Predictive analysis for Alzheimer's diagnosis through data mining techniques

Stefano Coscarella[2], Tommaso Ruga[3,1], Ester Zumpano[3,4], Eugenio Vocaturo* [1,3]

[1] CNR-NANOTEC National Research Council, Rende, Italy

[2] DESF - Department of Economic Statistics and Finance, University of Calabria, Rende, Italy

[3] DIMES - Department of Computer Engineering, Electronic Modelling, and Systems Engineering, University of Calabria, Rende, Italy

{cscsfn93s04d086o@studenti.unical.it; tommaso.ruga@dimes.unical.it; e.zumpano@dimes.unical.it; eugenio.vocaturo@cnr.it} * Corresponding Author

---

Abstract: Alzheimer's disease represents one of the most significant challenges for contemporary healthcare, as there is still no effective cure for it. Magnetic Resonance Imaging (MRI) techniques have contributed to the diagnosis and prediction of its progression, but they require time and specialized skills for image analysis. Therefore, the use of deep learning techniques is crucial in analyzing large amounts of MRI images with high accuracy for early detection and prediction of the disease progression. In the following work, we focused on feature extraction from multiple sources and their integration to improve the accuracy of Alzheimer's diagnosis. Three distinctive methodologies have been developed. The first one utilizes a Feed Forward Neural Network (FFNN) with features extracted from models like ResNet50 and DenseNet201 with and without the application of Principal Component Analysis (PCA). The second methodology combines features from both models, with and without the application of PCA. Finally, the third methodology combines features from the models with those extracted from manual techniques like Discrete Wavelet Transform (DWT), Local Binary Pattern (LBP), and Gray Level Co-occurrence Matrix (GLCM) by applying PCA before or after feature combination.

Keywords: Alzheimer Detection; Magnetic Resonance Imaging; Deep learning; Features extraction.

Introduction

Alzheimer's disease (AD) is a neurodegenerative condition characterized by cognitive decline and abnormal protein accumulation in the brain. This pathology affects millions of individuals worldwide, emerging as the most common form of dementia among the elderly. AD prevails due to its incidence, manifesting through a gradual deterioration of memory, cognitive ability, and daily activities. According to the World Health Organization (WHO), dementia affects more than 50 million people worldwide, with approximately 60% of cases attributable to AD. This increasing prevalence has significant social and economic impacts. Although there is still no effective cure, early diagnosis is essential for adopting timely interventions and slowing the progression of the disease. Brain imaging techniques, such as magnetic resonance imaging (MRI), play a crucial role in the diagnosis and prediction of AD, although they require time and specialized skills.

The use of artificial intelligence techniques is fundamental to overcome these challenges, offering the possibility of analyzing data efficiently and comprehensively, thus improving the diagnosis and management of the disease. In this work, through the application of Deep Learning (DL) techniques, various methodologies will be examined to obtain a set of meaningful metrics to distinguish AD and predict its different stages. Due to the similarity of the biological and clinical features of Alzheimer's stages, it is difficult to distinguish between the stages of AD development. In order to optimize the effectiveness of classification, various complementary techniques have been adopted and integrated. In particular, techniques aimed at improving the quality of images obtained through MRI and identifying specific regions of interest (ROI) through segmentation have been used. Additionally, a manual feature extraction process has been performed from the identified ROI. This approach aims to maximize the information extracted from diagnostic images, integrating the effectiveness of DL techniques with human expertise in identifying and extracting features relevant to the classification of different stages of AD.

Related work

In addition to the following study, there is a vast landscape of related research exploring similar approaches and addressing analogous challenges in the application of advanced techniques in artificial intelligence. Conducting an analysis of related works allows for a more comprehensive understanding of the current state of research in this field and identifies possible points of convergence and future development. Below is a review of some previous research, accompanied by a brief description of the techniques used and the results obtained.

In the study conducted by Minseok Song et al.[1] , an extensive collection of brain magnetic resonance imaging (MRI) data provided by ADNI was examined to compare the effectiveness of different machine learning approaches. Among these approaches, models such as Support Vector Machine (SVM), Multi-layer perceptron (MLP), and Convolutional Neural Network (CNN) were tested, in addition to the Random Forest (RF) model. The results showed that the Random Forest (RF) algorithm achieved the highest performance, with a precision of 92.43%, recall of 87.26%, and F1-score of 89.2%.

Gopi Battineni et al.[2] compare the effectiveness of various machine learning approaches in analyzing brain magnetic resonance imaging (MRI) data provided by ADNI. The techniques used include Random Forest (RF), Support Vector Machine (SVM), Gaussian Naive Bayes (GNB), Logistic Regression (LR), Gradient Boosting (GB), and AdaBoost. Model validation was performed

[1] Song M, Jung H, Lee S, Kim D, Ahn M. *Diagnostic Classification and Biomarker Identification of Alzheimer's Disease with Random Forest Algorithm*, 2021

[2] Battineni G, Hossain MA, Chintalapudi N, Traini E, Dhulipalla VR, Ramasamy M, Amenta F. *Improved Alzheimer's Disease Detection by MRI Using Multimodal Machine Learning Algorithms, 2021*

using the 10-fold cross-validation method to evaluate accuracy and prevent overfitting. The best results were achieved with the Gradient Boosting (GB) algorithm, with a precision of 98%, recall of 96%, and F1-score of 97%.

Haijing Sun et al.[3] propose a ResNet model based on Spatial Transformer Networks (STN) for early diagnosis of AD. The Mish activation function was set for the ResNet50 model, and STN was inserted into the ResNet-50 layers. The proposed method was validated using the brain magnetic resonance imaging (MRI) dataset provided by ADNI and compared with other baseline models such as ResNet50, ResNet50 + Mish, and STN + ResNet50 + Mish. The results showed a precision of 95.5%, recall of 95.3%, and F1-score of 95.4%.

Ahmed A. Abd El-Latif et al.[4] introduce a Lightweight Convolutional Neural Network (LCNN) model for the detection of Alzheimer's disease (AD) stages through an MRI image dataset provided by Kaggle. This model, designed for real-time applications, distinguishes itself from classic Convolutional Neural Network (CNN) models by its ability to achieve high detection performance without the need for deep layers. It eliminates the use of traditional feature extraction and classification methods, condensing them into a single phase. With only seven layers, the system is less complex compared to other CNN models. The results obtained from this model showed a precision of 95.93%, recall of 95.88%, and F1-score of 95.90%.

Due to the similar characteristics of MRI images in the early stages of Alzheimer's disease and the low contrast between diseased and healthy brain cells, it is crucial to integrate different feature extraction techniques to obtain a more complete and discriminative representation of the data. Compared to the listed works, which have mainly focused on single types of features extracted from AI models, the following work stands out for the integration of Convolutional Neural Networks (CNN) with handcrafted methods such as Discrete Wavelet Transform (DWT), Local Binary Models (LBP) and Gray Level Co-occurrence Matrix (GLCM). This integration allows capturing both relevant visual features and detailed information extracted manually, overcoming limitations related to low contrast and similarity of brain structures in the early stages of the disease. The inclusion of manually extracted features makes the following model more transparent and interpretable, allowing experts to better understand the model's decision-making process and increasing confidence in the accuracy of predictions in the clinical setting of early Alzheimer's diagnosis.

[3] Sun H, Wang A, Wang W, Liu C. *An Improved Deep Residual Network Prediction Model for the Early Diagnosis of Alzheimer's Disease, 2021*
[4] El-Latif AAA, Chelloug SA, Alabdulhafith M, Hammad M. *Accurate Detection of Alzheimer's Disease Using Lightweight Deep Learning Model on MRI Data*. Diagnostics, 2023

Table 1. Compares this work with the related work or previous research by other researchers.

Authors and Reference	Precision	Recall	F1-score
Minseok Song et al. [1]	92,43%	87,26%	89,2%
Gopi Battineni et al. [2]	**98%**	**96%**	**97%**
Haijing Sun et al. [3]	95,5%	95,3%	95,4%
Ahmed A. Abd El-Latif et al. [4]	95,93%	95,88%	95,90%
This work	92,93%	92,64%	92,62%

Key Contribution

1. **Development of Innovative Methodologies:** The work proposes three distinctive methodologies for the analysis of MRI images in the context of early diagnosis and prediction of Alzheimer's disease progression. These methodologies combine features extracted from deep learning models and manual techniques, offering an integrated approach to improve diagnostic accuracy.
2. **Integration of Features from Multiple Sources**: A significant aspect of this work is the integration of features extracted from various sources, such as pre-trained deep learning models and manual techniques. This approach aims to capture a broader range of information from MRI images, thereby enhancing the ability to discriminate between individuals affected by Alzheimer's and healthy individuals.
3. **Evaluation of PCA Effectiveness:** The use of Principal Component Analysis (PCA) to reduce the dimensionality of the extracted features is examined in detail in all three proposed methodologies. This allows for assessing the impact of dimensionality reduction on the diagnostic performance of the system, providing a deeper understanding of the utility of PCA in this context.
4. **Application of Feedforward Neural Networks:** The utilization of feedforward neural networks (FFNN), along with other deep learning techniques, underscores the significance of advanced machine learning models in Alzheimer's diagnosis. This work demonstrates how such models can be effectively integrated with other techniques to enhance overall diagnostic performance.

Method, Experiments and Results

1. FFNN Network According to the Features of CNN Models

In the first method, features extracted from ResNet and DenseNet, along with features optimized and reduced through PCA, were used as inputs for the FFNN. In this context, four distinct FFNN models were created: two models using the original features extracted from ResNet and DenseNet, and two models employing features reduced through PCA. This approach allows for assessing the impact of dimensionality reduction on the predictive performance of the model, enabling a direct comparison between the use of original features and reduced features.

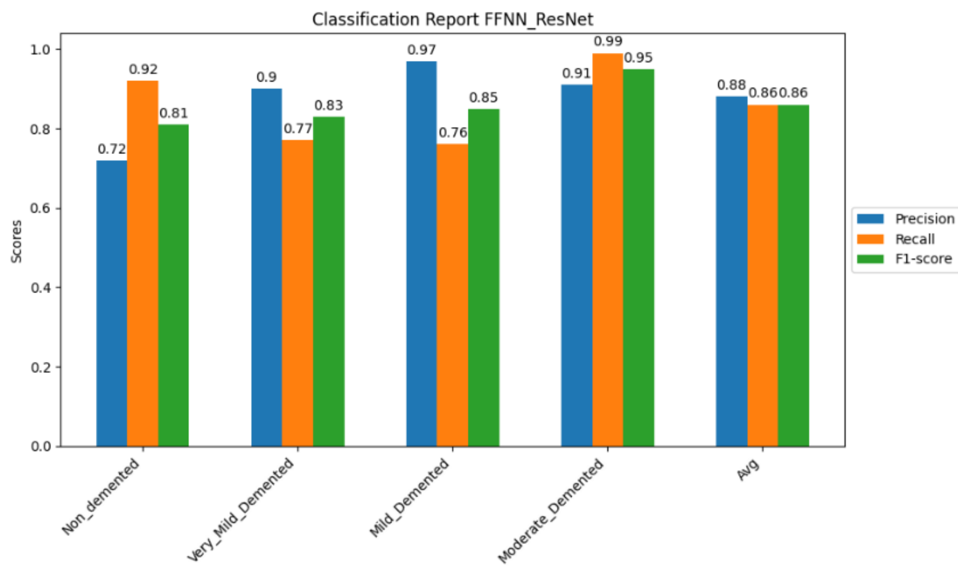

Figure 1 Classification report FFNN_ResNet

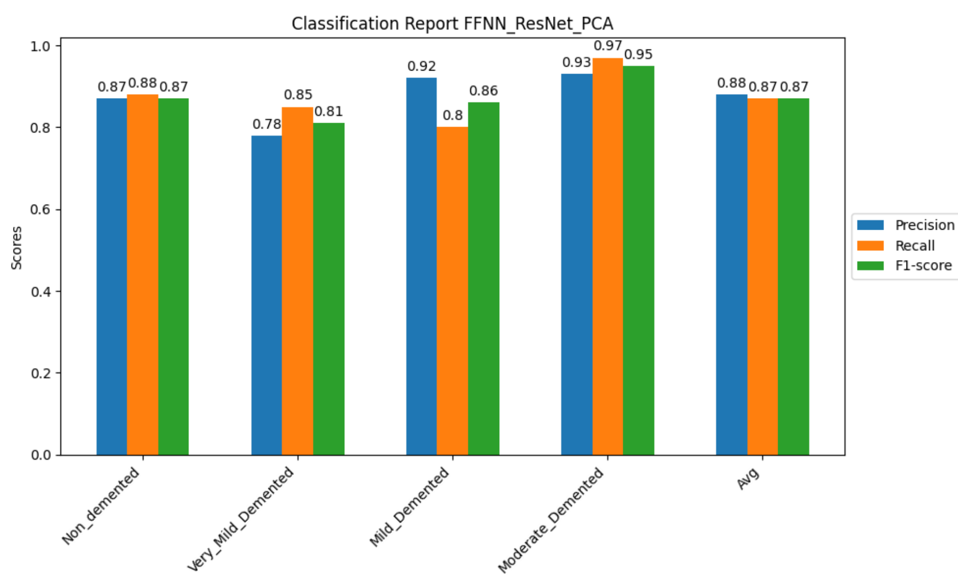

Figure 2 Classification report FFNN_ResNet_PCA

In the FFNN_ResNet model, we observe variations in precision, recall, and F1-score metrics across different classes. The "Mild Demented" class exhibits the highest precision (97%), while the "Non Demented" class shows the lowest precision (72%). Regarding recall, the "Moderate Demented" class achieves the highest value (99%), whereas the "Mild Demented" class has the lowest value (76%). The F1-score, which considers both precision and recall, ranges from 81% to 95% across different classes. Overall, the model demonstrates a good balance between precision and recall, with significantly high F1-score scores.

For the FFNN_ResNet_PCA model, there is a general trend of improvement in precision, recall, and F1-score metrics compared to the model without PCA. Variations between classes are less pronounced, indicating greater uniformity in the model's performance. In addition to increased uniformity, the higher average values of recall and F1-score suggest an overall improvement in the model's performance in correctly retrieving positive instances (recall) and achieving a balance between precision and recall (F1-score). This indicates that the model has become more effective in correctly recognizing the various classes of dementia, potentially reducing the number of false positives and false negatives. This improvement can be attributed to better data representation after applying PCA and a reduction in overfitting.

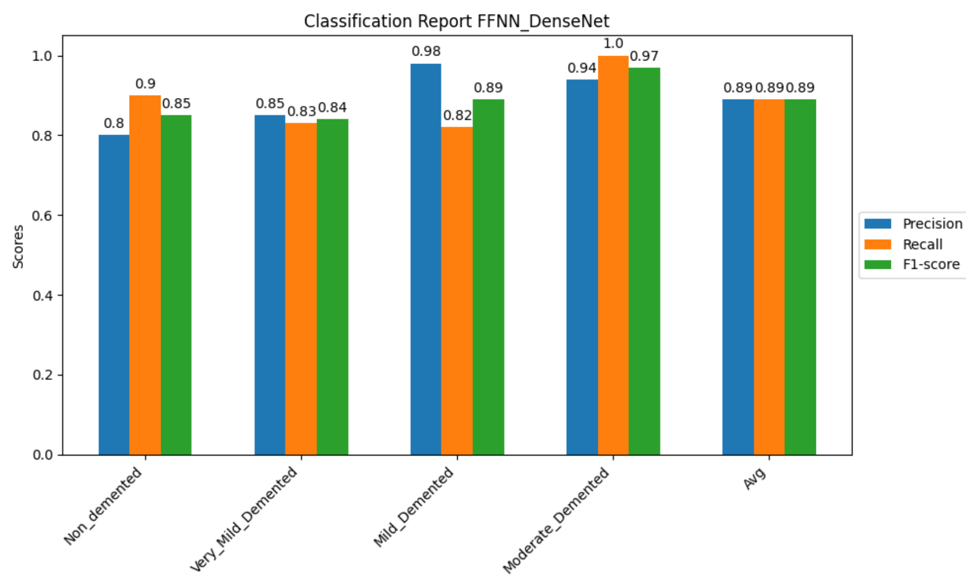

Figure 3 Classification report FFNN_DenseNet

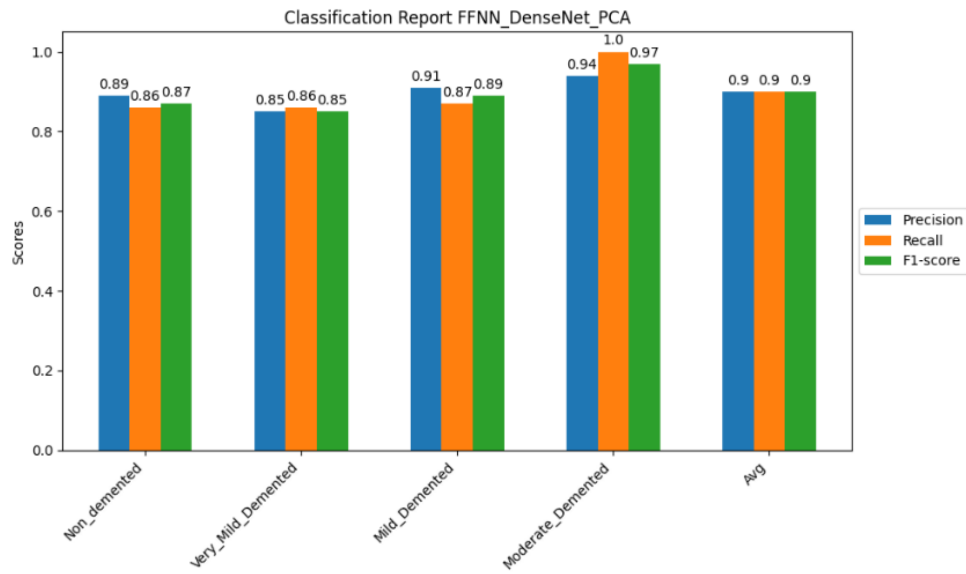

Figure 4 Classification report FFNN_DenseNet_PCA

Both models, FFNN_DenseNet and FFNN_DenseNet_PCA, demonstrate promising results in the classification report. The model with PCA shows higher average values of precision, recall, F1-score, and accuracy compared to the model without PCA, with each of the listed metrics increasing from 89% to 90%. This suggests that the PCA model is able to classify observations more accurately and sensitively, with greater consistency in average performance across all classes. Overall, these results suggest that the FFNN_DenseNet_PCA model is a better choice than the FFNN_DenseNet model, as it offers slightly superior performance in classifying dementia observations. PCA appears to have contributed to improving the stability and sensitivity of the model, enabling better generalization and greater adaptability to validation data.

2. FFNN Network According to Fusing Features of CNN Models

In the second method, features extracted from ResNet and DenseNet were utilized in two different modes. In the first mode, features extracted from both models were combined and then subjected to dimensionality reduction using PCA. The features obtained from this process were then used as input for the FFNN. In the second mode, the optimized and reduced features with PCA, extracted separately from ResNet and DenseNet, were concatenated and directly used as input for the FFNN. This approach allows for evaluating the effectiveness of feature combination before dimensionality reduction, comparing it with the direct use of reduced features.

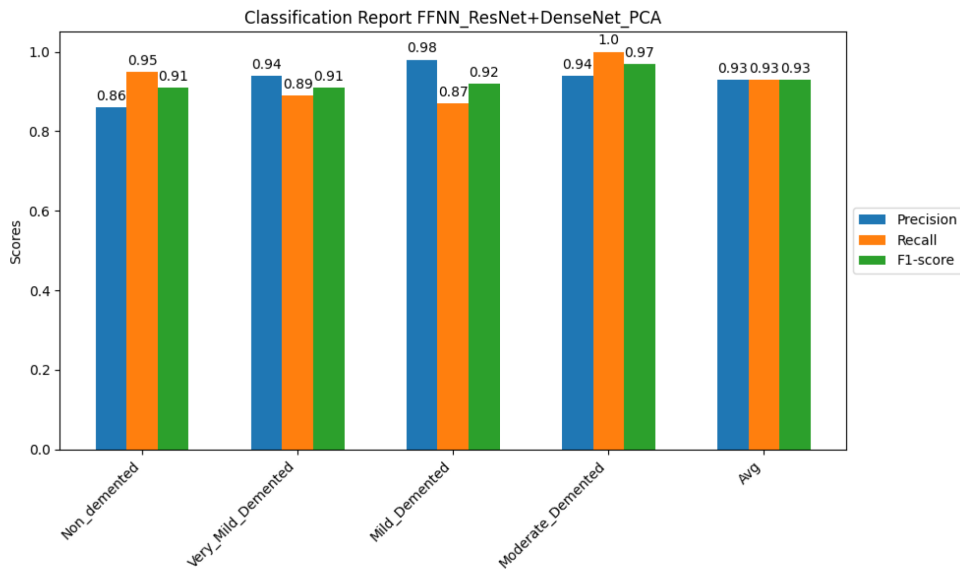

Figure 5 Classification report FFNN_ResNet+DenseNet_PCA

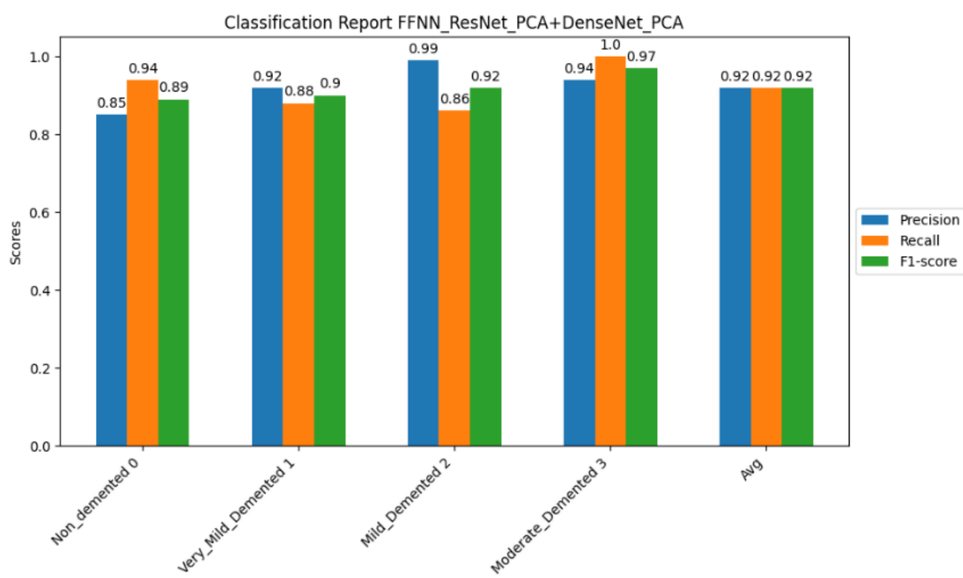

Figure 6 Classification report FFNN_ResNet_PCA+DenseNet_PCA

Both models show very similar results for the metrics of individual dementia classes, with precision, recall, and F1-score varying slightly between the two approaches. However, it is important to note that the FFNN_ResNet+DenseNet_PCA model achieves slightly higher average precision, recall, and F1-score values compared to the FFNN_ResNet_PCA+DenseNet_PCA model. Specifically, the former has an average precision of 93%, compared to 92% in the latter. Similarly, the average recall is 93% in the former, compared to 92% in the latter. This translates to an average F1-score of 93% for the former, compared to 92% for the latter.

These improvements in average metric values indicate a greater ability of the FFNN_ResNet+DenseNet_PCA model to discriminate between different dementia categories, improving precision, recall, and F1-score uniformly across all classes. This trend suggests that applying PCA after feature fusion has contributed to a better data representation, leading to an overall improvement in the model's performance.

3. FFNN Network According to Fusing Features of CNN Models and ROI

In the third method, features extracted from DWT, GLCM, and LBP were first normalized and then concatenated. Subsequently, two model configurations were created. In the first configuration, the concatenated features from ResNet and DenseNet were combined with the concatenated features from DWT, GLCM, and LBP. These aggregated features were then subjected to dimensionality reduction through PCA before being used as input for the FFNN. In the second configuration, the optimized and reduced features with PCA from ResNet and DenseNet were concatenated with the concatenated features from DWT, GLCM, and LBP, without further dimensionality reduction. This approach allows for evaluating the effectiveness of combining features extracted from CNN models with those obtained from other feature extraction techniques, such as DWT, GLCM, and LBP, both with and without PCA before use on FFNN.

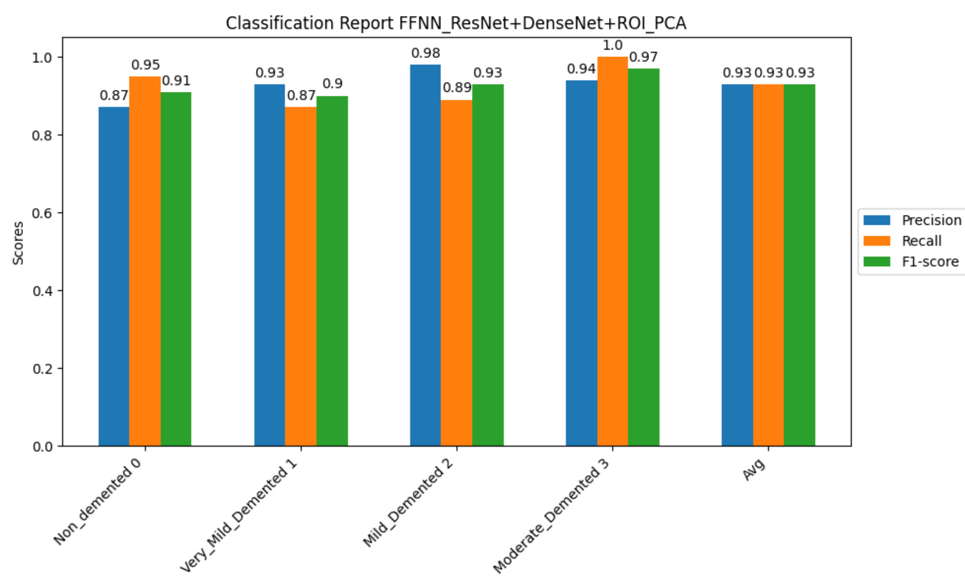

Figure 7 Classification report FFNN_ResNet+DenseNet+ROI_PCA

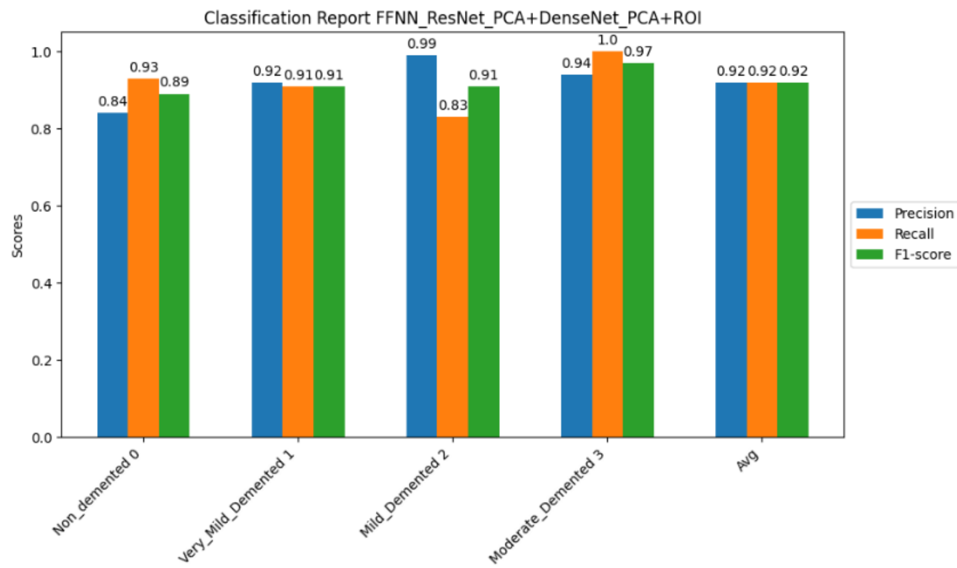

Figure 8 Classification report FFNN_ResNet_PCA+DenseNet_PCA+ROI

Both models exhibit overall excellent results, with high precision, recall, and F1-score scores for all dementia classes. However, there are some differences in metric values between the two models. In the FFNN_ResNet+DenseNet+ROI_PCA model, slightly higher precision, recall, and F1-score scores are observed for some classes compared to the FFNN_ResNet_PCA+DenseNet_PCA+ROI model. Specifically, in addition to the improvements in the metrics of individual classes, the former model has an average precision of 93%, while the latter model has an average precision of 92%.

The same applies to recall and F1-score, where the former model has slightly higher average values than the latter. These differences suggest that the approach used in the former model, with the application of PCA before feature fusion, may have led to better overall predictive capability compared to the latter model, where PCA was applied after feature fusion. The difference in approach in handling features may have influenced the model's ability to capture relevant information and generalize better to test data.

Discussions

Model	**Accuracy**	**Precision**	**Recall**	**F1-score**	**AUC**
FFNN_ResNet	86,07%	87,66%	86,07%	86,08%	96,19%
FFNN_ResNet_PCA	87,29%	87,59%	87,29%	87,29%	96,91%

FFNN_DenseNet	88,57%	89,14%	88,57%	88,57%	97,03%
FFNN_DenseNet_PCA	89,65%	89,66%	89,71%	89,65%	97,11%
FFNN_ResNet+DenseNet_PCA	**92,64%**	**92,96%**	**92,64%**	**92,62%**	**97,91%**
FFNN_ResNet_PCA+DenseNet_PCA	91,93%	92,33%	91,93%	91,91%	98,61%
FFNN_ResNet+DenseNet+ROI_PCA	**92,64%**	**92,93%**	**92,64%**	**92,62%**	**98,49%**
FFNN_ResNet_PCA+DenseNet_PCA+ROI	91,79%	92,24%	91,79%	91,76%	98,14%

The analysis of the performance of various approaches used in the classification of MRI images of Alzheimer's patients reveals a promising picture. The use of Convolutional Neural Networks (CNNs) such as ResNet and DenseNet for feature extraction has proven to be effective in capturing relevant information contained in the images. The subsequent application of Principal Component Analysis (PCA) has contributed to reducing the dimensionality of features, allowing for better generalization and increased robustness of the models. Furthermore, the integration of features extracted from regions of interest (ROI) through techniques such as Discrete Wavelet Transform (DWT), Gray-Level Co-occurrence Matrix (GLCM), and Local Binary Patterns (LBP) has further enriched the information useful for classification.

The most performing models, such as FFNN_ResNet+DenseNet_PCA and FFNN_ResNet+DenseNet+ROI_PCA, have demonstrated a significant capability in discriminating between the various dementia classes. In comparing the two models, there is notable consistency in accuracy, precision, recall, and F1-score values, indicating solid predictive capability in both cases. However, the Area Under the Curve (AUC) stands out as the only metric that shows a substantial difference between the two models.

The FFNN_ResNet+DenseNet_PCA model achieved an AUC of 97.91%, while the FFNN_ResNet+DenseNet+ROI_PCA model reached an AUC of 98.49%. These results suggest that, although both models demonstrate good overall classification capability, the integration of features from ROI appears to offer a slight advantage in the ability to discriminate between different dementia classes, as highlighted by the slightly higher AUC. Furthermore, examining the confusion matrix of the FFNN_ResNet+DenseNet+ROI_PCA model, we can observe that out of a total of 1400 images, 1297 were classified correctly. This highlights the robustness of the model in accurately distinguishing between the different classes. Its ability to minimize classification errors is crucial for ensuring the precision and reliability of the model.

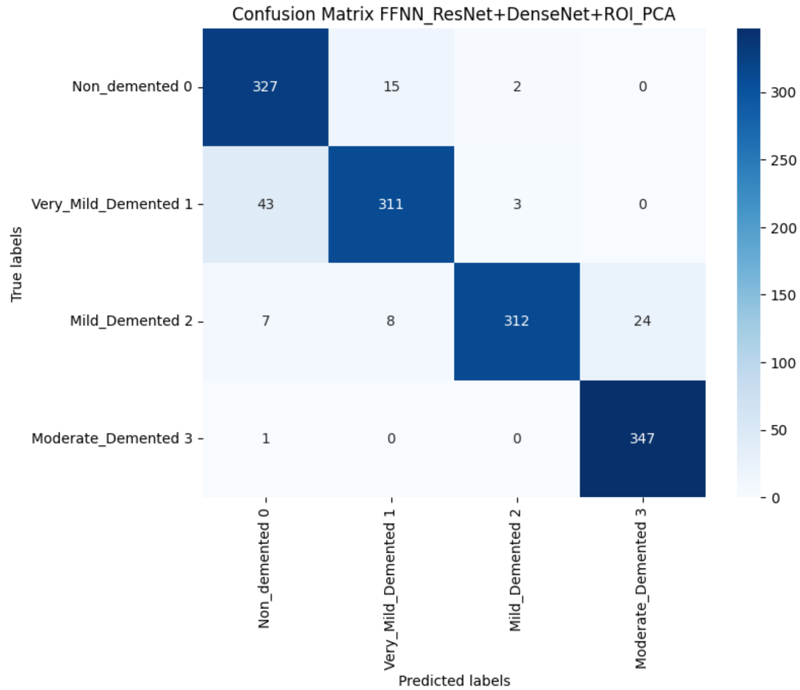

Figure 9 Confusion matrix FFNN_ResNet+DenseNet+ROI_PCA

The high rate of correct classifications underscores the validity of the adopted strategy, which integrates features extracted from ResNet and DenseNet, optimized through PCA, along with features extracted from regions of interest (ROI) through DWT, GLCM, and LBP. This synergistic combination of feature extraction techniques and modeling has allowed the system to effectively adapt to the complexity of the data and provide accurate and reliable results.

Conclusions

The present study has demonstrated the effectiveness of combining features extracted from deep learning models with specific information from Regions of Interest (ROI) in the context of Alzheimer's diagnosis, thus making a significant contribution to research in the field of personalized medicine and digital healthcare. Comparative analyses conducted with related works have not only confirmed the efficacy of the proposed method but also underscored the importance of pursuing further developments and enhancements to optimize diagnostic performance.

The FFNN_ResNet+DenseNet+ROI_PCA model emerges as the most effective among those examined for classifying stages of Alzheimer's disease based on MRI images. This model provides a solid foundation for improving diagnostic accuracy, enabling timely and targeted interventions for Alzheimer's patients. The results obtained provide a robust basis for future development. One promising avenue could be to enhance the approach to selecting Regions of Interest (ROI).

Introducing manual selection of ROI with expert support would represent a significant step towards more advanced and personalized diagnostics of neurological diseases such as Alzheimer's. This approach would allow for greater precision in identifying salient features in MRI images, providing the model with more detailed and specific information. Our future research directions involve the possibility of use evolved multiple instance models [5], that have showed to perform very well in medical contest [6]. The final aim is to realize structured framework able to support diagnosis [7], overcoming the "black box" vision that affect deep learning solutions [8].

References

1. Song, M., Jung, H., Lee, S., Kim, D., & Ahn, M. (2021). Diagnostic Classification and Biomarker Identification of Alzheimer's Disease with Random Forest Algorithm. *Brain sciences*, *11*(4), 453. https://doi.org/10.3390/brainsci11040453
2. Battineni, G., Hossain, M. A., Chintalapudi, N., Traini, E., Dhulipalla, V. R., Ramasamy, M., & Amenta, F. (2021). Improved Alzheimer's Disease Detection by MRI Using Multimodal Machine Learning Algorithms. *Diagnostics (Basel, Switzerland)*, *11*(11), 2103. https://doi.org/10.3390/diagnostics11112103
3. Sun, H., Wang, A., Wang, W., & Liu, C. (2021). An Improved Deep Residual Network Prediction Model for the Early Diagnosis of Alzheimer's Disease. *Sensors (Basel, Switzerland)*, *21*(12), 4182. https://doi.org/10.3390/s21124182
4. El-Latif, A. A. A., Chelloug, S. A., Alabdulhafith, M., & Hammad, M. (2023). Accurate Detection of Alzheimer's Disease Using Lightweight Deep Learning Model on MRI Data. *Diagnostics (Basel, Switzerland)*, *13*(7), 1216. https://doi.org/10.3390/diagnostics13071216
5. Astorino, A., Fuduli, A., Gaudioso, M., Vocaturo, E., Multiple instance learning algorithm for medical image classification, *CEUR Workshop Proceedings* 27th Italian Symposium on Advanced Database Systems, SEBD 2019 Grosseto 16 June 2019, vol. 2400, CEUR-WS.org, https://ceur-ws.org/Vol-2400/paper-46.pdf
6. Ester Zumpano, Antonio Fuduli, Eugenio Vocaturo, and Matteo Avolio. 2021. Viral pneumonia images classification by Multiple Instance Learning: preliminary results. In Proceedings of the 25th International Database Engineering & Applications Symposium (IDEAS '21). Association for Computing Machinery, New York, NY, USA, 292–296. https://doi.org/10.1145/3472163.3472170
7. Ester Zumpano, Pasquale Iaquinta, Luciano Caroprese, Francesco Dattola, Giuseppe Tradigo, Pierangelo Veltri, and Eugenio Vocaturo. 2020. SIMPATICO 3D Mobile for Diagnostic Procedures. In Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services (iiWAS2019). Association for Computing Machinery, New York, NY, USA, 468–472. https://doi.org/10.1145/3366030.3366087
8. Caroprese, L. , Vocaturo, E. , Zumpano, E., Argumentation approaches for explanaible AI in medical informatics, Intelligent Systems with Applications, V. 16, 200109, 2022, 2022/11/01 SN - 2667-3053, https://doi.org/10.1016/j.iswa.2022.200109