VFAST Transactions on Software Engineering
https://vfast.org/journals/index.php/VTSE@ 2026, ISSN(e): 2309-3978, ISSN(p): 2411-6246
Volume 14, Number 1, January-March 2026 *pp:301-317*

Lightweight Texture-Based Classification of Huffaz Status from Structural MRI Using GLCM Correlation and Brodmann-Area VOIs

Mohd Izzuddin Mohd Tamrin[1], **Iqbal Jamaludin**[2], **Abdul Halim Sapuan**[2], **Mohd Zulfaezal Che Azemin**[2*], **Asadullah Shah**[1]

[1]Kulliyyah of ICT, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia;
[2]Integrated Omics Research Group, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Kampus Kuantan, Jalan Sultan Ahmad Shah, 25200 Kuantan, Pahang Darul Makmur, Malaysia

Keywords: *Structural MRI Radiomics, Gray-Level Co-Occurance Matrix (GLCM) Correlation, Brodmann Area 46, Texture Analysis, Huffaz (Quran Memorization)*

Journal Info:
Submitted:
March 06, 2026
Accepted:
March 21, 2026
Published:
March 30, 2026

Abstract Previous neuroimaging studies suggest that intensive Quran memorization may be associated with structural brain differences, but simple and interpretable MRI-based biomarkers remain limited. This study investigated whether a single interpretable radiomics feature, gray-level co-occurrence matrix (GLCM) Correlation, extracted from predefined Brodmann Area (BA) volumes of interest (VOIs), is associated with Huffaz status.A cross-sectional case-control design was used involving 47 participants (23 Huffaz, 24 non-Huffaz). Structural MRI scans were pre-processed using a standard SPM pipeline to generate modulated, warped, and smoothed tissue-class images (smwc*.nii). Inferential analyses were restricted to five literature-driven candidate regions (BA22, BA24, BA32, BA40, and BA46) to reduce multiplicity. Best-subset logistic regression was performed with sex forced into all models and Bayesian Information Criterion (BIC) used for model selection. Bootstrap resampling was applied to assess feature-selection stability.BA46 GLCM Correlation emerged as the lowest-BIC predictor set. Bootstrap resampling showed higher selection stability for BA46 (approximately 0.68 selection frequency) than for the other candidate regions (approximately 0.13–0.15), supporting a reproducible single-region signal within the tested set. These findings suggest that BA46 texture organization, as captured by GLCM Correlation, may differentiate Huffaz from non-Huffaz after adjustment for sex.The modest sample size and lack of external validation limit generalizability. Nevertheless, the findings support the potential of an interpretable single-feature radiomics biomarker and motivate further validation in larger independent cohorts, as well as robustness analyses across preprocessing and discretization settings.

***Correspondence author email address:** zulfaezal@iium.edu.my
DOI: 10.21015/vtse.v14i1.2363

1 Introduction

Sustained cognitive training may occur as long-term intensive memorization which can be accompanied by an observable brain alteration. Structural MRI studies to date have largely relied on macrostructural results, in particular, regional volume and voxel-based morphometry (VBM), to compare Quran memorizers with non-memorizers in Huffaz (individuals who memorize the Quran), in this case. As an illustration, Jamaludin et al [1]. have provided group comparisons between VBM-derived gray matter measures in Quran memorizers, and the current use of morphometric or

tissue-density contrasts as the major final outcome. A similar memorization study by Sapuan et al. [2] also used VBM to determine the regional differences in gray matter volume between memorizers and controls. While informative, volume / VBM primarily quantifies "how much tissue" is present and may be less sensitive to subtle, distributed differences in within-region tissue organization.

In order to go beyond volume-only measures, Jamaludin et al. [3] proposed a fractal characterization of Huffaz brain, comprising of global and volume-of-interest (VOI) based fractal measures, which explicitly motivates imaging measures of structural complexity beyond standard morphometry. However, fractal measures reflect a specific notion of geometric complexity and do not directly quantify second-order spatial intensity relationships, a key aspect of tissue heterogeneity often targeted by radiomics texture analysis.

Radiomics has been widely applied in neuroimaging for disease characterization and prediction using MRI features and machine learning models [4–9]. In particular, texture radiomics summarizes spatial patterns of voxel intensities inside a VOI. One widely used texture representation is the Gray Level Co-occurrence Matrix (GLCM), which encodes how frequently pairs of discretized gray levels co-occur at a specified spatial offset within a region. From the GLCM, Haralick-type texture features can be computed, including GLCM Correlation, which captures the linear dependency between neighboring gray levels. Recent reproducibility-focused neuroimaging radiomics evidence supports prioritizing GLCM-derived features when robustness is important: a *NeuroImage* study by Azimi et al. [4] assessing reproducibility of radiomic features in brain imaging reported that GLCM and GLDM feature classes were among the most stable, providing specific justification for selecting GLCM-based measures (including correlation) as candidates for reliable modelling.

Radiomics combined with machine learning has shown strong predictive capability in several neuroimaging applications, including mild cognitive impairment diagnosis and disease progression prediction using MRI-based radiomic features [10–12].

At the same time, the radiomics literature emphasizes that stability depends strongly on methodological choices (e.g., discretization, scanner variability, and preprocessing). Multi-scanner test-retest work by Wennmann et al. [13] highlights the value of restricting analyses to reproducible features to improve generalizability across centers. Methodological study by Zhao et al. [14] also demonstrates that discretization strategy can materially affect feature reproducibility, motivating transparent reporting of key parameters that directly shape the GLCM. Robust feature selection strategies are therefore critical for radiomics modelling to ensure generalizable predictive performance across datasets [15].

In addition, modern standardization and reporting guidance, such as IBSI updates and radiomics auditing checklists, reinforces the need to document preprocessing and feature definitions to support reproducible, comparable texture biomarkers [16, 17]. Finally, recent work by Zhang et al. [18] comparing approaches to identify repeatable radiomic features (e.g., perturbation vs. est-retest) further underlines the importance of reliability-aware feature selection before making biological interpretations.

Therefore, building on Huffaz neuroimaging that has focused on volume/VBM and subsequently fractal characterization, the present study evaluates whether VOI-based GLCM Correlation provides an additional, texture-sensitive marker associated with Huffaz status, while adjusting for sex as an *a priori* covariate. By focusing on predefined Brodmann Area VOIs, this work aims to provide anatomically interpretable evidence using a radiomics feature family that is widely used and supported by recent stability findings. Radiomics has also been explored for multiple neurological conditions including glioma grading, Parkinson's disease detection, and cerebral small vessel disease characterization [19–22].

2 Related Work

Previous neuroimaging studies on Huffaz have primarily used voxel-based morphometry, regional volume, and fractal analysis to investigate structural brain differences associated with Quran memorization [1–3]. While these approaches are valuable, they mainly characterize tissue amount or geometric complexity rather than within-region texture organization [1–3].

In parallel, radiomics has become increasingly used in brain MRI for quantitative biomarker discovery, including classification and prediction tasks in neurological disease [4–12, 19–22]. The radiomics of texture, in particular, GLCM-based features, represents spatial relationships of intensity of a VOI, and thus is capable of identifying small-scale tissue heterogeneity that is not identified in traditional morphometric features. Recent research has been keen on the need to have feature reproducibility, methodological transparency, and sensitivity to preprocessing and discretization decisions in radiomics workflows. [4, 13–18].

It is in this context that the present study has been novel in its design that is deliberately minimalistic and interpretable. Instead of defining the data in high-dimensional black-box feature space, this analysis considers one GLCM texture feature of a small literature-based set of Brodmann-area VOIs. This approach was intended to reduce multiplicity, improve anatomical interpretability, and provide a transparent candidate biomarker that can be more readily replicated and refined in future work [1–4, 13–18].

3 Methodology

3.1 Study Design and Participants

A cross-sectional case-control analysis was performed to test whether predefined BA VOIs exhibit associations between MRI radiomics texture and Huffaz status. A total of 47 participants aged 20 to 25 years were enrolled after MRI screening (mean $\pm$ SD age: 22.55 ± 1.47 years). Recruitment was conducted through social media platforms, including Facebook and WhatsApp. The final sample comprised 23 Huffaz and 24 non-Huffaz participants, which satisfied the required sample size for statistical analysis. To reduce educational heterogeneity between groups, all participants were undergraduate students enrolled in science-based bachelor's degree programmes at IIUM Kuantan Campus. The Huffaz group consisted of students who had previously completed the Darul Quran–IIUM Tahfiz Certificate Programme, an approximately 18-month post-secondary memorization programme undertaken after completion of secondary school (Sijil Pelajaran Malaysia level). All participants were informed of the study objectives prior to participation.

To reduce potential educational confounding, both the Huffaz and non-Huffaz groups were recruited from the same university campus and were enrolled in undergraduate science-based degree programmes. This sampling strategy helped improve comparability in age range, current educational level, and academic environment. However, differences may still remain in prior training history, cognitive profile, bilingual or linguistic background, and other lifestyle-related factors that were not fully quantified in the present study.

The primary binary outcome for subject s was defined as:

$$y_s = \begin{cases} 1, & \text{Huffaz} \\ 0, & \text{non-Huffaz} \end{cases} \tag{1}$$

Sex was included as an *a priori* adjustment covariate and forced into all models:

$$\text{Sex}_s = \begin{cases} 1, & \text{female} \\ 0, & \text{male} \end{cases} \tag{2}$$

3.2 Data Acquisition and MRI Preprocessing

Structural T1-weighted MRI volumes were acquired for all participants (23 Huffaz vs. 24 non-Huffaz), and processed using a standard voxel-based morphometry (VBM)-style workflow implemented in Statistical Parametric Mapping (SPM) [1–3, 23]. The preprocessing goal was to convert the structural image of individual participants in native space to a normalized anatomical space, create tissue-class probability images, and create harmonized gray-matter images that can be used later in extracting region-of-interest texture.

First, all structural scans were converted to NIfTI format when needed and visually checked before formal preprocessing. Consistent with SPM guidance, the images were inspected for gross artifacts, orientation problems, and obvious misalignment relative to the template space. When necessary, images can be manually reoriented before segmentation so that the subsequent affine and nonlinear registration steps start from a reasonable initial position. This preliminary quality-control stage is important because poor initial alignment can degrade later tissue classification and normalization accuracy.

Second, each T1-weighted image underwent tissue segmentation in SPM to classify voxels into gray matter, white matter, cerebrospinal fluid, and other non-brain tissue classes. In SPM12, segmentation produces tissue probability maps such as c1 (gray matter), c2 (white matter), and c3 (CSF), together with deformation information used for normalization. In practical terms, this step estimates for each voxel the probability that it belongs to a given tissue type, thereby isolating the gray-matter image that is most relevant for morphometric and texture-based analyses.

Third, the segmented tissue maps were spatially normalized (warped) to a standard stereotactic reference space, typically MNI space, so that anatomically corresponding regions from different participants occupied comparable voxel locations. In SPM/VBM processing, this warping is achieved through affine plus nonlinear deformation fields estimated during segmentation or template-based inter-subject registration. The purpose of warping is not to alter biology, but to reduce gross anatomical variability across participants so that the same Brodmann-area masks can be applied in a common coordinate system.

Fourth, the normalized gray-matter images were modulated. Modulation scales voxel intensities by the local Jacobian determinants of the deformation field so that tissue amounts are preserved after expansion or contraction during spatial normalization. In other words, if a brain region is enlarged during warping, its voxel values are reduced proportionally, whereas if a region is compressed, its voxel values are increased proportionally. This step allows the resulting images to retain information related to local gray-matter volume after normalization, which is why modulated images are commonly used in VBM-style analyses.

Fifth, the modulated normalized tissue maps were smoothed using an isotropic Gaussian kernel, yielding the final smoothed, modulated, warped tissue-class images conventionally labeled `smwc*.nii` in SPM. Smoothing reduces high-frequency noise, compensates partly for small residual inter-subject misregistrations, and makes the data more spatially stable for group analysis. SPM documentation notes that VBM studies commonly use isotropic smoothing kernels in the range of approximately 8–12 mm, although the exact kernel can vary according to study design and the spatial scale of the expected effects.

Finally, the outputs of preprocessing were reviewed by visual inspection to confirm plausible tissue segmentation and acceptable anatomical normalization. One preprocessed gray-matter image per subject, following the SPM naming convention `smwc*.nii`, was then retained for radiomics processing. These standardized images formed the input for the subsequent Brodmann-area VOI masking and GLCM feature extraction pipeline used in the present study.

Let the resulting preprocessed 3D image for subject s be denoted

$$I_s(\mathbf{x}), \quad \mathbf{x} \in \Omega \subset \mathbb{Z}^3 \tag{3}$$

where $\mathbf{x}$ indexes voxels in the image domain Ω.

3.3 VOI Definition (Brodmann Areas)

Volumes of interest (VOIs) were defined using Brodmann-area atlas masks rather than manually drawn regions. In line with our earlier Huffaz VOI study, the atlas source was the Wake Forest University (WFU) PickAtlas, which provides Brodmann-area and other neuroanatomical masks based on the Talairach Daemon and makes them available in MNI space for use with SPM [3]. In atlas-based ROI workflows, the fundamental idea is that each anatomical label is first defined on a standard template brain and then transferred to the processed subject data through the shared coordinate system. VOIs corresponded to Brodmann Areas BA01–BA46. In the dataset, each BA VOI mask was encoded as:

$$\{rz01, rz02, \ldots, rz46\} \tag{4}$$

Structural MRI volumes were first preprocessed in SPM and spatially normalized to standard space, yielding the final modulated, warped, and smoothed tissue-class images used for radiomics. Because both the preprocessed subject images and the PickAtlas Brodmann masks are represented in standard MNI space, the first level of alignment was achieved through this common normalization framework. In other words, atlas localization was not performed in raw native space; instead, it relied on the fact that each subject image had already been warped into the same stereotactic reference system as the atlas. This is the same general logic used in many SPM/WFU PickAtlas ROI studies,

where atlas-defined Brodmann masks are interrogated after normalization to template space.

After selecting the required Brodmann area from WFU PickAtlas, the corresponding atlas mask was exported as a binary image, where voxels belonging to the selected BA were assigned value 1 and all other voxels were assigned 0. For the current study, although the atlas can define many Brodmann regions, inferential analyses were restricted to the pre-specified candidate VOIs BA22, BA24, BA32, BA40, and BA46. Each mask therefore functioned as a spatial selector identifying which voxels in the preprocessed MRI belonged to that anatomical region. To reduce model complexity and multiple testing, inferential analyses were restricted to five literature-driven candidate regions [3]:

$$\mathcal{B} = \{22, 24, 32, 40, 46\} \tag{5}$$

BA22 and BA40 were prioritized for language / phonological processing, BA24 and BA32 for cognitive control/monitoring, and BA46 for executive working-memory functions relevant to intensive memorization practice.

A second alignment step was then performed at the image-grid level. Even when two images are in the same anatomical coordinate space, their matrix dimensions, voxel sizes, or origin definitions can still differ slightly. Therefore, for each subject, the selected BA mask was checked against the preprocessed MRI volume. If the mask grid did not exactly match the subject image grid, the mask was resampled onto the subject image lattice using nearest-neighbor interpolation. Nearest-neighbor resampling was used because the BA mask is a label image; unlike linear or spline interpolation, nearest-neighbor preserves discrete region membership and avoids generating artificial fractional labels at VOI boundaries. This is standard practice whenever atlas label images are transferred to another voxel grid for ROI extraction. Each BA VOI b was represented by a binary mask:

$$M_b(\mathbf{x}) \in \{0, 1\} \tag{6}$$

Once the BA mask had been brought onto the same voxel grid as the subject's preprocessed MRI, VOI extraction was performed voxelwise. Following the logic described in our earlier VOI-based fractal study, the preprocessed subject image can be regarded as volume V_1 and the BA mask as volume V_2; the localized VOI image is then obtained by voxelwise multiplication of V_1 and V_2, or equivalently by retaining only voxels for which the mask equals 1. In this way, all subsequent regional measurements are calculated exclusively from intensities inside the selected Brodmann area, while voxels outside the region are excluded.

Operationally, this means that the BA mask served as an anatomical filter on the already normalized gray-matter image. After localization, the resulting BA-specific voxel set was passed to the radiomics pipeline, where intensity discretization and GLCM Correlation were computed only within that masked region. Thus, the radiomic value for a given subject and BA did not represent the whole brain, but the texture organization of the subset of voxels lying inside the aligned Brodmann-area VOI.

Finally, visual quality control was performed to confirm that the atlas mask overlaid the expected cortical territory on the normalized subject image. This verification step is important because atlas-based localization depends on both successful normalization and correct label transfer; misregistration at either stage can shift boundary voxels and alter the measured regional texture.

3.4 Radiomics Feature Extraction (Single-Feature Design)

This study used a deliberately simple radiomics design in which only one texture feature, GLCM Correlation, was extracted from each Brodmann-area VOI. The computation followed three sequential steps: intensity discretization, GLCM construction, and Correlation calculation. In radiomics, this ordering is important because GLCM-based texture features are not computed directly from raw continuous intensities; rather, voxel values are first converted into a limited set of gray-level bins, and the co-occurrence structure of those discretized values is then summarized mathematically.

3.4.1 Step 1: Intensity Discretization Inside the VOI

After a Brodmann-area mask had been aligned to the subject image grid, only voxel intensities lying inside that VOI were retained for texture analysis. Because GLCM methods require a finite set of gray levels, the continu-

ous MRI intensities within the VOI were discretized using a fixed bin width. Operationally, this means that the original intensity range was partitioned into equally sized intervals, and each voxel was reassigned to an integer gray-level category according to the interval into which its intensity fell. Thus, neighboring voxels with similar intensities were mapped to the same or nearby discrete gray levels, while large intensity differences were mapped to more distant levels.

GLCM Correlation was extracted from each BA VOI using a fixed-bin-width discretization scheme with bin width = 0.01. The GLCM was computed in 3D using a voxel distance of 1 (PyRadiomics), across 13 directions in 3D, with symmetric matrices and mean aggregation across directions. The matrix was normalized before feature calculation, and one scalar GLCM Correlation value was obtained for each subject and each BA VOI.

To compute GLCM-based texture, voxel intensities within the VOI were discretized using a fixed bin width w:

$$g_s(\mathbf{x}) = \left\lfloor \frac{I_s(\mathbf{x})}{w} \right\rfloor \tag{7}$$

where $g_s(\mathbf{x})$ is the discretized gray level at voxel $\mathbf{x}$.

In radiomics, discretization strongly affects downstream texture values because it determines how much local intensity variation is preserved versus compressed. Standard radiomics guidance and software documentation therefore emphasize transparent reporting of discretization settings, and fixed-bin-width approaches are commonly used because they keep bin size constant across subjects.

3.4.2 Step 2: Construction of the Gray-level Co-occurrence Matrix

Once all voxels within the VOI had been discretized, a gray-level co-occurrence matrix (GLCM) was constructed. The GLCM is a square matrix in which each entry records how often a voxel with gray level i occurs adjacent to a voxel with gray level j for a defined spatial relationship. In practical terms, the algorithm scans through voxel pairs inside the VOI, identifies the gray-level label of the first voxel and that of its neighbor, and increments the corresponding matrix cell. Repeating this for all eligible neighboring pairs yields a count table describing how often different gray-level combinations occur in the region.

After counting, the matrix was normalized, so that each matrix element represented a relative co-occurrence probability rather than a raw count. This normalization is important because it places subjects and VOIs on a comparable scale even when the number of contributing voxel pairs differs across regions or individuals. Conceptually, the normalized GLCM is therefore a probability distribution over neighboring gray-level pairs inside the VOI.

For each subject s and VOI b, the normalized GLCM $P_{s,b}(i,j)$ was constructed from discretized intensities $g_s(\mathbf{x})$ restricted to $M_{b\to s}$, such that:

$$\sum_i \sum_j P_{s,b}(i,j) = 1 \tag{8}$$

3.4.3 Step 3: Calculation of GLCM Correlation

Finally, GLCM Correlation was computed from the normalized matrix. Correlation is a Haralick-type texture feature that measures the degree of linear dependency between the gray levels of neighboring voxels. It uses the normalized co-occurrence probabilities together with the marginal gray-level means and standard deviations to quantify whether high gray levels tend to occur next to high gray levels, low next to low, or whether neighboring values vary more independently.

GLCM Correlation for subject s and BA b was defined as

$$\mathrm{Corr}_{s,b} = \frac{\sum_i \sum_j (i-\mu_i)(j-\mu_j)\, P_{s,b}(i,j)}{\sigma_i \sigma_j} \tag{9}$$

with

$$\begin{aligned} \mu_i &= \sum_i i \left(\sum_j P_{s,b}(i,j) \right), \\ \mu_j &= \sum_j j \left(\sum_i P_{s,b}(i,j) \right), \end{aligned} \tag{10}$$

$$\begin{aligned} \sigma_i &= \sqrt{\sum_i (i-\mu_i)^2 \left(\sum_j P_{s,b}(i,j) \right)}, \\ \sigma_j &= \sqrt{\sum_j (j-\mu_j)^2 \left(\sum_i P_{s,b}(i,j) \right)}. \end{aligned} \tag{11}$$

Thus, each subject contributed a single scalar value per BA VOI:

$$x_{s,b} = \mathrm{Corr}_{s,b}, \qquad b \in \mathcal{B}. \tag{12}$$

In practical interpretation, a higher Correlation value indicates that neighboring discretized intensities inside the VOI vary in a more predictable and linearly related way, whereas a lower value indicates weaker linear dependence and a less orderly pairwise intensity pattern. Importantly, this feature summarizes within-region spatial organization, not overall signal magnitude or regional volume. Thus, for each subject and each BA mask, the output of the feature-extraction stage was one scalar number representing the texture organization of that VOI.

3.5 Statistical Modelling

After feature extraction, statistical validation was performed using binary logistic regression to distinguish the two study groups. The dependent variable was group membership, and all candidate models included sex as a mandatory covariate. In addition, up to five Brodmann-area radiomics predictors corresponding to the selected top-five BA regions were considered as optional predictors in the model-building stage. Before modelling, each BA radiomics feature was standardized to zero mean and unit standard deviation so that effect sizes and odds ratios were comparable across regions and could be interpreted per one-standard-deviation increase in the feature value. Sex was coded as a binary covariate.

Rather than fitting only a single prespecified multivariable model, the analysis explored all possible subsets of the candidate BA predictors while always retaining sex in the design matrix. Because there were only five optional BA predictors, exhaustive evaluation was computationally feasible: all possible subsets were fitted, from the null subset containing no BA predictor up to the full subset containing all five BA predictors. In each case, a logistic regression model of the form

$$\text{group} \sim \text{sex} + \text{subset of BA predictors} \tag{13}$$

was estimated. Thus, the model-building procedure was iterative in the sense that each candidate predictor combination was fitted and scored separately before final selection.

For each fitted model, the Bayesian Information Criterion (BIC) was computed and used as the model-selection criterion. BIC balances model fit against model complexity by penalizing the inclusion of additional predictors; therefore, among competing candidate models, the model with the lowest BIC was selected as the preferred simplified model. In the present implementation, BIC was explicitly chosen as the ranking criterion, although the code infrastructure also allowed AIC as an alternative. The selected final model was then summarized by its estimated regression coefficients, Wald standard errors, p-values, and exponentiated coefficients reported as odds ratios (ORs) with 95% confidence intervals. Because BA predictors were standardized before modelling, each OR corresponds to the change in odds associated with a one-standard-deviation increase in that radiomic feature.

To assess the robustness of predictor selection, a bootstrap stability analysis was performed after identifying the best BIC-ranked model. In each bootstrap iteration, a new dataset of the same size as the original sample was generated by sampling subjects with replacement from the original cohort. The complete model-selection procedure was then rerun on that bootstrap sample: all predictor subsets were evaluated again, logistic regression models were refitted, and the subset with the lowest BIC was identified for that resampled dataset. This process was repeated a large number of times, and the selected BA predictors from each bootstrap replicate were recorded.

The main output of this bootstrap stage was the selection frequency of each BA predictor, defined as the proportion of successful bootstrap replicates in which that predictor appeared in the BIC-optimal model. A predictor selected very often across bootstrap samples was interpreted as more stable, whereas one selected only infrequently was regarded as less robust to sampling variability. The bootstrap procedure also recorded the distribution of selected model sizes and the number of failed fits, such as those affected by numerical instability or separation. This approach provided an empirical complement to single-sample model selection by showing whether the chosen predictors remained competitive under repeated perturbations of the dataset. Fig. 1 summarizes the overall methodological workflow of the study, from MRI acquisition and SPM preprocessing to Brodmann-area VOI localization, radiomics feature extraction, and statistical validation.

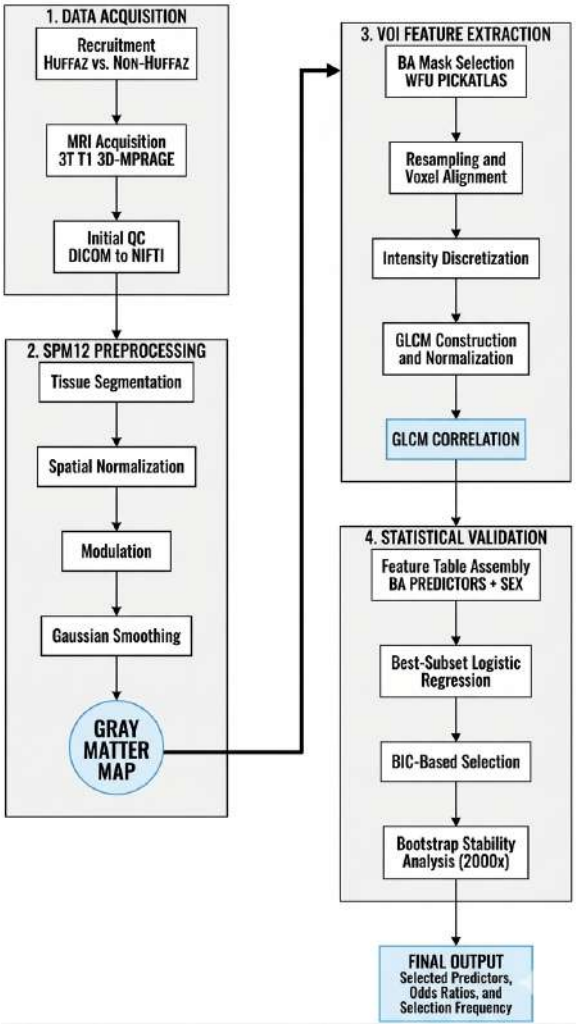

Figure 1. High-level workflow of the study, showing sequential progression from data acquisition and SPM12 preprocessing to VOI-based feature extraction, statistical validation, and final model output.

3.6 Comparative Benchmarking Against Previous Approaches

To place the proposed radiomics feature in context, an additional comparative benchmarking analysis was performed against previously used structural descriptors. Specifically, the discriminatory performance of GLCM Correlation was compared with two reference measures: regional gray-matter volume and fractal-based measurement. These benchmark features were selected because they reflect the principal analytical directions used in our previous work, namely voxel-based morphometry/volume analysis and fractal analysis of localized brain regions [1–3].

For this comparison, receiver operating characteristic (ROC) analysis was performed for the selected BA46 region using each feature separately as a univariable discriminator between the Huffaz and non-Huffaz groups. The ROC curve was generated by varying the decision threshold across the observed feature range, and the corresponding area under the curve (AUC) was used as the summary measure of classification performance. An AUC close to 0.5 indicates no discriminative ability beyond chance, whereas larger AUC values indicate better separation between the two groups.

This comparative step was intended to determine whether the proposed radiomics feature provides added discriminatory value beyond conventional regional descriptors. In this way, the analysis does not merely

report the performance of GLCM Correlation in isolation, but explicitly evaluates whether it improves upon earlier feature types previously used in related Huffaz brain studies.

4 Results

4.1 Best-subset Model Under BIC (sex forced in)

Best-subset selection was performed over five candidate Brodmann Areas (BA24, BA32, BA22, BA40, BA46) using the Bayesian Information Criterion (BIC) as the selection criterion, while forcing sex into every model. Results in Table 1 indicate that the lowest-BIC model included a single BA predictor, BA46, with BIC = 69.922 and AIC = 64.372.

$$\mathrm{BIC} = -2\,LL + p\ln(n) \tag{14}$$

Table 1. Top Ranked Candidate Models (Lower BIC Preferred)

subset	*k*_pred.	AIC	BIC	llf
BA46	1	64.3720	69.9224	-29.1860
(none)	0	68.1068	71.8071	-32.0534
BA40, BA46	2	65.5852	72.9858	-28.7926
BA32, BA46	2	65.9711	73.3717	-28.9855
BA24, BA46	2	66.0306	73.4312	-29.0153
BA22, BA46	2	66.1295	73.5301	-29.0648
BA24	1	69.2513	74.8017	-31.6256
BA40	1	69.3921	74.9426	-31.6961
BA32	1	69.5434	75.0939	-31.7717
BA22	1	69.8937	75.4442	-31.9469

4.2 Bootstrap Stability of Predictor Selection

As summarized in Fig. 2, to quantify selection stability, best-subset selection (BIC; sex forced in) was repeated across $B = 2000$ bootstrap resamples. Selection frequency for each BA was computed as the proportion of bootstrap fits in which the BA appeared in the chosen subset. BA46 showed the highest stability (selection frequency = 0.6795), whereas the remaining regions were selected less frequently (approximately 0.1295 to 0.1445).

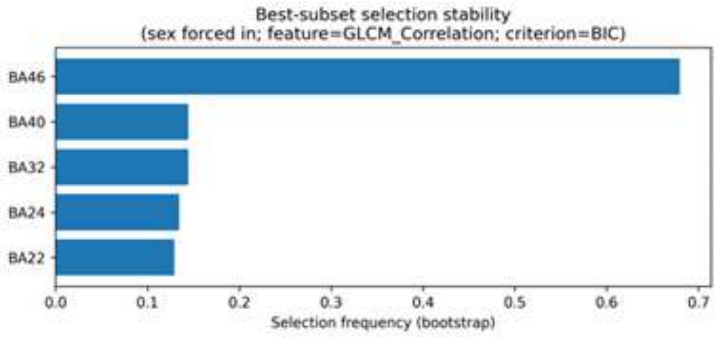

Figure 2. Best-subset selection stability under bootstrap resampling (BIC; sex forced in).

4.3 Best Model Coefficients

As evidenced in Table 2, the best-subset model (BIC) included BA46 and sex. Parameter estimates are reported as log-odds coefficients (β) with corresponding odds ratios (OR). The BA predictor was standardized; thus, the OR reflects a one-standard-deviation increase in BA46 GLCM Correlation.

In the adjusted model, BA46 GLCM Correlation was associated with lower odds of being Huffaz ($\beta = -0.923$, OR = 0.397, 95% CI = [0.174, 0.906], $p = 0.028$).

4.4 Comparative ROC Analysis for BA46

A comparative ROC analysis was performed for BA46 to evaluate the relative discriminatory ability of three feature types: volume, fractal measurement, and GLCM Correlation. As shown in Fig. 3, GLCM Correlation achieved the highest classification performance, with an AUC of 0.701, whereas volume and fractal features showed substantially lower discrimination, with AUCs of 0.540 and 0.525, respectively.

These results indicate that the conventional structural descriptors, namely volume and fractal measurement, provided only near-chance-level discrimination between the Huffaz and non-Huffaz groups in BA46. In contrast, the GLCM-based radiomics feature demonstrated a noticeably stronger ability to separate the two groups. This suggests that texture-based characterization captures discriminatory information not sufficiently represented by simple regional volume or fractal complexity alone.

From a methodological perspective, this finding supports the inclusion of radiomics analysis as a complementary approach to previous morphometric and

Table 2. Best-Subset Model Coefficients (Sex Forced In)

term	beta	SE	OR	CI95_OR_low	CI95_OR_high	*p*_wald
const	-0.2843	0.5634	0.7526	0.2495	2.2704	0.6139
sex(F=1)	0.3645	0.7665	1.4398	0.3205	6.4681	0.6344
BA46 (1SD)	-0.9231	0.4208	0.3973	0.1741	0.9064	0.0283

fractal methods. While the earlier methods remain useful as baseline descriptors, the present ROC comparison indicates that GLCM Correlation may offer superior sensitivity for detecting subtle structural differences within the selected Brodmann-area VOI.

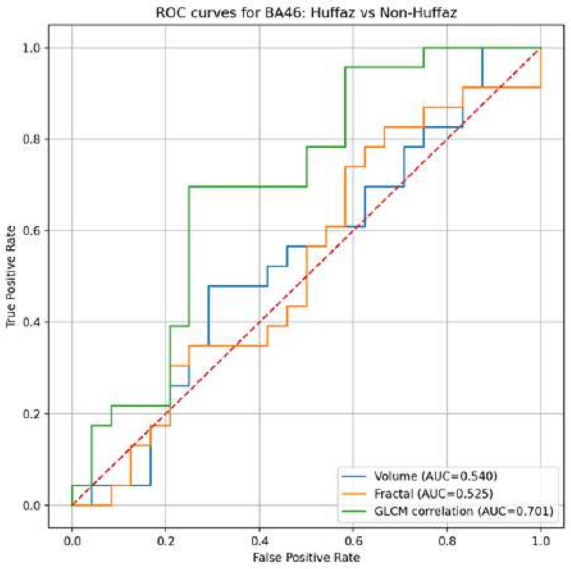

Figure 3. ROC curves for BA46 comparing three feature types for discrimination between Huffaz and non-Huffaz. GLCM Correlation showed the highest performance (AUC = 0.701), outperforming volume (AUC = 0.540) and fractal measurement (AUC = 0.525).

5 Discussion

5.1 Interpretation of the best-subset result in a hypothesis-driven, single-feature framework

The paper implemented a highly simplistic radiomics approach a single interpretable, second-order texture measure (GLCM Correlation) was computed on predefined Brodmann-area VOIs and inference was limited to a small set of candidate areas which were driven by previous neuroimaging studies of memorization, language, attention and executive control (BA22, BA24, BA32, BA40, BA46). In this restricted design, best-subset selection using BIC always found BA46 GLCM Correlation to be the best predictor offering the most favorable trade-off between model fit and complexity in the presence of sex as an a priori addition.

It is informative, in itself, that the null model (sex only) is only second in terms of BIC: it indicates that of the five candidate models, only BA46 contains additional predictive information sufficiently strong to overcome the penalty of BIC which is relatively severe on the addition of parameters.

To interpret this result, it is necessary to make a close distinction between (i) predictive utility in the candidate set being tested and (ii) biological specificity. The result of model selection is not that BA46 is the sole region that varies between Huffaz and non-Huffaz and that the changes in BA46 are the most potent brain correlate of memorization, in general. Instead it suggests that in the small range of regions and with the particular imaging preprocessing and discretization choices applied herein, BA46 texture organization when detected with GLCM Correlation provides a consistent signal that distinguishes the groups compared to the other candidate regions.

Such interpretation is consistent with the reasons why the analysis should be limited: multiplicity and degrees of freedom should be minimized in order to maximize interpretability and enhance reproducibility at small samples. The current work therefore contributes a "single-feature/single-region" candidate biomarker that is transparent enough to be interrogated, challenged, and refined. In radiomics and neuroimaging more broadly, such constrained models can be valuable as anchors for subsequent studies, particularly when prior literature suggests plausible involvement of frontal

executive circuitry in intensive memorization practice.

5.2 What does a BA46 "texture" signal mean? Conceptual meaning of GLCM Correlation in structural MRI

A key goal of this study was interpretability. GLCM Correlation is a Haralick texture measure that quantifies the linear dependency between neighboring discretized gray levels. In practical terms, within a VOI, higher GLCM Correlation typically reflects more predictable co-occurrence structure, where neighboring voxels' discretized intensities vary in a more linearly related manner; lower correlation reflects weaker linear dependency and potentially more complex, less linearly coupled local intensity relationships. Because the feature is computed from a discretized intensity field inside a region, it captures within-region organization rather than bulk volume.

In structural MRI (and especially tissue-class images produced by segmentation and modulation), voxel intensities can be interpreted as continuous-valued maps related to tissue composition or local density. Texture features computed on these images therefore summarize how these local intensity values arrange spatially. Importantly, texture features do not map one-to-one onto a single microstructural property. Instead, they may reflect a mixture of biological and technical sources: local cytoarchitecture and laminar organization, partial volume composition, segmentation properties, smoothing, registration and modulation effects, and discretization choices. That said, a stable, region-specific texture difference especially in a region with strong cognitive plausibility can be viewed as a candidate signature of distributed tissue organization differences associated with long-term training. Similar radiomics approaches have been used to characterize structural brain alterations in neurological disorders, including Alzheimer's disease and Parkinson's disease, where texture and intensity-derived features capture subtle tissue organization patterns beyond volumetric measures [24, 25].

In this study, the standardized BA46 coefficient was negative (OR $<$ 1), meaning that higher BA46 GLCM Correlation was associated with lower odds of being Huffaz. One interpretable hypothesis is that sustained memorization and related cognitive routines might be accompanied by subtle remodeling that changes the local heterogeneity and spatial dependency of intensity values in BA46. However, because preprocessing steps such as smoothing can alter fine-scale patterns, any biological interpretation must remain tentative until robustness checks across pipelines are performed.

5.3 Neurocognitive plausibility of BA46 in intensive memorization and recitation practice

BA46 is usually associated with dorsolateral prefrontal cortex (DLPFC), which plays an important role in executive functions such as working memory maintenance/manipulation, top-down attentional control, monitoring, and goal-directed behavior. This memorizing of high volumes of structured text does not just mean storage, but rather repetitive encoding, active retrieval, error checking, sequence, attentional control and such strategies that aid long-term memory. Huffaz practice may therefore engage a cognitive profile that overlaps with DLPFC functions: maintaining and updating sequences, sustaining attention during recitation, suppressing distractions, monitoring performance, and using structured retrieval cues.

In that sense, it is reasonable to consider that a BA46 signature exists and can be used to supplement previous research that has focused on temporal and parietal brain areas related to language understanding and phonological decoding (e.g., BA22 and BA40), and cingulate brain areas that are related to cognitive regulation and monitoring (e.g., BA24 and BA32). The existing outcome does not refute the potential of those areas being involved; on the contrary, it indicates that in the existing feature and analysis limitations, BA46 displays the most noticeable and repeatable texture organization difference.

There are at least three non-mutually exclusive interpretations of why BA46 might emerge in a texture-based case-control analysis:

- **Executive-control specialization hypothesis:** The long-term memorization reinforces the frontal executive systems of long-term, goal-oriented retrieval and sequencing. The structural adaptations may not be in the form of a massive volumetric expansion but may be in the form of subtle reorganization within the re-

gion that can be detected by the texture.

- **Strategy and metacognitive monitoring hypothesis:** Memorizers who are skilled might devise mechanisms that involve measuring, inspection and correction of errors, which are also often associated with the prefrontal circuitry. Alterations in local organization can indicate long-term involvement of such networks.
- **Distributed network compensation hypothesis:** Memorization can concern several areas, yet the fact that people have similarities despite variations in the approach might be the fact that they use the executive control systems. This might produce a more consistent BA46 signal across participants than language-related regions whose engagement may vary with recitation style, linguistic background, or training method.

These hypotheses motivate follow-up work that pairs the radiomics feature with behavioral measures (e.g., working memory capacity, retrieval speed, error rates during recall, attention control tasks) and explores whether BA46 texture correlates with training intensity, onset age, or cumulative practice hours.

5.4 Why did BA22, BA24, BA32, and BA40 not show comparable stability in this setting?

The lower selection frequency for BA22, BA24, BA32, and BA40 does not imply that these regions are irrelevant to memorization. Several methodological and substantive factors can reduce detectability.

5.4.1 Feature specificity mismatch

GLCM Correlation captures a specific kind of spatial dependency pattern. It is possible that other regions differ between groups but in ways better captured by other texture measures (e.g., contrast, homogeneity, entropy), intensity statistics, or morphometric indices rather than correlation.

5.4.2 Heterogeneity of functional roles and training strategies

Language comprehension and phonological processing (BA22, BA40) may be engaged differently depending on recitation practice, language proficiency, and teaching method. Such heterogeneity can dilute consistent structural signatures in small samples.

5.4.3 Preprocessing sensitivity

Spatial normalization, modulation, and smoothing can interact with regional anatomy differently across cortical areas. Some regions may be more susceptible to partial volume effects or segmentation variability, affecting texture features.

5.4.4 Candidate set restriction and penalized selection

BIC favors parsimonious models. If two regions share overlapping information (collinearity in texture measures), selection might favor the single region that most efficiently captures the signal, potentially pushing other regions out even if they contribute weakly.

Therefore, the present BA46-first result should be framed as a starting point for targeted replication rather than an exclusive localization claim.

5.5 Stability findings: what the bootstrap selection frequency adds beyond p-values

One of the objectives of the current study was to find a candidate predictor, but also to measure its resampling stability. In small to moderate datasets, classical p-values can be unstable, and variable selection can be sensitive to sampling noise. The bootstrap results provide a complementary view: BA46 was selected in approximately 0.68 of resamples, whereas other candidates were near 0.13–0.15. This pattern suggests that BA46 is not a one-off artifact of a single split or a single fit; instead, it is the region most consistently supported when the dataset is perturbed through resampling.

One should understand about 0.68 appropriately: it is not an absolute stability, and it does not imply that BA46 will be necessarily chosen. Instead, it shows a comparative advantage in the candidate set. In practice, when the sample size is small, a selection frequency that is significantly larger than that of the competitors can be an effective empirical measure of robustness, particularly when used together with a transparent single-feature method.

The interpretability and parsimony of the signal are also supported by the model-size distribution (one-predictor models are the most popular). This can

be advantageous for reproducibility: a single-region model is simpler to validate, easier to describe, and less likely to overfit than multi-region signatures with many correlated features.

5.6 Covariate adjustment, effect direction, and interpretational caution

All models were subjected to sex to correct any possible sex-related variation in brain structure and texture that could confound comparison of groups. The best BIC model also had no statistically significant sex difference with zero, and BA46 was significant under a Wald test. It implies that sex differences in the sample distribution do not explain the BA46 effect in a trivially important way. The failure to find a sex effect in this dataset, however, cannot be construed to mean that sex does not matter, but might be due to low power, feature employed or the difference in sex might be more manifested in other outcomes.

The trend of BA46 coefficient is worth discussing. The negative standardized coefficient means that the increase of BA46 GLCM Correlation is linked to a decrease in odds of Huffaz status. This direction is not easily interpreted biologically since increased correlation does not directly correspond to more healthy or more trained. Rather, it is a certain within-region intensity dependency trend. One reasonable narrative is that Huffaz status is associated with altered tissue organization in BA46 that changes local intensity relationships in a way that reduces linear dependency, consistent with increased heterogeneity or more complex local patterning. However, such interpretation has to be reserved till the robustness is tested with discretization decisions and preprocessing parameter, the details of feature computation are standardized and reported, and that the direction and strength is verified by independent replication.

5.7 Methodological contributions: why a single-feature, VOI-based radiomics design matters

The majority of radiomics studies use high-dimensional feature set which is usually in hundreds and even thousands of variables and then use regularization or feature selection. Even though it is powerful, it can decrease interpretability and raise the susceptibility to analytic flexibility. Conversely, the current paper focuses on a lean and auditable method:

- one feature that is defined (GLCM Correlation),
- a few anatomically determinable VOIs (Brodmann areas),
- a pre-specified candidate set to limit multiplicity,
- a model-selection criterion (BIC) that penalizes complexity, and
- a complexity-penalized model-selection criterion (BIC).

This design reduces degrees of freedom and makes it easier for other groups to replicate and test. In addition, focusing on Brodmann areas facilitates integration with cognitive neuroscience interpretations, since BA labels are commonly used to relate neuroimaging findings to known functional systems.

At the same time, the design also highlights trade-offs: by restricting the feature space, the approach may miss signals detectable by other features or regions. This trade-off is acceptable here because the goal is to establish a transparent candidate marker rather than maximize classification performance.

5.8 Technical considerations that can influence GLCM Correlation in structural MRI

A robust discussion of radiomics requires explicit acknowledgement of methodological sensitivities. Several factors can influence GLCM Correlation values.

5.8.1 Intensity scaling and harmonization

Even after preprocessing, intensity distributions can differ across subjects due to scanner factors, acquisition settings, or normalization choices. If intensity normalization is not consistent, discretization can map similar biological patterns to different gray-level bins.

5.8.2 Discretization strategy

Fixed bin width (as used here) versus fixed bin count can yield different gray-level distributions within VOIs. GLCM features are sensitive to this choice. Reporting bin width and any intensity normalization steps is essential for reproducibility.

5.8.3 GLCM parameters: offsets, directions, distance, symmetry, and averaging

GLCM Correlation can be computed with different spatial offsets (e.g., 1 voxel), across multiple directions (13 directions in 3D is common), and then averaged. Different implementations (2D slice-wise vs. 3D, symmetric vs. not, normalized vs. not) can change feature values. Explicit parameter reporting is necessary.

5.8.4 Smoothing and resampling

Smoothing can reduce high-frequency variation and may increase local predictability, potentially affecting correlation. Likewise, resampling VOI masks can alter boundary voxels and partial volume composition, impacting texture features.

Because these factors can create pipeline signatures, future work should include sensitivity analyses where BA46 remains the focus but preprocessing and discretization settings are varied systematically to test stability of the association.

5.9 Statistical considerations: small-sample inference, optimism, and generalizability

With $N = 47$, this study should be framed as exploratory/confirmatory within a constrained hypothesis set rather than definitive. Small samples increase uncertainty in effect estimates and can produce optimistic performance if classification accuracy is emphasized. The present study reduces overfitting risk by limiting predictors and using BIC, but several statistical issues remain relevant:

- **Coefficient uncertainty:** Confidence intervals can be wide even when p-values are below 0.05. Emphasizing effect sizes and uncertainty is important.
- **Selection-induced inference:** Inference after selection can inflate significance if not accounted for. The bootstrap stability partially addresses robustness, but formal post-selection inference is not performed here.
- **Class balance and calibration:** Although the groups are reasonably balanced (23 vs. 24), calibration of predicted probabilities and decision thresholds should be assessed in future work, especially if translation to screening is considered.
- **External validity:** Without an independent cohort, generalizability remains unknown.

To strengthen the statistical evidence, future studies could use nested resampling for performance estimation, preregistered analysis plans, and multi-site validation to test robustness across scanners and populations.

5.10 Integration with prior Huffaz neuroimaging and contribution beyond volume/VBM and fractal measures

Most work on neuroimaging by Prior Huffaz has been done on macrostructural measures like VBM-based gray matter measures and VOI volumes, and introduced fractal measures to quantify complexity. The present research is complementary to such methods because it addresses another aspect, which is second-order spatial intensity structure in predetermined cortical areas. This has a conceptual sense, in that two persons may be the same in the volume of the region, but differ in the internal structure of tissue-class intensities. The idea of texture can then give a potentially sensitive description of a distributed organizational difference which may not be represented as big changes in volume.

Instead of placing texture as an alternative to VBM or fractal descriptors, it can be presented as the other perspective of structural organization. These descriptors might be learned in a controlled way in an ideal future model (e.g. VBM, fractal measures, a small panel of stable texture features) to assess whether they represent similar or complementary changes in training-associated neuroplasticity.

5.11 Practical implications and how this could be used responsibly

The value of single-region, single-feature marker lies in the fact that a single-region, single-feature marker is easily implemented and interpreted compared to complex models. Should it be replicated, BA46 GLCM Correlation could be used as a hypothesis-based research biomarker to examine long-term memorization training, a research target in studying brain-behavior correspondence treatments involving imaging and performance measures, or

a feature candidate in multimodal models involving cognitive tests, training history, and imaging.

Nevertheless, there should be responsible interpretation. This paper does not suggest the use of MRI texture in identifying people in practice. Huffaz status is a learned skill and a social identity; translating imaging findings into individual-level classification would raise ethical concerns and would require rigorous validation, careful consideration of confounding variables (education, socioeconomic factors, bilingualism, religious practice patterns, sleep, stress), and clear benefit-risk justification.

5.12 Limitations and future work

The case-control design applied in this study was cross-sectional, and the sample size ($N = 47$) was rather small, which restricts the precision and the capacity to test interactions or non-linear effects. Only a single radiomics feature was analyzed by design; while this reduces multiplicity and improves interpretability, it may miss complementary information captured by other texture, shape, or intensity features. Future work should (i) validate the BA46 finding in an independent cohort, (ii) test whether the association generalizes across alternative preprocessing settings or discretization choices, and (iii) evaluate a broader radiomics panel under appropriate multiple-comparison control and preregistered hypotheses.

6 Conclusion

This paper has offered a purposely simplistic and interpretable MRI radiomics model to investigate the relationships between structural MRI texture and Huffaz status.Multiplicity was limited by using a single texture measure, GLCM Correlation, applied to literature-defined Brodmann Area VOIs on SPM-preprocessed smwc.nii images and inference was restricted to the five candidate regions (BA22, BA24, BA32, BA40, BA46). Best-subset selection under BIC with sex forced into all models consistently identified BA46 GLCM Correlation as the most informative regional predictor. Bootstrap resampling further supported this finding, with BA46 showing markedly higher selection stability than the other candidate regions, suggesting a reproducible single-region signal within the evaluated set.

Overall, the results indicate that BA46 texture organization, as captured by GLCM Correlation, may differentiate Huffaz from non-Huffaz after adjustment for sex. Future work should validate the BA46 association in larger independent cohorts, assess robustness to preprocessing and discretization settings, and evaluate whether combining complementary radiomics features or multimodal measures improves predictive performance while maintaining interpretability.

Declaration on Generative AI

ChatGPT 5.2 was used to assist with language editing and to improve clarity during manuscript preparation. The authors take full responsibility for the content, confirm that all scientific interpretations and conclusions are their own, and approve the final version of the manuscript.

Data Availability Statement

The structural MRI dataset used in this study is publicly available on the OpenNeuro repository (Accession Number: ds006521) and can be accessed at: https://doi.org/10.18112/openneuro.ds006521.v1.0.0

Acknowledgment

This work was supported by the Ministry of Higher Education of Malaysia under the Fundamental Research Grant Scheme (FRGS/1/2024/WAS12/UIAM/01/3).

Author Contributions

Conceptualisation, Mohd Izzuddin Mohd Tamrin, Mohd Zulfaezal Che Azemin and Asadullah Shah; methodology, Mohd Izzuddin Mohd Tamrin and Iqbal Jamaludin; software, Mohd Izzuddin Mohd Tamrin; validation, Iqbal Jamaludin, Mohd Zulfaezal Che Azemin and Asadullah Shah; formal analysis, Mohd Izzuddin Mohd Tamrin; investigation, Iqbal Jamaludin and Abdul Halim Sapuan; data curation, Mohd Izzuddin Mohd Tamrin and Iqbal Jamaludin; writing-original draft preparation, Mohd Izzuddin Mohd Tamrin; writing-review and editing, Mohd Izzuddin Mohd Tamrin, Iqbal Jamaludin, Abdul Halim Sapuan, Mohd Zulfaezal Che Azemin and Asadullah Shah; visualisation, Mohd Izzuddin Mohd Tamrin; supervision, Mohd Zulfaezal Che Azemin and Asadullah Shah All authors have read and agreed to the published version of the manuscript.

Compliance with Ethical Standards

The authors declare that they have no conflict of interest. Ethical approval for this study was obtained from the Institutional Ethics Committee of International Islamic University Malaysia (IIUM/504/14/11/3/IREC 654). Informed consent was obtained from all subjects involved in the study.

References

[1] I. Jamaludin, M. Z. Che Azemin, A. H. Sapuan, and R. Hassan, "Voxel-Based Morphometry Analysis of Gray Matter on Quran Memorizers," *IIUM Medical Journal Malaysia*, vol. 16, no. 1, 2017.

[2] A. H. Sapuan, N. S. Mustofa, M. Z. Che Azemin, et al., "Grey Matter Volume Differences of Textual Memorization: A Voxel Based Morphometry Study," in *International Conference for Innovation in Biomedical Engineering and Life Sciences (ICIBEL 2015)*, IFMBE Proceedings. Springer, 2015.

[3] I. Jamaludin, M. Z. Che Azemin, M. I. M. Tamrin, and A. H. Sapuan, "Volume of Interest-Based Fractal Analysis of Huffaz's Brain," *Fractal and Fractional*, vol. 6, no. 7, Art. 396, 2022, doi: 10.3390/fractalfract6070396.

[4] M. S. Azimi, M. Cheraghi, F. MahdiMaleki, et al., "Toward standardization: Assessing the reproducibility of radiomics features in partial volume-corrected brain PET images," *NeuroImage*, vol. 318, Art. 121398, 2025, doi: 10.1016/j.neuroimage.2025.121398.

[5] H. Fan, Y. Luo, F. Gu, et al., "Artificial intelligence-based MRI radiomics and radiogenomics in glioma," *Cancer Imaging*, vol. 24, Art. 36, 2024, doi: 10.1186/s40644-024-00682-y.

[6] J. Feng, Y. Huang, X. Zhang, et al., "Research and application progress of radiomics in neurodegenerative diseases," *Meta-Radiology*, vol. 2, no. 1, Art. 100068, 2024, doi: 10.1016/j.metrad.2024.100068.

[7] M. A. Pirozzi, F. Franza, M. Chianese, et al., "Combining radiomics and connectomics in MRI studies of the human brain: A systematic literature review," *Computer Methods and Programs in Biomedicine*, vol. 266, Art. 108771, 2025, doi: 10.1016/j.cmpb.2025.108771.

[8] M. Inglese, A. Conti, and N. Toschi, "Radiomics across modalities: A comprehensive review of neurodegenerative diseases," *Clinical Radiology*, vol. 85, Art. 106921, 2025, doi: 10.1016/j.crad.2025.106921.

[9] Q. Sun and F. Wang, "Using artificial intelligence and radiomics to analyze imaging features of neurodegenerative diseases," *Frontiers in Neurology*, vol. 16, Art. 1624867, 2025, doi: 10.3389/fneur.2025.1624867.

[10] A. Lin, Y. Chen, Y. Chen, et al., "MRI radiomics combined with machine learning for diagnosing mild cognitive impairment: A focus on the cerebellar gray and white matter," *Frontiers in Aging Neuroscience*, vol. 16, Art. 1460293, 2024, doi: 10.3389/fnagi.2024.1460293.

[11] Y. Li, P. Yi, M. Jin, et al., "A radiomics model predicts progression from mild cognitive impairment to Alzheimer's disease using structural MRI," *Scientific Reports*, vol. 15, Art. 25679, 2025, doi: 10.1038/s41598-025-19586-3.

[12] Z. Yuan, N. Qi, X. Chen, et al., "Magnetic resonance radiomics-based deep learning model for Alzheimer's disease diagnosis," *Digital Health*, vol. 11, 2025, doi: 10.1177/20552076251337183.

[13] M. Wennmann, L. T. Rotkopf, F. Bauer, et al., "Reproducible Radiomics Features from Multi-MRI-Scanner Test-Retest-Study: Influence on Performance and Generalizability of Models," *Journal of Magnetic Resonance Imaging*, vol. 61, no. 2, pp. 676–686, 2025, doi: 10.1002/jmri.29442.

[14] W. Zhao, Z. Hu, A. Kazerooni, et al., "Physics-Informed Discretization for Reproducible and Robust Radiomic Feature Extraction Using Quantitative MRI," *Investigative Radiology*, vol. 59, no. 5, pp. 359–371, 2024, doi: 10.1097/RLI.0000000000001026.

[15] S. Hong, S. Hong, E. Oh, et al., "Development of a flexible feature selection framework in radiomics-based prediction modeling: Assessment with four real-world datasets," *Scientific Reports*, vol. 14, Art. 29297, 2024, doi: 10.1038/s41598-024-80863-8.

[16] H. Aguirre-Meneses, P. Stoehr-Muñoz, M. Molina-Gonzalez, et al., "Radiomics and the Image Biomarker Standardisation Initiative (IBSI): A Narrative Review Using a Six-Question Map and Implementation Framework for Reproducible Imaging Biomarkers," *Cureus*, vol. 17, no. 10, 2025, doi: 10.7759/cureus.95335.

[17] B. Koçak, A. Borgheresi, A. Ponsiglione, et al., "Explanation and Elaboration with Examples for CLEAR (CLEAR-E3), an EuSoMII Radiomics Auditing Group Initiative," *European Radiology Experimental*, vol. 8, no. 1, Art. 72, 2024, doi: 10.1186/s41747-024-00471-z.

[18] J. Zhang, X. Teng, X. Zhang, et al., "Comparing effectiveness of image perturbation and test-retest imaging for radiomic feature repeatability," *Scientific Reports*, vol. 13, Art. 18263, 2023, doi: 10.1038/s41598-023-45477-6.

[19] F. Pignotti, T. Ius, R. Russo, et al., "Development and validation of a MRI-radiomics-based machine learning approach in High Grade Glioma to detect early recurrence," *Frontiers in Oncology*, vol. 14, Art. 1449235, 2024, doi: 10.3389/fonc.2024.1449235.

[20] S. Yin, J. Ming, H. Chen, et al., "Integrating deep learning and radiomics for preoperative glioma grading using multi-center MRI data," *Scientific Reports*, vol. 15, Art. 36756, 2025, doi: 10.1038/s41598-025-20711-5.

[21] Z. Cheng, L. Yang, C. Liang, et al., "Advancing cerebral small vessel disease diagnosis: Integrating quantitative susceptibility mapping with MRI-based radiomics," *Human Brain Mapping*, vol. 45, no. 13, Art. e70022, 2024, doi: 10.1002/hbm.70022.

[22] C. Li, D. Hui, F. Wu, et al., "Automatic diagnosis of Parkinson's disease using artificial intelligence based on T1-weighted MRI," *Frontiers in Medicine*, vol. 10, Art. 1303501, 2024, doi: 10.3389/fmed.2023.1303501.

[23] G. Antonopoulos, S. More, F. Raimondo, et al., "A systematic comparison of VBM pipelines and their application to age prediction," *NeuroImage*, vol. 279, Art. 120292, 2023, doi: 10.1016/j.neuroimage.2023.120292.

[24] S. Aghajanian, F. Mohammadifard, I. Mohammadi, et al., "Longitudinal structural MRI-based deep learning and radiomics features for predicting Alzheimer's disease progression," *Alzheimer's Research & Therapy*, vol. 17, no. 1, Art. 182, 2025, doi: 10.1186/s13195-025-01827-2.

[25] Z. Chen, S. Bi, Y. Shan, et al., "Multiparametric hippocampal signatures for early diagnosis of Alzheimer's disease using ^{18}F-FDG PET/MRI radiomics," *CNS Neuroscience & Therapeutics*, vol. 30, no. 4, Art. e14539, 2024, doi: 10.1111/cns.14539.