ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# MACHINE LEARNING FOR ALZHEIMER DETECTION: A COMPREHENSIVE APPROACH

**CHINNU MARY GEORGE[1],SANTHOSH MENON[2]**

[1]Lecturer, Computer Engineering and Informatics, Middlesex University, Dubai, UAE
[2]Lecturer, Computer Engineering and Informatics, Middlesex University, Dubai, UAE

Email: c.m.george@mdx.ac.ae, s.menon@mdx.ac.ae

**ABSTRACT**

Alzheimer's disease (AD) represents a significant and growing concern worldwide, particularly among older adults, as it remains the leading cause of dementia. The increasing incidence rates of AD, along with its profound impact on individuals, families, and healthcare systems, highlight the urgent need for effective diagnostic tools. AD is characterized by progressive neurodegenerative changes within the brain, making early detection critical for effective treatment and minimizing potential damage. Given the challenges of predicting AD in its initial stages, this research explores various Machine Learning (ML) models, including Support Vector Machine (SVM), Random Forest (RF), and Gradient Boosting (XGBoost), to develop accurate prediction models. Utilizing datasets from Kaggle, this study employs two distinct feature extraction methods: Local Binary Patterns (LBP) and Discrete Wavelet Transform (DWT). Both feature sets are fed into ML models, and the performance of these models is evaluated using essential metrics, including accuracy, precision, F1 score, True Positive Rate (TPR), True Negative Rate (TNR), False Positive Rate (FPR), and False Negative Rate (FNR).Among the six evaluated models, the combination of the XGBOOST model with DWT features stood out, proving to be the most effective in predicting Alzheimer's Disease emerging as the standout performer, achieving the highest accuracy rate of 97.88%. This research underscores the potential of ML in early AD detection, offering a promising avenue for improving patient outcomes and alleviating the societal, financial, and economic burdens associated with this devastating condition.

**Keywords:** *Alzheimer, Kaggle, Local Binary Pattern, Discrete Wavelet Transform, Machine Learning, XGBoost, Accuracy*

## 1. INTRODUCTION

The human brain, an incredibly intricate organ within the body, serves numerous crucial functions, such as generating ideas, solving problems, making decisions, fostering creative thinking, and storing and recalling memories. Of these functions, memory plays a particularly vital role as it acts as a repository for our life experiences, significantly influencing our character and identity. The experience of memory loss, often associated with conditions like dementia, can be profoundly distressing. Specifically, Alzheimer's disease (AD) stands out as the most prevalent form of dementia, and as individuals grow older, the fear of developing Alzheimer's becomes increasingly intense [1]. AD progressively harms brain cells, leading to a disconnection from one's surroundings, the loss of cherished memories, childhood recollections, the ability to recognize family members, and even basic skills like following instructions [2, 3].Advanced stages bring about the loss of abilities like swallowing, coughing, and breathing [4]. Dementia affects around 50 million individuals globally, with the associated healthcare and social care costs ranking as equivalent to the 18th largest economy in the world. Additionally, it is estimated that by the year 2050, the yearly count of newly diagnosed AD and dementia cases will increase threefold, reaching a total of 152 million cases [5]. This translates to a new case of dementia occurring approximately every 3 seconds. Diagnosing AD can be challenging due to symptoms that resemble those of typical aging or vascular dementia (VD) [6]. Early and accurate diagnosis is crucial for efforts related to prevention,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

treatment, and patient management, as it allows for the monitoring of disease advancement [7, 8].

Images play a crucial role across various scientific domains, with medical imaging emerging as a potent tool for comprehending brain functions. Magnetic Resonance Imaging (MRI) is employed in medical diagnostics to visualize the brain's structure and functionality [9]. Physicians assess AD symptoms and conduct various tests to diagnose dementia, including laboratory tests, brain imaging, and memory assessments. These tests are helpful in eliminating the possibility of other conditions with comparable symptoms. MRI scans can identify abnormalities in the brain associated with mild cognitive impairment (MCI) and forecast which MCI patients are at a higher risk of developing AD in the future. With the progression of technology and the increasing availability of brain imaging data, machine learning (ML) and deep learning (DL) are assuming an ever more significant role in extracting precise and relevant data. Consequently, this allows for accurate predictions of AD based on brain imaging information. Numerous research initiatives concentrate on utilizing brain imaging to detect AD, employing various ML techniques for AD classification. The research typically comprises three stages: 1) the collection and processing of images, 2) feature extraction from these processed images, and 3) the development and evaluation of classification models.

The format of the paper is as outlined below: Section 2 examines previous studies on AD diagnosis and categorization, Section 3 describes the theoretical concept of feature extraction and ML algorithm, Section 4 shows experimental results and evaluations, and Section 5 concludes the paper by addressing possible future areas of study.

## 2. LITERATURE SURVEY

In the realm of AD detection and classification using ML and DL techniques, several noteworthy studies have been conducted and presented in this section. According to a study [11], the primary objective was to create an effective computational approach to pre-process and categorize AD, particularly in its initial phases. The study utilized multifractal geometry to capture the most dynamic characteristics linked to AD. Following that, a machine learning algorithm known as K-Nearest Neighbour (KNN) was utilized to classify the four principal early stages of AD. The method achieved exceptional results, with a higher rate of accuracy and sensitivity than recently developed methods. Research work by [12] aimed to detect and classify AD using KNN classifiers. Two groups were involved, medium and Weighted KNN classifiers, each consisting of 30 samples. The data set was sourced from kaggle.com, containing Alzheimer's and normal conditions. The study measured classifier performance and Weighted KNN classifiers demonstrating substantial superiority in AD identification and classification. In the paper [13], an improved lightweight DL model for AD detection was proposed, utilizing MRI images. This model combined feature extraction and classification into a single stage, simplifying the system with only seven layers. The approach achieved high accuracy for binary and multi-classification tasks, surpassing previous models, using a Kaggle dataset with a limited size.

The research [14] presented a hybrid KNN and SVM for the early identification of Alzheimer and Parkinson's disease. This technique combined the best features of parametric and non-parametric techniques, resulting in superior classification accuracy. Testing on ADNI, OASIS, and NTUA PD datasets demonstrated superior accuracy and specificity compared to popular DL algorithms. In the study [15], a hybrid model integrating Particle Swarm Optimisation (PSO) and Convolutional Neural Networks (CNNs) was presented for AD detection. The accuracy for brain illnesses was enhanced by using PSO to optimize CNN hyper-parameters, and the loss function score was minimized. Experiments with benchmark datasets demonstrated the model's superior accuracy rates. Experiment [16] utilized MRI images and clinical data from the AD Neuroimaging Initiative dataset to detect different stages of Alzheimer's and predict the conversion from MCI to Alzheimer's. Various ML and DL techniques were applied, achieving binary and multi-classification of AD, Late MCI, Early Cognitive Impairment, and Cognitive Control. In the study [17], an advanced DL-based system was introduced to detect AD early. Using a substantial MRI sample with normal and affected subjects, the study successfully classified subjects into three classes: MCI, AD, and Normal. Various classification approaches, including SVM and Deep Neural Network (DNN) algorithms, achieved impressive accuracy levels, ranging from 80% to 90%, in predicting AD. The research emphasizes the potential of highly accurate computational-automated ML tools for early disease diagnosis.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Study [18] presented four systems designed to track various stages of AD development. These systems employed different methodologies and materials. The first system used feed-forward neural networks (FFNN)and artificial neural networks (ANNs) based on hybrid feature extraction methods. The second system employed two pre-trained DL models, AlexNet and ResNet-18. The subsequent system used a combination of ResNet-18 and AlexNet algorithms to retrieve features from a dataset, and SVM for categorization. Hybrid ANN/FFNN algorithms were utilized in the final system. All of these methods have shown impressive efficacy for the initial AD diagnosis. The study [19] investigated the use of DL architectures for the categorization of brain areas identified using Automated Anatomical Labelling (AAL). For training deep belief networks, images of grey matter (GM) were segmented into 3D patches using AAL-defined areas. These networks were then combined into an ensemble to construct a robust categorization framework. Using a large dataset from the Alzheimer's Disease Neuro imaging Initiative (ADNI), the approach was able to successfully categorize people with MCI and distinguish between normal and AD images. In the study [20], neuroimaging modalities were used to investigate the efficacy of longitudinal data analysis, AI, and ML techniques. The importance of extracting features from neuroimaging data, pinpointing sensitive brain areas, and determining biomarker cut-off values was highlighted. The study's primary objective was to refine automated methods for detecting the onset of AD disease and better understanding how the disease develops.

## 3. METHODLOGY

This section delves into the concept of feature extraction techniques such as LBP and DWT, as well as the operational principles of machine learning models including SVM, RF, and XGBoost.

### A. Feature Extraction

**LBP:** The LBP has gained prominence as a highly effective local feature descriptor [21], particularly in the context of image recognition. For the purpose of labelling individual pixels, the Local Binary Pattern (LBP) operator makes use of the intensity value as a threshold, comparing it to the pixel values within a 3 * 3 neighbourhood. The outcome is then understood as a binary numeral. Typically, LBP is calculated using $P$ sampling points ($x_p \in (0, \ldots\ldots, p-1)$) positioned at a radial distance denoted as $R$ from the central pixel $x_m(i_c, j_c)$.

$$LBP_{P,R} = \sum_{p=0}^{p-1} t_s \,.\, (x_p - x_m).\, 2^p \quad [1]$$

$$t_s(diff) = \begin{cases} 1, & (diff) \geq 1 \\ 0, & (diff) < 0 \end{cases} \quad [2]$$

The function $t_s(diff)$ in the formula represents a threshold function. Bilinear algorithms are used to interpolate sample points p that are not accurately located inside the immediate region of the central pixel. Ojala et al. [22] developed the idea of "uniform patterns," wherein a binary pattern is considered uniform if it displays not exceeding two consecutive transitions from 0 to 1 if viewed as a circle. This notion gave rise to the development of "uniform" patterns, which are regarded as fundamental patterns in local image textures.

$$U(LBP_{P,R}) = |S(x_{p-1} - x_m) - s(x_0 - x_m)| + \sum_{p=1}^{p-1} | s(x_p - x_m) - s(x_{p-1} - x_m)| \quad [3]$$

When $LBP_{P,R}$ is transformed into $LBP_{P,R}^{u2}$, the subscript $u2$ denotes that the uniform patterns $U(LBP)$ have maximum values of 2. There are $P * (p-1) + 2$ uniform patterns, whereas the other non-uniform patterns are grouped into one class, yielding a feature dimension of $P * (p-1) + 3$. The numerical values of all pixels in an input image $x_L(i,j)$ are gathered and organised into a histogram after the LBP labelling is applied to the image. This histogram can be written as:

$$H_l = \sum_{i,j} F\{x_L\,(i,j) = 1\}, \quad l = 0,1,2, \ldots\ldots, n-1 \quad [4]$$

$$F\{A\} = \begin{cases} 1, & if A is true \\ 0, & if A is false \end{cases} \quad [5]$$

The number of unique labels the LBP algorithm generates is denoted by $n$ here. In the case of $LBP_{8,1}^{u2}$, for instance, there are a grand total of 59 features to consider. An image's localized organization of dots and edges can be deduced from the LBP histogram, which is a collection of micro-patterns.

**DWT:** The DWT is a mathematical transformation that involves sampling wavelets at discrete intervals [23]. It provides a unique perspective by recording spatial and frequency domain data of an image at the same time. As part of the DWT procedure, an image is analysed using decimation and filter techniques. There are low-pass (LPF) and high-pass filters (HPF) integrated into the analytical filter library at each level of decomposition. While the HPF concentrates on sharper features like edges, the LPF pulls out more generalized features from the image [24]. Low-frequency information is stored in the 1D DWT's approximate coefficients, whereas high-frequency information is stored in the detail coefficients. In 2D DWT, the input image is broken down into 4 frequency bands: low-frequency vertical and horizontal elements, high-frequency vertical and horizontal elements, low-frequency vertical and high-frequency horizontal elements, and high-frequency vertical and low-frequency horizontal elements. It is also possible to use the notations LL, LH, HL, and HH to refer to these sub-bands. The sub-band depiction of an image I in the context of a 1-D DWT is given by

$$I = I_a^1 + \{I_h^1 + I_v^1 + I_d^1\} \quad [6]$$

where $I_a^1$ represents the approximation of the input image, and $I_h^1, I_v^1, I_d^1$ represents the horizontal, vertical, and diagonal details, respectively. For deeper decomposition levels, further steps involve successive decomposition of the LL sub-band, resulting in multiple bands. In the case of a 5-level DWT decomposition, the representation of the image is given by

$$I = I_a^5 + \sum_{i=1}^{5}\{I_h^1 + I_v^1 + I_d^1\} \quad [7]$$

*B. ML Model*

**XGBOOST:** XGBoost stands as one of the implementations of XGBoost machines, recognized for its exceptional performance in supervised learning tasks [25]. This versatile algorithm can be applied for classification tasks, making it a preferred choice among data scientists. XGBoost's popularity stems from its remarkable speed in execution, even when dealing with large datasets and out-of-core computation. Here's how XGBoost operates: Consider a dataset, DS, comprising m features and n examples, denoted as $= \{(x_i, y_i): i = 1 \dots n, x_i \in \mathbb{R}^m, y \in \mathbb{R}\}$. In this context, $\widehat{y_i}$ represents the predicted output generated by an ensemble of tree models. These tree models are defined as follows:

$$A_\iota = \emptyset(X_\iota) = \sum_{k=1}^{K} f_k(x_\iota), f_k \in \mathcal{F}$$

Here, $K$ denotes the number of trees within the model, and $f_k$represents the individual k-th tree.

To efficiently address the given equation, the objective is to identify the best collection of functions that minimizes both the loss and regularization criteria:

$$L(\emptyset) = \sum_i l(y_i, A_i) + \sum_k \Omega(f_k) \quad [8]$$

In this equation, $l$indicates the loss function, quantifying the disparity between the predicted $\widehat{y_i}$ and the actual output $y_i$. On the other hand, Ω serves as a measure of model complexity, helping prevent overfitting by assessing the model's intricacy. The complexity is determined by the following formula:

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda||w||^2 \quad [9]$$

In this equation, $T$ represents the count of leaves within the tree, and w stands for the weight assigned to each leaf. To enhance the performance of the model during training, boosting is utilized. This involves the incremental addition of new functions (trees) to the model. During each iteration ($t^{th}$), a new function (tree) is introduced in the following manner:

$$L^{(t)} = \sum_{i=1}^{n} l\left(y_i, A_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \quad [10]$$

$$L_{split} = \frac{1}{2}\left[\frac{(\sum_{i\epsilon I_L} g_i)^2}{\sum_{i\epsilon I_L} h_i+\lambda} + \frac{(\sum_{i\epsilon I_R} g_i)^2}{\sum_{i\epsilon\ R} h_i+\lambda} - \frac{(\sum_{i\epsilon I} g_i)^2}{\sum_{i\epsilon I} h_i+\lambda}\right] - \gamma \quad [11]$$

$$g_i = \partial_{A_i^{(t-1)}} l\left(y_i, A_i^{(t-1)}\right) \quad [12]$$

$$h_i = \partial^2_{A_i^{(t-1)}} l\left(y_i, A_i^{(t-1)}\right) \quad [13]$$

**SVM:** The SVM method, introduced by Vapnik et al. in 1995, has proven to be a highly successful predictive tool in both classification and regression tasks [26]. SVM consistently seeks the optimal global solution, ensuring it converges to the same result in each run. Its operation involves mapping training data into a high-dimensional space, where it endeavours to find a classifier capable of maximizing the margin between two distinct data classes. Essentially, SVM aims to locate the optimal separator function, often referred to as the best hyperplane among countless possibilities. The objective of SVM is to create a decision structure that cleanly divides the training data into the appropriate classes, in accordance with the idea of structural risk reduction. The key to SVM's decision-making lies in the selection of support vectors, which are the crucial elements within the training sample.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Imagine a collection of N points that are linearly separable, represented as $S = \{x_i \varepsilon R_n | i = 1, 2, \ldots, N\}$, where each point $x_i$ is assigned to one of two classes labeled as $y_i \in \{-1, +1\}$. A separating hyperplane partitions the set S into two regions, with each region exclusively containing points from a single class [27]. This separating hyperplane can be characterized by the pair $(w, b)$that meets the equation:

$$w.x + b = 0 \quad [14]$$

For each $i = 1, 2, \ldots, N$, the following conditions hold:

$$\begin{cases} w.x_i + b \geq +1, if\ y_i = +1 \\ w.x_i + b \leq -1, if\ y_i = -1 \end{cases} \quad [15]$$

Here, the dot product operation $(\cdot)$ is employed between $x$ and $w$vectors. Finding the optimal separating hyperplane (OSH) that maximizes the margin on both sides is the primary focus of SVM training. This optimization can be expressed as the minimization of ½ $w$. In the classification process, SVM relies on the OSH instead of the entire training dataset to make decisions. The OSH test pattern's location is only identified. This feature of SVM renders it highly viable with respect to computational effectiveness and predicted accuracy.

**RF:** RF comprises a collection of k classification trees, employing the concept of aggregating multiple weak classifiers into a robust classifier [28]. These classification trees are composed of different nodes, with the root node initially representing the training dataset. Each internal node serves as a weak classifier, tasked with dividing the samples based on a particular attribute. Meanwhile, each leaf node corresponds to labelled training or test data, which is utilized to classify input data into separate categories. RF's ultimate decision outcome is determined by aggregating the optimal decisions made by all the classification trees.The key to RF's operation lies in the utilization of the Gini Index (GI) to determine the optimal binary split point for a given feature [29]. The GI, denoted as $G_{gini}(D)$, quantifies the uncertainty present within the dataset $D$. In this classification problem involving $N$ classes and a set of samples $D$, the GI is defined as:

$$G_{gini}(D) = 1 - \sum_{n=1}^{N} \left(\frac{|C_n|}{D}\right)^2 \quad [17]$$

Here, $C_n$represents the group of samples within the dataset D that belong to the nth class. If we split the dataset $D$ into two parts, namely $D_1$ and $D_2$, it's done based on whether the feature A has a value of "a" or not.

$$D_1 = \{(x, y) \in D | A(x) = a\} \quad [18]$$

$$D_2 = D - D_1 \quad [19]$$

The conditional GI for feature A is defined as:

$$G_{gini}(D, A) = \frac{|D_1|}{|D|} G_{gini}(D_1) + \frac{|D_2|}{|D|} G_{gini}(D_2) \quad [20]$$

The value $G_{gini}(D, A)$represents the level of uncertainty within the dataset $D$ after dividing it by the attribute $A = a$. When creating a classification tree as part of the RF framework, the attribute with the lowest GI, along with the corresponding optimal binary split point, is chosen. The process of constructing a Random Forest involves the following steps:

- Using the bootstrap resampling method, a $k^{th}$ sample set, denoted as $D_k$, is drawn from the original dataset $D$. For each kth classification tree, a random vector $\theta_k$ is generated independently and identically distributed with the previously generated random vectors.
- For each of the $k$ samples, classification trees are built. This tree-building process is recursive, and it involves selecting the attribute with the smallest GI to split the binary tree.
- The final classification outcome is determined using a voting mechanism, which takes into account the results obtained from each of the classification trees.

## 4. RESULT AND DISCUSSION

This section offers a detailed and comprehensive overview of the outcomes achieved by ML models when employing features extracted through LBP and DWT. It provides insights into how these features contribute to the performance and results of the ML models.

### *A. Data Collection and Process*

The dataset used for Alzheimer's detection was sourced from Kaggle [30]. It consists of various categories, including Non-Dementia, Very Mild Dementia, Mild Dementia, and Moderate Dementia. Figure 1 provides a visual representation of sample images from each category. For the training phase, there were 2560 samples of Non-Dementia, 1792 samples of Very Mild Dementia, 717 samples of Mild Dementia, and 52 samples of Moderate Dementia. For testing, the dataset included 640 samples of Non-Dementia, 448 samples of Very Mild Dementia, 179 samples of

Journal of Theoretical and Applied Information Technology
29th February 2024. Vol.102. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Mild Dementia, and 12 samples of Moderate Dementia. The detailed data distribution can be found in Table 1, and Figure 2.

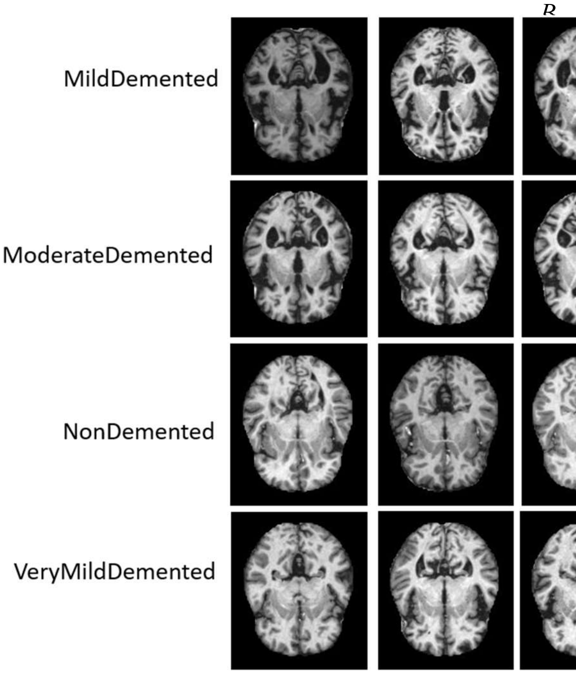



*Fig. 1. Sample Alzheimer Images From Kaggle*

*Table 1. Various AD Data Distribution*

| DISEASE | TRAIN | TEST |
|---|---|---|
| Non Dementia | 2560 | 640 |
| Moderate Dementia | 52 | 12 |
| Mild Dementia | 717 | 179 |
| Very Mild Dementia | 1792 | 448 |

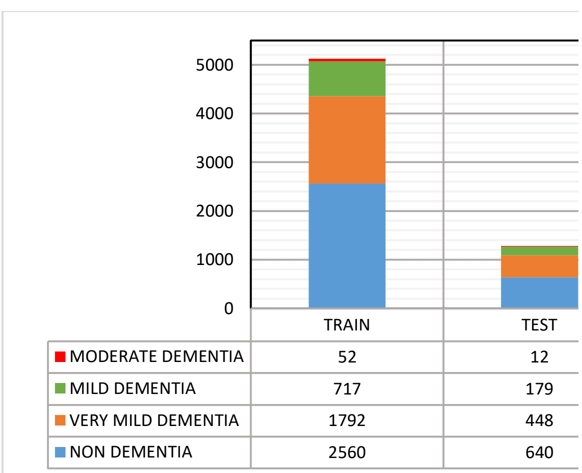



*Fig. 2. AD Data Count Plot*

*B. Model Evaluation*

Features were extracted from the processed data using LBP and DWT. Initially, the LBP features were utilized in the ML model for Alzheimer's classification. The metrics chosen are accuracy, precision, F1 score, TPR, TNR, FPR, and FNR. SVM achieved metrics of 95.15%, 96.26%, 95.22%, 94.2%, 96.14%, 3.85%, and 5.79%. RF yielded similar results with metrics values of 93.35%, 94.1%, 93.45%, 92.81%, 93.92%, 6.08%, and 7.18%. Meanwhile, XGBOOST exhibited metrics values of 95.54%, 96.42%, 95.6%, 94.8%, 96.32%, 3.68%, and 5.19%. The metrics values of the ML model using LBP data are presented in Figure 3.a.

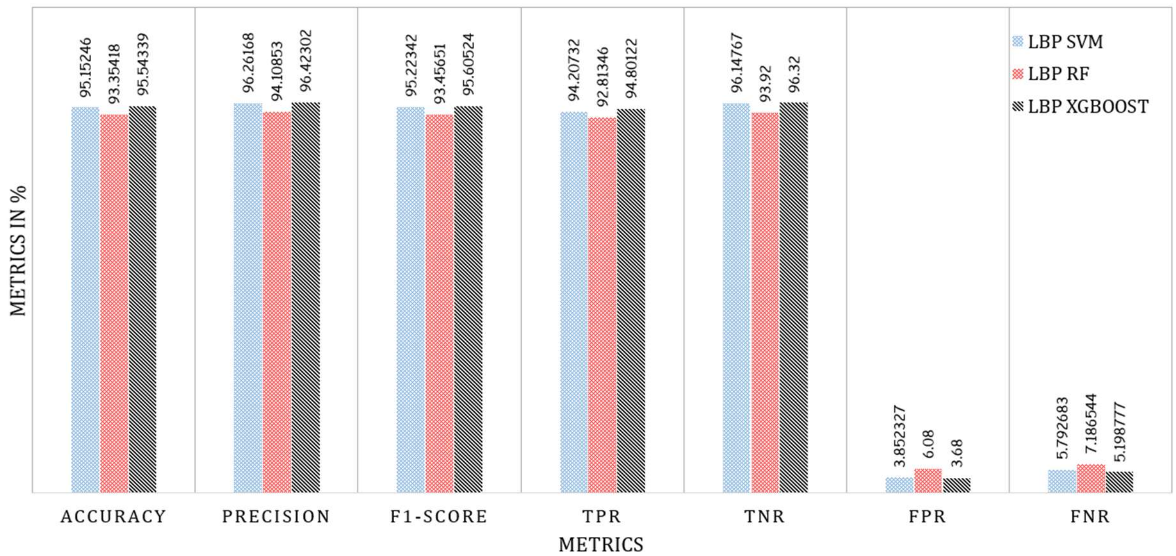
Subsequently, the ML model was provided with features obtained through DWT for Alzheimer's classification. SVM achieved metrics values of 96.63%, 97.19%, 96.66%, 96.14%, 97.14%, 2.85%, and 3.85%. RF showed metrics values of 94.13%, 95.18%, 94.4%, 93.63%, 94.69%, 5.3%, and 6.36%. Finally, XGBOOST attained metrics values of 97.88%, 98.13%, 97.91%, 97.68%, 98.09%, 1.9%, and 2.31%. The metrics values of the ML model using DWT data are depicted in Figure 3.b.

Table 2 provides a comprehensive comparison of performance metrics for ML models using both LBP and DWT features in AD classification. This analysis sheds light on the strengths and weaknesses of each model and feature extraction technique.

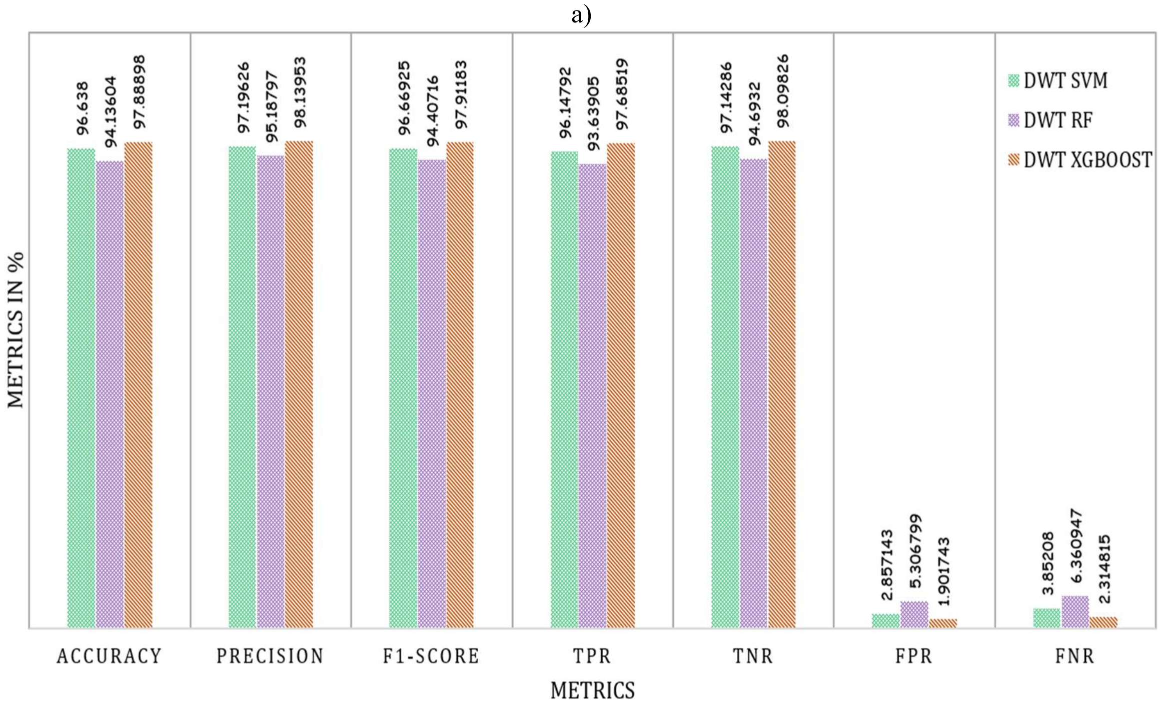
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

*Table 2. Performance Metrics Comparison*

| FEATURE EXTRACTION | MODEL | ACCURACY | PRECISION | F1-SCORE | TPR | TNR | FPR | FNR |
|---|---|---|---|---|---|---|---|---|
| LBP | SVM | 95.15246 | 96.26168 | 95.22342 | 94.20732 | 96.14767 | 3.852327 | 5.792683 |
| | RF | 93.35418 | 94.10853 | 93.45651 | 92.81346 | 93.92 | 6.08 | 7.186544 |
| | XGBOOST | 95.54339 | 96.42302 | 95.60524 | 94.80122 | 96.32 | 3.68 | 5.198777 |
| DWT | SVM | 96.638 | 97.19626 | 96.66925 | 96.14792 | 97.14286 | 2.857143 | 3.85208 |
| | RF | 94.13604 | 95.18797 | 94.40716 | 93.63905 | 94.6932 | 5.306799 | 6.360947 |
| | XGBOOST | 97.88898 | 98.13953 | 97.91183 | 97.68519 | 98.09826 | 1.901743 | 2.314815 |




a)




b)

*Fig. 3. ML Model Evaluation*

While SVM and RF proved their competence, XGBOOST, coupled with DWT features, emerged as the top-performing model. This finding holds significant promise for improving the accuracy and efficiency of AD detection, offering hope for better clinical outcomes and contributing to our understanding of this debilitating condition.

## 5. CONCLUSION

In this study, the power of Machine learning to tackle the critical challenge of AD detection. AD is a global healthcare concern, and early diagnosis is essential for improving patient care and outcomes. Leveraging a dataset sourced from Kaggle, we employed two distinct feature extraction methods: LBP and DWT. These features were then used as inputs for ML models, leading to the exploration of six different model-feature combinations. Our findings underscored the significance of these efforts in AD detection. Notably, the XGBOOST model, when combined with DWT features, emerged as the most reliable predictor among the six models. This research focuses the importance of machine learning in early Alzheimer's disease detection, offering a promising avenue for improving patient outcomes and alleviating the societal, financial, and economic burdens associated with this devastating condition.

## REFERENCE

[1]. French, Samantha L., Mark Floyd, Stacy Wilkins, and Sheryl Osato. "The fear of Alzheimer's disease scale: A new measure designed to assess anticipatory dementia in older adults." *International journal of geriatric psychiatry* 27, no. 5 (2012): 521-528.

[2]. Evans, Michelle Lynn. "Alzheimer’s Disease: More Than Just Memory Loss: An Awareness of the Disease and an Active Plan of Coping Strategies and Caregiving." (2016).

[3]. Powell, Lenore S. "Alzheimer's disease: a practical, psychological approach." *Women & Health* 10, no. 2-3 (1986): 53-62.

[4]. Mitchell, Susan L. "Advanced dementia." *New England Journal of Medicine* 372, no. 26 (2015): 2533-2540.

[5]. Mura, Thibault, J-F. Dartigues, and Claudine Berr. "How many dementia cases in France and Europe? Alternative projections and scenarios 2010–2050." *European journal of neurology* 17, no. 2 (2010): 252-259.

[6]. Kalaria, Raj. "Similarities between Alzheimer's disease and vascular dementia." *Journal of the neurological sciences* 203 (2002): 29-34.

[7]. Porsteinsson, A. P., R. S. Isaacson, Sean Knox, M. N. Sabbagh, and I. Rubino. "Diagnosis of early Alzheimer’s disease: clinical practice in 2021." *The journal of prevention of Alzheimer's disease* 8 (2021): 371-386.

[8]. Chu, L. W. "Alzheimer’s disease: early diagnosis and treatment." *Hong Kong Medical Journal* 18, no. 3 (2012): 228.

[9]. Ding, Zhaohua, Ran Xu, Stephen K. Bailey, Tung-Lin Wu, Victoria L. Morgan, Laurie E. Cutting, Adam W. Anderson, and John C. Gore. "Visualizing functional pathways in the human brain using correlation tensors and magnetic resonance imaging." *Magnetic resonance imaging* 34, no. 1 (2016): 8-17.

[10]. Gaser, Christian, Katja Franke, Stefan Klöppel, Nikolaos Koutsouleris, Heinrich Sauer, and Alzheimer's Disease Neuroimaging Initiative. "BrainAGE in mild cognitive impaired patients: predicting the conversion to Alzheimer’s disease." *PloS one* 8, no. 6 (2013): e67346.

[11]. Elgammal, Yasmina M., M. A. Zahran, and Mohamed M. Abdelsalam. "A new strategy for the early detection of alzheimer disease stages using multifractal geometry analysis based on K-Nearest Neighbor algorithm." *Scientific Reports* 12, no. 1 (2022): 22381.

[12]. Kaur, Manbir, Chintan Thacker, Laxmi Goswami, T. R. Thamizhvani, Imad Saeed Abdulrahman, and A. Stanley Raj. "Alzheimer’s Disease Detection using Weighted KNN Classifier in Comparison with Medium KNN Classifier with Improved Accuracy." In *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, pp. 715-718. IEEE, 2023.

[13]. El-Latif, Ahmed A. Abd, Samia Allaoua Chelloug, Maali Alabdulhafith, and Mohamed Hammad. "Accurate Detection of Alzheimer’s Disease Using Lightweight Deep Learning Model on MRI Data." *Diagnostics* 13, no. 7 (2023): 1216.

[14]. Shaffi, Noushath, ViswanVimbi, Mufti Mahmud, Karthikeyan Subramanian, and Faizal Hajamohideen. "Bagging the Best: A Hybrid SVM-KNN Ensemble for Accurate and Early Detection of Alzheimer’s and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Parkinson's Diseases." In *International Conference on Brain Informatics*, pp. 443-455. Cham: Springer Nature Switzerland, 2023.

[15]. Ibrahim, Rahmeh, Rawan Ghnemat, and Qasem Abu Al-Haija. "Improving Alzheimer's Disease and Brain Tumor Detection Using Deep Learning with Particle Swarm Optimization." *AI* 4, no. 3 (2023): 551-573.

[16]. Nagarathna, C. R., and M. M. Kusuma. "Early detection of Alzheimer's Disease using MRI images and deep learning techniques." *Alzheimer's & Dementia* 19 (2023): e062076.

[17]. Raees, PC Muhammed, and Vinu Thomas. "Automated detection of Alzheimer's Disease using Deep Learning in MRI." In *Journal of Physics: Conference Series*, vol. 1921, no. 1, p. 012024. IOP Publishing, 2021.

[18]. Abunadi, Ibrahim. "Deep and hybrid learning of MRI diagnosis for early detection of the progression stages in Alzheimer's disease." *Connection Science* 34, no. 1 (2022): 2395-2430.

[19]. Ortiz, Andres, Jorge Munilla, Juan M. Gorriz, and Javier Ramirez. "Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease." *International journal of neural systems* 26, no. 07 (2016): 1650025.

[20]. Aberathne, Iroshan, Don Kulasiri, and Sandhya Samarasinghe. "Detection of Alzheimer's disease onset using MRI and PET neuroimaging: longitudinal data analysis and machine learning." *Neural Regeneration Research* 18, no. 10 (2023): 2134-2140.

[21]. Pietikäinen, Matti, Abdenour Hadid, Guoying Zhao, and Timo Ahonen. *Computer vision using local binary patterns*. Vol. 40. Springer Science & Business Media, 2011.

[22]. Ojala, Timo, Matti Pietikainen, and Topi Maenpaa. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." *IEEE Transactions on pattern analysis and machine intelligence* 24, no. 7 (2002): 971-987.

[23]. Williams, John R., and Kevin Amaratunga. "A discrete wavelet transform without edge effects using wavelet extrapolation." *Journal of Fourier analysis and Applications* 3 (1997): 435-449.

[24]. Vishwanath, Mohan, and Robert Michael Owens. "A common architecture for the DWT and IDWT." In *Proceedings of International Conference on Application Specific Systems, Architectures and Processors: ASAP'96*, pp. 193-198. IEEE, 1996.

[25]. Chen, Tianqi, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mitchell, Ignacio Cano, and Tianyi Zhou. "XGBoost: extreme gradient boosting." *R package version 0.4-2* 1, no. 4 (2015): 1-4.

[26]. Vapnik, V. (1995). The nature of statistical learning theory. New York: SpringerVerlag.

[27]. Yue, Shihong, Ping Li, and Peiyi Hao. "SVM classification: Its contents and challenges." *Applied Mathematics-A Journal of Chinese Universities* 18 (2003): 332-342.

[28]. Rigatti, Steven J. "Random forest." *Journal of Insurance Medicine* 47, no. 1 (2017): 31-39.

[29]. SivaramKrishnan, M., Rajini, A. R., Logu, K., Kumarasamy, M., Jayaprakash, S., Gandhi, R. R., & Ramkumar, M. S. (2022, October). Leaf disease identification using machine learning models. In AIP Conference Proceedings (Vol. 2519, No. 1). AIP Publishing.