IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 13 May 2024, accepted 28 June 2024, date of publication 29 July 2024, date of current version 6 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3426085

RESEARCH ARTICLE

Deep Learning Applications in MRI-Based Detection of the Hippocampal Region for Alzheimer's Diagnosis

YORI PUSPARANI[1,2], CHIH-YANG LIN[3], (Senior Member, IEEE), YIH-KUEN JAN[4], FU-YU LIN[5], BEN-YI LIAU[6], PETER ARDHIANTO[7], ELVIN NUR FURQON[8], JOHN SAHAYA RANI ALEX[9], (Senior Member, IEEE), JEETASHREE APARAJEETA[9], AND CHI-WEN LUNG[4,10]

[1]Department of Visual Communication Design, Budi Luhur University, Jakarta 12260, Indonesia
[2]Department of Digital Media Design, Asia University, Taichung 413305, Taiwan
[3]Department of Mechanical Engineering, National Central University, Taoyuan 320317, Taiwan
[4]Rehabilitation Engineering Laboratory, Department of Kinesiology and Community Health, University of Illinois at Urbana–Champaign, Urbana, IL 61801, USA
[5]Department of Neurology, China Medical University Hospital, Taichung 404327, Taiwan
[6]Department of Automatic Control Engineering, Feng Chia University, Taichung 407102, Taiwan
[7]Department of Visual Communication Design, Soegijapranata Catholic University, Semarang 50234, Indonesia
[8]Department of Electrical Engineering, Yuan Ze University, Taoyuan 320315, Taiwan
[9]School of Electronics Engineering, Vellore Institute of Technology, Chennai 600127, India
[10]Department of Creative Product Design, Asia University, Taichung 413305, Taiwan

Corresponding author: Chi-Wen Lung (cwlung@asia.edu.tw)

This work was supported by the National Science and Technology Council in Taiwan under Grant NSTC 113-2221-E-035-014 and Grant MOST111-2923-E-155-004-MY3.

HIGHLIGHT

1. Detecting landmarks in MRI images has proven to be a valuable approach that helps neurologists save time when diagnosing Alzheimer's disease.
2. YOLOv5 is a suitable deep-learning model for detecting the hippocampal region in three views and categories and assisting neurologists in diagnosing Alzheimer's disease.
3. The sagittal view has been identified as the most dependable view for detecting the hippocampal region.
4. Diagnosis of Alzheimer's disease using MRI requires three distinct views accurately to visualize brain shrinkage.

ABSTRACT The hippocampal region is one of the most affected brain areas observed as a landmark in Magnetic Resonance Imaging (MRI) images for Alzheimer's disease (AD) diagnosis. The diminished alterations in the hippocampal and degeneration of cholinergic circuits have been conclusively correlated with a decline in memory and cognitive function. However, the hippocampal region may not appear as clearly defined as other brain regions, making it difficult for neurologists and researchers to identify by visual inspection. The application of deep learning models to pinpoint the hippocampal region was initially valued. We assessed the ability of a deep learning model, You Only Live Once (YOLO), to detect hippocampal regions in three MRI image views and categories. The Alzheimer's Disease Neuroimaging Initiative-first (ADNI−1) dataset was used with 220 subjects in three categories using the three YOLO models. We obtained the YOLO performance for hippocampal region detection with accuracy in three views and categories. The average mean Average Precision (mAP) performance accuracy for YOLOv3 was 0.87, YOLOv4 was 0.85, and YOLOv5 was 0.96, respectively. The high accuracy of the detection of the hippocampal region was remarkable. We found that the sagittal view was higher than the axial and coronal views. Simultaneously, the Mild Cognitive Impairment (MCI) in the coronal view was lower among the three models. The results showed that YOLOv5 is a suitable model for detecting the hippocampal region in MRI images, and the sagittal view is the most reliable for detecting the hippocampal region in diagnosing AD. Our findings demonstrate the importance of detecting the hippocampal region to diagnose AD and accurately analyzing the hippocampal

The associate editor coordinating the review of this manuscript and approving it for publication was Humaira Nisar.

area within the region. The YOLOv5 model substantially affected performance metrics and interpretability across the three views and categories.

INDEX TERMS Landmark, hippocampal region, MRI image, YOLO, object detection.

I. INTRODUCTION

Dementia attributed to Alzheimer's disease (AD) constitutes the most common cause of cognitive impairment globally. The global number of persons with AD dementia, prodromal AD, and preclinical AD were estimated at 32, 69, and 315 million, respectively [1]. Deterioration in the brain regions is critical for observing and exploring early AD diagnosis before disease progression. Gradual loss of memory and cognitive function progresses through a transient clinical stage of AD [2]. Identifying AD-related brain changes is paramount for optimizing clinical interventions and attenuating disease progression and onset [3]. Patients with AD are typically diagnosed by examining the hippocampal, which is crucial for memory and cognitive function [4]. As there is no cure for AD, previous studies have focused on exercise therapy to prevent hippocampal volume reduction [5]. However, symptoms of progressive AD can manifest as mild cognitive impairment (MCI). Moreover, it is challenging to differentiate and study the hippocampal in each category because the transition from normal control (CN) to AD has one intermittent stage, popularly known as MCI [6]. These categories are beneficial for differentiating disease progression.

In addition, the hippocampal region is a wider area encompasses various structures involved in memory and cognitive function. Dysfunction or damage in the hippocampal region is associated with memory disorders, as is often observed in patients with AD [7]. Structural alterations in the hippocampal regions are significant markers of AD progression. To do this, magnetic resonance imaging (MRI) is a valuable tool for brain imaging and represents a reliable modality for diagnosing AD [8]. MRI images have three standard views: axial, coronal, and sagittal [9]. Three views of MRI images contain valuable information regarding the hippocampal region, which is beneficial for AD diagnosis [10], and each view allows the neurologist to accurately identify and delineate the hippocampal boundaries in MRI images [11].

Furthermore, deep learning systems are more effective in many research areas. For instance, deep learning presents a potent methodology for analyzing alterations in the hippocampal associated with AD [12]. It also saves neurologists time in diagnosing AD [13]. Various types of deep learning models, namely, You Only Look Once (YOLO), are used in medical images [14], and used YOLOv3 [15], YOLOv4 [16], and YOLOv5 [17]. Other YOLO models, such as YOLOv6, YOLOv7, and YOLOv8, offer advanced object detection capabilities. However, these models are typically underutilized in medical imaging, where specialized algorithms or detection frameworks are more commonly applied [18]. Thus, we proposed three YOLO models (i.e., YOLOv3, YOLOv4, and YOLOv5), as recent studies have used these models to detect brain tumors [19], [20], [21]. Using medical image detection, These YOLO models showed an accuracy in the range of 0.80 to 0.98. In this study, our novelty tries to focus on the detection task of the hippocampal region that was initially valued. This task may be used to analyze the hippocampal area accurately to ascertain the initial alterations in the three views and categories.

As delineated in the preceding statements, MRI images contain three views; prior investigations solely utilized only one view of MRI images for hippocampal region detection [22]. Moreover, the hippocampal region is a complex area of several distinct structures, most of which are small and have boundaries that are difficult to visualize in only one view [23], [24]. However, this could result in occlusion or ambiguous delineation of the hippocampal regions. However, using three views ensures that the YOLO model better detects the hippocampal region. Additionally, it is important in medical applications, where precise delineation of structures can affect diagnoses, surgical planning, and treatment outcomes in the future. To our knowledge, no prior investigation has explored the use of three views of MRI images to detect the hippocampal region for better AD diagnosis and interpretability. Our study makes the following contributions that are summarized below:

- The three views of the MRI images may be used to obtain more interpretability of the hippocampal in diagnosing AD.
- Three categories of AD may help differentiate the hippocampal area for the disease progression.
- We used three YOLO models to determine which model is suitable for detecting the hippocampal region.

We hypothesized that YOLO models could automatically detect the hippocampal region and save neurologists time diagnosing AD. It may improve the ability to visualize and study the hippocampal, owing to its intricate anatomy and variability among the three views and categories. In addition, we compared the performance of the three YOLO models in detecting the hippocampal region to determine which YOLO model shows higher accuracy and is suitable for detecting the hippocampal region. Thus, the results for the three views of MRI images may indicate which view shows higher accuracy in detecting the hippocampal region in diagnosing AD.

II. MATERIAL AND METHODS

A. DATA COLLECTION

Data used in this study were obtained from the publicly available Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu) during the first phase (ADNI−1). We used the baseline ADNI−1 database from a 1.5T Tesla scanner, preprocessed with Magnetization

Prepared Rapid Gradient Echo (MP-RAGE) with a resolution of 256 × 256 × 170 voxels. ADNI−1 holds a preeminent position among researchers investigating AD due to its widespread usage and popularity [10], [25]. ADNI−1 provides three categories: AD, MCI, and NC. These categories were used to compare which category had higher accuracy. A total of 220 subjects were divided into three classes (AD (75), MCI (72), and NC (73)), and each class included 250 AD, 250 MCI, and 250 NC, with 750 with raw ADNI imaging data (.nii extension). Similar to the previous study that used more than 500 datasets and the raw ADNI imaging data [22], [26]. Raw imaging data were used, and specific details are available in the ADNI database [27].

In total, 2,250 raw ADNI imaging data were used. We then split the raw ADNI imaging data into an 80:20 ratio for training and validation. We used 1,800 raw ADNI imaging data for the training set, and each category and view had 200 raw ADNI imaging data. Next, 450 raw ADNI imaging data were used for the validation set; each category and view had 50 raw ADNI imaging data. In addition, one raw ADNI imaging data in each view and category has 160–170 MRI image slices [28], [29]. According to previous studies, using approximately five slices based on image atrophy can reduce computational time and achieve higher accuracy [30], [31]. Thus, the preprocessing data incorporates a select slices method for the hippocampal region detection. We then selected five slices with clearer images of the hippocampal region in three views (i.e., axial, coronal, and sagittal) of the three categories (i.e., AD, MCI, and NC). The slicing method ensures that the model is trained on the most relevant region information, which helps achieve better performance [31]. However, the annotation for training deep learning is challenging, and interpreting decisions play an important role, especially in MRI images, which may improve the interpretability of the model by focusing on clinically significant regions. Finally, we obtained, in total, 11,250 MRI slices from the three views of the three categories. Detailed information on the separated datasets is presented in **Table 1**.

TABLE 1. The balanced dataset in each view and category for training and validation.

MRI views	Category	Raw ADNI			
		Training		Validation	
		80% (imaging data)	Select five slices (image)	20% (imaging data)	Select five slices (image)
Axial	AD	200	1,000	50	250
	MCI	200	1,000	50	250
	NC	200	1,000	50	250
Coronal	AD	200	1,000	50	250
	MCI	200	1,000	50	250
	NC	200	1,000	50	250
Sagittal	AD	200	1,000	50	250
	MCI	200	1,000	50	250
	NC	200	1,000	50	250
Summary		1,800	9,000	450	2,250

Note: AD, Alzheimer's Disease; MCI, Mild Cognitive Impairment; NC, Normal Controls.

B. HIPPOCAMPAL REGION PROCEDURE

In the following step, we labeled the five slices using the labelImg software (https://github.com/tzutalin/labelImg). The procedure for selecting five slices of MRI images in three views and categories was used in our previous studies [31]. We used labelImg to label the landmarks of the MRI images, namely the hippocampal region, in three views and categories. The labeling used a bounding box and annotations. We used the same bounding box size for each view and category based on the hippocampal region. For this reason, the hippocampal region covers the hippocampal and other nearby structures, such as the amygdala and parahippocampal. In addition, landmarks in MRI images can be used to identify significant regions among millions of voxels [32]. Besides, using the landmarks of MRI images as an input image can simplify the network structure and facilitate training performance [33]. An example of the bounding box size used to label the hippocampal region in the sagittal view, including the marking sign of the hippocampal, is shown in **Figure 1**.

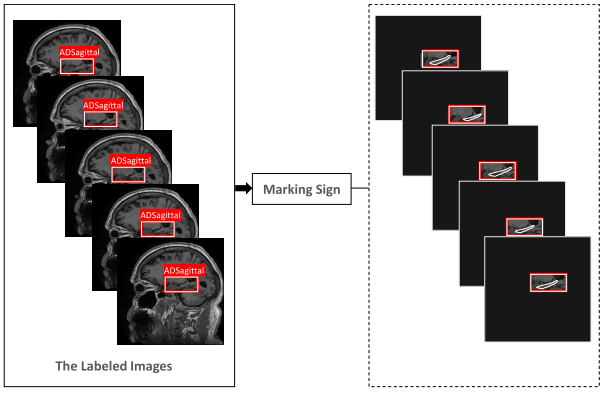

FIGURE 1. The bounding box size of the hippocampal region and marking the sign of the hippocampal by the region in sagittal view. AD, Alzheimer's Disease.

Furthermore, in the axial view, we labeled the hippocampal region with the bounding box with the annotation in AD (ADLeft, ADRight), MCI (MCILeft, MCIRight), and NC (NCLeft, NCRight). In the coronal view, we labeled the hippocampal region with the bounding box with the annotation in AD (ADLeft, ADRight), MCI (MCILeft, MCIRight), and NC (NCLeft, NCRight). In the sagittal view, we labeled the hippocampal region with the bounding box with the annotation in AD (ADSagittal), MCI (MCISagittal), and NC (NCSagittal). Our labeled dataset was validated by a neurologist from the China Medical University Hospital, a senior with experience in neurology, to ensure that the slices were labeled correctly with the hippocampal region. The detailed labeling images of the three views and categories are shown in **Figure 2**.

According to **Figure 2**, the marking sign provided by the advanced YOLO models improved the ability to visualize and study hippocampal volume loss. Therefore, the hippocampal is a complicated anatomy that is challenging to identify [34]. Furthermore, it can be used for quantitative analysis, similar to previous studies which found that the apparent water exchange rate (AXR_{BBB}) is significantly correlated with cognitive dysfunction and increases in the hippocampal [35].

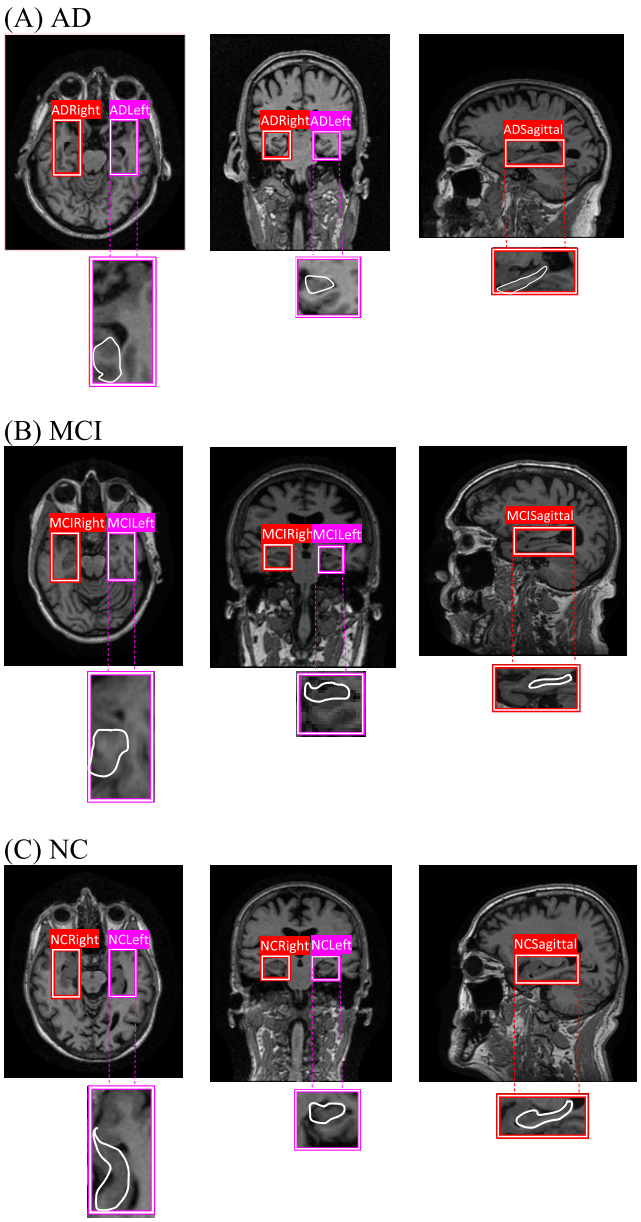

FIGURE 2. The labeled images in the hippocampal region of MRI images in axial, coronal, and sagittal views: (A) AD; (B) MCI; and (C) NC. AD, Alzheimer's Diseas; MCI, Mild Cognitive Impairment; and NC, Normal Control.

C. YOLO ARCHITECTURE

Deep learning is more effective in many research areas that associate hippocampal associated with AD [12]. You Only Look Once (YOLO), previously used in medical images, such as YOLOv3 [15], YOLOv4 [16], and YOLOv5 [17].

Finally, we used the labeled images from the proposed YOLO models (i.e., YOLOv3, YOLOv4, and YOLOv5). YOLO models have become popular owing to their high speed and accuracy, and this model detects bounding boxes from image pixels. The YOLO algorithm determines the bounding boxes of images [36]. YOLO is a state-of-the-art deep-learning framework for real-time object recognition [37]. The architecture employs 24 convolutional layers to extract image features and two fully connected layers for bounding box detection. The network was constructed using the Darknet framework [38].

This study used Darknet53 as the backbone to extract features from input images. Darknet53, a convolutional neural network (CNN), is our deep learning model's foundational architecture for feature extraction. The backbone of a deep neural network comprises a series of convolutional layers designed to identify and capture essential features from the input data. These layers perform critical functions such as detecting edges, textures, shapes, and other relevant patterns within the images, enabling robust feature extraction that underpins the subsequent stages of object detection and segmentation. It uses a feature pyramid network (FPN) as a neck [39]. YOLOv4, as a modified version of YOLOv3, used Cross Stage Partial Network (CSPNet) in Darknet, creating a new feature extractor backbone called CSPDarknet53. The convolution architecture was based on a modified DenseNet [40]. The image is fed into CSPDarknet53 for feature extraction. The neck component of the network introduces additional layers between the backbone and the dense prediction head. This section comprises a Spatial Pyramid Pooling (SPP) module and a Path Aggregation Network (PAN). The SPP module combines the max-pooling outputs from the low-resolution feature maps to identify the most representative features. This configuration enhances the model's ability to capture critical information, thereby improving object detection accuracy. This two-stage process enhances the model's ability to identify and merge pertinent features from the input image accurately. However, YOLOv5 differs significantly from its predecessors. This architecture leverages the strengths of CSPDarknet53 to enhance feature extraction, ensuring efficient processing and accurate detection. The integration of CSPDarknet53 within YOLOv5 provides a robust foundation that supports the model's superior performance in object detection tasks. Building upon the foundation laid by YOLOv4, YOLOv5 incorporates an adaptive anchor strategy and utilizes a refined architecture featuring a CSP backbone. This combination allows for more efficient processing and improved accuracy in object detection tasks [41]. The YOLOv5 network uses PANet as the neck to enhance the information flow, thereby improving the localization capabilities in the lower layers of the network. This enhancement in information flow significantly contributes to the accuracy of object localization, as it allows the model to better capture and integrate fine-grained details from various feature maps, leading to more precise detection and classification of objects within the images. Each detection head consists of convolutional layers, followed by two fully connected layers that output the final detections.

The model was implemented using Windows 10 with Python 3.7.6 on a machine with the following specifications: Core i7-11700 CPU, 32 GB RAM, and an NVIDIA GeForce RTX 3090 GPU with 24 GB of GDDR6X memory. The workflow process for detecting the hippocampal region in MRI images is shown in **Figure 3.**

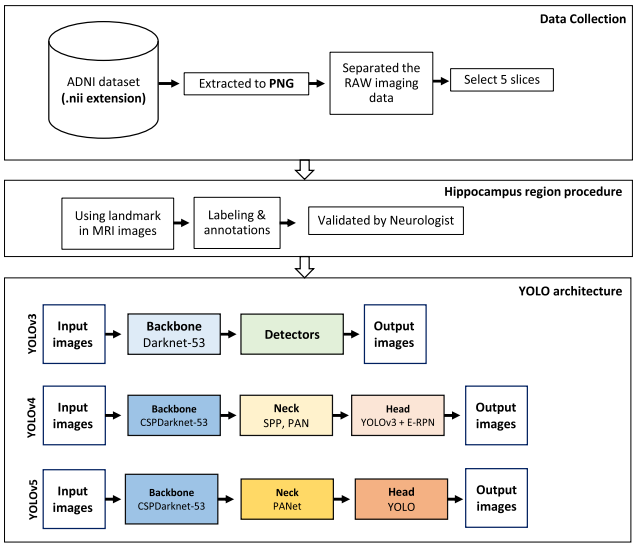

FIGURE 3. The workflow process for detecting the hippocampal region in MRI images. YOLO, You only look once; SPP, Spatial Pyramid Pooling; PAN, Path Aggregation Networ; E-RPN, Euler-Region Proposal Network; CSP, and Cross-Stage-Partial.

III. RESULTS

This study compared the performance of the YOLOv3, YOLOv4, and YOLOv5 models in detecting the hippocampal region in MRI images. In addition, we compared the accuracy of the three views based on the results, which demonstrated higher accuracy. In addition, comparisons between the three YOLO models were used to determine which model showed the highest accuracy in detecting the hippocampal region. The average mean Average Precision (mAP) value performance accuracy for YOLOv5 was 0.96, which was higher than those of the other YOLO models. While YOLOv3 showed 0.87 and YOLOv4 were 0.85 accuracies.

In addition, comparisons between the three views were used to determine which view showed the highest accuracy in detecting the hippocampal region using the three YOLO models. According to Table 2, the sagittal view showed 0.95 accuracies, higher than the axial view (0.93) and coronal views (0.81). The axial showed 0.93, and the coronal view showed 0.81 accuracies. In addition, we make comparisons in three categories among three views and categories.

TABLE 2. The mAP performance of three YOLO models in three views of MRI images.

Model	MRI view			Average mAP value
	Axial (mAP)	Coronal (mAP)	Sagittal (mAP)	
YOLOv3	0.96	0.75	0.91	0.87
YOLOv4	0.83	0.75	0.96	0.85
YOLOv5	0.99	0.92	0.98	**0.96**
Average mAP value	0.93	0.81	**0.95**	--

Note: mAP, mean Average Precision; YOLO, You Only Look Once.

Additionally, as shown in Table 3, we found that MCI in the coronal view was lower among the three YOLO models. YOLOv3 showed 0.50, YOLOv4 showed 0.56, and YOLOv5 showed 0.78 accuracies. Moreover, in YOLOv3, the axial view was higher in NC (0.97), the coronal view was higher in AD (0.90), and the sagittal view was higher in MCI (0.97). Then, in YOLOv4, the axial view was higher in AD (0.88), the coronal view was higher in NC (0.92), and the sagittal view was higher in MCI (0.99). In YOLOv5, the axial view was higher among the three categories (0.99). The coronal view was higher in NC (0.99). The sagittal view was higher in AD and MCI (0.99).

TABLE 3. The mAP performance of three YOLO models in three views and three categories of MRI images.

Model	MRI view	mAP		
		AD	MCI	NC
YOLOv3	Axial	0.96	0.96	0.97
	Coronal	0.90	**0.50**	0.84
	Sagittal	0.87	0.97	0.89
YOLOv4	Axial	0.88	0.80	0.80
	Coronal	0.78	**0.56**	0.92
	Sagittal	0.96	0.99	0.92
YOLOv5	Axial	0.99	0.99	0.99
	Coronal	0.98	**0.78**	0.99
	Sagittal	0.99	0.99	0.97

Note: mAP, mean Average Precision; YOLO, You Only Look Once; MRI, Magnetic Resonance Imaging; AD, Alzheimer's Disease; MCI, Mild Cognitive Impairment; NC, Normal Control.

IV. DISCUSSION

This research applied three YOLO models (i.e., YOLOv3, YOLOv4, and YOLOv5) to detect the hippocampal region in three views (i.e., axial, coronal, and sagittal) and categories (Alzheimer's disease (AD), Mild Cognitive Impairment (MCI), and Normal Control (NC)) of MRI images. This finding may support our hypothesis that using three YOLO models could automatically detect the hippocampal region and help medical experts save time in diagnosing AD. The average mAP accuracy of YOLOv5 was higher than those of YOLOv3 and YOLOv4. We also found that the sagittal view had a higher average mAP accuracy than the axial and coronal views. In addition, we found that MCI in the coronal view was lower among the three YOLO models.

Our study proposed three YOLO models: YOLOv3, YOLOv4, and YOLOv5. These three YOLO models are widely used for object detection in MRI images [42], [43]. The performance of YOLO models used the mean average precision (mAP), the current benchmark metric used by the computer vision research community, to evaluate the robustness and accuracy of object detection [44]. The average mAP for the YOLOv5 model performed better than YOLOv3 and YOLOv4 in detecting the hippocampal region in the MRI images. Therefore, YOLOv5 was developed in the Ultralytics PyTorch framework, which makes inferences faster than other YOLO models because it has a smaller structure [45]. Another advantage of YOLOv5 is mosaic augmentation in training, which combines four images into four blocks of random proportion. Mosaic augmentation is beneficial for object detection, helping the model to learn to detect objects [46]. Thus, we may say that YOLOv5 demonstrates the feasibility

and effectiveness of detecting the hippocampal region with a higher average mAP performance.

Our results were similar to those of Arunachalam et al. in that the YOLOv5 model has the benefit of finding the hippocampal region quickly, even in noisy, blurry, and foggy images [17]. Additionally, previous studies found that using YOLOv5 achieved higher accuracies than YOLOv3 and YOLOv4 using MRI images; for instance, Chen et al. in detecting stroke lesions [43], Arunachalam and Sethumathavan detect benign and malignant tumors [17]. Thus, we may say that the YOLOv5 model may be capable of detecting the hippocampal region in MRI images. In the future, it may be possible to analyze medical images immediately.

This study used three views of MRI images to detect the hippocampal regions. Considering that one view of an MRI image leads to the loss of 3D information, using three views of MRI images can obtain details from 2D images [47]. Thus, we may assume that three views of MRI images are used to detect the hippocampal region and may facilitate the interpretation of the diagnosis of AD. As shown in **Table 2,** the sagittal view showed higher accuracy than the axial and coronal views. The sagittal view shows clearer information regarding the hippocampal and one of the source sites for AD tangles and senile plaques, which is valuable for diagnosing AD [48].

Therefore, in a study by Cohen et al., the hippocampal volume determinations were based on manual outlining of the sagittal view aided by axial and coronal views [49]. Further, the sagittal view could expect to detect the hippocampal region to see the volume changes with additional information from the axial and coronal views. Thus, our funding may be used for measuring the hippocampal volume changes through the bounding box in the sagittal view and with additional information regarding the hippocampal region through the axial and coronal views.

The illustration simulations to analyze the hippocampal volume changes through MRI consist of three phases. In the first phase, a patient with a cognitive problem will undergo an MRI, and the results from MRI scans are 3D MRI images, which are then converted to 2D images for further analysis [50]. In the second phase, the patient visits a neurologist to check the MRI images. Then, to save time, neurologists may use YOLOv5 to detect the hippocampal region, as in previous studies [51]. In the last phase, the neurologist will use the hippocampal region to analyze the hippocampal volume changes [52]. Thus, our study can be an automatic tracing tool that is useful in diagnosing AD. An illustration of hippocampal volume changes in the MRI images is shown in **Figure 4**.

In addition, we found that MCI in the coronal view was lower among the three YOLO models. For this reason, the bounding box size of the hippocampal region in the coronal view was smaller than in the axial and sagittal views. Moreover, the hippocampal region in the coronal view is unreliable for observing other brain regions affecting AD,

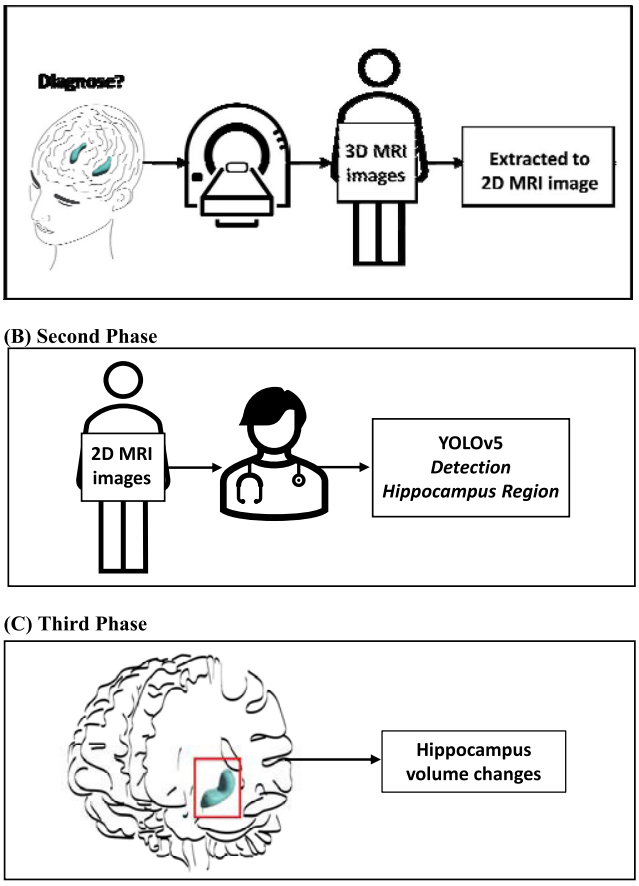

FIGURE 4. The illustration analyzes the hippocampal volume changes in MRI images; (A) MRI scans converted to 2D image; (B) YOLOv5 to detect the hippocampal region; and (C) Analyze the hippocampal volume changes. 2D, Two Dimensional; 3D, Three Dimensiona; YOLO, You Only Look Once; MRI, Magnetic Resonance Imaging.

such as the hippocampal, amygdala, and other regions within the landmark. This may have affected the detection result. Furthermore, we found that the MCI accuracy was lower than that of AD and NC among the three views and categories. MCI has a high probability of misdiagnosing AD, and the structural changes in MCI are relatively subtle [53]. Thus, these studies may provide evidence that MCI in the coronal view is more challenging in detecting the hippocampal region in MRI images to diagnose AD.

The current study had some limitations. First, we limited our study to detect the landmarks in MRI images, namely the hippocampal region and did not apply the multiple-class classification. Further studies may detect the biomarker of MRI images, allowing for a less invasive and more accurate AD diagnosis [54]. Biomarkers or biological markers refer to a broad subcategory of medical signs [55]. For instance, hippocampal volume has become the best-established imaging biomarker for AD diagnosis [56]. However, previous studies have shown the use of multiple class classification to distinguish among various stages for the early diagnosis of AD [57]. In further studies, we may use the multiple-class classification in detecting the

hippocampal region. Second, we used only object detection to detect the hippocampal region in three views of the MRI images. Further, we may use instance segmentation based on the hippocampal region labeled to segment AD biomarkers, such as hippocampal volume changes. Instance segmentation combines these classical computer vision tasks (detection and semantic segmentation), such as Mask R-CNN [58]. The benefit of using instance segmentation is that it saves time for delineation and reduces reproducibility, which is called into question [59]. Third, we may use various data augmentation methods to increase accuracy, similar to a recent study [60].

V. CONCLUSION

In conclusion, we demonstrated three YOLO models (i.e., YOLOv3, YOLOv4, and YOLOv5) and three views (i.e., axial, coronal, and sagittal) of MRI images to detect the hippocampal region. This study supports our hypothesis that using YOLO models would automatically detect the hippocampal region in MRI images and help neurologists save time. We found that YOLOv5 is a suitable model for detecting the hippocampal region. At the same time, we found that the sagittal view showed higher average mAP accuracy than the axial and coronal view among the three models. In addition, we found that MCI in the coronal view was lower among the three models. In conclusion, our study findings demonstrate the importance of detecting the hippocampal region to diagnose AD and accurately analyze changes in hippocampal volume loss using the marking sign. The YOLOv5 model substantially affects the performance metrics and interpretability across the MRI views and the three categories.

DECLARATION OF COMPETING INTEREST

All authors declare that there are no conflicts of interest regarding the publication of this article.

ACKNOWLEDGMENT

The funding agency did not have any involvement in data collection, data analysis, and data interpretation.

The authors wish to express gratitude to Ardha Ardea Prisilla, Fahni Harris, and Gilang Titah Ramadhan for their assistance. This article uses data from the publicly available Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). Thus, the ADNI investigators contributed to the design, implementation, and data collection but not to the analysis and writing of this article. A complete listing of ADNI investigators can be found at: https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

REFERENCES

[1] A. Gustavsson, N. Norton, T. Fast, L. Frölich, J. Georges, D. Holzapfel, T. Kirabali, P. Krolak-Salmon, P. M. Rossini, M. T. Ferretti, L. Lanman, A. S. Chadha, and W. M. van der Flier, "Global estimates on the number of persons across the Alzheimer's disease continuum," *Alzheimer's Dementia*, vol. 19, no. 2, pp. 658–670, Feb. 2023.

[2] K. Clifford, M. Moreno, and C. M. Kloske, "Navigating late-stage dementia: A perspective from the Alzheimer's association," *Alzheimer's Dementia, Diagnosis, Assessment Disease Monit.*, vol. 16, no. 1, Jan. 2024, Art. no. e12530.

[3] J. A. Hey, S. Abushakra, K. Blennow, E. M. Reiman, J. Hort, N. D. Prins, K. Sheardova, P. Kesslak, L. Shen, X. Zhu, A. Albayrak, J. Paul, J. F. Schaefer, A. Power, and M. Tolar, "Effects of oral ALZ-801/valiltramiprosate on plasma biomarkers, brain hippocampal volume, and cognition: Results of 2-year single-arm, open-label, phase 2 trial in APOE4 carriers with early Alzheimer's disease," *Drugs*, vol. 84, pp. 1–13, Jun. 2024.

[4] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, and P. M. Thompson, "The clinical use of structural MRI in Alzheimer disease," *Nature Rev. Neurol.*, vol. 6, no. 2, pp. 67–77, Feb. 2010.

[5] H. A. Pahlavani, "Exercise therapy to prevent and treat Alzheimer's disease," *Frontiers Aging Neurosci.*, vol. 15, Aug. 2023, Art. no. 1243869.

[6] R. A. Hazarika, A. K. Maji, D. Kandar, P. Chakrabarti, T. Chakrabarti, K. S. J. Rao, J. Carvalho, B. Kateb, and M. Nami, "An evaluation on changes in hippocampus size for cognitively normal, mild cognitive impairment, and Alzheimer's disease patients using fuzzy membership function," *Innov. Syst. Softw. Eng.*, pp. 1–15, Jul. 2023.

[7] A. J. McDonald and D. D. Mott, "Functional neuroanatomy of amygdalohippocampal interconnections and their role in learning and memory," *J. Neurosci. Res.*, vol. 95, no. 3, pp. 797–820, Mar. 2017.

[8] M. Odusami, R. Maskeliūnas, and R. Damaševičius, "An intelligent system for early recognition of Alzheimer's disease using neuroimaging," *Sensors*, vol. 22, no. 3, p. 740, Jan. 2022.

[9] M. T. Vlaardingerbroek and J. A. Boer, *Magnetic Resonance Imaging: Theory and Practice*. Berlin, Germany: Springer, 2013.

[10] A. Zandifar, V. S. Fonov, J. C. Pruessner, and D. L. Collins, "The EADC-ADNI harmonized protocol for hippocampal segmentation: A validation study," *NeuroImage*, vol. 181, pp. 142–148, Nov. 2018.

[11] X. Pan, T.-L. Phan, M. Adel, C. Fossati, T. Gaidon, J. Wojak, and E. Guedj, "Multi-view separable pyramid network for AD prediction at MCI stage by 18F-FDG brain PET imaging," *IEEE Trans. Med. Imag.*, vol. 40, no. 1, pp. 81–92, Jan. 2021.

[12] I. Sarasua, S. Pölsterl, and C. Wachinger, "Hippocampal representations for deep learning on Alzheimer's disease," *Sci. Rep.*, vol. 12, no. 1, p. 8619, May 2022.

[13] M. Mujahid, A. Rehman, T. Alam, F. S. Alamri, S. M. Fati, and T. Saba, "An efficient ensemble approach for Alzheimer's disease detection using an adaptive synthetic technique and deep learning," *Diagnostics*, vol. 13, no. 15, p. 2489, Jul. 2023.

[14] A. Dixit and P. Singh, "Brain tumor detection using fine-tuned YOLO model with transfer learning," in *Proc. Int. Symp. Artif. Intell. Med. Data*, 2023, pp. 363–371.

[15] M. M. Chanu, N. H. Singh, C. Muppala, R. T. Prabu, N. P. Singh, and K. Thongam, "Computer-aided detection and classification of brain tumor using YOLOv3 and deep learning," *Soft Comput.*, vol. 27, no. 14, pp. 9927–9940, Jul. 2023.

[16] J. I. Zena, E. Lucky, C. G. Ellaine, I. S. Edbert, and D. Suhartono, "Deep learning approach based classification of Alzheimer's disease using brain MRI," in *Proc. 5th Int. Seminar Res. Inf. Technol. Intell. Syst. (ISRITI)*, Dec. 2022, pp. 397–402.

[17] S. Arunachalam and G. Sethumathavan, "An effective tumor detection in MR brain images based on deep CNN approach: I-YOLOv5," *Appl. Artif. Intell.*, vol. 36, no. 1, Dec. 2022, Art. no. 2151180.

[18] E. Casas, L. Ramos, E. Bendek, and F. Rivas-Echeverría, "Assessing the effectiveness of YOLO architectures for smoke and wildfire detection," *IEEE Access*, vol. 11, pp. 96554–96583, 2023.

[19] M. Rahimi, M. Mostafavi, and A. Arabameri, "Automatic detection of brain tumor on MRI images using a YOLO-based algorithm," in *Proc. 13th Iranian/3rd Int. Mach. Vis. Image Process. Conf. (MVIP)*, Mar. 2024, pp. 1–5.

[20] B. J. D. Kalyani, K. Meena, E. Murali, L. Jayakumar, and D. Saravanan, "Analysis of MRI brain tumor images using deep learning techniques," *Soft Comput.*, vol. 27, no. 11, pp. 7535–7542, Jun. 2023.

[21] G. J. Jagtap, P. Peddi, and Y. D. Sinkar, "Application of YOLOv5 for brain tumor detection from MRI," in *Proc. Int. Conf. Emerg. Smart Comput. Informat. (ESCI)*, Mar. 2024, pp. 1–5.

[22] S. S. Abd-Aljabar, N. M. Basheer, and O. I. Alsaif, "Alzheimer's diseases classification using YOLOv2 object detection technique," *Int. J. Reconfigurable Embedded Syst.*, vol. 2089, no. 4864, p. 4864, 4864.

[23] K. Aderghal, J. Benois-Pineau, K. Afdel, and C. Gwenaëlle, "FuseMe: Classification of sMRI images by fusion of deep CNNs in 2D+ε projections," in *Proc. 15th Int. Workshop Content-Based Multimedia Indexing*, Jun. 2017, pp. 1–7.

[24] G. Chetelat and J.-C. Baron, "Early diagnosis of Alzheimer's disease: Contribution of structural neuroimaging," *NeuroImage*, vol. 18, no. 2, pp. 525–541, Feb. 2003.

[25] L. Cao, L. Li, J. Zheng, X. Fan, F. Yin, H. Shen, and J. Zhang, "Multi-task neural networks for joint hippocampus segmentation and clinical score regression," *Multimedia Tools Appl.*, vol. 77, no. 22, pp. 29669–29686, Nov. 2018.

[26] S. Liu, S. Liu, W. Cai, H. Che, S. Pujol, R. Kikinis, D. Feng, and M. J. Fulham, "Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, pp. 1132–1140, Apr. 2015.

[27] L. E. Gibbons, A. C. Carle, R. S. Mackin, D. Harvey, S. Mukherjee, P. Insel, S. M. Curtis, D. Mungas, and P. K. Crane, "A composite score for executive functioning, validated in Alzheimer's disease neuroimaging initiative (ADNI) participants with baseline mild cognitive impairment," *Brain Imag. Behav.*, vol. 6, no. 4, pp. 517–527, Dec. 2012.

[28] S. Tomassini, A. Sbrollini, G. Covella, P. Sernani, N. Falcionelli, H. Müller, M. Morettini, L. Burattini, and A. F. Dragoni, "Brain-on-Cloud for automatic diagnosis of Alzheimer's disease from 3D structural magnetic resonance whole-brain scans," *Comput. Methods Programs Biomed.*, vol. 227, Dec. 2022, Art. no. 107191.

[29] K. V. Kumari and S. S. Barpanda, "Investigating the volumetric biomarkers of Alzheimer disease: A state-of-the art review," in *Proc. ICADCML*. Singapore: Springer, 2022, pp. 581–591.

[30] N. M. Khan, N. Abraham, and M. Hon, "Transfer learning with intelligent training data selection for prediction of Alzheimer's disease," *IEEE Access*, vol. 7, pp. 72726–72735, 2019.

[31] Y. Pusparani, C.-Y. Lin, Y.-K. Jan, F.-Y. Lin, B.-Y. Liau, P. Ardhianto, I. Farady, J. S. R. Alex, J. Aparajeeta, W.-H. Chao, and C.-W. Lung, "Diagnosis of Alzheimer's disease using convolutional neural network with select slices by landmark on hippocampus in MRI images," *IEEE Access*, vol. 11, pp. 61688–61697, 2023.

[32] J. Zhang, Y. Gao, Y. Gao, B. C. Munsell, and D. Shen, "Detecting anatomical landmarks for fast Alzheimer's disease diagnosis," *IEEE Trans. Med. Imag.*, vol. 35, no. 12, pp. 2524–2533, Dec. 2016.

[33] F. Li and M. Liu, "A hybrid convolutional and recurrent neural network for hippocampus analysis in Alzheimer's disease," *J. Neurosci. Methods*, vol. 323, pp. 108–118, Jul. 2019.

[34] R. A. M. Haast, S. Kashyap, D. Ivanov, M. D. Yousif, J. DeKraker, B. A. Poser, and A. R. Khan, "Insights into hippocampal perfusion using high-resolution, multi-modal 7T MRI," *Proc. Nat. Acad. Sci. USA*, vol. 121, no. 11, Mar. 2024, Art. no. e2310044121.

[35] Y. Zhang, Y. Wang, Z. Li, Z. Wang, J. Cheng, X. Bai, Y.-C. Hsu, Y. Sun, S. Li, J. Shi, B. Sui, and R. Bai, "Vascular-water-exchange MRI (VEXI) enables the detection of subtle AXR alterations in Alzheimer's disease without MRI contrast agent, which may relate to BBB integrity," *NeuroImage*, vol. 270, Apr. 2023, Art. no. 119951.

[36] M. Mushtaq, M. U. Akram, N. S. Alghamdi, J. Fatima, and R. F. Masood, "Localization and edge-based segmentation of lumbar spine vertebrae to identify the deformities using deep learning models," *Sensors*, vol. 22, no. 4, p. 1547, Feb. 2022.

[37] G. Li, Z. Song, and Q. Fu, "A new method of image detection for small datasets under the framework of YOLO network," in *Proc. IEEE 3rd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Oct. 2018, pp. 1031–1035.

[38] M. Hammami, D. Friboulet, and R. Kechichian, "Cycle GAN-based data augmentation for multi-organ detection in CT images via YOLO," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2020, pp. 390–393.

[39] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 936–944.

[40] F. Iandola, M. Moskewicz, S. Karayev, R. Girshick, T. Darrell, and K. Keutzer, "DenseNet: Implementing efficient ConvNet descriptor pyramids," 2014, *arXiv:1404.1869*.

[41] G. Jocher, K. Nishimura, T. Mineeva, and R. Vilariño, *YOLOv5 Code Repository*, 2020.

[42] R. Mirchandani, C. Yoon, S. Prakash, A. Khaire, A. Naran, A. Nair, and S. Ganti, "Comparing the architecture and performance of AlexNet faster R-CNN and YOLOv4 in the multiclass classification of Alzheimer brain MRI scans," 2021.

[43] S. Chen, J. Duan, H. Wang, R. Wang, J. Li, M. Qi, Y. Duan, and S. Qi, "Automatic detection of stroke lesion from diffusion-weighted imaging via the improved YOLOv5," *Comput. Biol. Med.*, vol. 150, Nov. 2022, Art. no. 106120.

[44] A. Bhattacharya and S. G. Cloutier, "End-to-end deep learning framework for printed circuit board manufacturing defect classification," *Sci. Rep.*, vol. 12, no. 1, p. 12559, Jul. 2022.

[45] K. S. Chahal and K. Dey, "A survey of modern object detection literature using deep learning," 2018, *arXiv:1808.07256*.

[46] M. Rocha, M. Claro, L. Neto, K. Aires, V. Machado, and R. Veras, "Malaria parasites detection and identification using object detectors based on deep neural networks: A wide comparative analysis," *Comput. Methods Biomech. Biomed. Eng., Imag. Vis.*, vol. 11, no. 3, pp. 351–368, May 2023.

[47] S. Dharwada, J. Tembhurne, and T. Diwan, "Multi-channel deep model for classification of Alzheimer's disease using transfer learning," in *Proc. 18th Int. Conf. Distrib. Comput. Intell. Technol. (ICDCIT)*, Bhubaneswar, India: Springer, Jan. 2022, pp. 245–259.

[48] R. Sharma, T. Goel, M. Tanveer, and R. Murugan, "FDN-ADNet: Fuzzy LS-TWSVM based deep learning network for prognosis of the Alzheimer's disease using the sagittal plane of MRI scans," *Appl. Soft Comput.*, vol. 115, Jan. 2022, Art. no. 108099.

[49] R. M. Cohen, C. Small, F. Lalonde, J. Friz, and T. Sunderland, "Effect of apolipoprotein E genotype on hippocampal volume loss in aging healthy women," *Neurology*, vol. 57, no. 12, pp. 2223–2228, Dec. 2001.

[50] S. Liu, B. Yang, Y. Wang, J. Tian, L. Yin, and W. Zheng, "2D/3D multimode medical image registration based on normalized cross-correlation," *Appl. Sci.*, vol. 12, no. 6, p. 2828, Mar. 2022.

[51] T. Shelatkar, D. Urvashi, M. Shorfuzzaman, A. Alsufyani, and K. Lakshmanna, "Diagnosis of brain tumor using light weight deep learning model with fine-tuning approach," *Comput. Math. Methods Med.*, vol. 2022, Jul. 2022, Art. no. 2858845.

[52] K. Kwak, M. Niethammer, K. S. Giovanello, M. Styner, and E. Dayan, "Differential role for hippocampal subfields in Alzheimer's disease progression revealed with deep learning," *Cerebral Cortex*, vol. 32, no. 3, pp. 467–478, Jan. 2022.

[53] W. Kang, L. Lin, B. Zhang, X. Shen, and S. Wu, "Multi-model and multi-slice ensemble learning architecture based on 2D convolutional neural networks for Alzheimer's disease diagnosis," *Comput. Biol. Med.*, vol. 136, Sep. 2021, Art. no. 104678.

[54] R. Craig-Schapiro, A. M. Fagan, and D. M. Holtzman, "Biomarkers of Alzheimer's disease," *Neurobiol. Disease*, vol. 35, no. 2, pp. 128–140, 2009.

[55] K. Strimbu and J. A. Tavel, "What are biomarkers?" *Current Opinion HIV AIDS*, vol. 5, no. 6, p. 463, 2010.

[56] S. J. Teipel, M. Grothe, S. Lista, N. Toschi, F. G. Garaci, and H. Hampel, "Relevance of magnetic resonance imaging for early detection and diagnosis of Alzheimer disease," *Med. Clinics North Amer.*, vol. 97, no. 3, pp. 399–424, May 2013.

[57] R. Khan, Z. H. Qaisar, A. Mehmood, G. Ali, T. Alkhalifah, F. Alturise, and L. Wang, "A practical multiclass classification network for the diagnosis of Alzheimer's disease," *Appl. Sci.*, vol. 12, no. 13, p. 6507, Jun. 2022.

[58] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2980–2988.

[59] M. Vania and D. Lee, "Intervertebral disc instance segmentation using a multistage optimization mask-RCNN (MOM-RCNN)," *J. Comput. Des. Eng.*, vol. 8, no. 4, pp. 1023–1036, Jun. 2021.

[60] A. Anaya-Isaza, L. Mera-Jiménez, L. Verdugo-Alejo, and L. Sarasti, "Optimizing MRI-based brain tumor classification and detection using AI: A comparative analysis of neural networks, transfer learning, data augmentation, and the cross-transformer network," *Eur. J. Radiol. Open*, vol. 10, Jan. 2023, Art. no. 100484.

YORI PUSPARANI is currently pursuing the Ph.D. degree in digital media design with Asia University. She is also an Assistant Professor and a full-time Lecturer with the Faculty of Communication and Creative Design, Budi Luhur University, Jakarta, Indonesia. She has established a reputation as a quick learner, an effective communicator, a strong leader, and a valuable team player. Possessing outstanding flexibility, she has demonstrated exceptional responsibility, commitment, and presentation skills while showcasing proficiency in computer applications, including Adobe Illustrator, Adobe Photoshop, and Adobe InDesign. Her research interests include artificial intelligence, deep learning, plantar pressure, Alzheimer's disease, and other medical imaging fields. She understands the challenges inherent in analyzing medical images and recognizes AI technology's potential to contribute significantly to scientific articles in this area.

CHIH-YANG LIN (Senior Member, IEEE) is currently with the Department of Mechanical Engineering, National Central University, Taoyuan, Taiwan. He has contributed to over 150 papers that have been featured in a wide range of international conferences and journals. His primary research interests include computer vision, machine learning, deep learning, image processing, big data analysis, and the design of surveillance systems. He has been recognized as an IET Fellow.

YIH-KUEN JAN is currently the Director and an Associate Professor with the Rehabilitation Engineering Research Laboratory, University of Illinois at Urbana–Champaign (UIUC), Urbana, IL, USA. He is committed to developing rehabilitation technologies for the prevention and management of secondary conditions in people with impaired mobility. He has extensively published his research work to improve the understanding of the effect of various pathophysiological conditions (spinal cord injury, aging, diabetes, and physical inactivity) on musculoskeletal injury. His research interests include soft tissue biomechanics and its role in the development of musculoskeletal injury and pressure injury.

FU-YU LIN received the B.Sc. degree in post-baccalaureate medicine from Kaohsiung Medical University, Taiwan, and the residency degree in neurology from National Taiwan University Hospital, Taiwan. She is currently a Physician with the Neurology Department, China Medical University Hospital, Taichung, Taiwan. Her primary research interests include electrophysiology and neuroimaging.

BEN-YI LIAU received the B.S. degree in agriculture machinery engineering from National Chung Hsing University, Taichung, Taiwan, the M.S. degree in biomedical engineering from National Yang-Ming University, Taipei, Taiwan, and the Ph.D. degree from the Graduate Institute of Electrical and Communications Engineering, Feng Chia University, Taichung, in 2008. He was an Associate Professor of biomedical engineering with Hungkuang University, Taichung. Since February 2023, he has been an Associate Professor of automatic control engineering with Feng Chia University. His current research interests include biomedical signal processing and biomechanics.

PETER ARDHIANTO received the B.Arts. degree from Sebelas Maret University, Indonesia, the M.Arts. degree from Indonesian Institute of the Arts Yogyakarta, and the Ph.D. degree in digital media design from the College of Creative Design, Asia University, Taichung, Taiwan, in 2022. He is currently an Assistant Professor with the Department of Visual Communication Design and the Director of the International Undergraduate Program, Soegijapranata Catholic University, Indonesia.

ELVIN NUR FURQON received the B.Eng. degree in computer engineering from Telkom University, Indonesia, in 2021. He is currently pursuing the master's degree with Yuan Ze University, Taiwan. He is also a member of the CVIT Laboratory, Yuan Ze University. His research interests include image processing, computer vision, and deep learning.

JOHN SAHAYA RANI ALEX (Senior Member, IEEE) received the B.S. and M.S. degrees in electronics engineering from Anna University, MIT Campus, India, and the Ph.D. degree in electronics engineering from Vellore Institute of Technology, Chennai, India. She is currently a Professor and Associate Dean (Research) with the Division of Healthcare Advancement, Innovation and Research, Vellore Institute of Technology. Her research interests include applications of biosignal (speech/EEG), image (MRI/Fundus) processing, deep learning applications for healthcare, and embedded system design for healthcare.

JEETASHREE APARAJEETA received the B.S. degree in electronics and telecommunication engineering from the Biju Patnaik University of Technology, Odisha, and the M.S. degree in telecom system engineering and the Ph.D. degree in biomedical image processing from Siksha 'O' Anusandhan University. She is currently an Assistant Professor with the School of Electronics Engineering, Vellore Institute of Technology. Her research interest includes biomedical image analysis.

CHI-WEN LUNG received the bachelor's degree in industrial design from Dayeh University, Taiwan, the master's degree in industrial design from Tatung University, and the Ph.D. degree in biomedical engineering from National Yang-Ming University, Taiwan. He is currently a Professor with the Department of Creative Product Design and the Department of Bioinformatics and Medical Engineering, Asia University, Taiwan. His research interests include soft tissue biomechanics and its role in the development of musculoskeletal injuries and pressure injuries. In addition, he developed advanced rehabilitation research instruments.