Deep Learning-Based Automated Segmentation of the Parcellated Corpus Callosum in Brain MRI

Suliman Mohamed Fati
College of Computer and Information Sciences, Prince Sultan University, Riyadh, Saudi Arabia
sgaber@psu.edu.sa (corresponding author)

Omaia Al-Omari
Department of Information Systems, College of Computer and Information Sciences, Prince Sultan University, Riyadh, Saudi Arabia
oalomari@psu.edu.sa

Received: 31 May 2025 | Revised: 1 June 2025 and 12 June 2025 | Accepted: 14 June 2025

 DOI: https://doi.org/10.48084/etasr.11783

ABSTRACT

This study proposed PCcS-RAU-Net, an enhanced deep learning model that integrates Residual Blocks and Attention Gates (AGs) into a U-Net architecture for segmenting the parcellated Corpus Callosum (Cc) into five anatomical regions: rostrum, genu, mid-body, isthmus, and splenium. The model was evaluated across three prominent datasets: ABIDE, OASIS, and Real Clinical Images (RCI), achieving 97.10%, 96.88%, and 97.11%, in terms of Dice Similarity Coefficients (DSCs), and corresponding Mean Intersection over Union (MIoU) scores of 94.43%, 93.89%, and 94.19% respectively. Additionally, this approach outperformed the classical U-Net, Residual U-Net, and Attention U-Net models, providing robustness and an effective general applicability concerning the different imaging protocols.

Keywords-corpus callosum segmentation; deep learning; U-Net; residual learning; attention mechanisms; brain MRI

I. INTRODUCTION

Cc, the largest white matter structure in human brain, is responsible for interhemispheric binding. Many neurological and psychiatric disorders, including Alzheimer's Disease (AD), Autism Spectrum Disorder (ASD), and psychosis have long been linked to the anatomical abnormalities of Cc [1, 2]. The segmentation of Cc into rostrum, genu, mid-body, isthmus, and splenium is critical for the detailed morphologic analysis and treatment planning of these disorders. The manual delineation of Cc is time-consuming and requires much effort. Except this, there is a possibility of interobserver variability in clinical setups, where there is an urgent need for time efficiency and interception comparison.

Traditional methods, such as Witelson's method [3], Hofer's scheme [4], and other geometrical rule-based techniques, depend on manual tracing and assumptions that may not be generalized for different populations or image conditions. Developments in deep learning, particularly Convolutional Neural Networks (CNNs), appear promising for automating the volumetric medical image segmentation tasks.

The U-Net architecture has become the new benchmark for biomedical image segmentation through its encoder–decoder structure and skip connections [5]. However, the standard U-Net architecture faces challenges when dealing with complex anatomical variability and fine-grained regional segmentation. To address these limits, researchers have explored enhancements, such as residual U-Net models, which facilitate deeper network training, and Attention U-Net models that focus on relevant spatial regions during segmentation [6, 7]. Nevertheless, achieving accurate and reliable segmentation of the parcellated Cc across diverse datasets remains a significant challenge due to variations in patient populations, disease progression, imaging protocols, and anatomical structures.

This study proposes the PCcS – RAU – Net model, a novel deep learning framework that integrates residual blocks and AGs into a U-Net backbone, in order to improve feature learning focusing on the relevant anatomical regions. Autism Brain Imaging Data Exchange (ABIDE), Open Access Series of Imaging Studies (OASIS), and RCI datasets were evaluated to demonstrate its robustness across the patient demographics and imaging settings. This approach aims to achieve high segmentation accuracy and reliability across datasets and minimize the manual intervention, thereby supporting research and the clinical diagnostic processes.

II. METHODOLOGY

In this study, an enhanced deep learning mechanism was utilized for the algorithmic segmentation of a Parcellated Corpus callosum (PCc) from midsagittal registered brain Magnetic Resonance Imaging (MRI) images. This method followed three primary stages:

- Dataset collection and preprocessing
- Model architecture design
- Model training, evaluation, and comparison of baseline models

The entire pipeline is illustrated in Figure 1, showing the workflow from inputting data to obtaining segmentation output and performance evaluation.

A. Dataset and Preprocessing

The proposed PCcS-RAU-Net model has been considered across three distinct datasets: the ABIDE dataset [8], the OASIS dataset [9], and a novel dataset comprising 3,250 RCI, collected in collaboration with Al-Taawon Medical Hospital in Yemen. This dataset covered a six-month period (January to June 2023) and included anonymized scans from patients diagnosed with various neurological disorders. All data were acquired under approved ethical protocols.

- The ABIDE dataset included T1-weighted midsagittal brain MRI images of individuals diagnosed with ASD as well as healthy controls [10].
- The OASIS dataset consisted of 3D T1-weighted MRI scans primarily focused on aging and AD. For this study, midsagittal slices were extracted from the 3D volumes, which served as an appropriate input for the segmentation model and allowed the comparison of performance across different demographics and pathologies [11].
- The RCI dataset, containing anonymized midsagittal T1-weighted MRI scans, was utilized to validate the model's clinical setting.

All MRI images were subjected to preprocessing before feeding them into training:

- Initially, each image was resampled and resized to an equal resolution of 256 x 256 pixels to ensure the standardization of their consistency across all the datasets.
- Pixel intensity values were normalized to a fixed range to excite a faster convergence during the training process [12].
- To increase the model robustness to the anatomical variations and scanner-driven aberrations, prominent data augmentation techniques, like random horizontal flip, slight rotation (± 10 degrees), and minor scaling operations were applied [13].

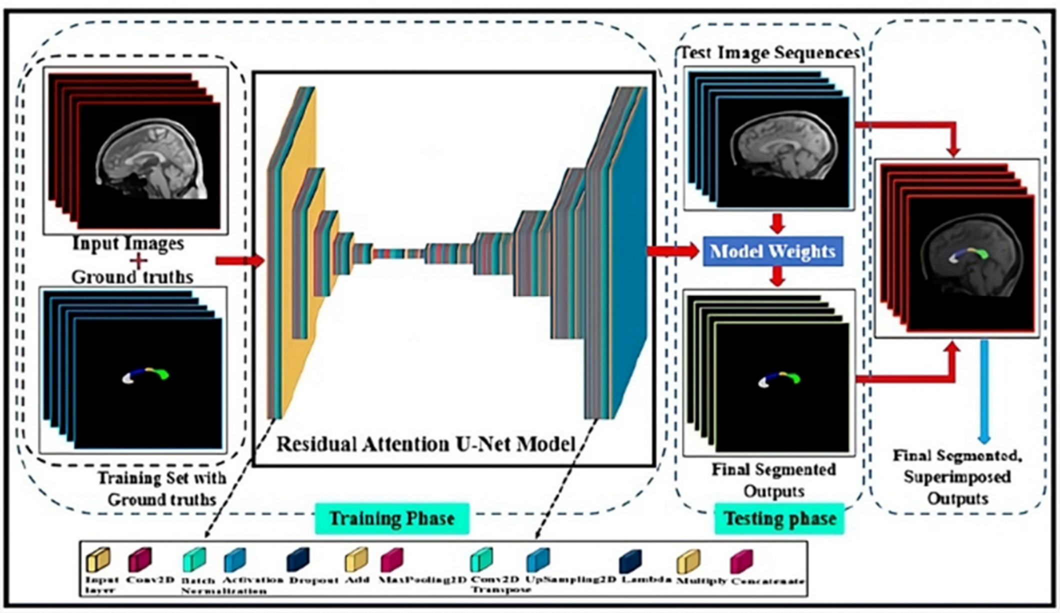

Fig. 1. Block diagram of PCcS-RAU-Net methodology.

B. Model Architecture and Feature Extraction

The proposed PCcS-RAU-Net architecture was based on the U-Net framework, specifically adapted for parcellation segmentation of Cc from midsagittal brain MRI images. U-Net's structure with synergy features -encoding by encoder paths and decoding by decoder paths along with a connection between two similar level- was effectively employed for medical image segmentation tasks [5].

To handle the variations of shape and anatomical diversity, residual blocks, which were inspired by the ResNet model introduced in [6], were added to either of the encoder or decoder paths and upon the practice of learning. These blocks enabled the network to learn identity mappings, applying improvements for training deeper networks, which could reduce the occurrence of gradient vanishing. Additionally, AGs were integrated with the sum skip connections, by attending critical regions in the feature maps and ignoring others irrelevant to the segmentation task, as proposed in [7]. This further hid the network's weakness in localizing and boosting the segmentation performance with minor structures as well as structures with notable variability, such as Cc subregions.

The PCcS-RAU-Net architecture consisted of a symmetric encoder-decoder structure. Specifically, the encoder captured contextual features through successive convolutional layers, residual connections, and down sampling operations. The bottleneck layer refined the feature representations before the decoder path progressively reconstructed the spatial resolution through up sampling, convolutional layers, and residual connections. Attention mechanisms were applied in the skip connections between corresponding encoder and decoder levels to emphasize critical anatomical regions of the Cc. Batch normalization layers were employed after each convolution operation to stabilize and accelerate the training process by normalizing the activations [12]. Activation functions, primarily Rectified Linear Unit (ReLU), introduced non-linearity into the model.

By integrating residual learning and attention mechanisms into the U-Net framework, PCcS-RAU-Net improved the feature representation, model convergence, and segmentation performance across different datasets with diverse anatomical and imaging characteristics. To better understand the methodological framework, Figure 1 presents a block diagram summarizing the main stages of the proposed method, including preprocessing, feature extraction through residual and attention mechanisms, and the segmentation process.

C. Training Strategy and Optimization

The proposed PCcS-RAU-Net model was trained using the Adam optimizer, which combined the advantages of both AdaGrad and RMSProp, due to its adaptive learning rate mechanism [14]. The initial learning rate was set to 10^{-4}, with a learning rate reduction strategy applied based on validation loss plateaus to ensure effective convergence. The model used the categorical cross-entropy loss function, suitable for multi-class segmentation tasks, where each pixel was assigned to one of the predefined classes (rostrum, genu, mid-body, isthmus, splenium, or background). The class imbalance was addressed by incorporating class weighting into the loss function to penalize underrepresented classes more heavily during training, enhancing the model's sensitivity to smaller regions of Cc.

Training was performed over 100 epochs with an early stopping mechanism monitoring the validation loss. If the validation loss did not improved for 10 consecutive epochs, the training was halted to prevent overfitting. A batch size of 16 was chosen, balancing the computational efficiency with the convergence stability. On-the-fly data augmentation was employed during training to enhance the variability while preventing overfitting. The augmentation techniques included horizontal flipping, small rotations, and slight zoom variations, all of which are standard practices to improve the model generalization when dealing with limited datasets [13]. The model weights were initialized using He normal initialization, which is particularly effective for networks utilizing ReLU activations [15]. During training, batch normalization was applied after each convolutional layer to stabilize the learning process by reducing the internal covariate shift [12]. When evaluating a trained model, the one with the best validation loss was selected for the final testing.

Advances in deep learning have contributed to medical image analysis, facial emotion recognition, sentiment classification, and bias-aware learning machine [16-19]. Surveys on the classification of brain tumors have also demonstrated that a great emphasis should be given to robust training strategies when dealing with tiny and imbalanced datasets [20].

III. RESULTS AND DISCUSSION

A. Model Performance Evaluation

The evaluation of the proposed PCcS-RAU-Net model was based on standard performance metrics, widely used in medical image segmentation tasks, including DSC, MIoU, Recall, and Precision [21].

$$\mathrm{DSC} = \frac{2\times|P\cap G|}{|P|+|G|} \quad (1)$$

$$\mathrm{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\mathrm{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\mathrm{MIoU} = \frac{1}{N}\sum_{i=1}^{N}\frac{|P_i\cap G_i|}{|P_i\cup G_i|} \quad (4)$$

where P and G denote the predicted and Ground Truth (GT) sets, respectively, while *TP, TN*, *FP*, and *FN* symbolize True Positives, True Negatives, False Positives, and False Negatives, respectively.

These metrics provided a comprehensive assessment of the segmentation quality and the model's ability to capture fine-grained anatomical structures accurately. The segmentation performance results for PCcS-RAU-Net on the ABIDE, OASIS, and RCI datasets are summarized in Table I.

TABLE I. SEGMENTATION PERFORMANCE OF PC_CS-RAU-NET ACROSS ABIDE, OASIS, AND RCI DATASETS

Dataset	DSC (%)	MIoU (%)	Recall (%)	Precision (%)
ABIDE	97.10	94.43	96.98	97.22
OASIS	96.88	93.89	96.72	97.05
RCI	97.11	94.19	97.00	97.23

To further validate the performance of the proposed PCcS-RAU-Net model, confusion matrices were computed for the ABIDE and OASIS datasets (Figure 2). Each matrix illustrated the distribution of *TP*, *TN*, *FP*, and *FN*.

From these results, the classification accuracy was calculated as:

$$\mathrm{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

yielding approximately 99.4% for both datasets.

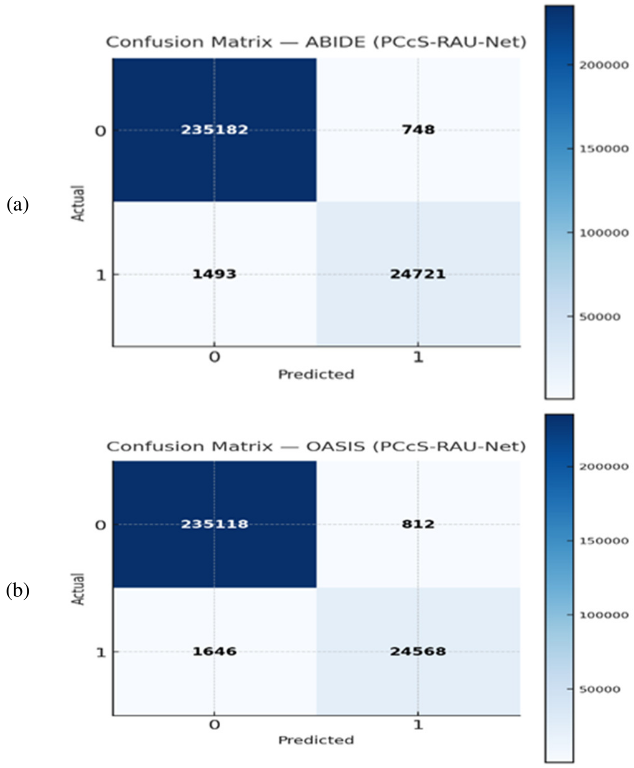

Fig. 2. Confusion matrices for the proposed PCcS-RAU-Net model: (a) ABIDE dataset, (b) OASIS dataset.

B. Report Analysis

The results of the segmentation derived by the proposed PCcS-RAU-Net model indicated optimal outcomes for the evaluated datasets. All the discussed metrics approached seamlessly above 93% (*DSC* and *MIoU*), pointing robustly toward the capture region of Cc subareas under anatomical variations along with the difference in imaging modalities. The Recall and Precision yields demonstrated a consistent model insight that reached 96%, finding almost all TPs.

Table II presents the comparison of PCcS-RAU-Net with other models (Standard U-Net, Residual U-Net, and Attention U-Net) on the ABIDE dataset. It is obvious that the proposed model exhibited the highest metrics values.

TABLE II. CLASSIFICATION REPORT FOR ALL MODELS

Model	DSC (%)	MIoU (%)	Recall (%)	Precision (%)
PCcS-RAU-Net	97.10	94.43	96.98	97.22
Residual U-Net	95.25	91.30	94.88	95.60
Attention U-Net	95.10	91.05	94.65	95.40
Standard U-Net	94.20	89.60	93.75	94.55

Figure 3 illustrates the sample's segmentation results across different subject groups (Healthy Controls, Autism, Asperger Syndrome, and Pervasive Developmental Disorder) from the ABIDE dataset, showing GT, predicted outputs, and their overlay along with DSC. The superior performance of PCcS-RAU-Net was attributed to the integration of residual blocks, which improved feature learning and network optimization, as well as AGs, which focused on relevant anatomical regions.

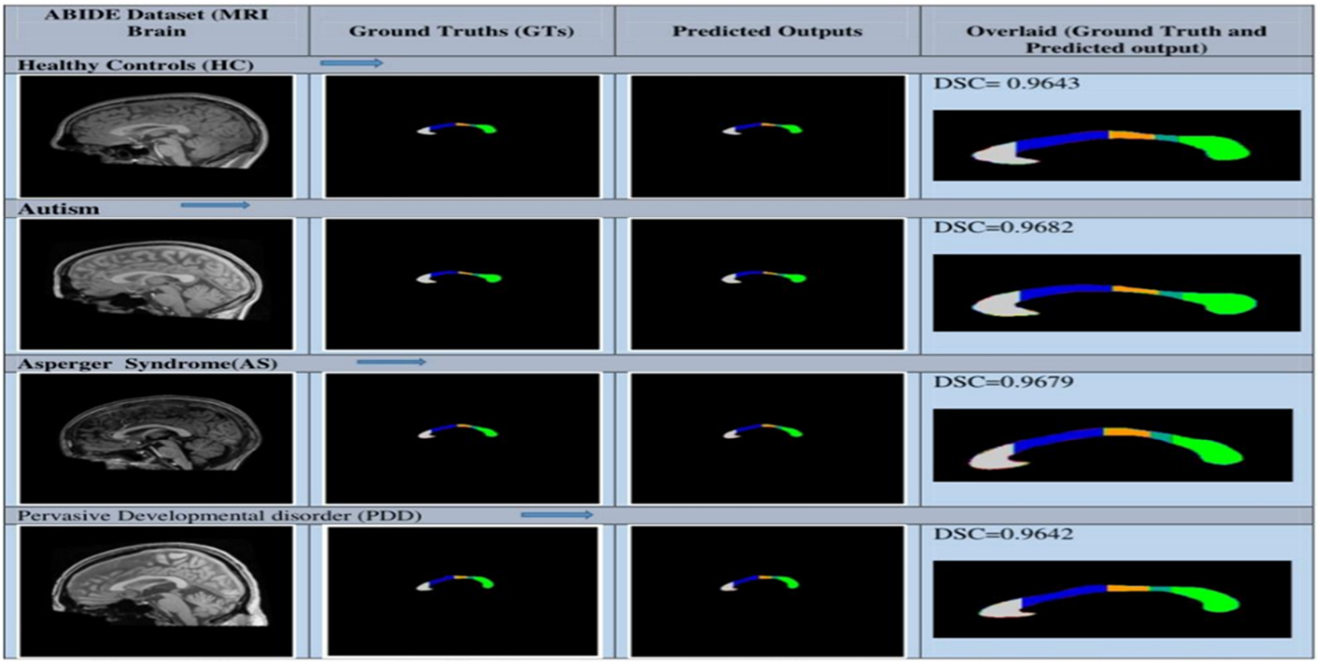

Fig. 3. Sample results of Cc segmentation.

C. Comparative Analysis with Existing Works

To further evaluate the effectiveness of the proposed PCcS-RAU-Net model, a comparative analysis was conducted against existing methods. Table III summarizes the segmentation performance reported in previous studies, which utilized both traditional machine learning techniques and deep learning models. The traditional approaches achieved DCS in the range of 86% - 90%, significantly lower than the results obtained by the proposed one. Even recent deep learning models that lacked residual or attention mechanisms generally reported DSC scores between 91% and 95%. Embedding the residual blocks and AGs into the U-Net structure offered three cardinal features to map the impressive up-scaling gadgets:

- It directed the feature spread to an enlarged bridging field of attention to the spatial pixels, so the attention module could target relevant spatial pieces
- It inspected and elucidated patterns at a very tiny scale useful for isolating the data damage
- It identified the most minuscule and intricate object boundaries in the Cc

Overall, this comparison further confirmed the strength of the proposed approach.

TABLE III. COMPARISON WITH OTHER STUDIES

Study	Approach	Dataset	DSC (%)
[3]	Geometrical Parcellation	MRI	~85
[4]	Fiber Tractography	MRI (DTI)	~88
[5]	Deep Learning	ABIDE	94.20
[7]	Deep Learning with Attention	ABIDE	95.10
This work	Deep Learning + Residual + Attention	ABIDE/OASIS/RCI	97.10

D. Medical and Technological Implications

The advancements introduced through the proposed PCcS-RAU-Net model exhibited significant medical and technological implications. From a medical perspective, the ability to automatically and accurately segment the Cc into anatomically meaningful subregions could support the early diagnosis, monitoring, and prognostic evaluation of neurological disorders, such as ASD and AD. Automating this process reduced the inter- and intra-observer variability and enabled large-scale morphological studies that were previously limited by manual segmentation efforts. From a technological viewpoint, the integration of residual blocks and AGs within the U-Net framework demonstrated a scalable strategy for enhancing the deep learning models applied to complex brain structure segmentation. The proposed enhancements improved the feature learning, spatial focus, and overall segmentation accuracy, which can be extended to other brain regions and medical imaging tasks. Furthermore, the model's robust performance across datasets with heterogeneous imaging protocols and patient demographics highlighted its potential for deployment in real-world clinical environments and multi-center studies, addressing a critical challenge of the data variability in medical imaging applications.

E. Study Limitations

Based on these findings, the model had few constraints to consider. Primarily, the training and evaluation of the model referred to midsagittal slices in 2D, enough for the Cc delineation, while the 3D anatomical context was not entirely captured. Furthermore, while filling the gap with the residual blocks and AGs, the increasing computational complexity and time of the model's training compared to simpler architectures limited its deployment in resource-constrained testing environments. Finally, the interpretations were largely anatomical, without associating the morphology results with the specific clinical outcomes or future possible longitudinal disease progression offering the opportunity for future research.

IV. CONCLUSION

This work introduced PCcS-RAU-Net, a deep learning model that integrated Residual Blocks and Attention Gates (AGs) within the U-Net architecture to segment the parcellated Corpus Callosum (Cc) from midsagittal brain Magnetic Resonance Imaging (MRI) scans. The proposed model demonstrated superior segmentation performance, with Dice Similarity Coefficients (DSCs) of 97.10%, 96.88%, and 97.11% on the ABIDE, OASIS, and RCI datasets, respectively. These results outperformed other conventional architectures, such as U-Net, Residual U-Net, and Attention U-Net, validating the effectiveness of the proposed approach. Overall, PCcS-RAU-Net presents a robust and fully automatic segmentation framework, toward support for large-scale neuroimaging studies alongside early neurological disorder detection and monitoring.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication. The authors would like to acknowledge the support of Prince Sultan University and for making this publication successful.

REFERENCES

[1] R. Kucharsky Hiess, R. Alter, S. Sojoudi, B. A. Ardekani, R. Kuzniecky, and H. R. Pardoe, "Corpus Callosum Area and Brain Volume in Autism Spectrum Disorder: Quantitative Analysis of Structural MRI from the ABIDE Database," *Journal of Autism and Developmental Disorders*, vol. 45, no. 10, pp. 3107–3114, Jun. 2015, https://doi.org/10.1007/s10803-015-2468-8.

[2] A. W. Russo *et al.*, "Associations between corpus callosum damage, clinical disability, and surface-based homologous inter-hemispheric connectivity in multiple sclerosis," *Brain Structure and Function*, vol. 227, no. 9, pp. 2909–2922, May 2022, https://doi.org/10.1007/s00429-022-02498-7.

[3] S. F. Witelson, "Hand and Sex Differences in the Isthmus and Genu of the Human Corpus Callosum: A Postmortem Morphological Study," *Brain*, vol. 112, no. 3, pp. 799–835, Jun. 1989, https://doi.org/10.1093/brain/112.3.799.

[4] S. Hofer and J. Frahm, "Topography of the human corpus callosum revisited—Comprehensive fiber tractography using diffusion tensor magnetic resonance imaging, " *NeuroImage*, vol. 32, no. 3, pp. 989–994, Sep. 2006, https://doi.org/10.1016/j.neuroimage.2006.05.044.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention*, Munich, Germany, Oct. 2015, pp. 234–241, https://doi.org/10.1007/978-3-319-24574-4_28.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition, " in *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 2016, pp. 770–778.

[7] O. Oktay *et al.*, "Attention U-Net: Learning Where to Look for the Pancreas." arXiv, May 20, 2018, https://doi.org/10.48550/arXiv.1804.03999.

[8] ABIDE dataset, https://sourceforge.net/projects/nidb/files/latest/download.

[9] OASIS dataset, https://download.nrg.wustl.edu/data/OAS2_RAW_PART1.tar.gz.

[10] A. Di Martino *et al.*, "The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism, " *Molecular Psychiatry*, vol. 19, no. 6, pp. 659–667, 2014, https://doi.org/10.1038/mp.2013.78.

[11] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults," *Journal of Cognitive Neuroscience*, vol. 19, no. 9, pp. 1498–1507, Sep. 2007, https://doi.org/10.1162/jocn.2007.19.9.1498.

[12] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in *Proceedings of the 32nd International Conference on Machine Learning*, Lille, France, Jun. 2015, pp. 448–456.

[13] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning, " *Journal of Big Data*, vol. 6, no. 1, Jul. 2019, Art. no. 60, https://doi.org/10.1186/s40537-019-0197-0.

[14] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization." arXiv, Jan. 30, 2017, https://doi.org/10.48550/arXiv.1412.6980.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification," in *Santiago, Chile*, 2015, pp. 1026–1034.

[16] A. Alyousef and O. Al-Omari, "Artificial Intelligence in Healthcare: Bridging Innovation and Regulation," *Journal of Ecohumanism*, vol. 3, no. 8, pp. 10582–10589, 2024, https://doi.org/10.62754/joe.v3i8.5673.

[17] A. Rehman, M. Mujahid, A. Elyassih, B. AlGhofaily, and S. A. O. Bahaj, "Comprehensive Review and Analysis on Facial Emotion Recognition: Performance Insights into Deep and Traditional Learning with Current Updates and Challenges," *Computers, Materials & Continua*, vol. 82, no. 1, 2025, Art. no. 41, https://doi.org/10.32604/cmc.2024.058036.

[18] N. A. Semary, W. Ahmed, K. Amin, P. Pławiak, and M. Hammad, "Improving sentiment classification using a RoBERTa-based hybrid model," *Frontiers in Human Neuroscience*, vol. 17, Dec. 2023, https://doi.org/10.3389/fnhum.2023.1292010.

[19] E. Othman and R. Mahafdah, "Recalibrating Human–Machine Relations through Bias-Aware Machine Learning: Technical Pathways to Fairness and Trust, " *Journal of Posthumanism*, vol. 5, no. 4, pp. 448–466, 2025.

[20] M. Rasool, A. Noorwali, H. Ghandorh, N. A. Ismail, and W. M. S. Yafooz, "Brain Tumor Classification using Deep Learning: A State-of-the-Art Review," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 16586–16594, Oct. 2024, https://doi.org/10.48084/etasr.8298.

[21] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," *BMC Medical Imaging*, vol. 15, no. 1, Aug. 2015, Art. no. 29, https://doi.org/10.1186/s12880-015-0068-x.