Available Online at http://www.journalajst.com

ASIAN JOURNAL OF
SCIENCE AND TECHNOLOGY

ISSN: 0976-3376

Asian Journal of Science and Technology
Vol. 12, Issue, 11, pp. 11929-11933, November, 2021

RESEARCH ARTICLE

A COMPARATIVE ANALYSIS OF MACHINE LEARNING ALGORITHMS TO CLASSIFY BRAIN DISEASES

[1,*]Saida D and [2]Kasuba Tulasi

[1]Research Scholar, Dept. of Computer Science & Engineering, College of Engineering, Osmania University, Hyderabad-500007

[2]Associate Professor, Dept. of Computer Science & Engineering, Siddhartha Institute of Engg & Technology, Ibrahim Patnam, Ranga Reddy-501510

ARTICLE INFO

Article History:
Received 17th August, 2021
Received in revised form
15th September, 2021
Accepted 20th October, 2021
Published online 28th November, 2021

Key words:
Brain Disease,
Classification,
Magnetic Resonance Imaging,
Neurodegenerative, Taxonomy.

ABSTRACT

Brain diseases are neurodegenerative diseases that cause nerves and brains to degenerate over time. The term "neurodegenerative" refers to the loss of brain cells in areas of the brain where dopamine is normally produced. It causes confusion, alters personality, and then destroys brain nerves and tissues. Magnetic Resonance Imaging (MRI) is a popular radiology technique for diagnosing brain diseases. The goal of this review is to present the main research methods of machine learning-based classification for brain diseases, as well as to provide relevant learning and reference for interested researchers. Meanwhile, it summarises the major flaws in existing methods and provides better direction for future research. The taxonomy of existing brain disease detection methods is also presented, as are the problems that have arisen as a result of the existing methods. The identification methods of machine learning that minimize the difficulty for humans can be used to understand the automatic classification and detection of brain diseases. This review lays the groundwork for effective clinical diagnosis, treatment planning, and precise quantitative evaluation of brain disease detection.

Citation: Saida D and Kasuba Tulasi, 2021. "A comparative analysis of machine learning algorithms to classify brain diseases", *Asian Journal of Science and Technology*, 11, 04, xxxx-xxxx.

INTRODUCTION

Brain diseases, also known as neurodegenerative diseases, are clinical manifestations of memory degradation and impairment in cognitive functions, which effectively endangers one's safety and quality of life (1). The term "neurodegenerative" refers to the loss of brain cells in areas of the brain that normally produce dopamine. Whereas, in affected brain 60% to 80% of cells that produce dopamine will be lost and brain cannot produce required amount of dopamine and results in brain diseases such as brain tumor, brain stroke, Parkinson's disease, Alzheimer disease, etc (2). Individuals will notice problems speaking, memory loss, and basic cognitive skills 20 years before the onset of diseases due to structural variations in the brain. The symptoms are caused by neuron damage in various areas of the brain that are responsible for thinking, cognitive function, memory, and learning (3).

***Corresponding author:* Saida D**
Research Scholar, Dept. of Computer Science & Engineering, College of Engineering, Osmania University, Hyderabad-500007.

Brain diseases are becoming more common around the world, and the symptoms that appear two years later tend to be lifelong persistent (4). A timely diagnosis can help to alleviate symptoms and provide significant functionality. As a result, developing an accurate and timely diagnosis system for brain diseases is essential However, brain disease diagnosis is difficult due to its highly heterogeneous symptoms and complex nature (5). The MRI is the most commonly used imaging technique for the brain in clinical practice, allowing doctors to determine the medical condition of the brain and identify the presence of specific diseases (6). The structural MRI provides non-invasive solutions for detecting abnormal changes in the brain and aids in the identification of brain disease imaging bio-makers for medical applications (7). The purpose of medical image analysis is to visualise and extract details from medical images, which is accomplished through image segmentation, matching, and classification analysis segmentation, matching and classification. Existing methods relied on either problem-specific or manual feature extraction or target-oriented feature extraction, resulting in larger and more complex medical image analyses (8).

Machine learning techniques have piqued the interest of researchers in investigating brain abnormalities in specific regions of patients (9). The classification of diseases in machine learning is a method of categorising a specific set of data into classes. The classification can be applied to both structured and unstructured data. The classification process begins with predicting the classes of specific data points. Labels, categories, and targets are the common names for these classes. Existing methods classify images as healthy or unhealthy, but methods that classify brain images into multiple stages are uncommon (10). As a result, an effective brain disease classification approach that categorises brain images into multiple categories is required.

TAXONOMY PROCEDURE: In this research study, there are four major steps involved namely collection of data, pre-processing, segmentation, feature extraction and finally, classification. The general block diagram for this research study is shown in Figure 1.

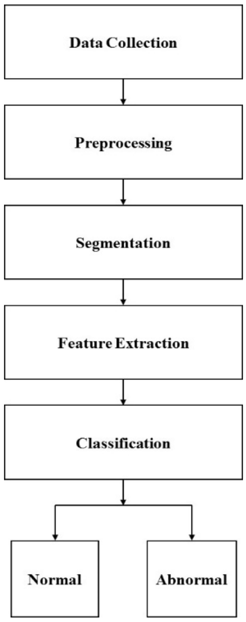

Figure 1. The general block diagram of brain disease classification method

Preprocessing: After collecting the datasets, the next step was preprocessing where the collected images were modified to enhance the image quality.
The preprocessing techniques such as noise reduction, brain extraction, bias field correction, image registration are used. The noise reduction is the process of decreasing the local variant region noise that is noticed in the MRI. The scripting of skull to extract the tissue of brain from the other tissue is carried out to determine the voxels as non-brain or brain. The contrast of brain MRI images is fine-tuned by using the field correction. The image registration can be used to transform the alignments of image from spatial into anatomical space.

Segmentation: After preprocessing the datasets, the process of segmentation was undergone. The objective of image segmentation is to divide the images into well-defined regions. The segmentation techniques such as Principle Component Analysis (PCA), K-Means, Fuzzy C-Means can be used. The PCA is used to carry out an linear orthogonal transformation of multiple sets of data that includes variables correlated with an larger dimensions and other which includes new variables which are not correlated with smaller dimensions.

The K-Means is the deterministic approach on the basis of unsupervised learning that makes data to divide into group of k clusters. The Fuzzy C-Means is also deterministic approach on the basis of unsupervised learning that represents an advanced version of K-Means. The Fuzzy C-Means is based on the subsets of fuzzy that give rise to the concept of adhesion by using membership function.

Feature Extraction: After segmenting the images, the process of feature extraction takes place. The feature extraction is the methods that combines or selects the variable into features that effectively reduce the amount of data need to be processed. The feature extraction techniques such as Grey Level Co-occurrence Matrix (GLCM), Scale Invariant Feature Transforms (SIFT), Wavelet transform can be used. The GLCM is the technique for extracting the second order statistical features of texture and it is the matric where number of columns and rows are equal to number of gray levels. The SIFT is the technique used to detect the features and describes the local features in the brain images. In SIFT, the key points are extracted first from the group of images and stores in the database. The wavelet transform is an image processing technique used for features detection and classification where the wavelets will be applied to analyze the images.

Classification: After extracting the features from the brain images, the classification brain disease takes place. The classification technique such as Support Vector Machine (SVM), Random Forest (RF), Naïve bayes, Decision Tree (DT), K-Nearest Neighbour (KNN) can be used to classify the brain diseases. The SVM is a supervised machine learning classifier that is used for classification of brain diseases from the images and it performs binary and multiple classifications of datasets.

The RF is a supervised learning algorithm for classification and regression. RF creates the decision trees on the selected random data and predicts the best solutions. The DT is the classification model that forms a structure of tree which breaks the data into smaller subsets. The KNN is the classification technique where it utilizes the similarity of features to identify the values of new data points. The classifiers algorithm classifies the brain disease images such as brain tumor, brain stroke, parkinson's disease, Alzheimer disease, etc, into normal or abnormal brain images.

LITERATURE REVIEW

The existing techniques of brain disease detection and classification is reviewed in this section. The advantages and limitation of existing techniques is also described in the following:

Salama A. Mostafa et.al.(11) developed a multiple feature evaluation approach to enhance the diagnosis performance of Parkinson's disease. The developed approach includes 5 agents, every agent evaluates and ranks the features then produced the subsets of feature vectors. The features were weighted and ordered the subsets of features on the basis of individual evaluation result. Then, features were filtered and produced the number of preliminary copies of optimized feature. In developed method, the classifiers algorithm shows the highest diagnosis accuracy by utilizing filtered datasets.

However, the developed method excluded the runtime from the process of evaluation.

Yongzhao Xu et.al.(12) developed an deep learning based improved method for analyzing the brain CT scan of hemorrhagic stroke patient. The enhanced deep learning method classifies the CT images of the skull by utilizing deep learning networks, classified CT images into uninjured stroke or hemorrhagic stroke. After the image classification, mask Region based Convolutional Neural Networks (RCNN) which segments the stroke by a learning transfer method and also by using machine learning methods. The advantage of the developed method was it efficiently classified and segmented the hemorrhagic stroke images. However, preprocessing step need to be included for accurate classification of images.

Jiahui Zhang et.al.(13) developed Surface Based Morphometric (SBM) method to detect the Parkinson's disease in patients having mild cognitive impairments. The 93 patients with Parkinson's disease data was collected and 276 medical image factors from MRI were analyzed and compared by SBM to construct the model. Then, support vector machine with grid identification method was used to find the Parkinson's disease. The support vector machine with SBM features reached a higher accuracy for identification. However, the multimodal features were not included for the analysis in the developed method.

Zehra Karapinar Senturk et.al.(14) developed machine learning based method for the early detection and diagnosis of parkinson's disease. The feature selection based decision support model was established by using features of voice signals from parkinson's patients and healthy people for diagnosis of Parkinson's early. Various feature selection techniques and various classification methods were utilized in the developed method. The developed method shown higher accuracy in classification and reduced the computational cost. However, the datasets collected in the developed method was only included vocal features.

Pavan Rajkumar Magesh et.al.(15) used local interpreTable model agnostic explainer method for the early detection of Parkinson's disease by Single Photon Emission Computed Tomography (SPECT) Dopamine Transporter (DaTSCAN) imaging method. The DaTSCAN were used from the Parkinson's progression markers initiative database and was trained on convolutional neural network by transfer learning method. The developed method increased the accuracy in early identification of disease and provided confidence for healthcare purpose in computer aided diagnosis. However, the developed method showed higher false positive rate and negative rate.

P. Sathish et.al.(16) Developed a clustering method called as Gaussian Hybrid Fuzzy Clustering (GHFC) for segmentation. Initially, the exponential cuckoo on the basis of radial neural network was utilized for the process of classification. Then, the images were taken for pre-processing and carried out for the segmentation. Further, the segmentation was carried out by establishing a hybridized clustering GHFC framework. The developed GHFC was created by combining the FCM and Sparse FCM to identify the effectiveness of centroid for classification. However, the computation time of the work is 6.09seconds, which requires deep learning strategy for classification.

Lu Li et.al.(17) developed a U-Net based deep learning approach for automatic identification and segmentation of hemorrhage strokes in brain CT images. The input from the network was constructed by combining the flipped images with the original CT slice which introduces the symmetry constraints of the brain images into the established model. To reduce the effect of the normal area in the brain image, the conditional Generative Adversarial Nets (GAN) were modified by using UNet as a generator. Then the discriminator was used to exploit higher-order inconsistencies in the samples synthesized by the generator. The advantage of the UNet based method was it showed higher accuracy for the detection of hemorrhage strokes. However, inspecting all the images was difficult to find the hemorrhage lesion part from the CT image.

U. Rajendra Acharya et.al.(18) developed automatic identification of ischemic strokes by utilizing the higher order spectrum features in the MRI images of brain. The higher order spectrum bi-spectrum entropy and phase feature were utilized to extract the salient details from the MRI images for the strokes estimation. The features were extracted by using Higher Order Spectra (HOS) and many entropies. Then, estimated feature were ranked by using the F-value during analysis of variance. The advantage of the developed method was effectively detected the stroke lesion and showed maximum accuracy and predictive value. But, the automatic detection of ischemic stroke method will not identify the high dimensional medical image data.

Asit Subudhi et.al.(19) developed automated classification and segmentation of brain strokes by utilizing expectation maximization and RF classifier on diffusion weighted images sequence of MRI. The stroke was classified into partial anterior circulation syndrome, lacunar syndrome and total anterior circulation stroke. The affected region of brain was due to stroke segmented by using expectation maximization algorithm and segmented regions was processed with fractional order Darwinian particle swarm optimization method to improve the rate of detection. The advantage of the automated segmentation and classification method was it efficiently detected the lesion structure on a large dataset. However, automated segmentation and classification was time consuming since it trains the larger sets of data to the model.

C. Narmatha et al. (20) implemented a Fuzzy Brain-Storm Optimization (FBSO) algorithm for segmentation and classification of brain tumors. The developed FBSO method includes four process such as image acquisition, image enhancement, feature extraction and classification. The developed FBSO was robust, efficient, and reduced the segmentation time of the optimization algorithm with accuracy of 93.85% and nearly 95% of F-score. However, the FBSO method didn't consider the important features, which consumes only 93.85% of accuracy. Therefore, the developed method requires effective feature extraction techniques.

COMPARATIVE ANALYSIS

Comparison analysis of existing brain disease classification and detection methods in terms of performance metrics is described in Table 1.

Brain disease Classi fication with performance metrics

Authors	Methodology	Datasets	Performance Measures
Chenjie Ge et.al.(21)	Developed an Generative Adversarial Network (GAN) for the classification of brain tumor.	Used 3 dimensional brain images such as TCGA-GBM and TCGA-LGG.	Accuracy= 88.825%, Sensitivity= 81.81% and specificity = 92.17%.
Hossam H. Sultan et.al.(22)	Implemented a deep learning approach based on Convolutional Neural Network (CNN)to classify types of brain tumors.	Utilized REMBRANDT dataset and TCIA dataset.	Accuracy= 97.54%, Sensitivity= 95.5%, Specificity= 98.7%, and precision = 95.8%.
Biao Jie et.al.(23)	Designed CNN method with functional connectivity network for the diagnosis of brain disease.	ADNI dataset was used.	Accuracy= 88.0%
Hossein Shahamat et.al.(24)	Developed an Genetic Algorithm based Brain Masking (GABM) for brain disease classification.	Utilized ADNI dataset and ABIDE dataset.	Accuracy= 85%.
M. Rohini et.al.(25)	Developed multivariate linear regression, Support Vector Machine (SVM) and logistic regression.	ADNI dataset was used in the developed method.	Accuracy = 89% and AUC score = 78%.

CONCLUSION

In this study, the existing brain disease detection methods by using machine learning approaches is reviewed. The accurate and early diagnosis system for the brain diseases is necessary to develop. But, the diagnosis of brain disease is difficult due to its highly heterogeneous symptoms and complex nature. The existing methods utilized either problem specific or manual feature extraction or target oriented feature extraction which resulted in larger and complex analysis of medical images. The existing methods automated segmentation and classification was time consuming since it trains the larger sets of data to the model. The existing methods classifies the images as healthy or unhealthy, the methods that classifies the brain images into multiple stages were rare. So, it is necessary to develop an effective brain disease classification approach that classifies the brain images into multiple categories. In order to solve such problems better model need to be developed for brain disease detection.

REFERENCES

(1) Yang, Z. and Liu, Z., 2020. The risk prediction of Alzheimer's disease based on the deep learning model of brain 18F-FDG positron emission tomography. *Saudi journal of biological sciences*, *27*(2), pp.659-665.

(2) Senturk, Z.K., 2020. Early diagnosis of Parkinson's disease using machine learning algorithms. *Medical hypotheses*, *138*, p.109603.

(3) Qiu, S., Joshi, P.S., Miller, M.I., Xue, C., Zhou, X., Karjadi, C., Chang, G.H., Joshi, A.S., Dwyer, B., Zhu, S. and Kaku, M., 2020. Development and validation of an interpreTable deep learning framework for Alzheimer's disease classification. *Brain*, *143*(6), pp.1920-1933.

(4) Smyser, C.D., Dosenbach, N.U., Smyser, T.A., Snyder, A.Z., Rogers, C.E., Inder, T.E., Schlaggar, B.L. and Neil, J.J., 2016. Prediction of brain maturity in infants using machine-learning algorithms. *NeuroImage*, *136*, pp.1-9.

(5) Bilgen, I., Guvercin, G. and Rekik, I., 2020. Machine learning methods for brain network classification: Application to autism diagnosis using cortical morphological networks. *Journal of neuroscience methods*, *343*, p.108799.

(6) Sathish, P. and Elango, N.M., 2020. Gaussian hybrid fuzzy clustering and radial basis neural network for automatic brain tumor classification in MRI images. *Evolutionary Intelligence*, pp.1-19.

(7) Liu, M., Zhang, J., Lian, C. and Shen, D., 2019. Weakly supervised deep learning for brain disease prognosis using MRI and incomplete clinical scores. *IEEE transactions on cybernetics*, *50*(7), pp.3381-3392.

(8) Choi, B.K., Madusanka, N., Choi, H.K., So, J.H., Kim, C.H., Park, H.G., Bhattacharjee, S. and Prakash, D., 2020. Convolutional neural network-based mr image analysis for Alzheimer's disease classification. *Current Medical Imaging*, *16*(1), pp.27-35.

(9) Tian, Z.Y., Qian, L., Fang, L., Peng, X.H., Zhu, X.H., Wu, M., Wang, W.Z., Zhang, W.H., Zhu, B.Q., Wan, M. and Hu, X., 2020. Frequency-Specific Changes of Resting Brain Activity in Parkinson's Disease: A Machine Learning Approach. *Neuroscience*, *436*, pp.170-183.

(10) Huang, G.H., Lin, C.H., Cai, Y.R., Chen, T.B., Hsu, S.Y., Lu, N.H., Chen, H.Y. and Wu, Y.C., 2020. Multiclass machine learning classification of functional brain images for Parkinson's disease stage prediction. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *13*(5), pp.508-523.

(11) Mostafa, S.A., Mustapha, A., Mohammed, M.A., Hamed, R.I., Arunkumar, N., Abd Ghani, M.K., Jaber, M.M. and Khaleefah, S.H., 2019. Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease. *Cognitive Systems Research*, *54*, pp.90-99.

(12) Xu, Y., Holanda, G., Fabrício, L., de F, S., Silva, H., Gomes, A., Silva, I., Ferreira, M., Jia, C., Han, T. and de Albuquerque, V.H.C., 2020. Deep Learning-Enhanced Internet of Medical Things to Analyze Brain CT Scans of Hemorrhagic Stroke Patients: A New Approach. *IEEE Sensors Journal*.

(13) Zhang, J., Li, Y., Gao, Y., Hu, J., Huang, B., Rong, S., Chen, J., Zhang, Y., Wang, L., Feng, S. and Wang, L., 2020. An SBM-based machine learning model for identifying mild cognitive impairment in patients with Parkinson's disease. *Journal of the Neurological Sciences*, *418*, p.117077.

(14) Senturk, Z.K., 2020. Early diagnosis of Parkinson's disease using machine learning algorithms. *Medical hypotheses*, *138*, p.109603.

(15) Magesh, P.R., Myloth, R.D. and Tom, R.J., 2020. An Explainable Machine Learning Model for Early Detection of Parkinson's Disease using LIME on DaTSCAN Imagery. *Computers in Biology and Medicine*, *126*, p.104041.

(16) Sathish, P. and Elango, N.M., 2020. Gaussian hybrid fuzzy clustering and radial basis neural network for automatic brain tumor classification in MRI images. *Evolutionary Intelligence*, pp.1-19.

(17) Li, L., Wei, M., Liu, B., Atchaneeyasakul, K., Zhou, F., Pan, Z., Kumar, S., Zhang, J., Pu, Y., Liebeskind, D.S. and Scalzo, F., 2020. Deep Learning for Hemorrhagic

Lesion Detection and Segmentation on Brain CT Images. *IEEE Journal of Biomedical and Health Informatics*.

(18) Acharya, U.R., Meiburger, K.M., Faust, O., Koh, J.E.W., Oh, S.L., Ciaccio, E.J., Subudhi, A., Jahmunah, V. and Sabut, S., 2019. Automatic detection of ischemic stroke using higher order spectra features in brain MRI images. *Cognitive systems research*, *58*, pp.134-142.

(19) Subudhi, A., Dash, M. and Sabut, S., 2018. Automated segmentation and classification of brain stroke using expectation-maximization and random forest classifier. *Biocybernetics and Biomedical Engineering*, *40*(1), pp.277-289.

(20) Narmatha, C., Eljack, S.M., Tuka, A.A.R.M., Manimurugan, S. and Mustafa, M., 2020. A hybrid fuzzy brain-storm optimization algorithm for the classification of brain tumor MRI images. *Journal of Ambient Intelligence and Humanized Computing*, pp.1-9.

(21) Ge, C., Gu, I.Y.H., Jakola, A.S. and Yang, J., 2020. Enlarged training dataset by pairwise gans for molecular-based brain tumor classification. *IEEE Access*, *8*, pp.22560-22570.

(22) Sultan, H.H., Salem, N.M. and Al-Atabany, W., 2019. Multi-classification of brain tumor images using deep neural network. *IEEE Access*, *7*, pp.69215-69225.

(23)Jie, B., Liu, M., Lian, C., Shi, F. and Shen, D., 2020. Designing weighted correlation kernels in convolutional neural networks for functional connectivity based brain disease diagnosis. *Medical image analysis*, *63*, p.101709.

(24) Shahamat, H. and Abadeh, M.S., 2020. Brain MRI analysis using a deep learning based evolutionary approach. *Neural Networks*, *126*, pp.218-234.

(25) Rohini, M. and Surendran, D., 2020. Toward Alzheimer's disease classification through machine learning. *Soft Computing*, pp.1-9.

****