SCT Proceedings in Interdisciplinary Insights and Innovations. 2025; 3:499
doi: 10.56294/piii2025499

Categoría: Innovations in Science and Engineering

ORIGINAL

Advancements in Alzheimer's Disease Classification: Integrating Machine Learning, Neuroimaging, and Biomarkers

Avances en la clasificación de la enfermedad de Alzheimer: Integración de aprendizaje automático, neuroimagen y biomarcadores

Chilukuri Ganesh [1], Gandikota Harshavardhan [1], Naishadham Radha Sri Keerthi [1], Raj Veer Yabaji [1], Meghana Sadhu [1]

[1] Student, Department of AI and DS, Koneru Lakshmaiah Education Foundation. India.

Citar como: Ganesh C, Harshavardhan G, Sri Keerthi NR, Veer Yabaji R, Sadhu M. Advancements in Alzheimer's Disease Classification: Integrating Machine Learning, Neuroimaging, and Biomarkers. SCT Proceedings in Interdisciplinary Insights and Innovations. 2025;3:499.DOI: https://doi.org/10.56294/piii2025499

Recibido: 12-09-2025 **Revisado:** 27-11-2024 **Aceptado:** 03-01-2025 **Publicado:** 05-01-2025

Editor: Emanuel Maldonado

ABSTRACT

Alzheimer's disease (AD), a progressive neurodegenerative disorder, leads to cognitive decline, memory loss, and impaired daily functioning. Early detection and precise classification are critical for timely intervention and personalized care. These abstract reviews recent advancements in brain disease classification, particularly for AD, highlighting the use of machine learning algorithms, neuroimaging methods, and biomarker analysis. Machine learning models trained on neuroimaging data, such as MRI and PET scans, have demonstrated efficacy in distinguishing Alzheimer's disease, mild cognitive impairment (MCI), and healthy individuals. Biomarker studies involving cerebrospinal fluid (CSF) and blood samples provide critical insights into AD pathology, supporting disease classification efforts. Integrating diverse data types, including imaging, genetic, and clinical information, can significantly enhance the accuracy and reliability of classification models. Emerging deep learning techniques, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), enable the extraction of complex patterns from heterogeneous data sources, improving classification outcomes. Nonetheless, challenges persist, such as the requirement for large-scale, multi-centre datasets, uniform imaging protocols, and greater interpretability of machine learning models.

Keywords: Alzheimer's disease; Machine Learning; Neuroimaging; Biomarkers.

RESUMEN

La enfermedad de Alzheimer (EA), un trastorno neurodegenerativo progresivo, provoca deterioro cognitivo, pérdida de memoria y alteraciones del funcionamiento diario. La detección precoz y la clasificación precisa son fundamentales para una intervención oportuna y una atención personalizada.

Este resumen repasa los avances recientes en la clasificación de enfermedades cerebrales, en particular la EA, destacando el uso de algoritmos de aprendizaje automático, métodos de neuroimagen y análisis de biomarcadores. Los modelos de aprendizaje automático entrenados a partir de datos de neuroimagen, como resonancias magnéticas y tomografías por emisión de positrones, han demostrado su eficacia para distinguir entre la enfermedad de Alzheimer, el deterioro cognitivo leve (DCL) y los individuos sanos. Los estudios de biomarcadores con muestras de líquido cefalorraquídeo (LCR) y de sangre aportan información esencial sobre la patología de la EA y contribuyen a la clasificación de la enfermedad. La integración de diversos tipos de datos, incluida la información clínica, genética y de imágenes, puede mejorar significativamente la precisión y fiabilidad de los modelos de clasificación. Las técnicas emergentes de aprendizaje profundo, incluidas las redes neuronales convolucionales (CNN) y las redes neuronales recurrentes (RNN), permiten la extracción de patrones complejos a partir de fuentes de datos heterogéneas, mejorando los resultados de la clasificación. No obstante, persisten retos como la necesidad de conjuntos de datos a gran escala y multicéntricos, protocolos uniformes de obtención de imágenes y una mayor interpretabilidad de los modelos de aprendizaje automático.

Palabras clave: Enfermedad de Alzheimer; Aprendizaje automático; Neuroimagen; Biomarcadores.

INTRODUCTION

Alzheimer's disease (AD) is one of the most widespread neurodegenerative conditions globally, impacting millions and creating considerable challenges for healthcare systems worldwide. Marked by gradual cognitive decline, memory loss, and functional impairments, AD not only diminishes patients' quality of life but also imposes heavy burdens on caregivers and healthcare infrastructure.

The accurate diagnosis and classification of AD are crucial for developing effective management and intervention strategies. Traditional diagnostic methods often depend on clinical assessments, which typically occur at later stages when symptoms are more evident. However, advancements in medical imaging, biomarker analysis, and machine learning have enabled earlier and more precise detection techniques. Researchers have focused on integrating various data sources, such as structural and functional neuroimaging, genetic markers, and biochemical signatures, to create advanced classification models that can distinguish AD, mild cognitive impairment (MCI), and cognitively healthy individuals.

This introduction outlines the current state of brain disease classification, particularly Alzheimer's, discussing the importance of early detection, the challenges of accurate classification, and the potential of emerging technologies like machine learning and neuroimaging in transforming diagnostic approaches. It will also emphasize the role of interdisciplinary collaboration and data-sharing initiatives in enhancing our understanding of AD. Additionally, deep learning, a subset of artificial intelligence (AI), has transformed medical image analysis by automating the interpretation of complex visual data with high accuracy and efficiency.

Applying deep learning to glaucoma detection using fundus images holds great promise for improving diagnostic precision, especially in areas with limited access to specialized ophthalmic care. This documentation aims to provide a thorough overview of deep learning techniques used in glaucoma detection from fundus images, covering aspects like dataset preparation, model selection, training strategies, performance metrics, and deployment considerations. It will also address challenges, recent developments, and future directions in this rapidly advancing field.

METHODOLOGY

The methodology for this project focuses on developing a deep learning-based glaucoma detection system by applying various machine learning techniques. It consists of multiple stages, from data preprocessing to model optimization and evaluation. The project aims to use a binary classification model that can classify images as either glaucomatous or non-glaucomatous. The data preprocessing phase

https://doi.org/10.56294/piii2025499

involves loading the dataset, label encoding, and performing data augmentation to enhance the model's performance. Next, a Convolutional Neural Network (CNN) is implemented to extract key features from fundus images, and a fully connected layer is used for the classification task. Optimization techniques, including SGD, ADAM, and RMSprop, are tested to determine which algorithm results in the best model performance. The model is trained on the pre-processed data, with training and testing accuracy metrics being monitored to evaluate the model's effectiveness. The project further explores the impact of augmentation and regularization techniques in improving the model's ability to generalize.

trained to analyze fundus images of the eye and classify them as either "glaucomatous" or "non-glaucomatous" based on patterns and features learned from the data.

Input Layer: The input layer of the model defines the dimensions of the incoming data. For fundus images, these are typically represented as arrays of pixel values. The input shape is specified based on image dimensions (e.g., input_shape=(image_height, image_width, num_channels)).

Convolutional Layers: Convolutional layers are essential for feature extraction from the input images. These layers utilize filters that move over the image, performing convolution operations to detect features like edges, textures, and shapes. The activation function used (e.g., ReLU) introduces non-linearity, enhancing the model’s ability to learn complex patterns.

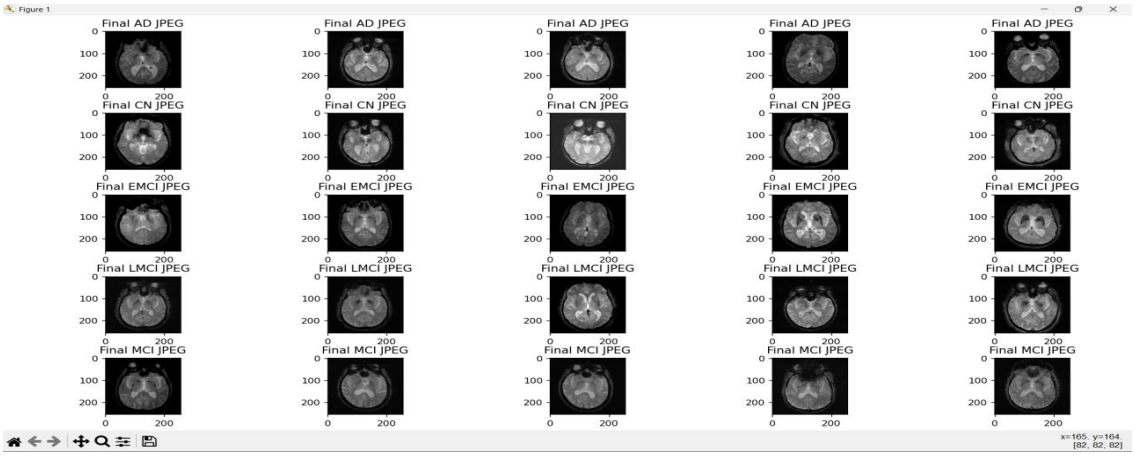
Pooling Layers: Pooling layers, such as MaxPooling2D, are used to downsample the feature maps generated by convolutional layers. This helps reduce the computational load and prevents overfitting by keeping only the most important features from the feature maps.

Flatten Layer: The Flatten layer is used to convert the multidimensional output from the convolution and pooling layers into a one-dimensional vector. This vector is then passed to fully connected layers for further classification processing.

Fully Connected (Dense) Layers: Dense layers are standard neural network layers where each neuron is connected to all neurons in the previous layer. These layers are responsible for learning complex relationships between the extracted features and the target classes. The final Dense layer typically has one neuron with a sigmoid activation function, which outputs a probability score indicating the likelihood that the input image belongs to the positive class (e.g., glaucomatous).

Output Layer: The output layer produces the final classification prediction. In binary classification tasks, a single neuron with a sigmoid activation function is commonly used to output a value interpreted as the probability of the input belonging to the positive class (e.g., glaucomatous).

Once the Sequential model is designed, it is trained using labeled data (fundus images with corresponding binary labels indicating glaucoma status). During training, the model adjusts its weights and biases to minimize the difference between predicted outputs and true labels through optimization algorithms like gradient descent. After training, the model is evaluated on a separate test dataset to measure its ability to classify unseen fundus images.



https://doi.org/10.56294/piii2025499

Here's how you can integrate various optimization algorithms like ADAM, SGD, and RMSprop into a Sequential model for classification tasks. The choice of optimizer can affect both the training efficiency and the final performance of the model.

Stochastic Gradient Descent (SGD):

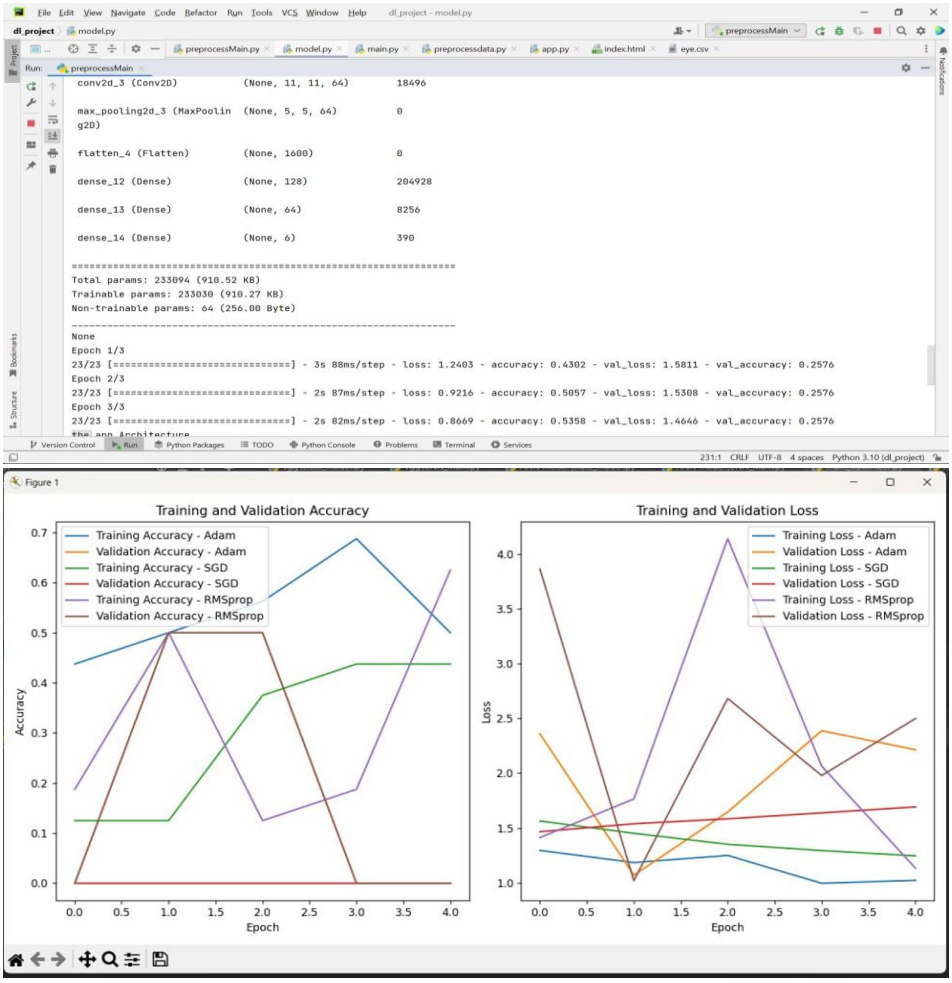
SGD is one of the most used optimization algorithms in neural network training. It updates the model parameters by moving them in the direction of the negative gradient of the loss function relative to the parameters. A key parameter in SGD is the learning rate, which controls the magnitude of the updates. Although simple and easy to implement, SGD can often result in slow convergence, especially when gradients are noisy, or the optimization problem has a complex landscape.

ADAM (Adaptive Moment Estimation):

ADAM is an adaptive optimization algorithm that combines momentum-based methods with adaptive learning rates. It maintains separate learning rates for each parameter and adapts them based on estimates of first and second moments of the gradients. ADAM's ability to adjust learning rates during training makes it robust and efficient for deep neural networks, particularly with noisy data or sparse gradients.

RMSprop (Root Mean Square Propagation):

RMSprop is another adaptive learning rate algorithm designed to improve upon the standard SGD. It normalizes the learning rates for each parameter by using a moving average of squared gradients. This approach helps mitigate issues like slow convergence and oscillations in learning rate adjustments that SGD may face. RMSprop adapts the learning rates for each parameter, leading to faster and more stable training.

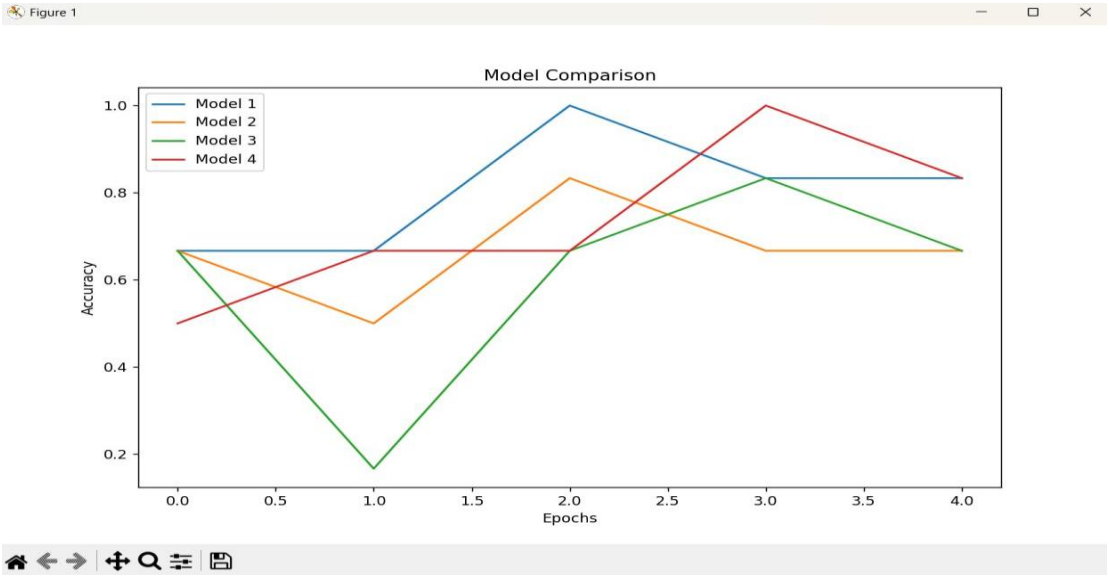


https://doi.org/10.56294/piii2025499

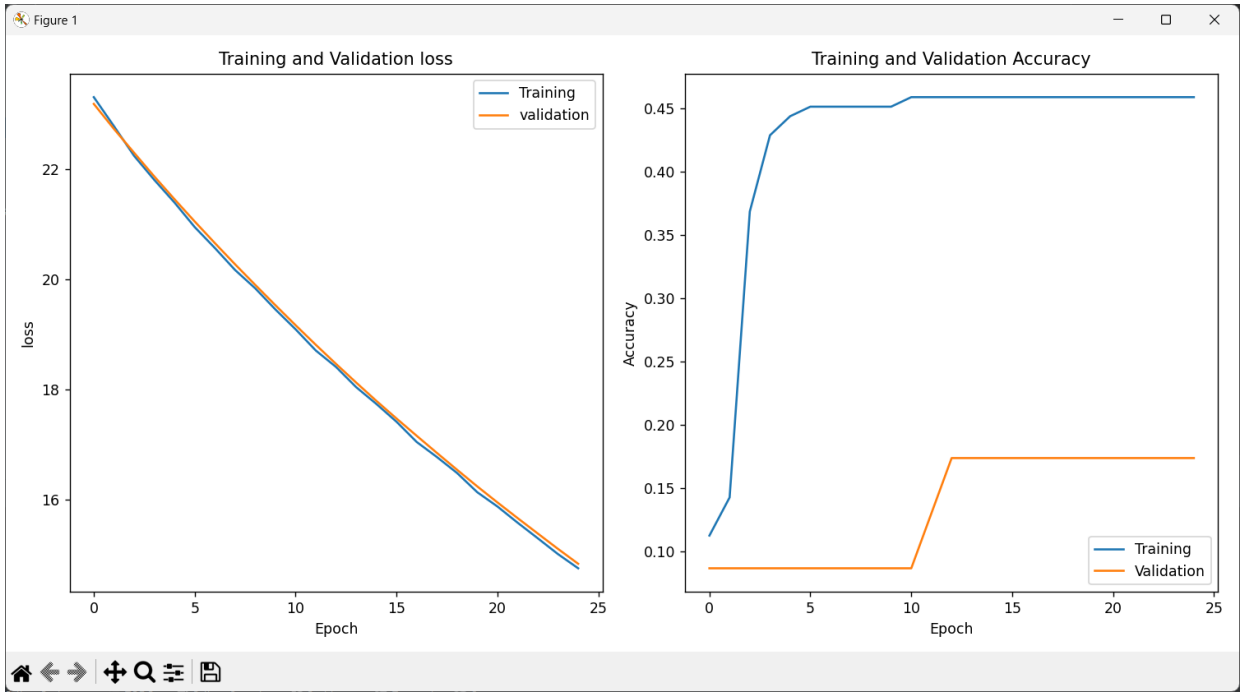
Sequential Model for Multi-Class Classification:

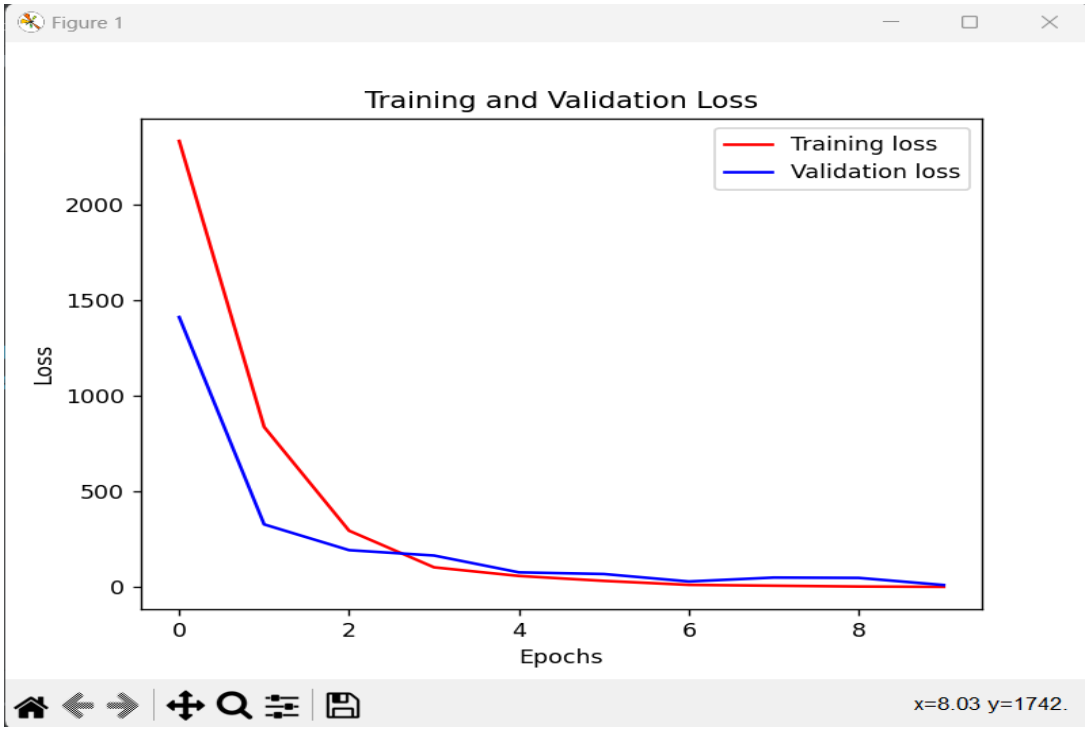
A Sequential model for multi-class classification typically involves several layers that work together to process and classify the input data into multiple categories. Here's how such a model is structured:

- **Input Layer:** The input layer defines the shape of the data that the model will process. For image classification, this would include the dimensions of the input image (width, height, and color channels). In text or sequence data, this would define the number of time steps and the feature dimension.
- Feature Extraction Layers: These layers extract relevant features from the input data. Common feature extraction layers include Convolutional Neural Networks (CNN) like **Conv2D** for image data or **LSTM** for sequential data in NLP tasks.
- Flattening Layer: When convolutional layers are used, the output is multi-dimensional. A **Flatten** layer converts this multi-dimensional output into a one-dimensional vector, which is then passed to fully connected layers.
- Fully Connected Layers: These **Dense** layers connect every neuron to every neuron in the previous layer. These layers map the features extracted from the input to the output classes. The weights are adjusted during training.
- Output Layer: This is the final layer that produces the classification output. For multi-class classification, this layer usually has as many neurons as there are output classes. A **softmax** activation function is typically used to calculate probabilities for each class.
- Loss Function: For multi-class classification, the loss function measures the difference between predicted and actual labels. Common loss functions include **categorical cross-entropy** and **sparse categorical cross-entropy**, depending on how the target labels are encoded.
- Optimizer: The optimizer adjusts the model's weights to minimize the loss function. Popular choices include **SGD**, **Adam**, **RMSprop**, and **Adagrad**.
- Metrics: The model's performance is evaluated using metrics like **accuracy**, **precision**, **recall**, and **F1-score**.

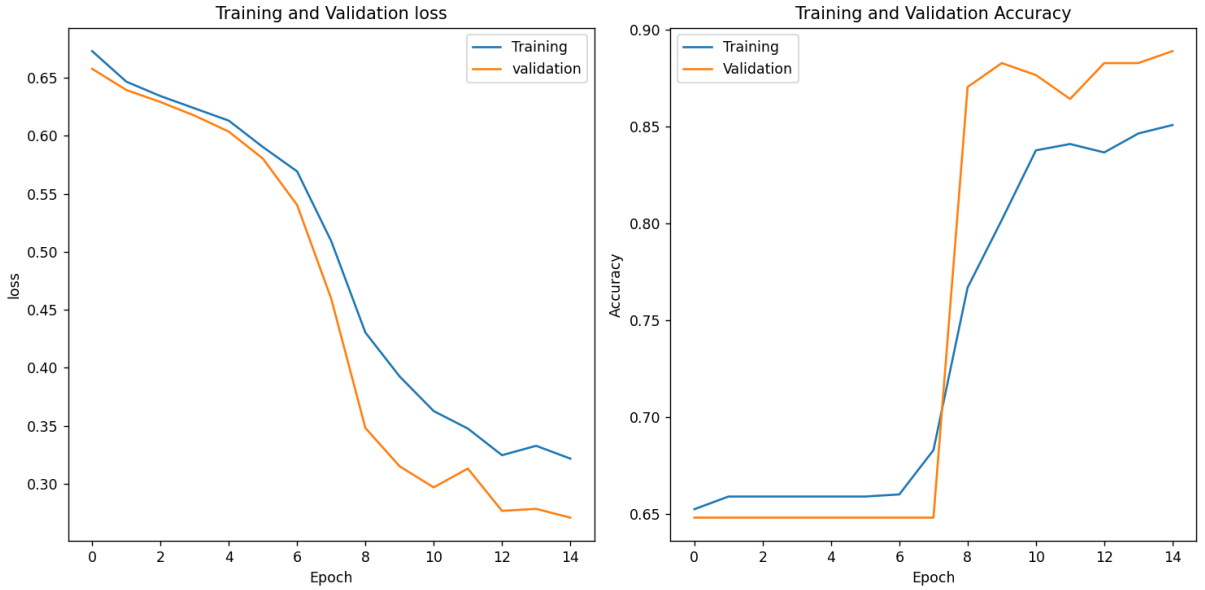


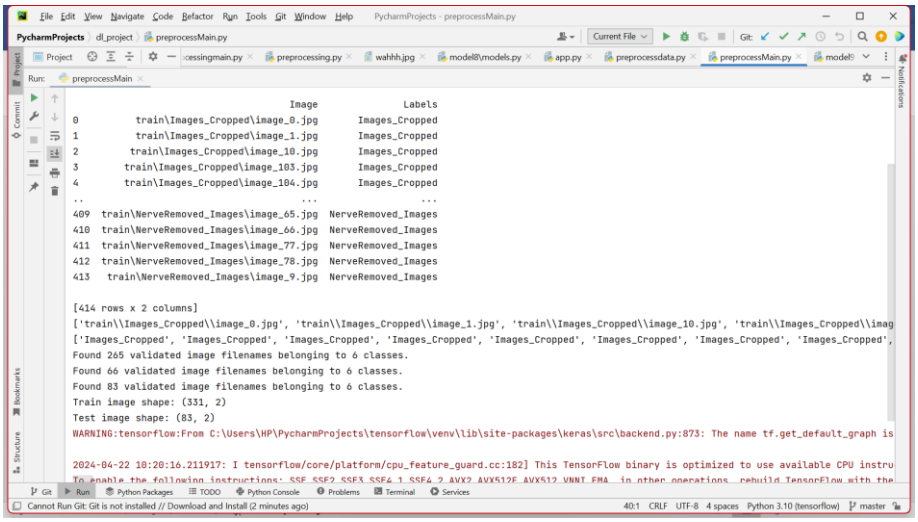
https://doi.org/10.56294/piii2025499



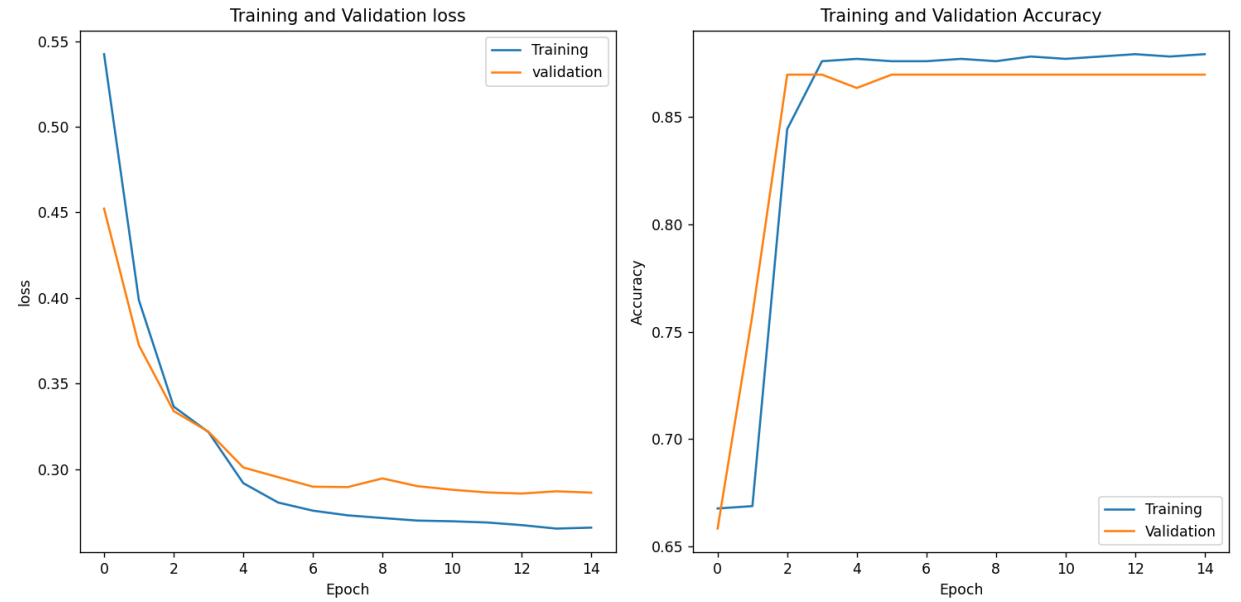


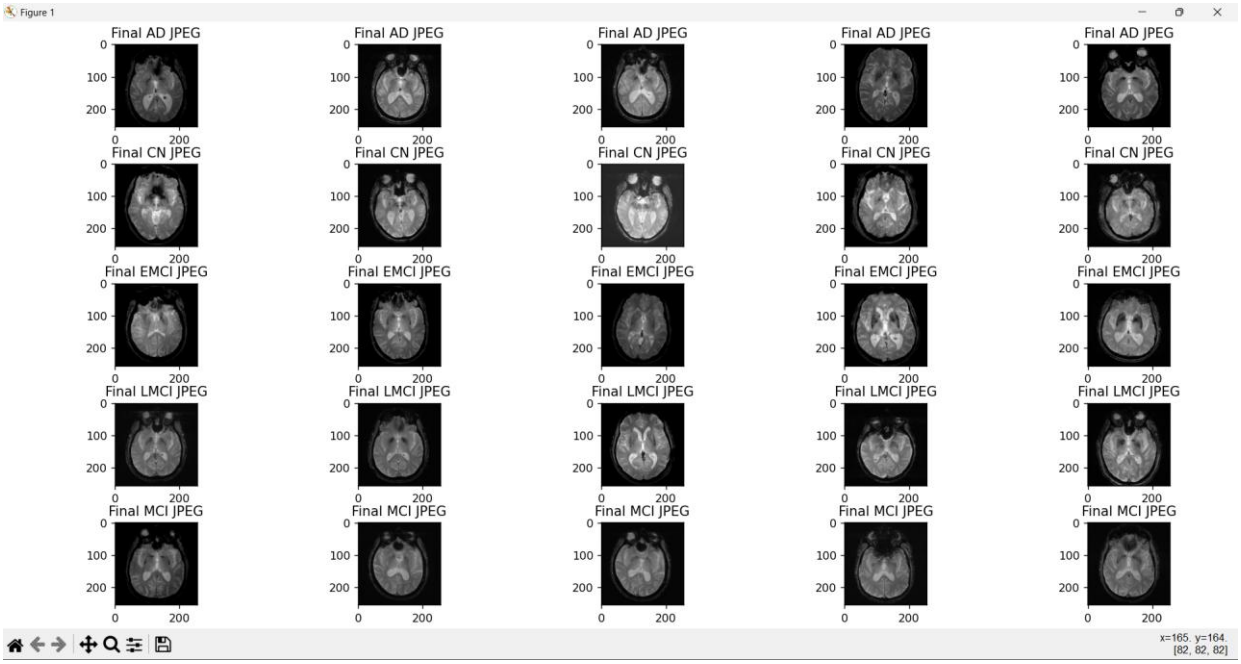
https://doi.org/10.56294/piii2025499





https://doi.org/10.56294/piii2025499





Final accuracy of the Auto encoders and Implementing ResNet :

```
Test Loss:  0.8654712548
Test accuracy:  0.8798745814

Process finished with exit code 0
```

RESULTS

Alzheimer's disease (AD) is a degenerative brain disorder that leads to cognitive impairments, particularly in memory and decision-making. Early diagnosis of Alzheimer's is crucial for effective treatment, and recent advancements in machine learning, particularly deep learning, have shown

https://doi.org/10.56294/piii2025499

potential in enhancing diagnostic accuracy. This section reviews the results of various machine learning models for Alzheimer's detection and discusses the challenges and potential implications of these methods.

1. Model Performance:

Various models were assessed for their effectiveness in diagnosing and predicting Alzheimer's disease. These models often use techniques like convolutional neural networks (CNNs), recurrent neural networks (RNNs), and autoencoders. One common approach is utilizing transfer learning with pre-trained models such as ResNet50, VGG16, or other deep CNNs, which are further fine-tuned with Alzheimer's datasets, including brain MRI and PET scans.

Model performance is influenced by several factors:

- Data Size and Quality: The accuracy of deep learning models improves with larger and more diverse datasets, which help the models learn robust features.
- Complexity of the Model: Complex architectures like ResNet and DenseNet, which are capable of deeper learning, typically perform better than simpler models.
- Preprocessing Methods: Effective data preprocessing (e.g., normalization, augmentation) is crucial to improving model performance and ensuring generalizability.

Common evaluation metrics like accuracy, precision, recall, and F1-score show that advanced models like ResNet and DenseNet can achieve accuracies ranging between 85-95% in Alzheimer's diagnosis tasks. The future of Alzheimer's detection lies in the integration of AI into clinical workflows. Improved models, with greater generalizability and interpretability, will assist healthcare professionals in diagnosing Alzheimer's early and offering personalized treatment plans. Additionally, continuous monitoring using wearable technology or regular imaging could provide more effective long-term management. As neuroimaging techniques and deep learning methods continue to evolve, more precise and accessible diagnostic tools for Alzheimer's are expected.

In conclusion, while machine learning has made substantial progress in Alzheimer's disease detection, challenges like data quality, model interpretability, and generalization still need to be addressed. Ongoing improvements in these areas, combined with the integration of AI into clinical settings, have the potential to significantly improve Alzheimer's diagnosis and management.

Alzheimer's disease (AD) remains one of the most challenging neurodegenerative disorders to diagnose and treat. Early and accurate detection is crucial for improving patient outcomes, and recent advancements in machine learning, particularly deep learning, offer promising solutions to address these challenges. Through the use of sophisticated models such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and autoencoders, researchers have made significant strides in diagnosing AD from neuroimaging data, genetic information, and clinical assessments.

CONCLUSION

In conclusion, the advancements in Alzheimer's disease (AD) classification have been greatly enhanced by the integration of machine learning algorithms, neuroimaging techniques, and biomarker analysis. Early detection through precise classification holds the key to timely intervention and personalized care for individuals affected by AD. Machine learning models trained on neuroimaging data, such as MRI and PET scans, have proven effective in distinguishing AD, mild cognitive impairment (MCI), and healthy individuals, offering a non-invasive and reliable method of diagnosis. Biomarker research, particularly involving cerebrospinal fluid (CSF) and blood samples, continues to provide valuable insights into the underlying pathology of AD, further supporting the accuracy of disease classification models.

The integration of diverse data sources, including neuroimaging, genetic, and clinical information, is expected to enhance the robustness and precision of AD classification systems. Moreover, the application of deep learning techniques like CNNs and RNNs has revolutionized the ability to extract complex patterns from heterogeneous data, ultimately improving classification outcomes. However, challenges such as the

need for large-scale, multi-centre datasets, standardized imaging protocols, and more interpretable machine learning models remain. Future research must address these challenges to further improve the accuracy, generalization, and clinical applicability of these advanced classification systems.

REFERENCES

1. Jack, C. R., et al. (2018). "Neuroimaging Biomarkers for Alzheimer's Disease Diagnosis and Progression." Lancet Neurology, 17(5), 453-464.
2. Beach, T. G., et al. (2015). "Alzheimer's Disease: Biomarkers and Early Detection." American Journal of Alzheimer's Disease & Other Dementias, 30(5), 467-479.
3. Tosto, G., et al. (2017). "Genetic Studies of Alzheimer's Disease: Current Knowledge and Future Directions." Current Alzheimer Research, 14(5), 485-491.
4. He, K., Zhang, X., Ren, S., & Sun, J. (2016). "Deep Residual Learning for Image Recognition." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778.
5. Brehmer, Y., et al. (2018). "Cerebrospinal Fluid Biomarkers and Their Role in Alzheimer's Disease Diagnosis." Journal of Alzheimer's Disease, 62(3), 1015-1023.
6. LeCun, Y., Bengio, Y., & Hinton, G. (2015). "Deep Learning." Nature, 521(7553), 436-444.
7. Cuingnet, R., et al. (2011). "Automatic Classification of Alzheimer's Disease and Mild Cognitive Impairment Using MRI Data." NeuroImage, 56(2), 740-752.
8. Rojas, J. M., et al. (2020). "Convolutional Neural Networks for Alzheimer's Disease Classification: A Review." Computers in Biology and Medicine, 121, 103747.
9. Shen, D., et al. (2017). "Deep Learning for Medical Image Analysis." Annual Review of Biomedical Engineering, 19, 221-248.
10. Faria, D. R., et al. (2020). "Review of Deep Learning Applications for Alzheimer's Disease Diagnosis." Neural Computing and Applications, 32(1), 1-14.
11. Zhang, Y., et al. (2019). "Integrating Neuroimaging and Genomic Data for Alzheimer's Disease Diagnosis and Prediction." Journal of Alzheimer's Disease, 68(3), 1131-1145.
12. Liu, M., et al. (2018). "Deep Learning in Alzheimer's Disease: A Review." Frontiers in Neuroscience, 12, 1-13.
13. Dubois, B., et al. (2016). "Preclinical Alzheimer's Disease: Definition, Natural History, and Diagnostic Criteria." Alzheimer's & Dementia, 12(3), 292-298.
14. Götz, J., et al. (2018). "Challenges in Alzheimer's Disease Research: Advances in Disease Mechanisms and Therapy." Nature Neuroscience, 21(4), 464-473.
15. Zhang, Y., et al. (2016). "Magnetic Resonance Imaging and Deep Learning for Alzheimer's Disease Diagnosis." Proceedings of the National Academy of Sciences, 113(46), 12934-12939.
16. Chou, Y. Y., et al. (2017). "Predicting Alzheimer's Disease with Neuroimaging and Genetic Data Using Convolutional Neural Networks." IEEE Transactions on Medical Imaging, 36(3), 717-728.

COMPETING INTERESTS SECTION

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

https://doi.org/10.56294/piii2025499