Transactions on Materials, Biotechnology and Life Sciences
ISSN: 2960-1932, eISSN: 2960-2319 | Volume 3
BLSME 2024

Alzheimer's MRI Recognition Based on Convolutional Neural Networks

Mingjun Wang *

School of Economics, Xiamen University, Xiamen, China, 361005

* Corresponding Author Email: 15220222201931@stu.xmu.edu.cn

Abstract. Alzheimer's disease (AD) is a progressively worsening neurodegenerative disorder, where early diagnosis is crucial for implementing effective disease management strategies. In multimodal diagnostic pathways, Magnetic Resonance Imaging (MRI) plays a pivotal role as a non-invasive imaging technique in disease identification and progression monitoring. Deep learning, particularly Convolutional Neural Networks (CNN), can extract key biomarkers from complex imaging data. By training CNN models to automatically interpret MRI scans, radiology experts can utilize these advanced analytical tools for more efficient and consistent pathological assessments, thereby enhancing clinical decision support systems. This study develops a deep learning-based method for the automatic classification of Alzheimer's brain MRI images, using convolutional neural networks to categorize MRI images into four stages: no dementia, very mild dementia, mild dementia, and moderate dementia. A CNN model is constructed, learning distinctive features of the images through multi-level feature extraction and performing feature map visualization. Early stopping is employed to prevent overfitting. The model is trained on a training set and evaluated on a test set, with performance metrics including confusion matrix, accuracy, precision, recall, F1 score, Kappa coefficient, Matthew's coefficient, ROC curve with AUC value, and PQ curve with AP value. The results show that the proposed model effectively differentiates between various categories of MRI images, providing valuable tools for early diagnosis and condition monitoring.

Keywords: Computer Vision; Deep Learning; CNN; Alzheimer's MRI Recognition.

1. Introduction

Alzheimer's disease is a major health challenge faced by modern society, and its incidence is expected to increase with the aging global population. As a chronic neurodegenerative disease, early diagnosis of AD is crucial for slowing disease progression, improving patient quality of life, and reducing treatment costs. Traditional diagnostic methods rely on cognitive testing and neuroimaging assessments, but these methods are often time-consuming and dependent on the professional judgment of the diagnostician. Not only are they laborious and time-intensive, but they may also be influenced by subjective judgment, limiting the efficiency and accuracy of the diagnosis.

Automated image analysis, especially methods based on deep learning, has garnered widespread attention in the research community. Deep learning, particularly Convolutional Neural Networks (CNN), is capable of learning hierarchical features in images and identifying complex patterns associated with AD. It demonstrates exceptionally outstanding performance in image recognition and classification tasks, which is crucial for interpreting MRI images and early detection of pathological changes in AD. Bae et al. (2020) developed a CNN model based on T1-weighted magnetic resonance imaging, focusing on cross-ethnic validation and emphasizing its effectiveness across different ethnic groups [1]. Islam et al. (2018) discussed a deep CNN ensemble system for brain MRI analysis in diagnosing Alzheimer's disease, highlighting the superior performance of deep learning models in medical imaging [2]. Rieke et al. (2018) focused on visualizing the role of CNNs in MRI-based Alzheimer's diagnosis, identifying the most important brain regions for AD diagnosis through relevant heat maps [3]. Amir Ebrahimi et al. (2021) covered methods for detecting Alzheimer's from MRI images using CNNs, providing detailed information on various CNN model architectures [4]. Feng W et al. (2020) introduced an automated deep learning model based on MRI data for detecting

Alzheimer's, demonstrating the potential of CNN-based methods and advancements in AI-driven diagnostic approaches in neurology [5]. These studies collectively showcase the significant role and promising future of deep learning technologies, especially CNNs, in medical diagnostics and neuroimaging analysis.

Building on previous research, this study further deepens the exploration and proposes an innovative deep learning framework that significantly enhances recognition accuracy. By visualizing feature mappings and employing early stopping to prevent model overfitting, the model's interpretability and generalizability are improved. Additionally, a variety of evaluation metrics are used to comprehensively assess the model's performance, such as confusion matrix, accuracy, precision, recall, F1 score, Kappa coefficient, Matthew's coefficient, ROC curve with AUC value, and PQ curve with AP value, providing new insights and methods for a more comprehensive evaluation of the model [6].

2. Basic Principles of Convolutional Neural Networks

2.1. Convolutional Layer

The convolutional layer is the core of CNN, responsible for extracting features from input data. Its basic operation involves applying a convolution kernel to the input data, defined as follows: Given an input feature map I with dimensions $H \times W$ and a convolution kernel K with dimensions $h \times w$, the convolution operation can be defined as:

$$(I * K)(i,j) = \sum_{m=0}^{h-1} \sum_{n=0}^{w-1} I(i+m, j+n) \cdot K(m,n) \tag{1}$$

Where (i,j) denotes the position on the output feature map, and m and n correspond to the positions within the convolution kernel. The convolution process in a CNN model is illustrated in Fig 1.

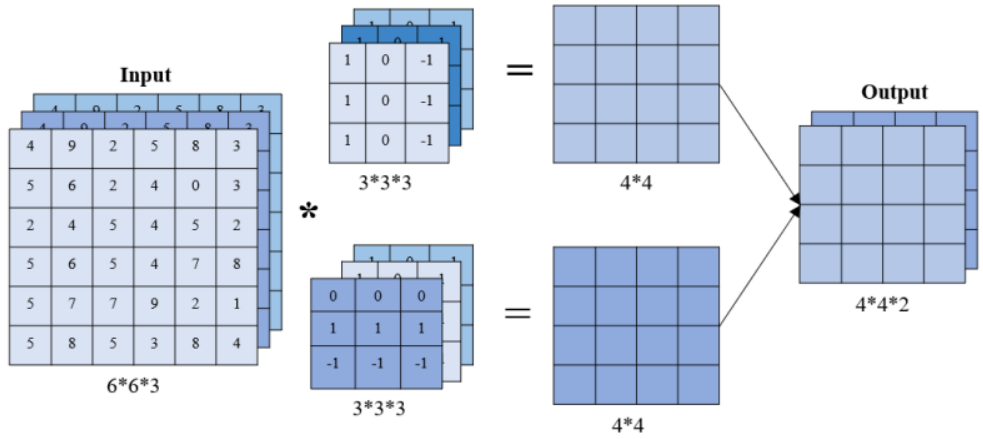

Figure 1. CNN Model Convolution Process

In practice, multiple convolution kernels are usually applied to the input data to extract different features, with padding and various strides used during convolution. Padding involves adding extra zero-value boundaries around the input data to control the output size, and the stride is the distance the convolution kernel moves on the input feature map. The output size is determined by the following formula:

$$O = \frac{W-K+2P}{S} + 1 \tag{2}$$

Where O is the output size, W is the input size, K is the size of the convolution kernel, P is the amount of padding, and S is the stride. The padding process in a CNN model is illustrated in Fig 2.

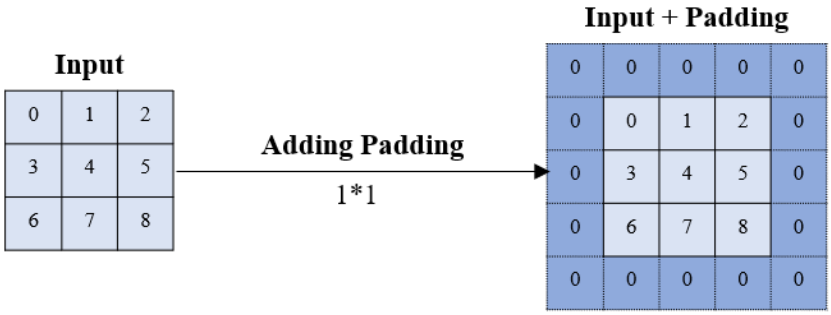

Figure 2. CNN Model Padding Process

2.2. Activation Function

The activation function is used to increase the non-linearity of the network. ReLU is one of the most commonly used activation functions, defined as:

$$f(x) = \max(x, 0) \tag{3}$$

This function sets all negative values to 0, while positive values remain unchanged.

2.3. Pooling Layer

The pooling layer reduces the dimensionality and number of parameters of the feature map, thereby reducing computational load and the risk of overfitting. Max pooling is a pooling operation that selects the maximum value within a region. For the output at position (i, j), max pooling is defined as:

$$P(i,j) = max_{\substack{0 \le m < w \\ 0 \le n < h}} I(s \cdot i + m, s \cdot i + n) \tag{4}$$

Where w and h are the width and height of the pooling window, and s is the stride. The pooling process in a CNN model is illustrated in Fig 3.

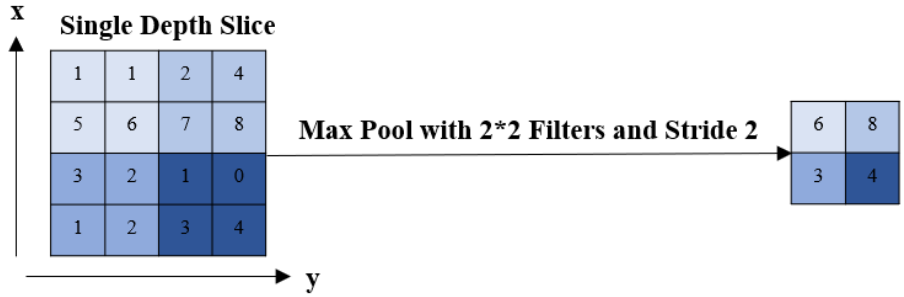

Figure 3. CNN Model Pooling Process

2.4. Fully Connected Layer

The fully connected layer takes the output from the previous layer and computes a weighted sum for each output unit, adding a bias, and then passes it through a non-linear activation function. For neuron i's output o_i, it can be expressed as:

$$o_i = f\left(\sum_j w_{ij} x_j + b_i\right) \tag{5}$$

Where w_{ij} is the weight connecting input unit j to output unit i, x_j is the activation value of input unit j, b_i is the bias, and f is the activation function.

2.5. Softmax Function

In classification problems, the output layer often uses the Softmax function to convert the neural network's output into a probability distribution. For the probability of the i-th category P_i, the Softmax function is defined as:

$$P_i = \frac{e^{z_i}}{\sum_{k=1}^{K} e^{z_k}} \tag{6}$$

Where z_i is the input to the i-th output unit, and K is the total number of categories.

2.6. Loss Function

CNN training is accomplished through gradient descent and backpropagation algorithms, where the difference between the network's output and the true label is quantified through a loss function in each iteration. The loss function used here is cross-entropy loss:

$$L = -\sum_i y_i log\,(\hat{y}_i) \tag{7}$$

Where y_i is the true label and $\hat{y}_i$ is the predicted probability. The model's objective is to minimize the loss function.

2.7. Gradient Descent and Backpropagation

The loss function's gradient with respect to the network parameters is calculated through the backpropagation algorithm, and these gradients are used to update the network's weights and biases:

$$w_{ij}^{(new)} = w_{ij}^{(old)} - \eta \frac{\partial L}{\partial w_{ij}} \tag{8}$$

$$b_i^{(new)} = b_i^{(old)} - \eta \frac{\partial L}{\partial b_i} \tag{9}$$

Where w_{ij} represents the weight connecting neuron j from the previous layer to neuron i of the current layer, b_i represents the bias of neuron i of the current layer, and η is the learning rate controlling the step size of the weight update. Through this iterative process, the network gradually learns the mapping from input data to output labels, optimizing its internal parameters to minimize the loss on the training set.

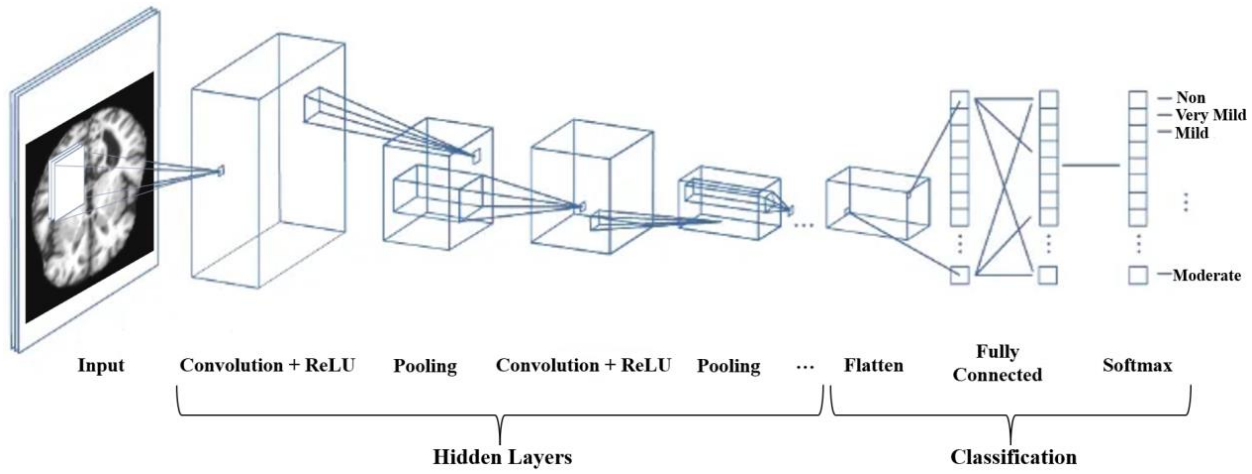

Figure 4. Complete Process of Feature Construction and Classification in CNN Model

The complete process of feature construction and classification prediction in a CNN model is illustrated in Fig 4.

3. Model Construction

To automatically classify the MRI images of Alzheimer's disease patients into stages, a deep convolutional neural network was constructed using the TensorFlow and Keras framework. The network first employs a convolutional layer with 16 3x3 filters and uses the ReLU activation function to extract basic features from the images. This is followed by a max pooling layer that downsamples the feature map using a 2x2 window, aiming to reduce computational complexity and enhance the model's spatial hierarchy.

The network's second convolutional layer contains 32 filters, again applying the ReLU activation function to enhance non-linear feature extraction. This is followed by another max pooling layer to further reduce feature dimensions. Subsequently, a flattening layer transforms the two-dimensional feature map into a one-dimensional feature vector, preparing it for processing by a fully connected layer. Next, a fully connected layer with 128 neurons processes the features further and uses the ReLU activation function to increase the network's learning capability. Finally, a fully connected layer with four output neurons, representing the four different stages of dementia, uses the Softmax activation function to output prediction probabilities [7].

When compiling the model, the Adam optimizer was chosen for its significant efficiency with large datasets, and the loss function was set to sparse categorical cross-entropy to handle multi-class labels. The model's performance was evaluated using the accuracy metric. To enhance the model's generalizability and prevent overfitting, early stopping was introduced as part of the training process. This early stopping callback monitors the loss on the validation set, and if the improvement is less than 0.001 over five consecutive training epochs, the training is halted.

During the model training phase, training was conducted over a predefined 100 epochs, processing 128 images per batch, with 20% of the training set data randomly extracted as a validation set. This method not only provides real-time feedback on model performance but also ensures self-correction during the training process.

3.1. Training Process

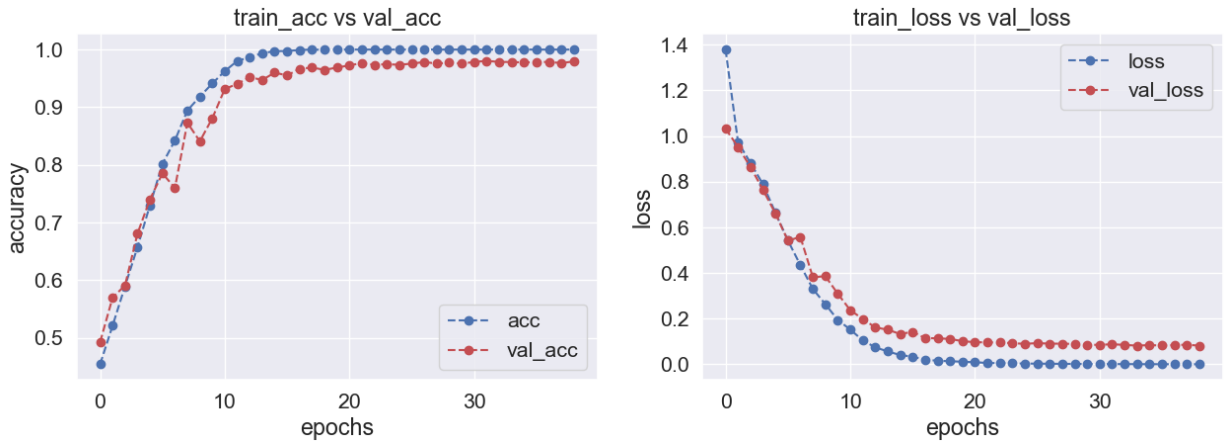

Figure 5. Variation of Accuracy and Loss in the CNN Model Over Training Epochs

From Fig 5, we can observe the changes in accuracy and loss of the convolutional neural network model during the training process on the Alzheimer's disease MRI image dataset. The left half of the graph depicts the training accuracy (blue dotted line) versus the validation accuracy (red dotted line) over the epochs. It is observed that after about 10 training epochs, both reach a high plateau, with the model's performance on the unseen validation data typically being slightly lower than on the training data. The right half of the graph shows the changes in training loss (blue dotted line) versus validation

loss (red dotted line) over the epochs. In the initial few epochs, the loss decreases rapidly, indicating the model's effectiveness in learning data features. As the number of training epochs increases, the decrease in loss stabilizes, indicating that the model is beginning to converge. Similarly, the small gap between training and validation loss suggests that there is no significant overfitting in the model.

These two graphs provide an intuitive view of the changes in model performance during the training process. The trends in training and validation accuracy indicate that the model has learned well from the training data and maintains good generalization capability on the validation data. The shape of the loss curves further confirms the stability and effectiveness of the model's training. These observations suggest that the chosen model and training strategy are suitable for handling the task of classifying MRI images for Alzheimer's disease.

3.2. Visualization of Convolutional Neural Network Feature Maps

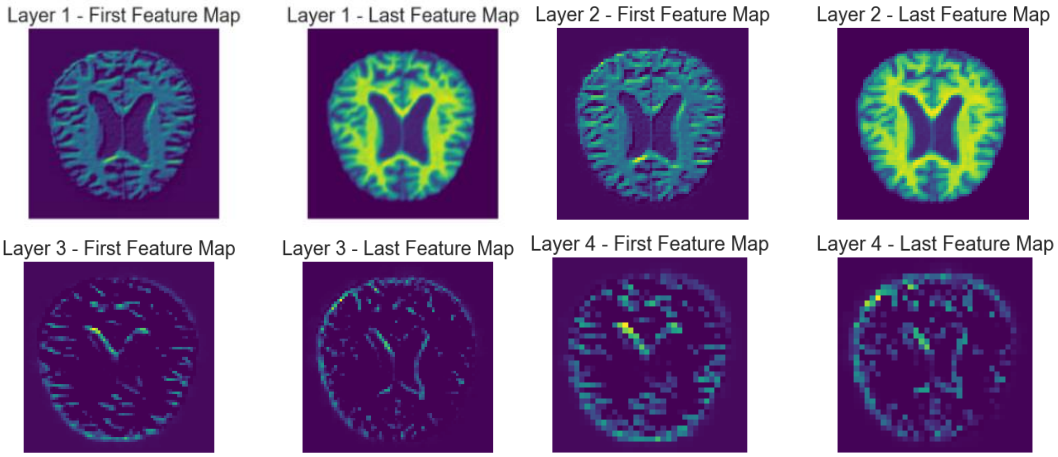

Figure 6. Visualization of Different Layers in the Convolutional Neural Network

To gain a deeper understanding of the performance of the convolutional neural network in identifying key features in Alzheimer's disease MRI images, this paper visualizes the first and last feature maps of the model's four layers. The first and last feature maps of each layer were selected to demonstrate the network's response to input data.

As seen in Fig 6, in the first layer, the feature maps reveal how the model detects simple edges and texture information. As the network's layers increase, the feature maps begin to capture more complex and abstract image characteristics. For instance, in the fourth layer, the network detects not only edges and textures but also starts responding to higher-level image features. The contrast in the feature maps shows how the network gradually transitions from raw pixels to feature representations that can signify more advanced concepts. The feature maps of the first layer present more intuitive physical features, while subsequent layers reveal the model's understanding of complex patterns [8]. These visualizations validate that the model is indeed learning useful features contributing to classification decisions.

4. Model Evaluation

4.1. Confusion Matrix

The confusion matrix is a key tool for evaluating the performance of classification algorithms. It details the correspondence between model predictions and actual labels, reflecting the model's accuracy and types of misclassifications in each category. The true positives (TP) and true negatives (TN) in the matrix reflect the number of correct predictions, while false positives (FP) and false negatives (FN) represent the number of misclassifications [9]. Since the number of instances in each category varies, a normalized confusion matrix was included in the evaluation.

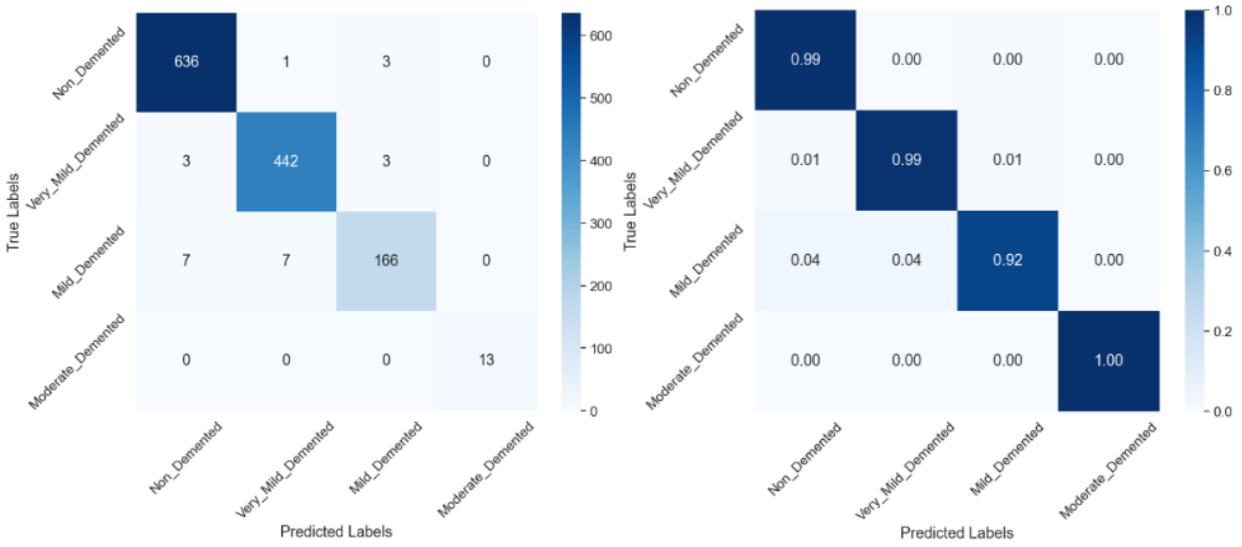

Figure 7. Heatmap of the Confusion Matrix for the CNN Model

The left half of Fig 7 shows the absolute numbers of the model's prediction results, highlighting the consistency of most predictions with actual labels, especially in the categories of no dementia and very mild dementia. However, the prediction accuracy for mild dementia and moderate dementia categories is slightly lower, which might reflect the overlap of features between these categories. In the right half of Fig 7, the normalized view, each cell's value represents the percentage of correct predictions made by the model in a given actual category. For example, the accuracy for the very mild dementia category is 99%, meaning the model's prediction accuracy is very high in this category. The normalized view provides the proportion of correct predictions in each category, helping to overlook the impact of sample number imbalance and more fairly evaluate the model's performance.

4.2. Accuracy

Accuracy is the number of correctly classified samples divided by the total number of samples, and its formula is:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

4.3. Precision

Precision measures the proportion of samples correctly identified as positive by the model among all samples identified as positive. Its calculation formula is:

$$Precision = \frac{TP}{TP+FP} \tag{11}$$

4.4. Recall

Recall measures the proportion of samples correctly identified as positive by the model among all actual positive samples. Its calculation formula is:

$$Recall = \frac{TP}{TP+FN} \tag{12}$$

4.5. F1 Score

The F1 Score is the harmonic mean of precision and recall. Its calculation formula is:

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{13}$$

4.6. Kappa Coefficient

The Kappa coefficient compares the observed accuracy with the random accuracy. Its calculation formula is:

$$Kappa = \frac{P_0 - P_e}{1 - P_e} \tag{14}$$

Where $P_0 = \frac{TP+TN}{TP+TN+FP+FN}$ represents the observed agreement rate, and $P_e = \frac{(TP+TN)\times(TP+FN)}{(TP+TN+FP+FN)^2}$ represents the chance agreement rate.

4.7. Matthews Coefficient

The Matthews coefficient is a more balanced measure, especially in datasets with imbalanced classes. Its calculation formula is:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{15}$$

Table 1. Multidimensional Evaluation Results of the CNN Model

Evaluation Metric	Value
Accuracy	0.98126
Precision	0.98296
Recall	0.97564
F1 Score	0.97918
Kappa Coefficient	0.96910
Matthews Coefficient	0.96915

As shown in Table 1, the CNN model performs very well in classification tasks, with high accuracy, precision, recall, F1 score, Kappa coefficient, and Matthew's coefficient. It demonstrates good consistency between its predictions and the actual labels, achieving high scores in multiple dimensions of evaluation metrics, both partially and as a whole.

4.8. ROC Curve and AUC Value

Value The ROC curve is a tool for describing classifier performance, depicting the relationship between true positive rate and false positive rate at different decision thresholds. $TPR = \frac{TP}{TP+FN}$ represents the proportion of positive cases correctly predicted as positive, and $FPR = \frac{FP}{FP+PN}$ represents the proportion of negative cases incorrectly predicted as positive. The AUC value represents the area under the ROC curve.

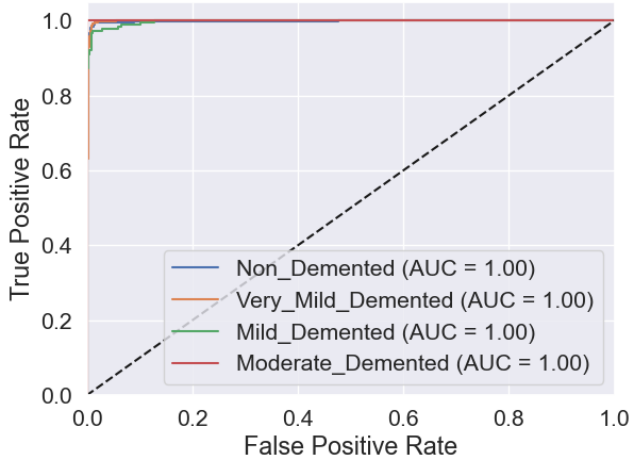

Figure 8. ROC Curve and AUC Value of the CNN Model

As shown in Fig 8, the ROC curve of the CNN model is above the curve of a random classifier (the line at AUC = 0.5) and the given AUC values are all 1, further confirming the model's good performance.

4.9. PQ Curve and AP Value

The PQ curve focuses on the relationship between precision and recall. The AP value under the PQ curve provides a measure of the overall performance of a classifier, regardless of specific decision thresholds. The higher the AP value, the better the model's performance [10].

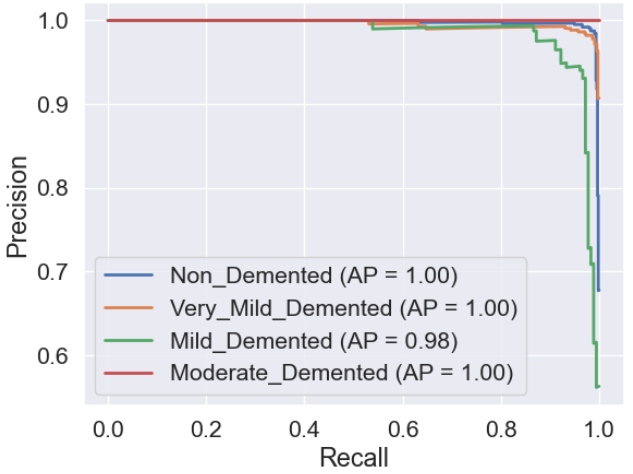

Figure 9. PQ Curve and AP Value of the CNN Model

As shown in Fig 9, the PQ curve and AP values of the CNN model are 1 for all categories except mild dementia, further confirming the model's good performance.

5. Conclusion

This study, based on the deep learning approach of Convolutional Neural Networks (CNN) technology, has successfully developed an automated Alzheimer's disease brain MRI image classification system. Through in-depth exploration and application of CNN, the proposed model excelled in classifying MRI images across four stages: no dementia, very mild dementia, mild dementia, and moderate dementia. The model employed multi-level feature extraction and validated its effectiveness in learning useful features contributing to classification decisions through feature mapping visualization. The proposed CNN model demonstrated high accuracy and reliability in the task of Alzheimer's disease MRI image recognition. It performed well on the training set and was favorably evaluated on the test set. The performance evaluation included various metrics such as confusion matrix, accuracy, precision, recall, F1 score, Kappa coefficient, Matthew’s coefficient,

ROC curve with AUC value, and PQ curve with AP value. Through these comprehensive evaluations, the model showed high efficiency across multiple dimensions.

This research provides valuable tools for early diagnosis and condition monitoring, while also demonstrating the potential and importance of deep learning technology in medical diagnostics and neuroimaging analysis. It offers medical professionals an efficient tool to assist in the diagnosis of Alzheimer's disease and provides a new perspective for research in deep learning technology in the field of medical image processing. Future research can build on this foundation to further explore early diagnosis methods for Alzheimer's disease in different ethnic and age groups, as well as further optimize the CNN model for broader medical applications.

References

[1] Bae J B, Lee S, Jung W, et al. Identification of Alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging [J]. Scientific reports, 2020, 10 (1): 22252.

[2] Islam J, Zhang Y. Brain MRI analysis for Alzheimer’s disease diagnosis using an ensemble system of deep convolutional neural networks[J]. Brain informatics, 2018, 5: 1 - 14.

[3] Rieke J, Eitel F, Weygandt M, et al. Visualizing convolutional networks for MRI-based diagnosis of Alzheimer’s disease[C]//Understanding and Interpreting Machine Learning in Medical Image Computing Applications: First International Workshops, MLCN 2018, DLF 2018, and iMIMIC 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16 - 20, 2018, Proceedings 1. Springer International Publishing, 2018: 24 - 31.

[4] Ebrahimi A, Luo S, Disease Neuroimaging Initiative A. Convolutional neural networks for Alzheimer’s disease detection on MRI images [J]. Journal of Medical Imaging, 2021, 8 (2): 024503 - 024503.

[5] Feng W, Halm-Lutterodt N V, Tang H, et al. Automated MRI-based deep learning model for detection of Alzheimer’s disease process [J]. International Journal of Neural Systems, 2020, 30 (06): 2050032.

[6] Ban Y, Zhang X, Lao H. Diagnosis of Alzheimer's disease using structure highlighting key slice stacking and transfer learning [J]. Medical Physics, 2022, 49 (9): 5855 - 5869.

[7] Menagadevi M, Mangai S, Madian N, et al. Automated prediction system for Alzheimer detection based on deep residual autoencoder and support vector machine [J]. Optik, 2023, 272: 170212.

[8] Venugopalan J, Tong L, Hassanzadeh H R, et al. Multimodal deep learning models for early detection of Alzheimer’s disease stage [J]. Scientific reports, 2021, 11 (1): 3254.

[9] Dyrba M, Hanzig M, Altenstein S, et al. Improving 3D convolutional neural network comprehensibility via interactive visualization of relevance maps: evaluation in Alzheimer’s disease [J]. Alzheimer's research & therapy, 2021, 13: 1 - 18.

[10] Patil V, Madgi M, Kiran A. Early prediction of Alzheimer's disease using convolutional neural network: a review [J]. The Egyptian Journal of Neurology, Psychiatry and Neurosurgery, 2022, 58 (1): 1 - 10.