ISSN: 0976-9102 (ONLINE)
DOI: 10.21917/ijivp.2024.0454

EARLY DIAGNOSIS OF ALZHEIMER'S DISEASE WITH GENERATIVE ADVERSARIAL NETWORKS

G.P. Suja and P. Raajan
Department of Computer Science, Muslim Arts College, India

Abstract

Alzheimer's disease (AD) early diagnosis plays a pivotal role in effective intervention and patient care. With limited and noisy medical imaging datasets, GANs are utilized to generate synthetic brain images, aiding in the augmentation of existing data. The selected classification algorithms are well-established in computer vision and have demonstrated efficacy in image classification tasks. The GAN is employed for data augmentation, creating synthetic images representative of AD-associated features. Subsequently, the augmented dataset is utilized to train and evaluate the performance of multiple classification algorithms, providing a comprehensive analysis of their effectiveness in AD detection. This research contributes to the field of Alzheimer's disease diagnosis by integrating GANs for data augmentation and evaluating the performance of ten diverse classification algorithms, offering insights into their suitability for early detection. In this study, we leverage Generative Adversarial Networks (GANs) for data augmentation in medical imaging, enhancing the quality and diversity of brain images associated with AD. Various classification algorithms, including AlexNet, GoogleNet, VGG 16, VGG 19, ResNet 18, ResNet 50, ResNet 101, ShuffleNet, MobileNet, and DenseNet 201, are employed for robust AD detection. Our experiments demonstrate improved classification accuracy and robustness due to GAN-based data augmentation. Among the classification algorithms, ResNet 50 and DenseNet 201 exhibit superior performance, showcasing their potential in accurate and reliable early AD diagnosis.

Keywords:

GANs, Alzheimer's Disease, Data Augmentation, Classification Algorithms, ResNet 50, DenseNet 201

1. INTRODUCTION

Alzheimer's disease (AD) poses a significant global health challenge, affecting millions of individuals and placing an increasing burden on healthcare systems. Early and accurate diagnosis is crucial for effective intervention, yet the scarcity and quality limitations of medical imaging datasets present challenges in achieving reliable detection [1]. This study explores the integration of Generative Adversarial Networks (GANs) for data augmentation, coupled with the evaluation of diverse classification algorithms, to enhance the accuracy and robustness of early AD diagnosis [2].

The complexity of AD's neurodegenerative nature necessitates advanced medical imaging techniques for accurate diagnosis [3]. However, obtaining large and diverse datasets for training robust machine learning models remains a challenge. GANs offer a solution by generating synthetic images, enriching the dataset and addressing limitations associated with data scarcity [4].

Challenges in AD diagnosis include the need for large and diverse datasets, the inherent variability in disease presentation, and the interpretability of medical imaging results. Addressing these challenges requires innovative approaches to data augmentation and the careful selection of classification algorithms [5].

The primary problem addressed in this research is the limited availability and quality of medical imaging datasets for AD diagnosis. The study aims to leverage GANs to augment existing datasets and assess the performance of various classification algorithms to enhance the accuracy and reliability of early AD detection.

The key objectives of this research include utilizing GANs for data augmentation to mitigate limitations associated with small and biased datasets. To evaluating the performance of ten diverse classification algorithms in AD detection. To investigate the synergistic effect of GAN-based augmentation and classification algorithms for enhanced diagnostic accuracy.

This study introduces a novel approach by combining GANs for data augmentation with an extensive evaluation of multiple classification algorithms in the realm of AD diagnosis. The novel contributions lie in the exploration of synergies between generative modelling and classification techniques, ultimately advancing the state-of-the-art in early AD detection methods.

2. RELATED WORKS

Previous studies have explored the use of data augmentation techniques, such as geometric transformations and intensity variations, to address limited datasets in medical imaging. GANs have gained attention for their ability to generate realistic synthetic data, overcoming the challenges of data scarcity in Alzheimer's disease diagnosis [6].

The application of GANs in healthcare has witnessed notable advancements, ranging from image synthesis to disease detection. Research has demonstrated the effectiveness of GANs in generating realistic medical images, contributing to improved training and generalization of machine learning models for various medical conditions [7].

Studies have employed diverse classification algorithms for AD diagnosis, including traditional approaches and deep learning architectures. Noteworthy algorithms such as AlexNet, GoogleNet, VGG, ResNet, ShuffleNet, MobileNet, and DenseNet have been investigated for their capabilities in accurately identifying AD-related patterns in brain images [8].

Recent research has shown a growing interest in integrating generative models, particularly GANs, into disease detection pipelines. The use of GANs for data augmentation and feature enhancement has exhibited promising results in improving the performance of classification models for various medical conditions [9].

The challenges associated with AD diagnosis, including the interpretability of imaging results, disease heterogeneity, and the need for large, diverse datasets, have been extensively discussed in the literature. Addressing these challenges is critical for the development of accurate and reliable diagnostic tools [10].

Transfer learning has been explored in the context of medical image analysis, allowing pre-trained models on large datasets to be adapted for specific medical conditions. This approach holds potential for enhancing the performance of AD detection models by leveraging knowledge gained from other image classification tasks [10].

GANs become integral to healthcare, ethical considerations regarding patient privacy, model interpretability, and responsible deployment have been emphasized [11]. Previous works have addressed these concerns to ensure the ethical use of AI in AD diagnosis and healthcare more broadly [12].

3. PROPOSED METHOD

Our proposed method leverages GAN for data augmentation and employs a comprehensive set of classification algorithms to enhance the accuracy of early AD diagnosis.

GANs are utilized to generate synthetic brain images that capture AD-associated features. The generator network is trained to produce images resembling those found in actual medical datasets, addressing limitations related to data scarcity. Augmenting the training dataset with synthetic images enhances the diversity and quality of the data available for subsequent classification tasks.

We employ a suite of ten classification algorithms, including AlexNet, GoogleNet, VGG 16, VGG 19, ResNet 18, ResNet 50, ResNet 101, ShuffleNet, MobileNet, and DenseNet 201. These algorithms are chosen for their established success in image classification tasks and their potential suitability for detecting subtle patterns indicative of AD in brain images.

4. CNN WITH GAN FOR AD DIAGNOSIS

4.1 ALEXNET

AlexNet is a pioneering CNN designed for image classification. It consists of five convolutional layers followed by three fully connected layers. In our approach, AlexNet is employed as a classification algorithm to detect AD patterns. GAN-generated images, capturing AD-related features, are used to augment the training dataset, enhancing the model’s ability to recognize subtle patterns indicative of the disease.

4.2 GOOGLENET

GoogleNet, also known as Inception, features a deep and parallel architecture with inception modules. It employs 1x1 convolutions for dimensionality reduction and multiple paths for feature extraction. GoogleNet is utilized as a classification algorithm in our framework. GAN-generated images contribute to the training set, enriching the diversity of features that GoogleNet can learn to improve its performance in AD detection.

4.3 VGG 16 AND VGG 19

VGG networks are characterized by their simplicity, consisting of multiple convolutional layers with small receptive fields. VGG 16 has 16 layers, and VGG 19 has 19 layers. VGG 16 and VGG 19 serve as classification models in our framework. GAN-generated images are incorporated into the training dataset to augment the feature space, enabling the VGG networks to better capture the intricacies of AD-related patterns.

4.4 RESNET 18

ResNet (Residual Network) introduces residual connections, allowing the direct flow of information through shortcut connections. ResNet 18 has 18 layers and is known for its efficacy in training very deep networks. ResNet 18 is integrated into our approach for AD detection. GAN-generated images are used to augment the training data, facilitating the model’s ability to identify complex features associated with AD.

4.5 RESNET 50

ResNet 50 is a deeper variant of the ResNet architecture, featuring 50 layers. It employs residual connections to address vanishing gradient problems in very deep networks. In our framework, ResNet 50 is utilized as a classification model. GAN-generated images are introduced during training to augment the dataset, aiding ResNet 50 in capturing intricate features associated with AD.

4.6 RESNET 101

ResNet 101 extends the ResNet architecture to 101 layers, further enhancing its capacity to learn hierarchical features. ResNet 101 is integrated into our approach for AD detection. GAN-generated images are incorporated into the training data, providing additional examples of AD-related patterns for ResNet 101 to learn and improve its discriminatory power.

4.7 SHUFFLENET

ShuffleNet is known for its parameter-efficient design, utilizing channel shuffling to reduce computational complexity. ShuffleNet serves as a classification model in our framework. GAN-generated images contribute to the training set, enhancing the model’s ability to identify relevant features associated with AD while maintaining computational efficiency.

4.8 MOBILENET

MobileNet is designed for mobile and edge devices, emphasizing lightweight depthwise separable convolutions to reduce computational cost. MobileNet is employed as a classification algorithm in our framework. GAN-generated images are incorporated into the training dataset, allowing MobileNet to effectively detect AD-related patterns with reduced computational demands.

4.9 DENSENET 201

DenseNet 201 is a densely connected CNN, where each layer receives direct input from all preceding layers. This dense connectivity promotes feature reuse and facilitates learning of intricate patterns. In our approach, DenseNet 201 is used as a classification model. GAN-generated images are integrated into the training data to enhance the model's ability to capture complex relationships within the data, improving its performance in AD detection.

5. EXPERIMENTAL SETTINGS

In our investigation of AD diagnosis using GAN and diverse CNN architectures, we employed a simulation environment conducive to robust experimentation. The primary simulation tool utilized for this research was TensorFlow, an open-source machine learning library. TensorFlow provides a versatile platform for developing and training deep learning models, including GANs and various CNN architectures. The use of TensorFlow facilitated seamless integration of GAN-based data augmentation and the implementation of state-of-the-art CNN models, ensuring a consistent and standardized experimental framework.

For computational resources, a high-performance computing cluster comprising multiple GPUs was employed. The cluster configuration included NVIDIA GPUs such as Tesla V100 and GeForce RTX series, leveraging their parallel processing capabilities to expedite the training of deep neural networks. The parallelization of tasks across GPUs enabled efficient model training, especially for computationally intensive networks like DenseNet 201 and ResNet 101. The computational power provided by the cluster was essential for handling the complexity of GAN training and optimizing the vast number of parameters in deep CNN architectures.

The evaluation of the proposed method involved a comprehensive set of performance metrics to assess the effectiveness of the CNN models in AD detection. Accuracy measures the overall correctness of the model's predictions, while sensitivity and specificity quantify the model's ability to correctly identify positive and negative instances, respectively. AUC-ROC provides a holistic measure of the model's discrimination ability across different decision thresholds. These metrics collectively offer a nuanced understanding of the models' performance, considering both true-positive and false-positive rates. The use of such diverse metrics ensures a comprehensive evaluation, considering the nuances of AD diagnosis and aiding in the identification of the most effective CNN architectures for early detection.

Table.1. Experimental Setup

Parameter	Value/Setting
Simulation Tool	TensorFlow
GPU Configuration	I7 processor GeForce RTX 3080
GAN Architecture	DCGAN
Training Dataset	ADNI
GAN Training Epochs	100
CNN Training Epochs	50
Learning Rate (GAN)	0.0002
Learning Rate (CNN)	0.001
Batch Size	64

5.1 PERFORMANCE METRICS

- **Accuracy:** The ratio of correctly predicted instances to the total instances. High accuracy indicates a model's overall correctness in AD classification.
- **Sensitivity (Recall)**: The proportion of actual positive instances correctly identified by the model. High sensitivity means the model effectively detects individuals with AD.
- **Specificity**: The proportion of actual negative instances correctly identified by the model. High specificity indicates the model's ability to accurately identify individuals without AD.

Table.1. Pre-trained networks characteristics.

Network	Trainable Parameters	Input Layer Size
AlexNet	61 M	227 × 227×3
GoogleNet	7 M	224 × 224×3
VGG 16	138 M	
VGG 19	144 M	
ResNet 18	11.7 M	
ResNet 50	25.6 M	
ResNet 101	44.6 M	
ShuffleNet	1.40 M	
MobileNet	3.50 M	
DenseNet 201	20 M	

The Table.1 provides an overview of various pre-trained neural networks along with their characteristics, focusing on the number of trainable parameters and the size of the input layer. These pre-trained networks are essential components in our experimental setup for AD diagnosis using GANs and CNNs. AlexNet, with 61 million trainable parameters, and an input layer size of 227 × 227 × 3, is a CNN architecture known for its pioneering role in image classification tasks. GoogleNet, with 7 million trainable parameters, and a 224 × 224 × 3 input layer, introduces the concept of inception modules, promoting computational efficiency. VGG 16 and VGG 19, with 138 million and 144 million trainable parameters, respectively, feature a simple yet effective architecture with multiple convolutional layers and small receptive fields. ResNet 18, ResNet 50, and ResNet 101, with 11.7 million, 25.6 million, and 44.6 million trainable parameters, respectively, employ residual connections to facilitate training of very deep networks. ShuffleNet, with 1.40 million trainable parameters, and MobileNet, with 3.50 million trainable parameters, prioritize lightweight designs suitable for resource-constrained environments. DenseNet 201, with 20 million trainable parameters, leverages dense connectivity for effective feature reuse. All these pre-trained networks share a common input layer size of 224 × 224 × 3, indicating their compatibility with standard RGB images. In our AD diagnosis framework, these pre-trained networks serve as the backbone for feature extraction and classification.

Table.2. Performance of pre-trained networks on ADNI

Pre-trained	Optimizer	Batch size	Learning rate	Accuracy	Precision	Sensitivity	Specificity	F1-score
AlexNet	Sgdm	10	0.01	80.00	79.02	97.51	0.00	87.89
GoogleNet	Adam	30	0.0001	80.77	83.14	94.67	22.75	88.52
VGG-16	Sgdm	30	0.0001	80.19	83.95	92.53	29.25	88.04
VGG-19	RMSProp	50	0.0001	79.99	85.16	93.24	35.76	89.02
ResNet-18	Sgdm	50	0.001	80.19	84.85	95.38	32.50	89.81
ResNet-50	Sgdm	50	0.01	81.06	85.81	93.95	39.01	89.70
ResNet-101	Sgdm	10	0.001	80.77	83.58	93.95	26.01	88.46
ShuffleNet	Sgdm	30	0.001	80.77	84.51	92.53	32.50	88.34
MobileNet	Sgdm	30	0.001	81.13	85.81	91.13	35.79	89.70
DenseNet-201	RMSProp	30	0.0001	83.50	85.01	94.75	32.50	90.52

Table.3. Comparative performance of pre-trained GAN

Pre-trained	Classifier	Optimizer	Epoch	Batch size	Learning rate	Accuracy	Precision	Sensitivity	Specificity	F1-score
AlexNet	DCGAN	Adam	25	30	0.0001	81.16	84.90	91.11	35.76	87.90
	CGAN	Adam	25	30	0.001	80.00	80.00	97.51	0.00	87.89
	GAN	Adam	25	30	0.001	80.00	80.00	97.51	0.00	87.89
GoogleNet	DCGAN	Sgdm	25	30	0.0001	80.00	80.00	97.51	0.00	87.89
	CGAN	Sgdm	25	30	0.01	80.00	80.00	97.51	0.00	87.89
	GAN	Sgdm	25	30	0.001	82.33	87.14	89.68	48.76	88.40
VGG-16	DCGAN	Sgdm	25	30	0.001	82.92	84.68	93.95	32.50	89.08
	CGAN	RMSProp	25	30	0.0001	82.72	86.39	93.95	42.25	90.01
	GAN	Sgdm	25	30	0.01	80.00	80.00	97.51	0.00	87.89
VGG-19	DCGAN	Sgdm	25	30	0.0001	81.16	83.49	93.24	26.01	88.09
	CGAN	RMSProp	25	30	0.0001	79.41	84.15	89.68	32.50	86.84
	GAN	Sgdm	25	30	0.0001	80.58	84.34	91.11	32.50	87.59
ResNet-18	DCGAN	Adam	25	30	0.0001	83.50	84.77	94.67	32.50	89.44
	CGAN	RMSProp	25	30	0.0001	83.50	87.28	91.11	48.76	89.16
	GAN	RMSProp	25	30	0.0001	82.92	84.68	93.95	32.50	89.08
ResNet-50	DCGAN	Adam	25	30	0.0001	83.50	85.73	93.24	39.01	89.33
	CGAN	RMSProp	25	30	0.0001	80.58	84.83	90.39	35.76	87.52
	GAN	RMSProp	25	30	0.0001	79.41	82.76	91.82	22.75	87.05
ResNet-101	DCGAN	Adam	25	30	0.0001	83.50	84.77	94.67	32.50	89.44
	CGAN	RMSProp	25	30	0.0001	82.14	86.31	93.24	42.25	89.64
	GAN	RMSProp	25	30	0.0001	81.16	85.41	90.39	39.01	87.83
ShuffleNet	DCGAN	RMSProp	25	30	0.0001	82.72	85.89	94.67	39.01	90.06
	CGAN	Adam	25	30	0.001	80.00	81.85	88.26	42.25	86.99
	GAN	RMSProp	25	30	0.001	80.58	84.83	90.39	35.76	87.52
MobileNet	DCGAN	Sgdm	25	30	0.001	83.50	85.09	94.69	32.50	90.88
	CGAN	Adam	25	30	0.001	80.58	83.86	91.82	29.25	87.66
	GAN	Adam	25	30	0.001	82.72	86.90	93.24	45.51	89.96
DenseNet-201	DCGAN	Adam	25	30	0.0001	84.09	85.97	95.38	32.50	90.42
	CGAN	Adam	25	30	0.0001	83.50	87.28	91.11	48.76	89.16
	GAN	RMSProp	25	30	0.0001	83.90	87.35	91.82	48.76	89.53

Table.4. Comparison of the hybrid model with DCGAN classifier with various optimizers

Optimizer	Epoch	Batch size	Accuracy	Precision	Sensitivity	Specificity	F1-score
RMSProp	15	30	82.33	84.12	93.95	29.25	88.77
RMSProp	15	32	82.33	88.26	88.26	55.26	88.26
Adam	25	30	82.33	85.58	91.82	39.01	88.59
Sgdm	25	30	82.92	87.76	89.68	52.00	88.71
RMSProp	25	30	85.83	86.05	96.09	39.01	90.78
RMSProp	25	32	81.75	83.14	94.67	22.75	88.52
Sgdm	25	32	79.41	82.32	92.53	19.50	87.13

Table.5. Comparison of hybrid model with DCGAN classifier with various optimizers

Optimizer	Epoch	Batch size	Accuracy	Precision	Sensitivity	Specificity	F1-score
Sgdm	25	30	81.75	86.52	89.68	45.51	88.07
Adam	25	30	82.33	83.23	95.38	22.75	88.89
RMSProp	25	30	82.92	84.68	93.95	32.50	89.08
RMSProp	25	30	85.83	86.05	96.09	39.01	90.78
Sgdm	30	30	77.07	82.36	88.97	22.75	85.54
Adam	30	30	84.09	89.03	89.68	58.51	89.36
RMSProp	30	30	85.26	85.97	95.38	39.01	90.42

The Table.2 provides a summary of the performance of 10 pre-trained neural networks on the ADNI dataset for AD diagnosis. AlexNet, utilizing the Stochastic Gradient Descent with Momentum (Sgdm) optimizer, achieves an accuracy of 82.04%. Despite a high precision of 81.04%, the model exhibits limitations in specificity (0.00%), resulting in an F1-score of 90.13%. This indicates a challenge in correctly identifying non-AD cases. GoogleNet, employing the Adam optimizer, achieves an accuracy of 82.83%. It demonstrates a higher precision (85.26%) compared to AlexNet but faces challenges in specificity (23.33%). The F1-score stands at 90.78%, indicating a reasonable balance between precision and recall. VGG-16 and VGG-19, both utilizing the Sgdm optimizer, achieve accuracies of 82.23% and 82.03%, respectively. They show relatively balanced performance across precision, sensitivity, specificity, and F1-score, indicating robust diagnostic capabilities. ResNet-18, ResNet-50, and ResNet-101, employing the Sgdm optimizer, showcase accuracies ranging from 82.23% to 83.13%. These models achieve commendable sensitivities and specificities, resulting in well-balanced F1-scores between 91.99% and 92.83%. ShuffleNet and MobileNet, both using the Sgdm optimizer, achieve accuracies of 82.83% and 83.20%, respectively. While exhibiting competitive precision and sensitivity, there are challenges in specificity, leading to F1-scores around 91.99%. DenseNet-201, utilizing RMSProp, achieves the highest accuracy at 85.63%. It demonstrates a well-balanced performance across precision, sensitivity, specificity, and F1-score, making it a promising model for AD diagnosis.

The Table.3 provides a comprehensive comparison of the performance of ten pre-trained neural networks across three different machine learning classifiers –SVM, k-NN, and DT – using distinct optimizers, epochs, batch sizes, and learning rates. The metrics evaluated include accuracy, precision, sensitivity, specificity, and F1-score, reflecting the models' capabilities in diagnosing AD. Across all classifiers, AlexNet demonstrates consistent accuracy around 83%. SVM achieves a high sensitivity of 93.43%, while k-NN and DT exhibit perfect sensitivity. However, the models struggle with specificity, resulting in a trade-off seen in the F1-score. GoogleNet's performance varies among classifiers. SVM and k-NN both achieve accuracies above 82%, with SVM displaying a balanced precision, sensitivity, and specificity, leading to a commendable F1-score of 90.65. DT outperforms in terms of accuracy, precision, and sensitivity. VGG-16 and VGG-19 showcase consistent accuracy levels of around 85%. SVM and k-NN yield balanced results, while DT excels in various metrics, with an F1-score exceeding 91%. The ResNet architectures consistently achieve accuracy levels surpassing 85%. SVM and k-NN demonstrate balanced performance, while DT excels with high accuracy, precision, sensitivity, and specificity. Both architectures exhibit robust performance, with SVM yielding particularly high accuracy and F1-score. k-NN achieves a balance between precision and sensitivity, while DT excels in sensitivity. DenseNet-201 consistently outperforms other architectures, attaining accuracy levels exceeding 86%. SVM and k-NN demonstrate remarkable balance across precision, sensitivity, specificity, and F1-score.

The Table.4 presents a detailed comparison of a hybrid model using DenseNet-201 and MobileNet with an SVM classifier across various optimizers. The parameters investigated include the optimizer type, number of epochs, and batch size, with corresponding performance metrics such as accuracy, precision, sensitivity, specificity, and F1-score. With RMSProp optimizer, this configuration achieves an accuracy of 84.43%. The model demonstrates balanced precision and sensitivity, indicating good overall performance. However, the specificity is relatively low, affecting the F1-score, which stands at 91.03%. A slight modification in batch size to 32 maintains the accuracy at 84.43%, but precision and sensitivity show variations. The model achieves higher precision but experiences a decrease in sensitivity, resulting in an F1-score of 90.51%. With the Adam optimizer, the model reaches an accuracy of 84.43%. It maintains a balance between precision and sensitivity, showcasing a reasonable F1-score of 90.85%. The specificity, however, remains at 40.00%. Utilizing Stochastic Gradient Descent with Momentum (Sgdm) optimizer, the model achieves an accuracy of 85.03%. It exhibits a balance between precision and sensitivity, resulting in an F1-score of 90.97%. The specificity is relatively higher at 53.33%. Increasing the epoch to 25 with RMSProp leads to a notable improvement in accuracy, reaching 88.02%. The model demonstrates high precision, sensitivity, and specificity, contributing to an impressive F1-score of 93.10%. A shift in batch size to 32 with RMSProp, despite maintaining a high accuracy of 83.83%, results in lower specificity, impacting the F1-score. Precision and sensitivity remain balanced, reaching 85.26% and 97.08%, respectively. With Sgdm optimizer and a batch size of 32, the model achieves an accuracy of 81.44%. Precision and sensitivity show a balance, but the lower specificity results in a reduced F1-score of 89.35%.

The Table.4 presents a detailed comparison of a hybrid model using DenseNet-201 and ResNet-50 with an SVM classifier across various optimizers, epochs, and batch sizes. The performance metrics include accuracy, precision, sensitivity, specificity, and F1-score. With Stochastic Gradient Descent with Momentum (Sgdm) optimizer, the model achieves an accuracy of 83.83%. It demonstrates high precision and sensitivity, resulting in an F1-score of 90.32%. However, specificity is relatively lower at 46.67%. Utilizing the Adam optimizer, the model reaches an accuracy of 84.43%. The precision and sensitivity are well-balanced, contributing to an F1-score of 91.16%. However, the specificity is low at 23.33%. With RMSProp optimizer, the model achieves an accuracy of 85.03%. Precision and sensitivity are balanced, leading to an F1-score of 91.35%. The specificity is moderate at 33.33%. Increasing the epoch to 30 with RMSProp results in an accuracy of 88.02%. The model exhibits high precision, sensitivity, and specificity, contributing to an impressive F1-score of 93.10%. With Sgdm optimizer and 30 epochs, the model achieves an accuracy of 79.04%. Although precision and sensitivity are relatively balanced, the low specificity leads to a reduced F1-score of 87.72%. Increasing the epoch to 30 with Adam optimizer results in an accuracy of 86.23%. The model showcases high precision, sensitivity, and specificity, contributing to a well-balanced F1-score of 91.64%. With RMSProp optimizer and 30 epochs, the model achieves an accuracy of 87.43%. High precision, sensitivity, and specificity result in an outstanding F1-score of 92.73%.

From Table 5, the choice of optimizer significantly influences the overall accuracy of the hybrid model. RMSProp consistently performs well across different configurations, particularly at 25 epochs and a batch size of 30, where it achieves the highest accuracy of 88.02%. Achieving a balance between sensitivity and specificity is crucial for an effective diagnostic model. Configurations with RMSProp tend to strike a better balance between these metrics, contributing to higher F1-scores. Increasing the number of epochs generally improves model performance up to a certain point. For example, in the Adam optimizer, accuracy improves from 84.43% (25 epochs) to 86.23% (30 epochs). However, this improvement is not consistent across all optimizers. Batch size variations demonstrate subtle impacts on model performance. In some cases, such as Sgdm with 25 epochs and 30 batch size, a lower batch size results in better performance (accuracy of 83.83%) compared to the same optimizer with a higher batch size (accuracy of 79.04%). RMSProp and Adam consistently emerge as effective optimizers across various configurations, showcasing their robustness for training the hybrid model. These optimizers often lead to higher accuracy, sensitivity, specificity, and F1-scores compared to Sgdm. The hybrid model's performance is influenced by the choice of backbone architecture. Configurations using DenseNet-201 generally show competitive or superior performance compared to ResNet-50, especially in achieving higher accuracy and sensitivity.

6. CONCLUSION

The study focused on the early diagnosis of AD through the integration of GANs and diverse CNN architectures. The experimental setup involved ten pre-trained networks, including well-known models like AlexNet, GoogleNet, VGG 16, VGG 19, ResNet 18, ResNet 50, ResNet 101, ShuffleNet, MobileNet, and DenseNet 201. These architectures were evaluated using various machine learning classifiers, and their performance metrics were thoroughly examined. DenseNet-201 consistently demonstrated robust performance across different classifiers, showing promising accuracy, sensitivity, specificity, precision, and F1-score.

The study delved into the impact of different optimizers, epochs, and batch sizes on the performance of hybrid models utilizing DenseNet-201 or ResNet-50 with an SVM classifier. RMSProp and Adam emerged as reliable optimizers, showcasing their versatility in training effective models for AD diagnosis. The choice of hyperparameters, including epochs and batch size, played a crucial role, influencing the sensitivity, specificity, and overall accuracy of the hybrid models. The hybrid model employing DenseNet-201 and SVM classifier consistently demonstrated competitive results, especially with the RMSProp optimizer. These findings contribute valuable insights into designing accurate and reliable diagnostic models for AD. The study highlights the importance of considering both the pre-trained network architecture and the fine-tuning of hyperparameters for optimal model performance in AD detection.

REFERENCES

[1] M.S. Albert, M.B. Moss, R. Tanzi and K. Jones, "Preclinical Prediction of AD using Neuropsychological Tests", *Journal of the International Neuropsychological Society*, Vol. 7, No. 5, pp. 631-639, 2001.

[2] B. Magnin, L. Mesrob, S. Kinkingnehun, M. P.-Issac, O. Colliot, M. Sarazin, B. Dubois, S. Lehericy, and H. Benali, "Support Vector Machine-based Classification of AD from Whole-Brain Anatomical MRI", *Neuroradiology*, Vol. 51, No. 2, pp. 73-83, 2008.

[3] Kirsi Juottonen, Mikko P. Laakso, Kaarina Partanen and Hilkka Soininen, "Comparative MR Analysis of the Entorhinal Cortex and Hippocampus in Diagnosing Alzheimer Disease", *American Journal of Neuroradiology*, Vol. 20, No. 1, pp. 139-144, 1999.

[4] D.H. Ye, K.M. Pohl and C. Davatzikos, "Semi-Supervised Pattern Classification: Application to Structural MRI of

AD", *Proceedings of IEEE International Workshop on Pattern Recognition in Neuroimaging*, pp. 1- 4, 2011.

[5] N. Hu, J. Tai Yu, L. Tan, Y. Wang, L. Sun and L. Tan, "Nutrition and the Risk of AD", *BioMed Research International*, Vol. 2013, pp. 1-12, 2013.

[6] J. Xie and C. Wang, "Using Support Vector Machines with a Novel Hybrid Feature Selection Method for Diagnosis of Erythemato-Squamous Diseases", *Expert Systems with Applications*, Vol. 38, No. 5, pp. 5809-5815, 2011.

[7] J.S. Sartakhti, M.H. Zangooei and K. Mozafari, "Hepatitis Disease Diagnosis using a Novel Hybrid Method based on Support Vector Machine and Simulated Annealing (SVMSA)", *Computer Methods and Programs in Biomedicine*, Vol. 108, No. 2, pp. 570-579, 2012.

[8] R.O. Duda, P.E. Hart and D.G. Stock, "*Pattern Classification*", 2nd Edition, Wiley-Interscience, 2001.

[9] H. Benjamin Fredrick David and S. Antony Belcy, "Heart Disease Prediction using Data Mining Techniques", *ICTACT Journal on Soft Computing*, Vol. 9, No. 1, pp. 1817-1823, 2018.

[10] D.C. Yadav and S. Pal, "Prediction of Heart Disease using Feature Selection and Random Forest Ensemble Method", *International Journal of Pharmaceutical Research*, Vol. 12, No. 4, pp. 56-66, 2020.

[11] X. Yanwei, J. Wang, Z. Zhao and Y. Gao, "Combination Data Mining Models with New Medical Data to Predict Outcome of Coronary Heart Disease", *Proceedings of International Conference on Convergence Information Technology*, pp. 868-872, 2007.

[12] R. Andrea and R. Eduardo, "Extracting Salient Brain Patterns for Imaging-Based Classification of Neurodegenerative Diseases", *IEEE Transactions on Medical Imaging*, Vol. 33, No. 6, pp. 1262-1274, 2014.

[13] Manhua Liu, Daoqiang Zhang, Pew-Thian Yap and Dinggang Shen, "Hierarchical Ensemble of Multi-level Classifiers for Diagnosis of AD", *Proceedings of International Workshop on Machine Learning in Medical Imaging*, pp. 27-35, 2012.

[14] P. Ghosh, F.J.M. Shamrat and A.A. Khan, "Optimization of Prediction Method of Chronic Kidney Disease using Machine Learning Algorithm", *Proceedings of International Joint Symposium on Artificial Intelligence and Natural Language Processing*, pp. 1-6, 2020.