Machine learning-based detection of ischemic stroke in older adults using CT brain

Siti Aishah Abdul Aziz[1], Nur Amirah Atikah Sharuddin[1], Nur Asma Sapiai[2], Norfataha Mohd Daud[3], Muhammad Akmal Remli[4]

[1]School of Health Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia, [2]School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia, [3]Radiology Department, Hospital Pakar Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia, [4]Institute for Artificial Intelligence and Big Data, Universiti Malaysia Kelantan, 16150 Kubang Kerian, Kelantan, Malaysia

ABSTRACT

Introduction: Ischemic stroke is a leading cause of morbidity and mortality among older adults. This study proposes a machine learning (ML)-driven approach to assist in the early detection of ischemic stroke using non-contrast computed tomography (CT) brain images. The optimized model aims to assist radiologists in diagnosing ischemic strokes more accurately, thereby reducing misdiagnosis and treatment delays, and ultimately improving patient outcomes. **Materials and Methods:** CT brain images from 184 patients aged between 18 to 60 years old were studied retrospectively. Data consist of 25 normal (N) CT brain scans from Hospital Pakar USM and 159 ischemic strokes (IS) from the Acute Ischemic Stroke Detection (AISD) dataset. All data were manually verified by a consultant radiologist and annotated as either ischemic or non-ischemic. The images were anonymised and organised into class-specific folders for batch processing during model training (90% of the dataset: 22 N & 143 IS) and testing (10% of the dataset: 3 N & 16 IS). Initial preprocessing steps includes intensity normalisation, denoising, and resolution adjustment were done after standardising to MNI space via SPM12. Radiomic features related to texture, shape, and intensity were extracted using the PyRadiomics to distinguish between ischemic and healthy brain tissue. The region of interest (ROI) for each image was identified using Otsu's thresholding, followed by morphological operations applied to refine the brain mask. Gray-Level Co-occurrence Matrix (GLCM) and Local Binary Patterns (LBP) were used to enhance the performance of stroke detection models. Three supervised ML classifiers: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF) were compared based on their proven effectiveness in medical image classification. Model performance was assessed using 10-fold cross-validation and evaluated based on accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC). **Results:** Random Forest achieved the highest diagnostic accuracy (accuracy = 93.75%, AUC = 0.973) with a training time of 0.208 seconds, followed by KNN (accuracy = 94.6%, AUC = 0.94, training time = 0.011 seconds) and SVM (accuracy = 87.5%, AUC = 0.984, training time = 0.015 seconds). The Friedman test confirmed statistically significant differences in classification performance among the models ($\chi^2 = 10.75$, $p = 0.00463$, $\alpha = 0.05$). Post-hoc Wilcoxon signed-rank tests with Bonferroni correction showed a significant difference between SVM and KNN ($p = 0.0156$), indicating KNN significantly outperformed SVM. However, the differences between SVM and RF ($p = 0.0313$) and between KNN and RF ($p = 0.75$) were not statistically significant, indicating comparable performance between KNN and RF. These findings align with existing literature, where KNN and RF have demonstrated similar performance in stroke detection tasks (Wang et al., 2023). Statistical validation from 10-fold cross-validation showed that KNN (mean accuracy = 93.93%) and RF (mean accuracy = 92.76%) exhibited high and comparable performance, outperforming SVM (mean accuracy = 87.87%). **Conclusion:** These promising results demonstrated that ML models can produce consistent outcomes across varied presentations of ischemic lesions. This suggests strong potential for automated stroke detection in geriatric medicine primary care and emergency settings.

Keywords: Ischemic stroke, computed tomography, machine learning, geriatric medicine