DEEP LEARNING–BASED MORPHOMETRIC ASSESSMENT OF CEREBRAL HEMISPHERE STRUCTURES USING MRI

G.E. Raxmonova — Professor of the Department of Medical Radiology No. 1, Tashkent State Medical University, Tashkent, Uzbekistan.

M.D. Rahmanova — Tashkent State Medical University, 1st Doctoral Student in Medical Radiology, Tashkent, Uzbekistan.

I.I. Izbasarov — Student of Group 202, Faculty of General Medicine No. 2, Tashkent State Medical University, Tashkent, Uzbekistan.

Introduction. Magnetic resonance imaging (MRI) remains the gold standard non-invasive technique for detailed morphological visualization of central nervous system structures in vivo. Its superior soft-tissue contrast, multiplanar capability, and absence of ionizing radiation make MRI indispensable for neurological assessment in both clinical and research settings. Morphometric analysis of cerebral hemispheres — encompassing quantitative evaluation of cortical thickness, gray matter volume, white matter integrity, and hemispheric symmetry indices — plays a pivotal role in the early detection, differential diagnosis, and longitudinal monitoring of a wide spectrum of neurological and psychiatric disorders, including Alzheimer's disease, schizophrenia, multiple sclerosis, focal epilepsy, and traumatic brain injury.

Conventional morphometric methods, including voxel-based morphometry (VBM) and manual delineation of regions of interest (ROI), are labor-intensive and highly susceptible to inter- and intra-observer variability. Manual segmentation of a single high-resolution volumetric MRI dataset may require several hours of expert neuroanatomist time, rendering population-scale studies logistically and economically impractical. Semi-automated approaches, while partially alleviating the time burden, still depend on operator-defined initialization parameters and frequently require manual correction of segmentation errors, particularly at tissue boundaries and in pathologically altered brain regions where signal intensity may be atypical.

The rapid advancement of artificial intelligence — specifically deep learning — has fundamentally transformed the landscape of medical image analysis over the past decade. Convolutional neural networks (CNNs), inspired by the hierarchical architecture of the mammalian visual cortex, have demonstrated remarkable capacity for learning complex spatial feature representations directly from raw imaging data. Encoder-decoder architectures, exemplified by the U-Net model originally developed for biomedical image segmentation, have proven particularly effective for dense pixel- and voxel-wise classification tasks. These models learn to map input image volumes to anatomical label maps through supervised training on large annotated datasets, subsequently generalizing to unseen cases with high accuracy and consistency. The integration of such deep learning frameworks into neuroimaging pipelines presents an opportunity to standardize morphometric assessment, reduce processing time to clinically viable intervals, and enable automated analysis at the population level — a transformation with profound implications for both diagnostic radiology and translational neuroscience.

Objective. To develop, train, and evaluate a deep learning–based model for fully automated morphometric assessment of cerebral hemisphere structures using high-resolution MRI data; to quantify its performance against expert manual segmentation using established statistical metrics; and to assess the clinical utility of the automated morphometric parameters derived from the model for characterizing structural brain differences in healthy individuals and patients with neurological conditions.

Materials and Methods.

Study design and dataset. The study was designed as a retrospective analytical investigation utilizing archival MRI data from Tashkent State Medical University's Radiology Center and affiliated neurological facilities. The total dataset comprised 480 high-resolution T1-weighted 3D brain MRI examinations: 210 from neurologically healthy adult volunteers (mean age 38.4 ± 12.1 years; 108 female, 102 male) and 270 from patients with confirmed neurological diagnoses, including Alzheimer's disease ($n = 85$), relapsing-remitting multiple sclerosis ($n = 72$), focal epilepsy ($n = 68$), and post-traumatic encephalopathy ($n = 45$). All participants provided written informed consent in accordance with the Declaration of Helsinki. The study protocol was approved by the Ethics Committee of Tashkent State Medical University (Protocol No. 14/2023).

MRI acquisition. All MRI examinations were performed on 1.5 T and 3.0 T scanners (Siemens Magnetom Aera and Prisma, respectively) using standard brain neuroimaging protocols. Volumetric T1-weighted magnetization-prepared rapid gradient-echo (MPRAGE) sequences were acquired with the following parameters: repetition time (TR) = 2300 ms, echo time (TE) = 2.98 ms, inversion time (TI) = 900 ms, flip angle = 9°, voxel size = $1 \times 1 \times 1$ mm^3, field of view (FOV) = 256×256 mm^2, 176 sagittal slices. FLAIR and T2-weighted sequences were additionally acquired for clinical characterization of pathological cases.

Image preprocessing. All MRI volumes underwent a standardized preprocessing pipeline prior to model training and evaluation. Steps included: (1) automated skull stripping using the Brain Extraction Tool (BET, FSL v6.0); (2) bias field correction using the N4 iterative algorithm to remove low-frequency intensity inhomogeneity; (3) spatial registration to MNI152 standard space using affine registration (FLIRT, FSL); (4) voxel intensity normalization to the range [0, 1] via z-score standardization computed over the intracranial volume. These preprocessing steps were applied uniformly across all datasets to ensure consistency of input data for the neural network.

Deep learning architecture. A modified 3D U-Net convolutional neural network architecture was implemented as the core segmentation framework. The model comprised an encoder pathway with four successive blocks of 3D convolution (kernel 3×3×3), batch normalization, and rectified linear unit (ReLU) activation, each followed by max-pooling for spatial downsampling. The bottleneck layer contained 512 feature channels. The symmetric decoder pathway used bilinear upsampling with skip connections from the corresponding encoder levels to preserve spatial detail and facilitate gradient flow during backpropagation. The final output layer employed a softmax activation function producing probabilistic maps for 34 cerebral structures, including cortical and subcortical regions of both hemispheres, defined according to the Desikan–Killiany–Tourville (DKT) atlas. The model contained approximately 19.4 million trainable parameters.

Training procedure. The dataset was partitioned into training (n = 336, 70%), validation (n = 72, 15%), and independent test (n = 72, 15%) sets using stratified random sampling to ensure balanced representation of diagnostic groups. Data augmentation strategies applied during training included random rotation (±15°), elastic deformation, random flipping along the sagittal axis, and Gaussian noise injection to improve model generalization and robustness to acquisition variability. The model was trained using the Adam optimizer with an initial learning rate of 1×10^{-4}, reduced by a factor of 0.5 upon validation loss plateau (patience = 10 epochs). A combined loss function of Dice loss and cross-entropy loss was used to address class imbalance between small subcortical structures and large cortical regions. Training was conducted over 120 epochs with a batch size of 2 on an NVIDIA A100 GPU (80 GB VRAM), with early stopping applied to prevent overfitting.

Performance evaluation. Segmentation accuracy was quantified using the Dice similarity coefficient (DSC) and 95th percentile Hausdorff distance (HD95) for spatial overlap and boundary accuracy, respectively. Volumetric reliability of morphometric parameters was assessed using the intraclass correlation coefficient (ICC, two-way mixed model, absolute agreement) and Bland–Altman analysis comparing automated versus expert manual measurements. Statistical analyses were performed in Python (SciPy v1.9, Pingouin v0.5) with a significance threshold of $p < 0.05$.

Results.

Segmentation performance. The deep learning model achieved high segmentation accuracy across all evaluated cerebral structures on the independent test set. Mean DSC was 0.912 ± 0.031 across all 34 regions, with the highest values observed for large cortical structures: frontal lobe gray matter (DSC = 0.941 ± 0.012), parietal lobe gray matter (DSC = 0.938 ± 0.014), and total hemispheric white matter (DSC = 0.951 ± 0.009). Subcortical structures demonstrated slightly lower but still clinically acceptable accuracy: hippocampus (DSC = 0.887 ± 0.028), caudate nucleus (DSC = 0.901 ± 0.021), and thalamus (DSC = 0.919 ± 0.018). Mean HD95 across all structures was 1.84 ± 0.63 mm, confirming precise delineation of structural boundaries. Processing time per complete volumetric brain scan was 3.2 ± 0.4 minutes on GPU, compared to 4.5–6.0 hours for expert manual segmentation, representing a reduction in analysis time of over 98%.

Morphometric reliability. Intraclass correlation analysis demonstrated excellent agreement between AI-derived and manually obtained morphometric parameters. ICC values were: total brain volume 0.978 (95% CI: 0.963–0.988), total gray matter volume 0.961 (95% CI: 0.942–0.974), total white matter volume 0.956 (95% CI: 0.935–0.971), mean cortical thickness 0.948 (95% CI: 0.924–0.965), and hemispheric asymmetry index 0.934 (95% CI: 0.906–0.955). Bland–Altman plots revealed no systematic bias in volumetric estimates, with limits of agreement within ±3.8% of mean values for all major morphometric parameters, confirming the absence of proportional error across the measurement range.

Group differences in morphometric parameters. Comparative morphometric analysis revealed statistically significant structural differences between diagnostic groups. Patients with Alzheimer's disease demonstrated bilateral hippocampal volume reduction (left: −24.3 ± 5.1%, right: −21.8 ± 4.7% relative to age-matched controls; p

< 0.001), entorhinal cortex thinning ($-18.6 \pm 3.9\%$; $p < 0.001$), and total gray matter volume deficit of $-11.4 \pm 2.8\%$ ($p < 0.001$). In the multiple sclerosis group, white matter volume was reduced by $8.7 \pm 2.3\%$ ($p < 0.001$), with cortical thinning most pronounced in frontal and parietal regions. Patients with focal epilepsy showed ipsilateral hippocampal sclerosis with volumetric asymmetry index of 0.21 ± 0.06 (healthy controls: 0.04 ± 0.02; $p < 0.001$). Post-traumatic encephalopathy cases exhibited heterogeneous patterns of focal cortical thinning and regional gray matter loss corresponding to injury sites. Hemispheric symmetry indices significantly differentiated pathological from healthy groups ($F(4,475) = 38.7$, $p < 0.001$, $\eta^2 = 0.246$).

Conclusion. The developed deep learning–based morphometric framework demonstrates robust, accurate, and reproducible automated segmentation and quantitative analysis of cerebral hemisphere structures on MRI, with performance metrics that meet or exceed established clinical and research thresholds. Segmentation accuracy (mean DSC > 0.91) and volumetric reliability (ICC > 0.93 for all primary morphometric parameters) were maintained across healthy controls and diverse neurological patient populations, validating the generalizability of the model. The dramatic reduction in processing time — from hours to minutes — without sacrificing measurement quality represents a clinically meaningful advancement that makes high-throughput morphometric neuroimaging practically feasible.

The ability of the system to detect and quantify disease-specific patterns of structural brain change — including hippocampal atrophy in Alzheimer's disease, white matter volume loss in multiple sclerosis, hippocampal asymmetry in focal epilepsy, and focal cortical thinning in post-traumatic encephalopathy — demonstrates its potential as a diagnostic decision support tool. Elimination of inter-observer variability further strengthens the reliability of longitudinal monitoring, where detection of subtle progressive changes is essential for therapeutic efficacy assessment.

Future work will focus on expanding the training dataset to include pediatric populations, rare neurological conditions, and multi-site multi-vendor MRI data to further improve generalizability. Prospective clinical validation studies are planned to assess the impact of AI-assisted morphometric reporting on diagnostic accuracy and clinical decision-making. Integration of the framework into a clinical PACS-compatible workflow with automated report generation is currently under development at Tashkent State Medical University. These efforts aim to translate the demonstrated technical capabilities of deep learning–based neuroimaging analysis into tangible improvements in patient care within the Central Asian healthcare context.

Keywords: deep learning; convolutional neural network; neuroimaging; morphometry; MRI segmentation; cortical thickness; gray matter; hemispheric symmetry; U-Net; Alzheimer's disease; multiple sclerosis.