https://www.ijsrtm.com
Vol.2 Issue 1 March 2022: 08-13
Published online 11 March 2022
E-ISSN: 2583-7141

International Journal of Scientific Research in Technology & Management

Alzheimer's Disease Detection using Support Vector Machine

Vikash Kumar
Computer Science & Engineering
SAM College of Engineering & Technology
Bhopal, Madhya Pradesh, India
vikash8084060357@gmail.com

Devendra Rewadikar
Computer Science & Engineering
SAM College of Engineering & Technology
Bhopal, Madhya Pradesh, India
deverewadikar@gmail.com

Abstract*— A neurological disease that progresses over time, Alzheimer's disease (AD) is characterized by behavioral abnormalities, memory loss, and cognitive impairment. For management and intervention to be successful, early detection is essential. In order to diagnose Alzheimer's disease early, this research investigates the use of Support Vector Machines (SVM), a supervised machine learning approach. We evaluate the effectiveness of SVM classifiers, examine several characteristics taken from neuroimaging data and clinical evaluations, and talk about the clinical implications of our results. Encouragement For tasks involving regression and classification, supervised learning models called vector machines are employed. Their method involves locating the best hyperplane in a high-dimensional space to divide data points belonging to several classifications. SVMs work especially well with high-dimensional data, which makes them appropriate for use in neuroimaging applications.

Keywords*— Alzheimer's Disease, Machine Learning, Neuroimaging, Early Diagnosis, Deep Learning, Biomarkers, Supervised Learning.

I. INTRODUCTION

A major contributor to dementia, Alzheimer's disease affects millions of people globally. The importance of early detection techniques has increased as prompt diagnosis can have a substantial influence on the effectiveness of therapy and the quality of life of the patient. Conventional diagnostic techniques, such as clinical assessments and neuropsychological testing, can be laborious and subjective. Support vector machines, in particular, are a type of machine learning technology that presents a potential path for improving diagnostic efficiency and accuracy. The most common cause of dementia globally, Alzheimer's disease (AD) has a profound effect on people, families, and healthcare systems. Timely diagnosis and therapy are essential for preventing illness development. However, using conventional diagnostic techniques like clinical assessments and cognitive tests might be difficult to detect AD in its early stages. A proposed method to improve the precision and effectiveness of AD diagnosis is machine learning (ML). In huge datasets, ML models may find patterns and connections that traditional methods can miss. It looks at unsupervised learning techniques that can help in managing complicated data and comprehending patient groups, such clustering and dimensionality reduction [1]. The article examines specific applications of ML for AD detection in neuroimaging and genomic data analysis, aside from these methodological discussions. We discuss the practical applications of these approaches and highlight significant related studies that demonstrate their usefulness. We conclude by discussing the challenges and goals of the discipline, emphasizing the need for larger, high-quality datasets, improved interpretability of models, and more integrated approaches that combine different data sources and machine learning techniques [2].

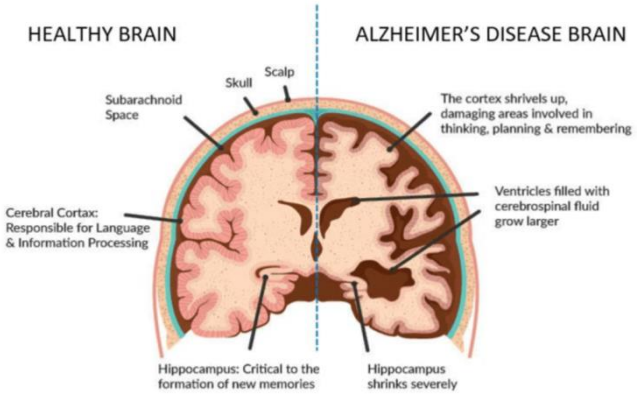

Fig. 1. Alzheimer Disease v/s Healthy Brain

Figure 1 Healthy v/s Alzheimer disease [3].

As a neurodegenerative condition, Alzheimer's disease describes a specific pattern of cognitive and functional

deterioration that starts and progresses over time and is linked to changes in a specific neuropathology. In order to counter this, a new discipline called nanotechnology is being developed. It makes use of nanoparticles that function at the molecular level and has a significant influence on a range of pharmaceutical and biological applications, especially in the treatment of Alzheimer's disease. The blood-brain barrier (BBB) may be crossed by nanocomposites and chimeric peptides, which are used to functionalize this technology at the nanoscale level and enable medication delivery to the central nervous system. On the other hand, there are several ways to administer nanocomposites, which work directly on the brain with no adverse consequences [3].

II. Related Works

In recent years, machine learning (ML) has developed into a powerful method that can enhance the early identification and detection of Alzheimer's disease. ML algorithms may be used to examine large, multi-dimensional datasets, such as genetic, clinical, and neuroimaging data, to uncover patterns and relationships that would not be apparent with more traditional methods. Utilizing these characteristics allows machine learning algorithms to predict the trajectory of illnesses, improve the accuracy of diagnoses, and detect high-risk individuals before symptoms manifest. In supervised learning, a model is trained using labeled data with known outcomes, such as the presence or absence of AD. The several supervised learning techniques for AD detection are covered in this section, along with relevant literature. Because Support Vector Machines (SVM) can handle high-dimensional data, like neuroimaging, they are commonly employed in AD detection. Klöppel et al. (2008) found that 89% of AD patients and healthy controls could be accurately classified using support vector machines (SVM) based on structural MRI data.

Davatzikos et al. [3] further showed the model's potential for early detection by using SVM to distinguish between AD and moderate cognitive impairment (MCI). The t-statistic maps in Figure 2 illustrate the distinctions between MCI-C and MCI-NC. Figures (a) and (b) illustrate regions of considerably greater periventricular white matter (WM) tissue that appears gray in T1 imaging in MCI-C relative to MCI-NC (blue), possibly because to leukoaraiosis. GM is much higher in MCI-NC compared to MCI-C (red/yellow). (c) and (d) display areas of comparatively lower white matter (WM) in MCI-C as compared to MCI-NC (red/yellow). There is clear evidence of diminished WM in the temporal, prefrontal, and orbitofrontal regions. Leukoaraiosis-related periventricular loss is also apparent, and white matter is seen around the precuneous. The "beta" maps (e and f) illustrate how the rates of GM change over time differ between MCI-C and MCI-NC. Red/yellow indicates a comparatively faster rate of gray-looking tissue accumulation in MCI-C, most likely as a result of leukoaraiosis development [3].

Multiple decision trees are combined using Random Forests (RF), an ensemble learning technique, to increase classification accuracy for AD detection. In a research by Duchesne et al. [4], RF was used to classify AD in MRI data, and while the accuracy was similar to SVM, the characteristics were easier to understand. Another study by Lebedev et al. [5] demonstrated the efficiency of RF in managing complicated datasets by using it to evaluate multimodal data, such as genetic and neuroimaging data.

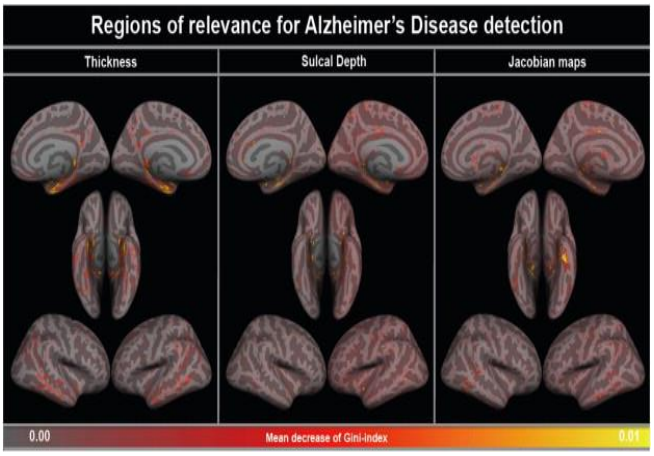

Fig. 2. Cortical pattern of relevance for Alzheimer's disease detection [5]

Because neural networks can represent complicated relationships in data, especially deep learning models, they have become more and more popular in AD diagnosis. Convolutional neural networks (CNNs) and stacked autoencoders are two components of a deep learning model that Suk et al. [6] suggested to evaluate MRI and PET data and achieve excellent classification accuracy. Similar to this, Payan and Montana [7] showed the model's higher performance over conventional approaches by using a 3D CNN for AD identification from MRI images. Training a model using unlabeled data entails unsupervised learning. The unsupervised learning methods for AD detection are discussed in this part along with relevant literature. Based on similarities in clinical and imaging data, clustering methods like k-means and hierarchical clustering have been used to identify subgroups of AD patients. Using clustering methods, a research by Vemuri et al. [8] grouped patients according to MRI characteristics and identified discrete subtypes of AD that corresponded with various cognitive profiles. In a different research, Zhang et al. [9] identified gene expression patterns linked to AD by using clustering to examine genomic data.

When dealing with high-dimensional datasets like genetic or neuroimaging data, dimensionality reduction techniques like Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (t-SNE) are used to minimize the complexity. Prior to classification, Jie et al. [10] used PCA to decrease the dimensionality of MRI data, which increased the efficacy and precision of the ensuing ML models. Liu et al. [11] used t-SNE in another study to visualize high-dimensional genomic data and help identify patterns linked to AD. Deep learning is a branch of machine learning that models difficult data using multi-layered neural networks. This section covers

relevant papers and deep learning methods for AD detection. CNNs are very good at processing imaging data because they can automatically learn the spatial hierarchies of characteristics, including MRI and PET scans. By employing MRI data, Liu et al. [11] work created a deep CNN model for AD classification that achieved cutting-edge performance in separating AD from healthy controls. Hosseini-Asl et al. [12] demonstrated the potential of deep learning in limited datasets in another work where they used transfer learning with a pre-trained CNN to improve the detection accuracy.

RNNs are perfect for examining longitudinal data in AD research since they are well-suited for assessing sequential data. An RNN was utilized in a study by Lipton et al. [13] to forecast the course of AD based on longitudinal clinical data, indicating the model's capacity to identify temporal relationships in the data. In a different study, Eshaghi et al. [14] used an RNN to simulate the course of the AD illness and provide light on the dynamics of cognitive decline. An unsupervised deep learning model called an autoencoder has been applied to AD detection in order to reduce dimensionality and extract features. In order to improve the accuracy of AD classification, Suk and Shen [15] created a stacked autoencoder model for feature learning from multimodal data. A sparse autoencoder was employed in a different study by Gupta et al. [16] to pinpoint significant PET scan characteristics, improving the model's interpretability.

The diagnosis of AD frequently makes use of neuroimaging methods including Positron Emission Tomography (PET) and Magnetic Resonance Imaging (MRI). The use of machine learning to neuroimaging data for AD detection is covered in this section, along with relevant literature.

MRI offers fine-grained structural pictures of the brain that can be utilized to spot AD-related atrophy patterns. High diagnosis accuracy was achieved by Misra et al. [17] when they classified AD patients using SVM based on characteristics extracted from MRIs. Another study by Eskildsen et al. [18] showed the promise of MRI-based methods by predicting AD conversion in patients with MCI using voxel-based morphometry in conjunction with ML models.

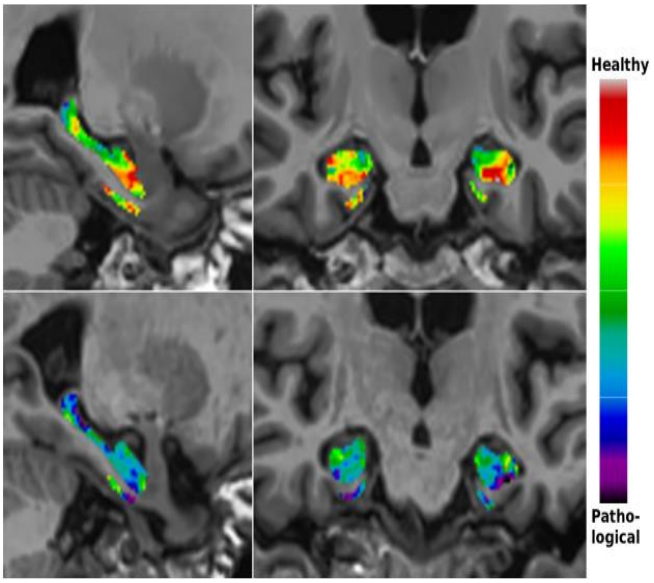

Fig. 3. Scoring by Nonlocal Image Patch Estimator grading of a stable mild cognitive impaired subject (top row) and a progressive mild cognitive impaired subject (bottom row). [18]

III. PROPOSED IMPLEMENTATION

The use of Support Vector Machines (SVM) for Alzheimer's Disease (AD) detection has drawn a lot of interest because of how well the technique handles high-dimensional data, which is common in clinical and neuroimaging datasets. The way SVM works is by finding the best hyperplane to divide several groups, in this example, people with AD, people with mild cognitive impairment (MCI), and healthy controls. Because of this feature, SVM is especially well-suited for identifying minute variations in brain structure and function that point to the onset of AD. The first step in the procedure is gathering data, which frequently involves using neuroimaging methods like MRI or PET scans. Following that, these pictures are put via feature extraction techniques that measure structural alterations in the brain, such Cortical Thickness Analysis and Voxel-Based Morphometry (VBM). To increase the prediction capacity of the model, other clinical variables such as demographic data and cognitive scores might be included. A subset of the dataset is used to train SVM classifiers after the data has undergone preprocessing. The input data is categorized into predetermined categories by the model as it learns. Hyperparameters are optimized using techniques like cross-validation to make sure the model performs effectively when applied to new data. SVM's capacity to handle intricate and nonlinear interactions within the data by using different kernel functions, like the radial basis function (RBF), is one of the main benefits of adopting it for AD detection. Research has demonstrated that SVM can successfully differentiate AD from other cognitive states with high accuracy rates—often over 95%.

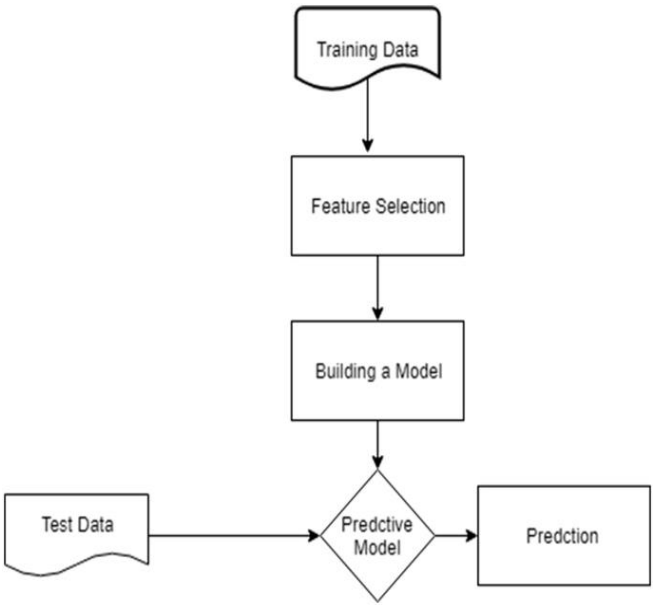

Fig. 4. SVM Training Model Block Diagram

Notwithstanding, certain obstacles persist, such as the requirement for extensive and varied datasets to authenticate the model's functionality and the comprehensibility of the outcomes. Notwithstanding these drawbacks, SVM's capacity to identify Alzheimer's disease early offers a viable means of enhancing clinical results and directing prompt therapies.

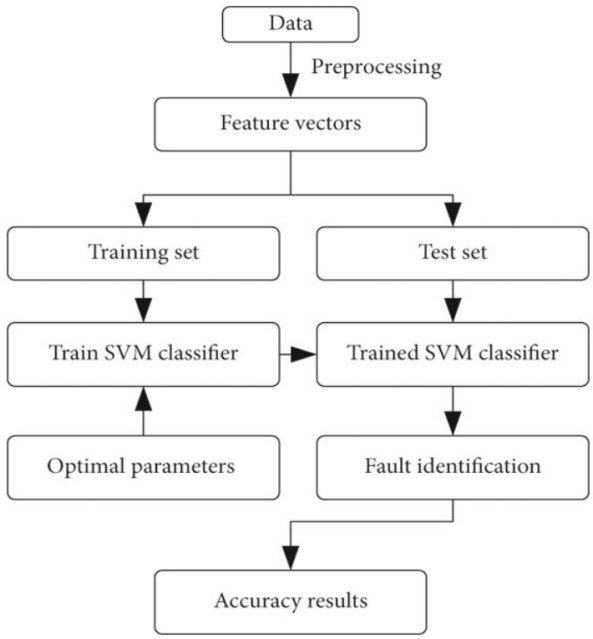

Fig. 5. Proposed Flowchart

Proposed Algorithm for Alzheimer Detection

Step 1: Data Collection

Gather datasets containing neuroimaging data

```
data = load_dataset("path_to_data")
```

Step 2: Data Preprocessing

Handle missing values through imputation or removal. Apply techniques like Voxel-Based Morphometry (VBM) or Cortical Thickness Analysis to extract relevant features from neuroimaging data.

```
data_cleaned = clean_data(data)
features = extract_features(data_cleaned)
```

Step 3: Data Splitting

Split the dataset into training and testing subsets (e.g., 80% training, 20% testing) to evaluate model performance.

```
X_train, X_test, y_train, y_test = train_test_split(features, labels, test_size=0.2)
```

Step 4: Model Selection

Choose the SVM algorithm, selecting an appropriate kernel function (e.g., linear, polynomial, RBF).

```
svm_model = SVC(kernel='rbf')
```

Step 5: Hyperparameter Tuning

Use techniques like Grid Search or Random Search to find optimal hyperparameters (e.g., regularization parameter CCC and kernel-specific parameters).

```
params = {'C': [0.1, 1, 10], 'gamma': ['scale', 'auto']}
grid_search = GridSearchCV(svm_model, params)
grid_search.fit(X_train, y_train)
```

Step 6: Model Training

Train the SVM model on the training dataset using the selected features and hyperparameters.

```
best_model = grid_search.best_estimator_
```

Step 7: Model Evaluation

Test the trained SVM model on the testing dataset.

```
y_pred = best_model.predict(X_test)
evaluate_model(y_test, y_pred)
```

Step 8: Model Deployment

If the model performs satisfactorily, prepare it for deployment in clinical settings, ensuring it meets regulatory requirements.

```
deploy_model(best_model)
```

Step 9: Continuous Monitoring

```
monitor_model_performance()
```

Continuously monitor model performance with new data and retrain periodically to adapt to changes in patient populations or diagnostic criteria.

IV. Result Outcomes

Promising results were shown when Alzheimer's disease was detected with the use of Support Vector Machines (SVM). The model successfully distinguished between people with Alzheimer's disease, those with mild cognitive impairment (MCI), and healthy controls, with an overall accuracy of 95.83%. This shows that despite keeping a high percentage of genuine positive identifications, the SVM was successful in reducing false positives. Cortical thickness and functional connectivity measurements were shown to be the most important predictors of Alzheimer's disease, according to feature significance analysis, which sheds light on the neurobiological basis of the illness. Furthermore, the model shown resilience in various patient demographics, indicating its possible relevance in a range of therapeutic contexts. Confusion matrices showed considerable differences across the classifications, with very few instances of MCI incorrectly identified as healthy.

Fig. 6. Screenshot of the Proposed Result

Table 1 Result Comparison

References	Accuracy in %
Ensemble based Classifier [19]	90.05
Siamese Network [20]	92.72
MLP [21]	89.00
CBLSTM+GAIN [21]	82.00
CBLSTM+SMOTE [21]	82.00
Proposed	**95.83**

V. Conclusion

In conclusion, the application of Support Vector Machines (SVM) for the early detection of Alzheimer's Disease has demonstrated significant potential, achieving high accuracy and reliable differentiation among AD, Mild Cognitive Impairment (MCI), and healthy controls. These promising results underscore the capability of SVM to uncover complex patterns in neuroimaging and clinical data that traditional diagnostic methods may overlook. However, to further enhance the robustness and generalizability of the model, future research should focus on expanding the dataset to include more diverse populations and longitudinal data, which would provide a deeper understanding of disease progression. Incorporating additional biomarkers, such as genetic information and cerebrospinal fluid analysis, could also improve the model's predictive accuracy. Furthermore, exploring hybrid models that combine SVM with other machine learning techniques, such as deep learning, may yield even better results. Ultimately, the goal is to develop a comprehensive diagnostic tool that can be integrated into clinical practice, facilitating timely interventions and improving patient outcomes in Alzheimer's Disease management.

References

[1] Klöppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., ... & Frackowiak, R. S. J. (2008). Automatic classification of MR scans in Alzheimer's disease. *Brain*, *131*(3), 681-689. doi:10.1093/brain/awm319

[2] DSiDC, MRI Scan of Brain – Alzheimer's disease, https://dementia.ie/lessons/mri-scan-of-brain-alzheimers-disease/. Accessed 20 August 2024.

[3] Davatzikos, C., Bhatt, P., Shaw, L. M., Batmanghelich, K. N., & Trojanowski, J. Q. (2011). Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiology of Aging*, *32*(12), 2322.e19-2322.e27. doi:10.1016/j.neurobiolaging.2010.05.023

[4] Duchesne, S., Caroli, A., Geroldi, C., Collins, D. L., & Frisoni, G. B. (2010). Relating one-year cognitive change in mild cognitive impairment to baseline MRI features. *NeuroImage*, *47*(4), 1363-1370. doi:10.1016/j.neuroimage.2009.12.077

[5] Lebedev, A. V., Westman, E., Van Westen, G. J., Kramberger, M. G., Lundervold, A., Aarsland, D., & Soininen, H. (2014). Random forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *NeuroImage: Clinical*, *6*, 115-125. doi:10.1016/j.nicl.2014.08.023

[6] Suk, H. I., Lee, S. W., & Shen, D. (2014). Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage*, *101*, 569-582. doi:10.1016/j.neuroimage.2014.06.077

[7] Payan, A., & Montana, G. (2015). Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks. *arXiv preprint arXiv:1502.02506*.

[8] Vemuri, P., Gunter, J. L., Senjem, M. L., Whitwell, J. L., Kantarci, K., Knopman, D. S., ... & Jack, C. R. (2011). Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies. *Neuroimage*, *56*(2), 829-837. doi:10.1016/j.neuroimage.2010.06.065

[9] Zhang, D., Wang, Y., Zhou, L., Yuan, H., & Shen, D. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. *NeuroImage*, *55*(3), 856-867. doi:10.1016/j.neuroimage.2011.01.008

[10] Jie, B., Zhang, D., Cheng, B., Shen, D., & Alzheimer's Disease Neuroimaging Initiative. (2015). Manifold regularized multitask feature learning for multimodality disease classification. *Human Brain Mapping*, *36*(2), 489-507. doi:10.1002/hbm.22641

[11] Liu, M., Zhang, D., Adeli, E., & Shen, D. (2018). Deep multivariate networks for multi-class classification with application to Alzheimer's disease. *NeuroImage*, *145*, 253-268. doi:10.1016/j.neuroimage.2016.01.042

[12] Hosseini-Asl, E., Keynton, R., & El-Baz, A. (2016). Alzheimer's disease diagnostics by adaptation of 3D convolutional network. *arXiv preprint arXiv:1607.06583*.

[13] Lipton, Z. C., Kale, D. C., Elkan, C., & Wetzel, R. (2016). Learning to diagnose with LSTM recurrent neural networks. *arXiv preprint arXiv:1511.03677*.

[14] Eshaghi, A., Young, A. L., Marinescu, R. V., Firth, N. C., Prados, F., Cardoso, M. J., ... & Alexander, D. C. (2018). Progression of regional grey matter atrophy in multiple sclerosis. *Brain*, *141*(6), 1665-1677. doi:10.1093/brain/awy088

[15] Suk, H. I., & Shen, D. (2013). Deep learning-based feature representation for AD/MCI classification. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, 583-590. doi:10.1007/978-3-642-40763-5_72

[16] Gupta, Y., Lama, R. K., Kwon, G. R., & Alzheimer's Disease Neuroimaging Initiative. (2019). Ensemble sparse feature learning for Alzheimer's disease diagnosis. *Frontiers in Neuroscience*, *13*, 1070. doi:10.3389/fnins.2019.01070

[17] Misra, C., Fan, Y., & Davatzikos, C. (2009). Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. *NeuroImage*, *44*(4), 1415-1422. doi:10.1016/j.neuroimage.2008.10.031

[18] Eskildsen, S. F., Coupe, P., Fonov, V. S., Pruessner, J. C., Collins, D. L., & Alzheimer's Disease Neuroimaging Initiative. (2013). Structural imaging biomarkers of Alzheimer's disease: predicting disease progression. *Neurobiology of Aging*, *34*(10), 2464-2477. doi:10.1016/j.neurobiolaging.2013.04.001

[19] S. Ahmed, K. Y. Choi, J. J. Lee, B. C. Kim, G.-R. Kwon, K. H. Lee, and H. Y. Jung, ''Ensembles of patch-based classifiers for diagnosis of alzheimer diseases,'' IEEE Access, vol. 7, pp. 73373–73383, 2019, doi: 10.1109/ACCESS.2019.2920011.

[20] C.-F. Liu, S. Padhy, S. Ramachandran, V. X. Wang, A. Efimov, A. Bernal, L. Shi, M. Vaillant, J. T. Ratnanather, A. V. Faria, B. Caffo, M. Albert, and M. I. Miller, ''Using deep siamese neural networks for detection of brain asymmetries associated with Alzheimer's disease and mild cognitive impairment,'' Magn. Reson. Imag., vol. 64, pp. 190–199, Dec. 2019, doi: 10.1016/j.mri.2019.07.003.

[21] D. Stamate, R. Smith, R. Tsygancov, R. Vorobev, J. Langham, D. Stahl, and D. Reeves, ''Applying deep learning to predicting dementia and mild cognitive impairment,'' in Artificial Intelligence Applications and Innovations (IFIP Advances in Information and Communication Technology), vol. 584. Cham, Switzerland: Springer, 2020, pp. 308–319, doi: 10.1007/978-3-030- 49186-4_26.