CHAPTER 6

ENSEMBLE BOOTSTRAP AGGREGATING WITH SUPPORT VECTOR MACHINE BASED HYBRID TECHNIQUES FOR CLASSIFYING STAGES OF AD

6.1 Introduction

This chapter discusses the various classification techniques for AD classification. The ensemble bootstrap aggregating with SVM is proposed in this study. A CNN is used for improving the performance in classifying the stages of AD. A pre-trained AlexNet has been used as a base network. The traditional feature extraction methods, such as the GLCM, GLDM, and GLRLM are used for comparison. A detailed description related to the AlexNet, which is used for feature extraction, is discussed in Chapter 5. The various classifiers and the proposed BAGGING_SVM classifier are defined in the subsequent sections of this chapter. Performance measures are used to analyze the experimental outcomes, and a final evaluation of the suggested method is provided.

6.2 Classifiers for classifying the stages of AD

The different classifiers, including KNN, J48, BAGGING, SVM, and the suggested BAGGING_SVM, are covered in this section.

6.2.1 KNN classifier

The input's class is determined by the common class of its neighbors. The KNN classifier gathers objects according to dominating classes. Here, K a positive integer specifies the neighbors count after creating the classification choice (Srinivasan et.al 2020). Equation (6.1) is used to estimate the distance to identify the nearest neighbors.

$$D_j = \sqrt{\sum_{\mathrm{i=1}}^{\mathrm{ndim}} (x_i - y_i)^2} \tag{6.1}$$

Euclidean distance estimates parameter space points distances. Here, 'd' denotes entire distance from j^{th} detected data point expected. x_i and y_i denotes the forecaster values in the new locations. L1-norm estimates the distance between vector pairs, while Euclidean distances deal with analytical solution that are liable. It is applied to the dataset to clear the irregularities. The predictor values of mean and variance are set zero. The estimation of distance at each data point in k is determined by using the smallest distance. The projected value is the average of the closest neighbor, which is subjective to the opposite of its distance.

6.2.2 J48 classifier

The classifier uses divide and conquer technique to create the initial tree by beginning with attribute that has maximum ratio as its root node. To enhance the accuracy, pruning technique is applied to eliminate the unwanted branches from the tree. The data is separated into two types for managing endless properties. To maintain DT model flexible across varied datasets, pruning is vital to avoid overfitting (Singaravelan et al., 2016).

$$E(S) = \sum_{i=1} -p \log_2 P_i \qquad (6.2)$$

Where 'i' denotes the class count, P_i denotes the proportion of S fitting to class 'i'.

$$Gain(S, A) = E(S) - value \sum (A) \frac{S_v}{S} E(S_v) \qquad (6.3)$$

Sv is a subset of "S" in the function "A" with value "v," where A is the set of all possible values. By predicting the entropy value, "S" is represented by the original collection entropy.

6.2.3 BAGGING classifier

The bagging classifier is an effective technique for creating classifiers which involves repeated sampling of the training datasets to crate many models. With a variety of datasets, it enhances the ensemble model's robustness while lowering overfitting. The steps involved in bagging classifier are given below:

i)	Input: |D| = N training dataset
ii)	For i between 1 and m

iii)	Use bootstrapping to sample from D. Using |Di| = N~ to obtain Di
iv)	Fit Di to create the model fi(x).
v)	Ensemble of the models {fi(x) | i = 1,2, …, m} and obtain the final model F(x). Perform binary classification F(x) = sign(∑i=1,2,…,mfi(x)).

It is motivated from the random forest model, which is robust. Every decision tree is trained using bootstrapping, contains sampling subsets. A random forest model consist of random features is added in the training procedure. The model randomly chooses a feature subset for each split, instead considering all features. A generalized forecast is performance with varied datasets which surges the variety among the trees (Liaw and Wiener 2002).

6.2.4 SVM classifier

SVM is a dominant ML algorithm for linear or nonlinear classifications like, text, image, and anomaly detection etc. SVMs are flexible and effective in various applications due to ability to manage high-dimensional data. Consider a training set $T = \{(x_i, y_i) \mid y_i = 1\, or\, -1\}$. A hyperplace is defined in the input feature which splits into two classes $\{i = 1,..., N\} T = \{x_i, y_i \mid y_i = 1\, or\, -1\}$. The margin between classes are maximized to enhance the classifier's capacity (Vapnik 1998). SVM is a robust learning model with an capacity to exploit margin among classes for better overview. SVM has the ability to balance between the margin and classification error, to improve the overall performance and flexibility (Rosenblatt et al., 1958). The SVM is represented as:

$$\min_{W>R} \frac{1}{2} W^T \cdot I_n \cdot W + C(1_N^T \cdot \zeta)$$
$$s.t \qquad A\begin{pmatrix} W \\ \xi \\ b \end{pmatrix} \geq 1_{N^p} \tag{6.4}$$

where A denotes the corresponding matrix

$$A = \begin{pmatrix} y_1 \phi(x_1)^T, e_1, 1 \\ y_2 \phi(x_2)^T, e_2, 1 \\ \vdots \\ y_N \phi(x_N)^T, e_N, 1 \end{pmatrix} \tag{6.5}$$

Where N-dimension vector is dented by ϵ k with k^{th} coordinate =1.

Support hyperplanes are assessed by slack vector variables ζ=(ζ_1,ζ_2,...,ζ_N)^T. Few solution optimality necessities are achieved from the Lagrangian function. The primary issue is changed into dual issue (Boyd et al., 2004).

Discriminative function: $f(x) = \text{sign}\left(\sum_{\alpha'_n>0} y_n \alpha_n^* \phi(x_n)^T * \phi(x) + b\right)$ (6.6)

Lagrange multiplier for i^{th} sample is α. Kernel algorithms are used to estimate the feature vector's inner product.

ϕ(x_n)^T+ϕ(x_n)=K(x_n,x_n) (6.7)

The classifier is represented by a sparse subset of training data.

6.2.5 The proposed BAGGING_SVM classifier

The efficacy of the bagging and SVM ensemble used in this study to identify AD is shown by refining accuracy, avoiding overfitting, and bolstering the stability of the BAGGING_SVM. It aims to train numerous SVM models on numerous training data subsets and integrate their forecasts to develop complete performance. Bagging with the SVM model has made promising results, especially in the healthcare domain, where exact classification is critical. An ensemble of bagging and SVM models is needed to minimize the overfitting issue in ML. The model's focus on a variety of features reduces the possibility of overfitting in the training data. This changeability in model forecasts supports the growth of a robust and firm classifier. The ensemble of bagging and the SVM model enhances the recall and lowers the missing rate, which is appropriate for classification (Yu 2023; Olfa Ben Ahmed et al., 2017).

The ensemble bagging algorithm is given below:

i) Select:

X: Data (training) characteristics

y: Labels for data categorization

B: SVM parameters (kernel type, regularization ‘C’) - Bootstrap sample count

SVM_models = [] Set up SVM model ensemble.

With reference to b = 1 to B

ii) A bootstrap sample

Create a bootstrap: X_b, y_b = bootstrap_sample(X, y)

iii) Use the specified parameters to train the SVM on X_b and Y_b.

SVM_model_b = Train (SVM_parameters, X_b, y_b)

iv) Store ensemble's SVM

Add SVM_model_b to SVM_models list.

v) Combining Forecasts

Compile all of the SVM_models' predictions for x_new for each new instance.

grouping: majority vote, total forecasts

majority_vote(forecasts) = projected_label

Save Expected_label

Regression: averaging predictions to produce a total.

Average(predictions) = predicted_value

Store Rewritten and predicted_value

6.3 Experimental result and analysis for BAGGING_SVM classifier

The performance measures of the BAGGING_SVM classifier, including accuracy, precision, recall, and F1-score, are covered in this section. The ADNI and OASIS datasets are used to evaluate the proposed classifier. A comparison of proposed and current classification methods is explained.

6.3.1 Performance evaluation

Using the following metrics, the BAGGING_SVM classifier's performance is evaluated:

i) Accuracy: fraction of precisely projected interpretations compared to all forecasts produced by the model.

$$Accuracy = (TP+TN)/(TP+TN+FP+FN) \quad (6.8)$$

ii) Precision: Ratio of appropriately planned, optimistically expected observations to positively expected observations overall

$$Precision = TP/(TP+FP) \quad (6.9)$$

iii) Recall: Referred to as sensitivity, is the proportion of accurately identified positive cases (true positives) among all actual positive instances.

$$Recall = TP/(TP+FN) \quad (6.10)$$

iv) F-measure: accuracy and recall weighted average. It takes into consideration all false negatives and erroneous positives. It is not as straightforward as precision and has an unequal distribution of classes.

$$FM = (2*Precision*Recall)/(Precision+Recall) \quad (6.11)$$

6.3.2 Result analysis of BAGGING_SVM classifier on the ADNI dataset

The performance metrics of different classification techniques using BAGGING_SVM to categorize AD stages on ADNI dataset are displayed in Table 6.1. On ADNI dataset, the suggested AlexNet + BAGGING_SVM classification approach outperformed the other classification techniques in terms of accuracy. The classification accuracy of the proposed method, which involves GLCM, GLDM, and GLRLM with BAGGING_SVM, obtained an accuracy of 90.8% and classification accuracy involving a pre-trained AlexNet with BAGGING_SVM, obtained an accuracy of 91.6%, respectively. The accuracy percentage obtained by the pre-trained AlexNet with the BAGGING_SVM method is slightly more compared to the accuracy percentage obtained by GLCM, GLDM and GLRLM with BAGGING_SVM method. This shows the efficiency of the pre-trained AlexNet model with BAGGING_SVM on brain neuron images on ADNI dataset.

Table 6.1 Performance comparison of the various classification methods with BAGGINNG_SVM used for AD stage classification on the ADNI dataset

Classification Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
J48	61.7	71.4	61.5	57.6
KNN	63.2	75	58.2	63.2
SVM	86.3	88.1	85.4	85.5
GLCM, GLDM and GLRLM + BAGGING_SVM	90.8	93.6	90.3	89.8
AlexNet + BAGGING_SVM	**91.6**	**94.9**	**91.2**	**90.5**

The performance of the BAGGING_SVM and the current classification techniques on ADNI dataset are compared in Table 6.2 under the identical conditions. Accuracy metric is the frequently used metric, and the other metrics that are not stated in the literature are represented as '-.' The classification accuracy of method which involves

GLCM, GLDM and GLRLM with BAGGING_SVM obtained an accuracy of 90.8% and the classification accuracy of method which involves the pre-trained AlexNet with BAGGING_SVM obtained an accuracy of 91.6% respectively. The accuracy percentage obtained by the pre-trained AlexNet with the BAGGING_SVM method is slightly more when compared to the accuracy percentage obtained by GLCM, GLDM and GLRLM with BAGGING_SVM method. This shows the efficiency of the pre-trained AlexNet model with BAGGING_SVM on brain neuron images on ADNI dataset.

Table 6.2 Performance comparison of BAGGING_SVM classification methods with the existing classification methods for AD stage classification on ADNI dataset

Authors	**Classification Method**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Lin et al.,(2018)	CNN	88.79	-	-	-
Albright et al., (2019)	CNN	86.60	-	-	-
Alickovic and Subasi (2020)	KNN	84.27	-	-	-
Shahbaz et al., (2019)	DT	74.22	-	-	-
Alickovic and Subasi (2020)	SVM	83.15	-	-	-
Chithra & Vijayabhanu (2024) (Proposed)	GLCM, GLDM and GLRLM + BAGGING_SVM	90.8	93.6	90.3	89.8
	AlexNet + BAGGING_SVM	**91.6**	**94.9**	**91.2**	**90.5**

6.3.3 Result analysis of BAGGING_SVM classifier on the OASIS dataset

The performance metrics of different classification techniques using BAGGING_SVM to categorize AD stages on the OASIS dataset are displayed in Table 6.3. On the OASIS dataset, the proposed AlexNet + BAGGING_SVM classification method outperformed the other classification techniques in terms of accuracy. The outcomes of performance indicators like as F1-score, recall, accuracy, and precision are displayed. The classification accuracy of the approach using pre-trained AlexNet with BAGGING_SVM was 89.2%, while the method using GLCM, GLDM, and GLRLM with BAGGING_SVM achieved an accuracy of 87.51%. The accuracy percentage obtained by pre-trained AlexNet with BAGGING_SVM method is maximum compared to the accuracy percentage obtained by GLCM, GLDM and GLRLM with BAGGING_SVM method. This shows the efficiency of the pre-trained AlexNet model with BAGGING_SVM on brain neuron images on OASIS dataset.

Table 6.3 Performance comparison of the various classification methods with BAGGING_SVM for AD stage classification on the OASIS dataset

Classification Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
J48	75.1	69.2	67.3	67.7
KNN	69.8	68.4	68.5	74.9
SVM	82.3	82.8	83.2	78.3
GLCM, GLDM and GLRLM + BAGGING_SVM	87.51	81.3	82.9	81.6
AlexNet + BAGGING_SVM	**89.2**	**82.7**	**85.7**	**81.8**

Table 6.4 compares the performance of BAGGING_SVM with the existing classification methods under the same setting on the OASIS dataset. Accuracy metric is the frequently used metric, and the other metrics that are not stated in the literature are represented as '-.' The classification accuracy of method which involves GLCM, GLDM and GLRLM and BAGGING_SVM obtained accuracy of 87.51% and the classification accuracy of method which involves the pre-trained AlexNet with BAGGING_SVM obtained accuracy of 89.2% respectively. The accuracy percentage obtained by the pre-trained AlexNet with BAGGING_SVM method is maximum compared to the accuracy percentage obtained by GLCM, GLDM and GLRLM with BAGGING_SVM method. This shows the efficiency of the pre-trained AlexNet model with BAGGING_SVM on brain neuron images on OASIS dataset.

Table 6.4 Performance comparison of BAGGING_SVM classification methods with the existing classification methods for AD stage classification on the OASIS dataset.

Authors	Classification Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Alroobaea et al., (2021)	Random Forest	83.92	-	-	-
Alroobaea et al., (2021)	Logistic Regression	84.33	-	-	-
Islam and Zhang (2018)	DNN	93.18	-	-	-
Islam and Zhang (2017).	SVM	71.25	-	-	-
Chithra & Vijayabhanu (2024) (Proposed)	GLCM, GLDM and GLRLM + BAGGING_SVM	87.51	81.3	82.9	81.6
	AlexNet + BAGGING_SVM	**89.2**	**82.7**	**85.7**	**81.8**

6.3.4 Performance comparison of the proposed AlexNet with BAGGING_SVM method on ADNI and OASIS dataset

Table 6.5 compares the performance of AlexNet with BAGGING_SVM method on ADNI and OASIS datasets. On ADNI dataset, AlexNet with BAGGING_SVM approach produced 90.5% F1-Score, 91.6% accuracy, 94.9% precision, and 91.2% recall. On the OASIS dataset, the AlexNet with BAGGING_SVM approach achieved 89.2% accuracy, 82.7% precision, 85.7% recall, and 81.8% F1-Score. In comparison to the OASIS, the ADNI performs better.

Table 6.5 Performance comparison of AlexNet with BAGGING_SVM on ADNI and OASIS dataset

Datasets	Classification Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ADNI	AlexNet + BAGGING_SVM	**91.6**	**94.9**	**91.2**	**90.5**
OASIS	AlexNet + BAGGING_SVM	**89.2**	**82.7**	**85.7**	**81.8**

6.3.5 Performance comparison of stage-wise analysis of AD detection using the proposed AlexNet with BAGGING_SVM method on ADNI and OASIS dataset

Figure 6.1 illustrates the stage wise accuracy analysis of the proposed AlexNet with BAGGING_SVM method on ADNI and OASIS datasets. Normal, stage 1, stage 2, and stage 3 are the labels assigned to the dataset classification. The accuracy attained for the classifications normal, stage 1, stage 2, and stage 3 in the ADNI dataset is 90.0%, 91.40%, 89.20%, and 94.40%, respectively. The accuracy attained for the classifications normal, stage 1, stage 2, and stage 3 in the OASIS dataset is 91.70%, 89.20%, 86.90%, and 88.50%. The graphic indicates that, for the ADNI dataset, stage 3 performs better than the other stages. In the normal stage, the OASIS dataset's performance accuracy is superior than that of other stages.

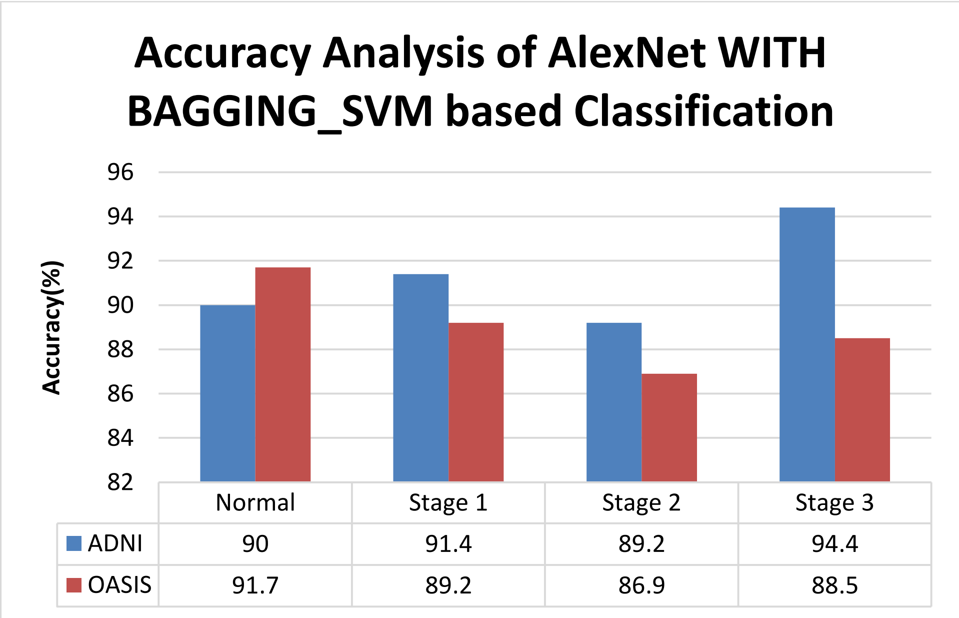

Figure 6.1 Stage wise accuracy analysis of classification using AlexNet with BAGGING_SVM on ADNI and OASIS datasets

Figure 6.2 shows the confusion matrix for AlexNet with BAGGING_SVM using ADNI dataset. It represents the performance of a classification model in predicting AD stages based on MRI data. The instances of actual classes are represented by each row in the matrix, whereas the instances of predicted classes are represented by each column. True positives (TP) include 1,508 for Normal, 819 for Stage 1, 58 for Stage 2, and 1,095 for Stage 3, suggesting high identification of Normal and Stage 3 cases, according to the confusion matrix, which demonstrates that the classification model successfully detects AD phases. Indicating areas for improvement, the false negatives (FN) were 170 for Normal, 80 for Stage 1, 12 for Stage 2, and 35 for Stage 3.

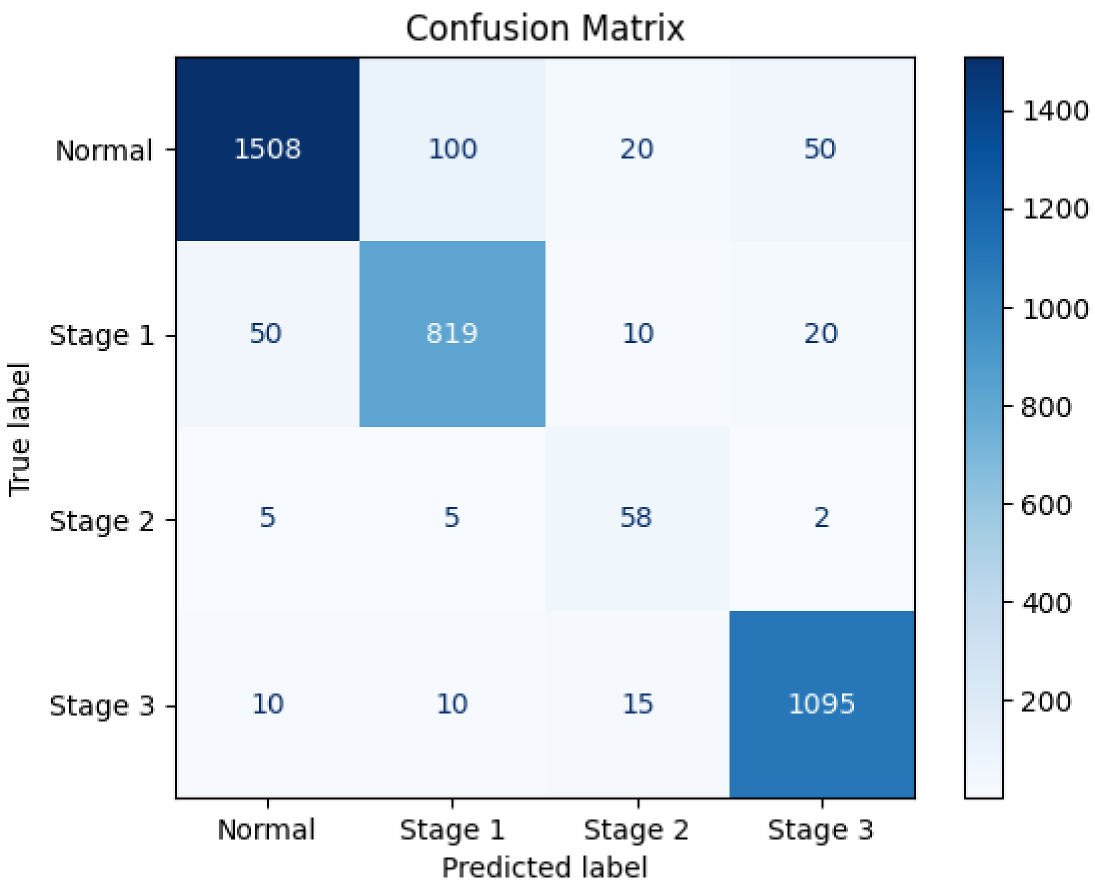

Figure 6.2 Confusion matrix for AlexNet with BAGGING_SVM using ADNI dataset

Figure 6.3 shows the confusion matrix for AlexNet with BAGGING_SVM using OASIS dataset. The confusion matrix indicates strong classification performance, with true positives (TP) being 1100 for Normal, 1070 for Stage 1, 1042 for Stage 2, and 173 for Stage 3. Nevertheless, there were false positives (FP): 60 for Stage 3, 70 for Stage 1, and 75 for Normal. Normal has 100 false negatives (FN), Stage 1 has 130, Stage 2 has 30, and Stage 3 has 25. It offers a thorough analysis of the model's performance, making it possible to evaluate both its advantages and disadvantages.

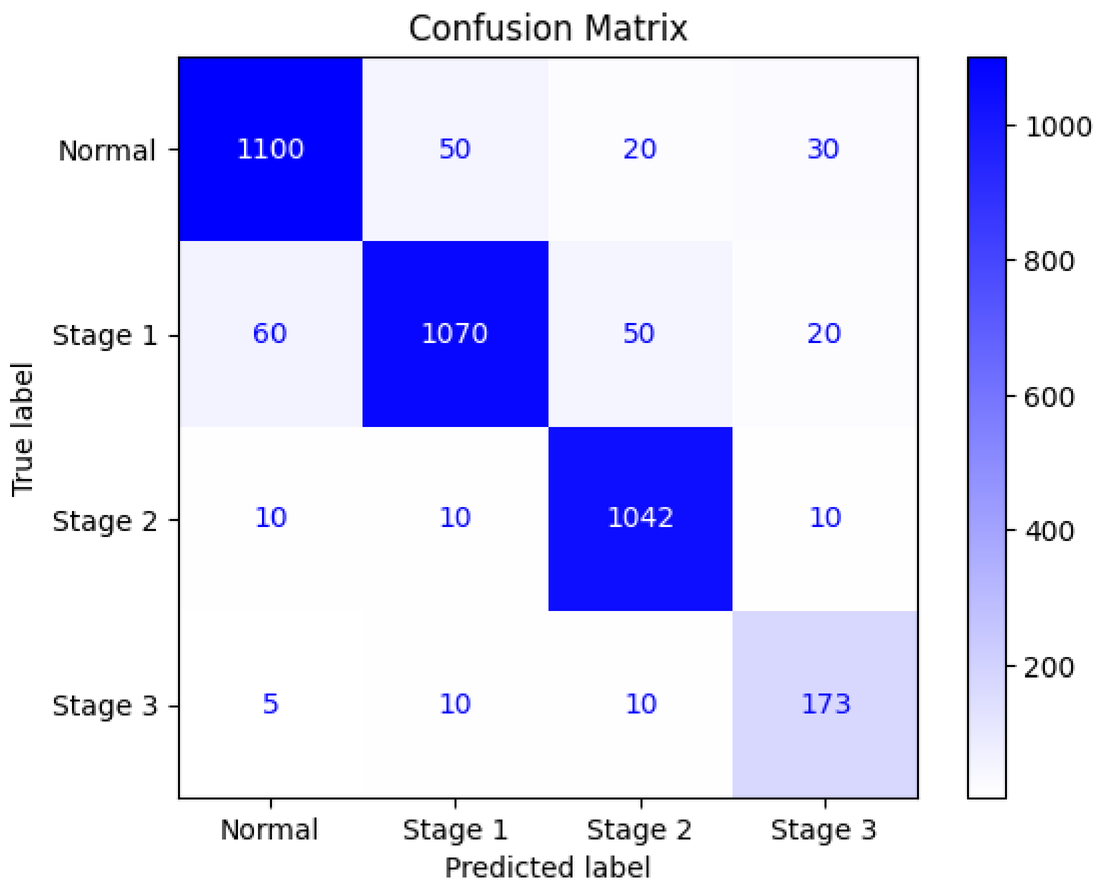

Figure 6.3 Confusion matrix for AlexNet with BAGGING_SVM using OASIS dataset

6.4 Summary

The BAGGING_SVM approach was presented in this chapter as a way to categorize AD phases according to their severity. In the study, the ensemble of bagging with SVM is proposed. To improve the performance in categorizing the stages of AD, a CNN model that has already been trained is used for feature extraction. For comparison, conventional feature extraction techniques such the GLCM, GLDM, and GLRLM are employed. Experimental results are validated and analyzed based on the performance metrics, and the outcome analysis and assessment of the proposed techniques are compared with the literature work. Thus, the accuracy of 91.6% and 89.2% is achieved in the AlexNet with BAGGING_SVM method on ADNI and OASIS dataset, respectively.