformed standard 3D models, the method remained confined to single-center datasets, reducing generalizability, and again, did not address classification into clinical atrophy grades. In our experiments, the average processing time per case was approximately 36 Minutes and 16 seconds on a GPU, indicating feasibility for a near real-time application. The system could be incorporated into existing radiology pipelines as a decision-support tool, where segmented hippocampi and automated MTA scores are generated alongside MRI scans and reviewed by radiologists. With a suitable graphical interface or integration into PACS systems, such a framework has the potential to reduce reporting time and improve consistency without disrupting standard clinical practice. What is evident across these studies is that the research community has primarily concentrated on segmentation precision, while overlooking the clinically essential step of translating segmentation outputs into interpretable atrophy levels (MTA scores). This omission creates a gap between algorithmic segmentation and clinical decision-making, where radiologists require not just a mask of the hippocampus but also an objective grading of atrophy severity. The proposed U-Net with ResNet-50 backbone addresses this critical gap by combining robust segmentation with automated classification into MTA grades (0–4). By embedding residual connections in the encoder path, the model ensures stable gradient flow and rich feature representation, preventing the rapid accuracy degradation observed in conventional U-Nets. This architectural synergy enables the system to achieve a classification accuracy of 99.9%, which not only surpasses related segmentation-only approaches but

Table 2. Comparative summary of related studies and proposed method.

[16]	CNN-based segmentation with enhanced preprocessing and training strategies	Accuracy: 96.5%	Limited generalization to heterogeneous MRI data; focuses only on segmentation, without classification into clinical MTA scores	Our study integrates segmentation with automated classification (MTA 0–4), bridging pixel-level analysis with clinically interpretable grading
[17]	U-Net with attention gates to emphasize hippocampal regions and suppress irrelevant background	Accuracy: 99.7%	Performance decreases with severe atrophy or low-quality MRI; it addresses segmentation only, without atrophy grading.	The proposed U-Net with ResNet-50 backbone performs both segmentation and classification, enabling direct MTA scoring for clinical utility.
[18]	3D U-Net combined with a novel filling approach for continuity in hippocampal segmentation	Dice: 0.8674, IoU: 0.7668	The dataset limited to single center; lacks adaptability to multi-modal or diverse clinical datasets; no classification into MTA levels	Our model demonstrates robust segmentation and extends to automated classification, achieving 99.9% accuracy on MTA grading.
Proposed Study	U-Net with ResNet-50 backbone for segmentation and classification	Accuracy: 99.9%	Dataset limitation: model requires re-adjustment when applied to different datasets or MRI acquisition protocols	Provides an end-to-end pipeline: segmentation of hippocampal structures and classification into MTA 0–4, filling the gap left by prior segmentation-only approaches

Corresponding author: Aldienannisa Devin Salsabila, aldienannisadevins@poltekkes-smg.ac.id, Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Jl.Tirto Agung, Pedalangan, Banyumanik, Semarang City, Central Java. **DOI:** https://doi.org/10.35882/ijeeemi.v7i4.263

also provides direct clinical interpretability. In practical terms, this means the system bridges the divide between low-level pixel analysis and high-level diagnostic reasoning. The main limitation of our study lies in dataset dependency.

The model was trained and validated on a single dataset, meaning that its application to multi-center or multi-protocol MRI data would require fine-tuning or transfer learning. Despite this constraint, the contribution of this study is substantial: it proposes the first fully automated framework that integrates segmentation and MTA grading, thereby extending beyond the boundaries of prior segmentation-focused research and offering genuine clinical relevance for dementia diagnosis.

Despite these limitations, the findings have important implications of this study demonstrate that the integration of U-Net with a ResNet-50 backbone offers a promising solution for objective and automated hippocampal segmentation and MTA grading. The implication of this research lies in its potential to transform both clinical practice and research in dementia diagnosis. Clinically, the proposed model reduces inter- and intra-rater variability inherent in manual MTA assessment, thereby providing radiologists with consistent and reproducible results. This standardization is crucial for enhancing diagnostic accuracy, particularly in distinguishing borderline MTA scores (e.g., grade 2 vs. grade 3), which are often prone to misinterpretation in routine practice. Moreover, by enabling automatic analysis, the model can significantly reduce reporting time and assist in large-scale dementia screening programs, especially in resource-limited settings where trained neuroradiologists may be scarce.

From a research perspective, the implementation of deep learning-based segmentation frameworks contributes to the growing body of evidence that artificial intelligence can bridge the gap between image processing and diagnostic decision-making. The end-to-end nature of this model allows seamless translation of image features into clinically interpretable scores, setting a foundation for future studies to incorporate multi-modal MRI or expand into longitudinal analysis of hippocampal atrophy progression. Ultimately, the adoption of such models supports the broader move toward precision medicine in neurology, where objective biomarkers are integrated into clinical workflows to improve patient outcomes.

V. Conclusion

This study aimed to develop an automated framework for hippocampal segmentation and Medial Temporal Atrophy (MTA) grading by integrating U-Net with a ResNet-50 backbone. Unlike prior works that emphasized segmentation alone, the proposed model combines pixel-level hippocampal delineation with clinically interpretable classification. The system achieved remarkable performance, with a training accuracy of 99.9%. In comparative testing against radiology specialists, it correctly predicted 26 out of 32 samples, misclassifying only 6 cases. These results demonstrate that the proposed architecture successfully bridges the gap between deep learning–based segmentation and clinical decision-making in dementia diagnosis. The primary limitation of this work lies in its reliance on a single dataset, which may restrict generalizability to broader clinical settings. As future work, this study will explore the use of 3D U-Net architectures to capture richer volumetric information from MRI scans, thereby improving the robustness and accuracy of hippocampal atrophy assessment also focus on validating the proposed framework on larger, multi-center cohorts and conducting prospective clinical evaluations to establish its robustness across diverse clinical environments. Additionally, integration into radiology workflows through PACS deployment and user-interface development will be explored to facilitate real-world applicability.

References

[1] S. Nugraha *et al.*, “Knowledge about Dementia and its Associated Factors: Study among the Middle-aged Population in Indonesia,” *Open Access Maced J Med Sci*, vol. 10, no. E, pp. 783–789, May 2022, doi: 10.3889/oamjms.2022.8892.

[2] S. V. Birajdar, M. Mulchandani, F. Mazahir, and A. K. Yadav, “Dementia and neurodegenerative disorder: An introduction,” in *Nanomedicine-Based Approaches for the Treatment of Dementia*, Elsevier, 2023, pp. 1–36. doi: 10.1016/B978-0-12-824331-2.00007-8.

[3] Y. L. Rao, B. Ganaraja, B. V. Murlimanju, T. Joy, A. Krishnamurthy, and A. Agrawal, “Hippocampus and its involvement in Alzheimer’s disease: a review,” *3 Biotech*, vol. 12, no. 2, p. 55, Feb. 2022, doi: 10.1007/s13205-022-03123-4.

[4] S. Ribarič, “Detecting Early Cognitive Decline in Alzheimer’s Disease with Brain Synaptic Structural and Functional Evaluation,” *Biomedicines*, vol. 11, no. 2, p. 355, Jan. 2023, doi: 10.3390/biomedicines11020355.

[5] A. Brown *et al.*, “Measuring the quality of life of family carers of people with dementia: development and validation of C-DEMQOL,” *Quality of Life Research*, vol. 28, no. 8, pp. 2299–2310, Aug. 2019, doi: 10.1007/s11136-019-02186-w.

[6] A. Molinder, D. Ziegelitz, S. E. Maier, and C. Eckerström, “Validity and reliability of the medial temporal lobe atrophy scale in a memory clinic population,” *BMC Neurol*, vol. 21, no. 1, p. 289, Dec. 2021, doi: 10.1186/s12883-021-02325-2.

[7] E. S. C. Korf, L.-O. Wahlund, P. J. Visser, and P. Scheltens, “Medial temporal lobe atrophy on MRI predicts dementia in patients with mild cognitive impairment,” *Neurology*, vol. 63, no. 1, pp. 94–100, Jul. 2004, doi: 10.1212/01.WNL.0000133114.92694.93.

Corresponding author: Aldienannisa Devin Salsabila, aldienannisadevins@poltekkes-smg.ac.id, Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Jl.Tirto Agung, Pedalangan, Banyumanik, Semarang City, Central Java. **DOI:** https://doi.org/10.35882/ijeeemi.v7i4.263

[8] A. Rau and H. Urbach, "The MTA score—simple and reliable, the best for now?," *Eur Radiol*, vol. 31, no. 12, pp. 9057–9059, Dec. 2021, doi: 10.1007/s00330-021-08340-8.

[9] G. Cipriani, S. Danti, L. Picchi, A. Nuti, and M. Di Fiorino, "Daily functioning and dementia," *Dement Neuropsychol*, vol. 14, no. 2, pp. 93–102, Jun. 2020, doi: 10.1590/1980-57642020dn14-020001.

[10] E. Nichols *et al.*, "Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019," *Lancet Public Health*, vol. 7, no. 2, pp. e105–e125, Feb. 2022, doi: 10.1016/S2468-2667(21)00249-8.

[11] E. J. W. Van Someren *et al.*, "Medial temporal lobe atrophy relates more strongly to sleep-wake rhythm fragmentation than to age or any other known risk," *Neurobiol Learn Mem*, vol. 160, pp. 132–138, Apr. 2019, doi: 10.1016/j.nlm.2018.05.017.

[12] N. R. D. Cahyo, C. A. Sari, E. H. Rachmawanto, C. Jatmoko, R. R. A. Al-Jawry, and M. A. Alkhafaji, "A Comparison of Multi Class Support Vector Machine vs Deep Convolutional Neural Network for Brain Tumor Classification," in *2023 International Seminar on Application for Technology of Information and Communication (iSemantic)*, IEEE, Sep. 2023, pp. 358–363. doi: 10.1109/iSemantic59612.2023.10295336.

[13] M. Tsuneki, "Deep learning models in medical image analysis," *J Oral Biosci*, vol. 64, no. 3, pp. 312–320, Sep. 2022, doi: 10.1016/j.job.2022.03.003.

[14] P. Subarkah, W. R. Damayanti, and R. A. Permana, "Comparison of Correlated Algorithm Accuracy Naive Bayes Classifier and Naive Bayes Classifier for Classification of heart failure," *ILKOM Jurnal Ilmiah*, vol. 14, no. 2, pp. 120–125, Aug. 2022, doi: 10.33096/ilkom.v14i2.1148.120-125.

[15] L. A. M. Zaki *et al.*, "Comparing two artificial intelligence software packages for normative brain volumetry in memory clinic imaging," *Neuroradiology*, vol. 64, no. 7, pp. 1359–1366, Jul. 2022, doi: 10.1007/s00234-022-02898-w.

[16] R. A. Hazarika, A. K. Maji, R. Syiem, S. N. Sur, and D. Kandar, "Hippocampus Segmentation Using U-Net Convolutional Network from Brain Magnetic Resonance Imaging (MRI)," *J Digit Imaging*, vol. 35, no. 4, pp. 893–909, Aug. 2022, doi: 10.1007/s10278-022-00613-y.

[17] N. Sohail and S. M. Anwar, "A Modified U-Net Based Framework for Automated Segmentation of Hippocampus Region in Brain MRI," *IEEE Access*, vol. 10, pp. 31201–31209, 2022, doi: 10.1109/ACCESS.2022.3159618.

[18] Q. Yang, C. Wang, K. Pan, B. Xia, R. Xie, and J. Shi, "An improved 3D-UNet-based brain hippocampus segmentation model based on MR images," *BMC Med Imaging*, vol. 24, no. 1, p. 166, Jul. 2024, doi: 10.1186/s12880-024-01346-w.

[19] N. R. D. Cahyo and M. M. I. Al-Ghiffary, "An Image Processing Study: Image Enhancement, Image Segmentation, and Image Classification using Milkfish Freshness Images," *IJECAR) International Journal of Engineering Computing Advanced Research*, vol. 1, no. 1, pp. 11–22, 2024.

[20] I. P. Kamila, C. A. Sari, E. H. Rachmawanto, and N. R. D. Cahyo, "A Good Evaluation Based on Confusion Matrix for Lung Diseases Classification using Convolutional Neural Networks," *Advance Sustainable Science, Engineering and Technology*, vol. 6, no. 1, p. 0240102, Dec. 2023, doi: 10.26877/asset.v6i1.17330.

[21] M. M. I. Al-Ghiffary, N. R. D. Cahyo, E. H. Rachmawanto, C. Irawan, and N. Hendriyanto, "Adaptive deep learning based on FaceNet convolutional neural network for facial expression recognition," *Journal of Soft Computing*, vol. 05, no. 03, pp. 271–280, 2024, doi: https://doi.org/10.52465/joscex.v5i3.450.

[22] A. Susanto, I. U. W. Mulyono, C. A. Sari, E. H. Rachmawanto, D. R. I. M. Setiadi, and M. K. Sarker, "Handwritten Javanese script recognition method based 12-layers deep convolutional neural network and data augmentation," *IAES International Journal of Artificial Intelligence*, vol. 12, no. 3, pp. 1448–1458, Sep. 2023, doi: 10.11591/ijai.v12.i3.pp1448-1458.

[23] A. Abedalla, M. Abdullah, M. Al-Ayyoub, and E. Benkhelifa, "Chest X-ray pneumothorax segmentation using U-Net with EfficientNet and ResNet architectures," *PeerJ Comput Sci*, vol. 7, p. e607, Jun. 2021, doi: 10.7717/peerj-cs.607.

[24] A. Salas-Espinales, E. Vélez-Chávez, R. Vázquez-Martín, A. García-Cerezo, and A. Mandow, "U-Net/ResNet-50 Network with Transfer Learning for Semantic Segmentation in Search and Rescue," 2024, pp. 244–255. doi: 10.1007/978-3-031-59167-9_21.

[25] J. Zhu, R. Zhang, and H. Zhang, "An MRI brain tumor segmentation method based on improved U-Net," *Mathematical Biosciences and Engineering*, vol. 21, no. 1, pp. 778–791, 2023, doi: 10.3934/mbe.2024033.

[26] M. M. I. Al-Ghiffary, C. A. Sari, E. H. Rachmawanto, N. M. Yacoob, N. R. D. Cahyo, and R. R. Ali, "Milkfish Freshness Classification Using Convolutional Neural Networks Based on Resnet50 Architecture," *Advance Sustainable Science Engineering and Technology*, vol. 5, no. 3, p.

Corresponding author: Aldienannisa Devin Salsabila, aldienannisadevins@poltekkes-smg.ac.id, Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Jl.Tirto Agung, Pedalangan, Banyumanik, Semarang City, Central Java. **DOI:** https://doi.org/10.35882/ijeeemi.v7i4.263

0230304, Oct. 2023, doi: 10.26877/asset.v5i3.17017.

[27] D. Kilichev and W. Kim, "Hyperparameter Optimization for 1D-CNN-Based Network Intrusion Detection Using GA and PSO," *Mathematics*, vol. 11, no. 17, Sep. 2023, doi: 10.3390/math11173724.

[28] L. Liao, H. Li, W. Shang, and L. Ma, "An Empirical Study of the Impact of Hyperparameter Tuning and Model Optimization on the Performance Properties of Deep Neural Networks," *ACM Transactions on Software Engineering and Methodology*, vol. 31, no. 3, pp. 1–40, Jul. 2022, doi: 10.1145/3506695.

[29] M. Heydarian, T. E. Doyle, and R. Samavi, "MLCM: Multi-Label Confusion Matrix," *IEEE Access*, vol. 10, pp. 19083–19095, 2022, doi: 10.1109/ACCESS.2022.3151048.

[30] I. Markoulidakis and G. Markoulidakis, "Probabilistic Confusion Matrix: A Novel Method for Machine Learning Algorithm Generalized Performance Analysis," *Technologies (Basel)*, vol. 12, no. 7, p. 113, Jul. 2024, doi: 10.3390/technologies12070113.

Author Biography

Aldienannisa Devin Salsabila, S.Tr.Kes. (Rad) is an academic and researcher specializing in radiology and diagnostic imaging. She earned her Bachelor of Applied Health in Radiology in 2023 and is currently pursuing a Master of Applied Health in Diagnostic Imaging at the Poltekkes Kemenkes Semarang, with an expected graduation in 2025. Throughout her academic journey, she has demonstrated a strong interest in advancing radiological science, particularly in its application to the early detection and diagnosis of degenerative diseases. Her most recent research, conducted between 2023 and 2025, focuses on dementia, reflecting her commitment to exploring the role of imaging technologies in cognitive disorder diagnostics. With a solid academic foundation and a deep dedication to scientific inquiry, she actively contributes to the development of radiological knowledge through research, academic collaborations, and scholarly publications.

Dr. Fatimah, S.ST., M.Kes. is an academic and practitioner in the field of radiology with extensive experience in education, management, and professional organizations. She earned her Doctorate in Medical and Health Sciences from Diponegoro University Semarang (2023), preceded by a Master of Applied Science (2014), a Bachelor of Applied Science in Radiology (2008), and a Diploma in Radiodiagnostic and Radiotherapy (1996). Her career spans from serving as a Radiographer at Dr. Kariadi Hospital Semarang (1998–2011) to her current role as Head of the Radiodiagnostic and Radiotherapy Department at Poltekkes Kemenkes Semarang (2018–2027). In addition, Dr. Fatimah has been actively involved in the Indonesian Association of Radiographers (PARI), holding key positions in research, publication, professional development, advocacy, and organizational leadership at both national and regional levels.

Darmini, S.Si., M.Kes. is a lecturer and currently serves as the Head of the Diploma III Radiology Study Program at Poltekkes Kemenkes Semarang, Indonesia. She earned her Bachelor's degree in Physics from Diponegoro University in 2002 and completed her Master's degree in Public Health at the same university in 2009. Her academic and professional interests lie at the intersection of radiology education, public health, and applied health sciences. With over a decade of experience in teaching and academic leadership, she has been actively involved in curriculum development and educational quality improvement for vocational radiology programs. Currently pursuing her doctoral degree at Universitas Negeri Semarang, she is committed to advancing radiology education by integrating evidence-based teaching approaches and promoting interdisciplinary collaboration. Her work contributes to the development of skilled radiology professionals capable of responding to evolving healthcare challenges in both clinical and public health settings.

Selamet Budi Kurniawan, S.Tr.Kes. is a radiographer currently serving at the National Brain Center Hospital (RS Pusat Otak Nasional) in Jakarta, Indonesia. He holds a Bachelor of Applied Health in Radiology and specializes in diagnostic imaging and clinical radiology services, particularly in the context of neurological and brain-related conditions. In addition to his clinical duties, he has actively contributed to academic research, particularly in supporting projects focused on radiographic techniques and diagnostic accuracy. His involvement in research supervision and implementation reflects a commitment to bridging clinical practice with scientific inquiry. With a strong foundation in radiologic technology, he plays an important role in ensuring high standards of patient care and imaging quality. His contributions bring valuable clinical insight into the development of evidence-based radiology practices.

Corresponding author: Aldienannisa Devin Salsabila, aldienannisadevins@poltekkes-smg.ac.id, Imaging Diagnostic Study Program of Postgraduate Program, Poltekkes Kemenkes Semarang, Jl.Tirto Agung, Pedalangan, Banyumanik, Semarang City, Central Java. **DOI:** https://doi.org/10.35882/ijeeemi.v7i4.263