International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(Research Article)

Check for updates

Hybrid-based predictive model for early detection of Myopia

Nkechi Grace C Udensi *, Ogochukwu C Okeke and Ike Joseph Mgbemfulike

Department of Computer Science, Faculty of Physical Sciences, Chukwuemeka Odumegwu Ojukwu University, Uli, Anambra State, Nigeria.

International Journal of Science and Research Archive, 2025, 15(02), 001-011

Publication history: Received on 18 March 2025; revised on 29 April 2025; accepted on 01 May 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.15.2.1273

Abstract

This study presents a hybrid-based predictive model for early detection of Myopia to enhance ophthalmic diagnostics. The proposed system was developed using a hybrid neural network in order to improve the early identification of myopia condition in patients. The model was trained and tested using a concept that combines several CNN learners that improved model prediction accuracy. The introduction of penalty terms and a user notification mechanism improved the model's ability to deal with complexity problems. A penalty term was introduced in order to make the model converge more quickly with better accuracy because it gives the user control over the layer's output. The hybrid framework was discouraged from utilizing larger weights by adding a penalty term that was based on the network weights' values. The hybrid CNN input and output layers were invariably fitted with a penalty term. The existing single CNN model achieved an accuracy of 84.89%, while the hybrid model outperformed it with a 95.91% detection accuracy. The current CNN had the lowest detection accuracy, and the system was never made better by adding more training examples. These results demonstrate the effectiveness of the proposed approach in improving early detection of Myopia, offering a scalable and accurate solution for medical diagnosis and intervention.

Keywords: Hybrid-Based; Myopia; Predictive model; Early detection; Penalty term; Convolutional Neural Network; Detection accuracy

1. Introduction

Nearsightedness, is an eye condition where objects that are near appear clear while objects that are far appear blurry (Chia et al., 2020). The medical term for nearsightedness is myopia. Myopia typically manifests with symptoms such as blurred distance vision, squinting, eyestrain, and headaches when trying to see distant objects. The word "myopia" comes from the Greek words myein (to shut) and ops (eye). This describes the squinting that people with myopia do to see clearly. Children, teenagers and adults often experience progressive myopia, where their vision worsens over time. Myopia can be influenced by both genetic and environmental factors. Individuals with a family history of myopia are at a higher risk.

The global prevalence of myopia has been rising at an alarming rate in recent decades. Studies indicate that the rapid increase is largely due to lifestyle changes, such as increased time spent on near work activities and reduced time spent outdoors (Holden et al., 2020). Myopia affects approximately 2.6 billion people globally, and projections estimate that nearly half of the world's population may be myopic by 2050 (Wong et al., 2021). The rapid urbanization and adoption of digital technology in daily life, particularly among younger populations, are driving factors behind this surge.

Myopia progression refers to the gradual increase in refractive error, primarily due to axial elongation of the eye, with the most rapid progression occurring in children and adolescents between ages 6 and 18 (Li et al., 2021). Various factors contribute to this progression, including environmental influences such as prolonged near-work activities (reading, digital device use) and reduced outdoor time, alongside genetic predispositions (Wu et al., 2020; Tedja et al., 2021). The

* Corresponding author: Nkechi Grace C Udensi

heterogeneity of myopia progression means that while some individuals experience rapid increases in refractive error, others progress more slowly. High myopia can lead to severe complications, including myopic macular degeneration and retinal detachment (Lanca & Saw, 2020). Understanding these factors is crucial for developing effective intervention strategies.

Early detection of myopia progression is very important for implementing timely interventions to slow or halt its advancement. Recent advances in optical treatments, such as orthokeratology (overnight contact lenses) and atropine eye drops, have shown promise in managing myopia progression. However, these interventions are most effective when applied early in the disease process, underscoring the importance of developing accurate and reliable predictive models for early detection (Liu & Chen, 2020).

Traditional myopia detection relies on optometric assessments such as visual acuity tests and refractive error measurements, focusing mainly on existing conditions rather than future progression. Techniques like cycloplegic refraction and keratometry enhance diagnostic accuracy, while axial length measurement offers insights into structural eye changes (Zhu et al., 2020). However, these methods often miss early-stage progression risk factors, such as lifestyle behaviors and environmental influences.

AI and machine learning have emerged as transformative tools in healthcare, enhancing diagnostic accuracy and patient outcomes. In ophthalmology, AI can analyze retinal images for early disease detection and is now being applied to predict myopia progression. AI models can incorporate diverse datasets, identifying risk factors and enabling personalized predictions, which are essential for effective myopia management (Liu et al., 2020).

Conventional machine learning algorithms encounter specific difficulties when applied to medical datasets, including imbalanced class problems and outliers. These issues can adversely affect the learning ability of these algorithms and reduce their predictive accuracy. Ensemble learning is a viable solution for managing class imbalance and detecting outliers (Mienye & Sun, 2021). Ensemble learning methods have achieved excellent performance in numerous applications, attracting significant attention in various research fields. These methods enhance the generalization ability of individual classifiers.

Ensemble techniques combine multiple models to create a more robust and accurate predictive model. These methods are particularly valuable in healthcare, where high prediction accuracy is essential. Ensemble methods improve predictive accuracy by aggregating the outputs of several models, each contributing its strengths and offsetting the weaknesses of the others.

2. Related Works

Fitriyani et al. (2019) developed an ensemble model for the early detection of hypertension and type 2 diabetes using patients' risk factor information. The study utilized four imbalanced datasets: hypertension, prehypertension, type 2 diabetes, and chronic kidney disease (CKD). The SMOTE-Tomek method was employed to create balanced datasets, and the Isolation Forest (iForest) algorithm was used for outlier detection. The proposed ensemble achieved accuracies of 96.7%, 85.8%, 75.8%, and 100% for predicting type 2 diabetes, hypertension, prehypertension, and CKD, respectively.

An et al. (2020) introduced an ensemble approach coupled with deep learning techniques to predict Alzheimer's disease. The methodology involved using sparse autoencoders for feature learning and various machine learning algorithms to create multiple models from the learned data. A Deep Belief Network (DBN) was trained as a meta-model to combine the predictions of these base models, achieving accuracy 4% higher than six widely used ensemble algorithms.

Kazemi and Mirroshandel (2020) proposed a novel ensemble method for the early detection of kidney stones, employing decision trees, naïve Bayes, and artificial neural networks (ANN) to learn the relationships among biological features. The combination method utilized a genetic algorithm to assign weights to the classifiers. The proposed ensemble achieved a classification accuracy of 97.1%, identifying critical features associated with kidney stone disease.

Mohammed et al. (2021) utilized an ensemble of convolutional neural network (CNN) classifiers for COVID-19 detection, attaining an accuracy of 77.0%. Ragab et al. (2021) employed a deep learning ensemble that included recurrent neural networks (RNN), gated recurrent units (GRU), and long short-term memory (LSTM) networks, achieving a classification accuracy of 97.2%.

Velusamy and Ramasamy (2021) proposed an ensemble classifier for heart disease detection, combining SVM, KNN, and Random Forest. Various ensemble combination schemes, including majority voting and weighted average voting (WAV), were employed. The study utilized the SMOTE technique for data resampling and the Boruta feature selection technique for optimal feature set selection. The weighted average voting method yielded the highest accuracy of 100%, demonstrating the superior performance of this technique.

Aljame et al. (2020) utilized ensemble learning to detect COVID-19, employing the XGBoost algorithm for early diagnosis. The dataset, containing 5644 instances with 559 positive cases, was preprocessed to handle missing values and outliers. The SMOTE technique was used to balance the dataset. The XGBoost classifier, trained with the preprocessed data, achieved a classification accuracy of 99.9%, outperforming other models in the literature.

Similarly, Abayomi-Alli et al. (2022) applied ensemble learning methods, including AdaBoost and Random Forest, for COVID-19 prediction, with AdaBoost achieving an accuracy of 99.3%. Mohammed et al. (2021) utilized an ensemble of convolutional neural network (CNN) classifiers for COVID-19 detection, attaining an accuracy of 77.0%. Ragab et al. (2021) employed a deep learning ensemble that included recurrent neural networks (RNN), gated recurrent units (GRU), and long short-term memory (LSTM) networks, achieving a classification accuracy of 97.2%.

Mienye and Sun (2021) utilized ensemble learning for heart disease prediction, employing the SMOTE-ENN technique to resample the dataset. They applied recursive feature elimination (RFE) to identify significant features for building the prediction model, where XGBoost achieved an accuracy of 95.6%. In another study, Mienye et al. (2021) proposed a weighted aging classifier ensemble (WAE) to combine predictions from multiple classification and regression tree (CART) models, achieving a classification accuracy of 93% on the Cleveland heart disease dataset.

Gao et al. (2021) examined the performance of the random forest ensemble classifier and other single machine learning algorithms for detecting heart disease. The study identified informative features using principal component analysis (PCA) and linear discriminant analysis (LDA). The ensemble classifier outperformed other algorithms, achieving an accuracy of 98.6%. Similarly, Prakash and Karthikeyan (2021) developed a heart disease prediction method combining genetic algorithms and LDA for robust feature selection. This approach used bagging-based predictive models with SVM, decision tree, and multilayer perceptron (MLP) as base learners, resulting in an accuracy of 93.7%.

3. Research methodology

The proposed system adopts an ensemble technique that combines several neural network models to diagnose myopia disorders among hospital patients. This involves integrating multiple network models at the base level and a single model at the meta-learning level, which learns from the outputs of the base models as depicted in figure 1. A neural network ensemble, comprising three different Convolutional Neural Networks (CNNs), is used to provide more accurate predictions. The three CNN models include one 1D CNN and two 2D CNNs. The 1D and 2D CNNs are stacked on top of a 2D CNN, thereby forming a hybrid system as can be seen figure 2 below.

The ensemble's final prediction is determined through majority voting among all classifiers at the base (zero) level. The class that receives the highest number of votes is selected as the final output of the ensemble, as estimated by the meta-learning model. The model is trained on a dataset that includes both eye images and numerical clinical data.

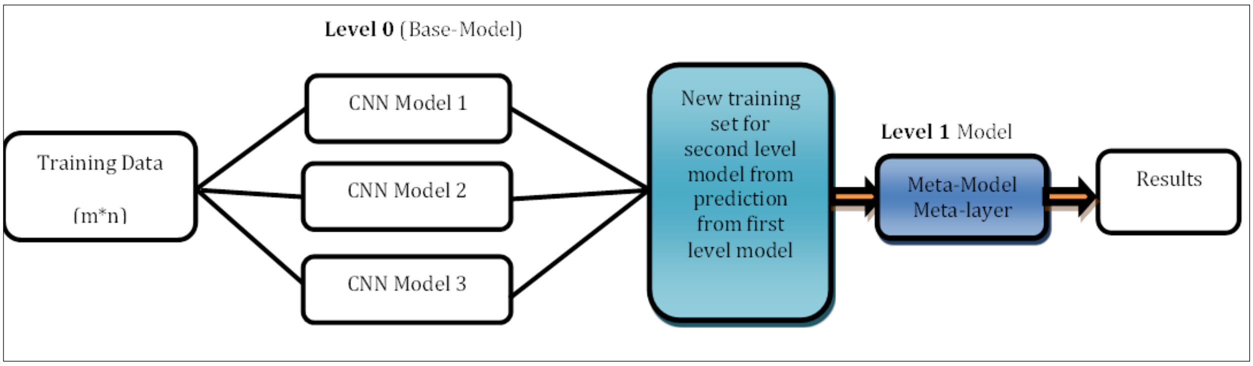

Figure 1 The Ensemble Model (Stefano et al. 2020)

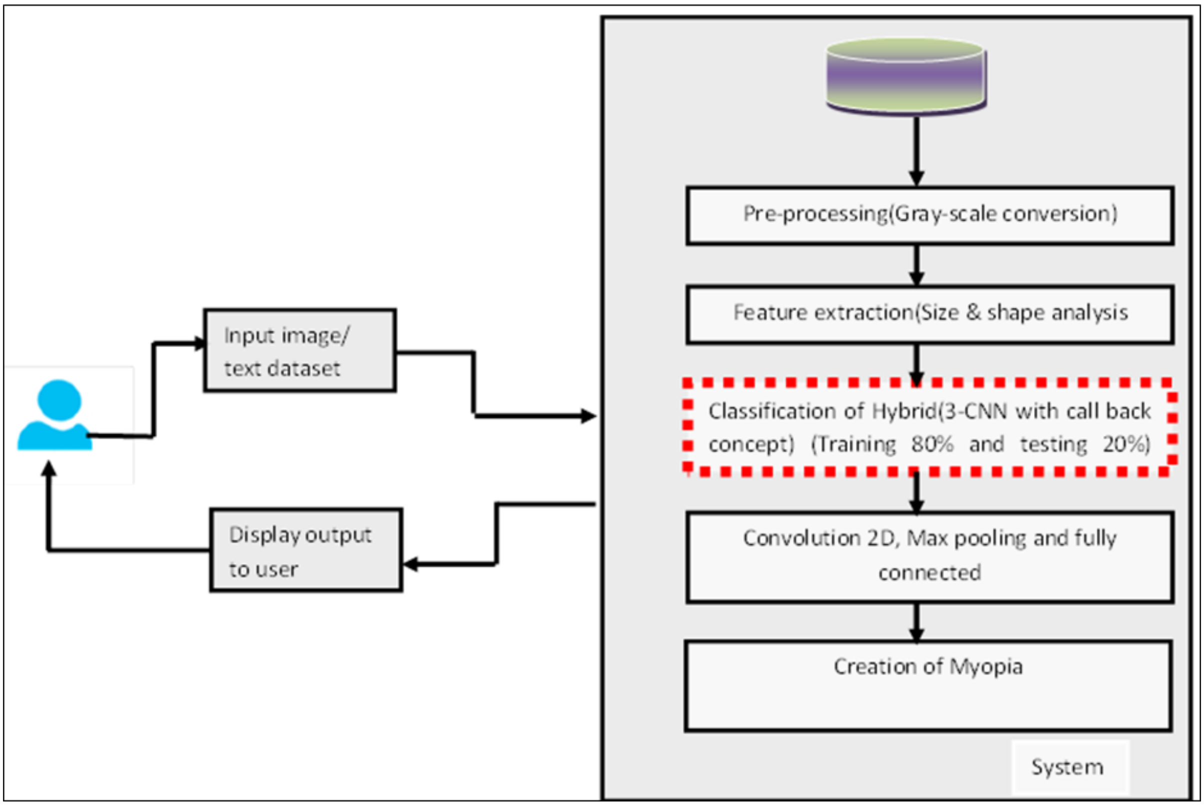

Figure 2 Architectural Diagram of the Proposed System

3.1. Data Collection

Two different datasets were obtained from the Kaggle site, containing both myopic images encoded in JGP format and quantitative data stored in CSV excel file format, with over 742 testing and 1730 training sets making up the 2472 CSV dataset. About 5078 eye images acquired, 2439 images showed myopia instances and 2839 normal images. Myopia image datasets were grouped into folders labeled train, test, and validation, each containing myopia and normal images. To divide our dataset in a half, we employ the train test split function. The train set variable is configured to retrieve Myopia images from the training folder using its file path, whereas the testing set variable is configured to extract images from the testing folder. The suggested collection provides a sizable dataset for building the proposed neural network ensemble.

3.2. Training and Testing Data Split

Three primary subsets were created from data split: the training set, which is employed to train the algorithm; the validation set, which is used to monitor the parameter settings and prevent over-fitting; and the testing set, which is used to assess the effectiveness of the model on newly collected data. This is taking an 80% random sample of the rows and adding them to our set of training data without replacing them. Included also, is the balance of twenty percent (20%) of the sample in the test set. Both numerical and image dataset is used in training the ensemble model.

During training, each level-0 model receives the real labels and passes its predictions as input features to the meta-learning model. Data preprocessing steps include normalization and feature selection. The eye images are preprocessed and passed through 2D CNNs for feature extraction, while the clinical numerical data is processed using a 1D CNN. The ensemble model is then trained and evaluated using performance metrics such as accuracy, precision, recall, and F1-score.

3.3. Machine Learning Module

A convolutional neural network (CNN) is an example of machine learning, specifically a deep learning technique used largely for analyzing images and text processing/classification. The CNN is an appropriate technique to analyze a stream of data with high accuracy. A CNN model was designed to assist with the difficult and time-consuming task of altering weights during each training cycle. The weights that are included in the ordering of inputs to the CNN's layers constitute the factors that cause its weights to change. The neural net weights vary at each of the layers in addition to the activating function. The activation processes change with each subsequent cycle since they serve as the data inputs for the subsequent CNN layer. The resulting shift in distribution requires each and every CNN layer to adjust to the changing data inputs, and that is the reason why the deep learning duration for training increases.

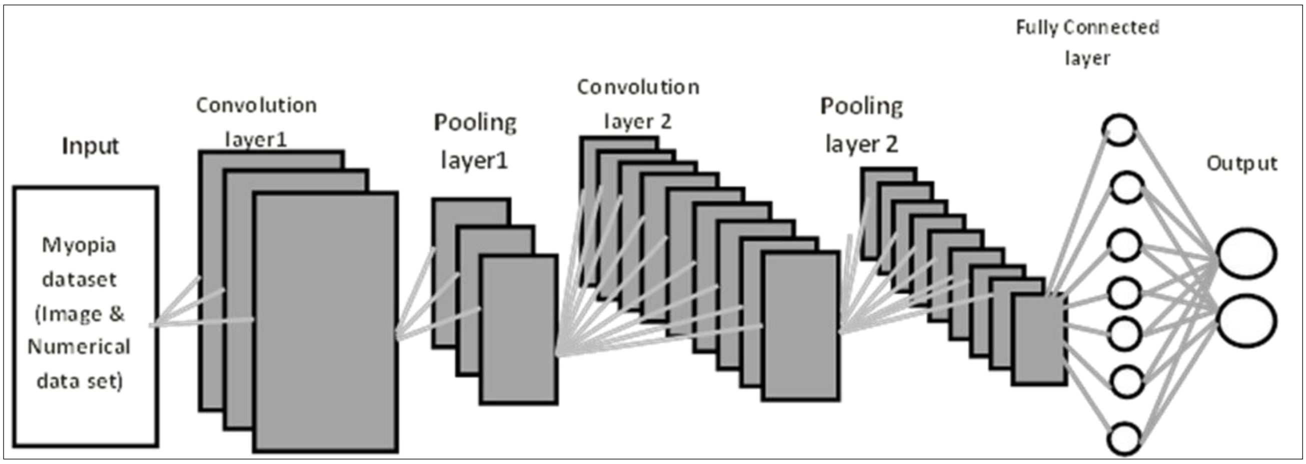

Figure 3 The Analysis of CNN Model

The CNN structure in figure 3, uses a convolutional technique to identify and differentiate between the numerous features on myopia dataset for analysis. The network has multiple pairs of pooling or convolution-based layers, and the layers that are completely linked with the output features from the previous layer. The goal of the CNN design is to overcome model complexity and improve detection accuracy. The proposed CNN analysis is made up of three important layers such as the convolutional layers, pooling and fully connected layers.

3.4. Ensemble Model Architecture

The ensemble model architecture in figure 4 combined three different CNN techniques placed at the 0-level often referred to as the base model. The average of functionalities of three different CNN models are used as at the base level to build a meta-learning model called the stack.

- **New Training Data:** The training dataset is divided into k-folds cross validation and fitted using the base models on k--1 path of the whole training set to compute its performance using the test dataset with predictions made for the k-th part. This process is repeated and predictions from training set are used as features for training the ensemble model and used to visualize the testing dataset.
- **Meta model** at level-1 is a meta-layer that accepts output from the base models (0-level) as the new training data. Three different neural network models stacked on top the based model called meta-learning model is producing the final predictions. The connections of the neural network ensembles used concatenated operations.

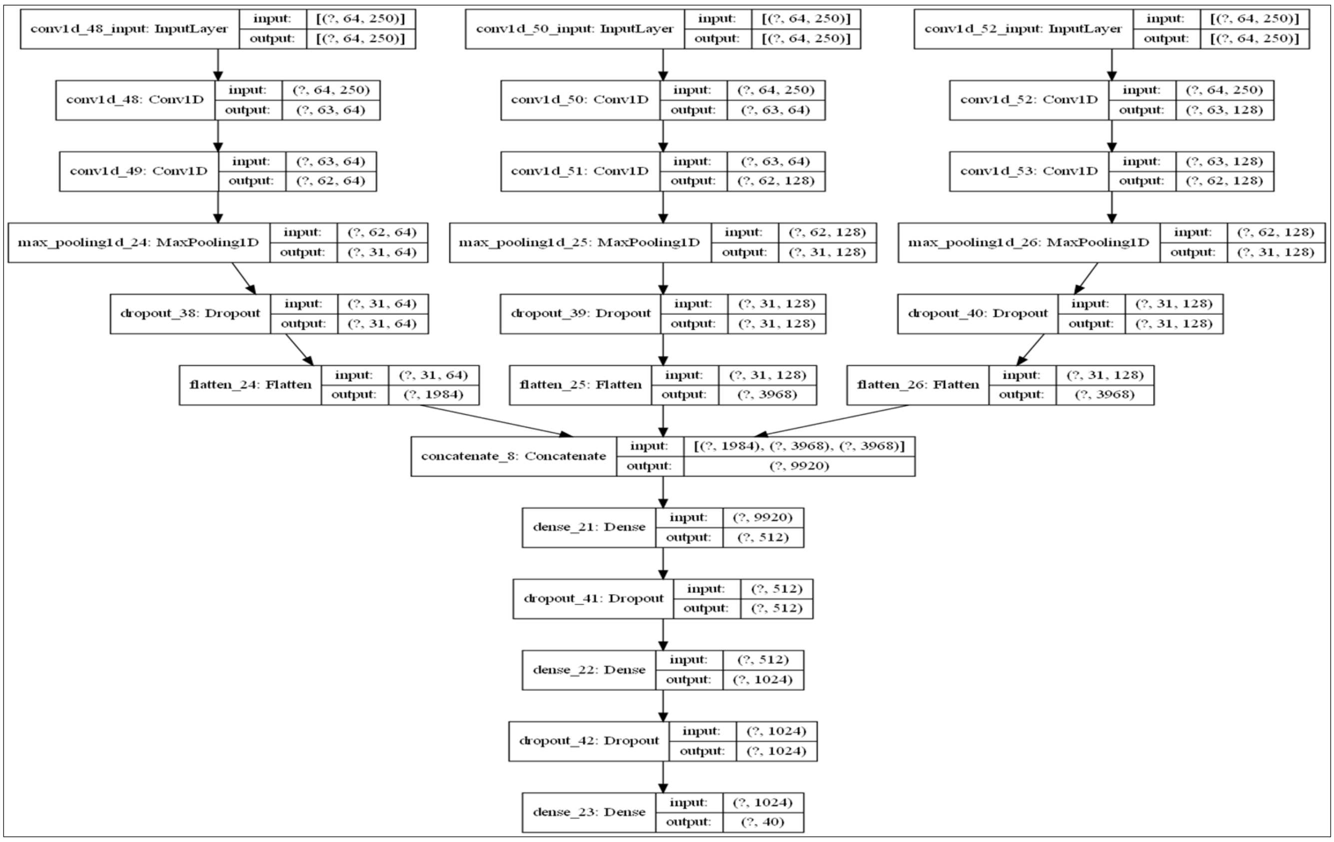

Figure 4 The Architectural Diagram of the Ensemble Model

3.5. Research Novelty

- Introducing penalty term to the activated weights of neural network ensemble: A penalty term was introduced in order to make the ensemble model results to be smaller (or more like 0); this is to make the model converge more quickly and with better accuracy because it gives the user control over the layer's output. The framework is discouraged from using larger weights by this penalty term, which is determined by the value of the weights. The above concept is built into CNN's input and output layers in all cases.
- Adding penalty term to bias weights: Penalty term was added to the layer's bias weights. The learning algorithm's bias weights are modified to promote the use of small weights by the network. Also modified was the loss computation utilized in the network optimization process to take the weight sizes into account. This is done to reduce model weights in the optimization process.
- Adding penalty term to the CNN kernel: A term for L1-penalty is included to the layer's bias vector, which proves helpful at times although the bias usually has less impact on the model's complexity. A penalty term equal to 0.01 times the square root of the normalized weights was introduced. It produced a neural network with several hidden layers when applied to convolutional or dense layers.

4. System Implementation

A cross validation set consisting of ten folds for each training set was utilized to train the system and assess its performance. Also the sample data set was partitioned into training set (80%) and test set (10%). The 80% of training set and 10% of test set were done using CNN and the hybrid with call-back concept. The modules provide a space for variety of activities such as sample dataset, pre-processing dataset/label training set, test set, training to predict normal and myopia features as well as performance evaluation analysis.

4.1. Convolutional Neural Network (CNN) Implementation

2D-CNN(neural network classifier) was launched in Python and imported from the sklearn.neural network library as shown in the screen shot (Fig. 4). The parameters used to define the 2D-CNN model were solver='lbfgs', alpha set to 1e-5, max iteration=200, activation function='relu,' and three hidden layers (first, second and third layers has 10, 30 and 10 nodes). Shuffle=True, and the random_state is set to 1. The model was created using the codes 2DCNN_Model.fit(X train,y test), which was used to train the model, and CNN_Pred=NN_Model.predict(X_test) to predict target using the testing dataset. The pre-processed training data was used to train the CNN classifier that could predict target class images.

5. Results and Discussion

The model's performance was tested using, classification reports and confusion matrix in order to predict normal and myopia features and analyze performance. The system used hybrid-based ML-drive concept that combined multiple predictions. A single meta-learning model that combines the predictions of three distinct CNN classifiers outperformed using only one model. Every model receives the predictions of the models that came before it as features in addition to the features in the dataset. Every model automatically receives the real labels as features at training.

- **Classification (prediction) Accuracy**: The Classification accuracy is the ratio of correctly classified data points to the total no. of points in the dataset which ranges from 0-100%.

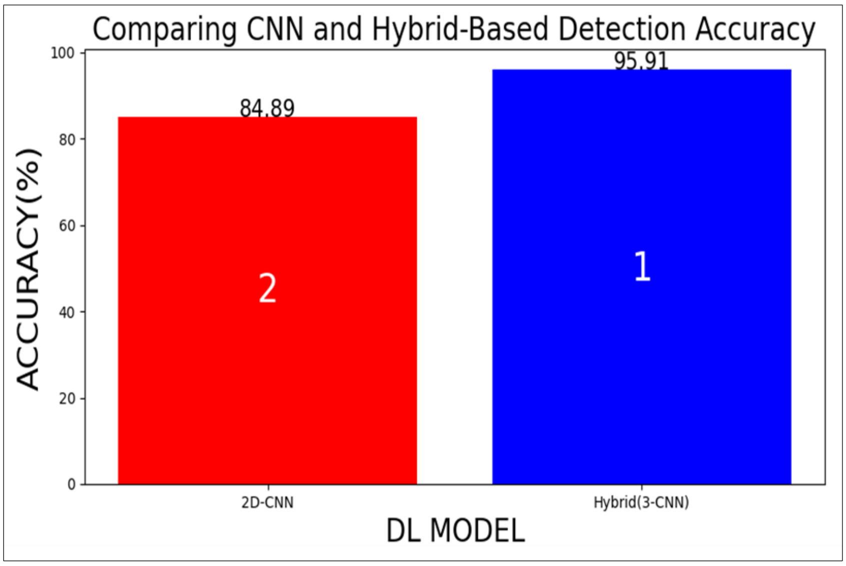

Figure 5 Prediction accuracy of CNN and Hybrid

Figure 5 depicts the performance accuracy of existing CNN and proposed Hybrid. The hybrid model produced the best results, with 95.91% accuracy and the existing CNN model yielded 84.89%. The existing CNN had the lowest detection accuracy and the addition of more training samples never improved the existing system.

Table 1 shows some of the formulas used in calculating the true positive rate, false positive rate, true negative rate and false negative rate to confirm the result shown in figure 5 above.

Table 1 Existing CNN and proposed Hybrid-based model

Metrics	**Existing CNN**	**HYBRID-BASED MODEL**
True Positive(TP)	427	409
False Negative(FN)	3	21
False PositiveFP)	51	25
True Negative(TN)	14	40
Correct predictions	441	449
Incorrect predictions	54	46
Prediction Accuracy	84.89	95.91

Table 2 Classification report of CNN

	Precision	Recall	F1-Score	Support
Myopia	0.90	0.99	0.94	430
Normal	0.73	0.25	0.37	65
Accuracy			0.89	495
Macro avg	0.81	0.62	0.65	495
Weighted avg	0.87	0.89	0.86	495

The 2D-CNN classification report for myopia and normal times are shown in Table 2. It includes the precision, recall, and f1-score accuracy of the existing system model. For myopia, the precision accuracy recall and f1-score produced results of 0.90, 0.99, and 0.94 for each. The normal cases yielded a precision accuracy of 0.73. f1-score of 0.37 and a recall score of 0.25. The classification accuracy recorded 0.89 score.

Table 3 Classification report of the hybrid model

	Precision	Recall	F1-Score	Support
Myopia	0.93	0.96	0.95	430
Normal	0.69	0.55	0.62	65
Accuracy			0.91	495
Macro avg	0.81	0.76	0.78	495
Weighted avg	0.90	0.91	0.90	495

The classification report of hybrid-based model is presented in Table 3, with precession, recall, and f1-score classification for normal and myopia cases. The highest score in the precision metrics is 0.93, recall for myopia recorded 0.96, and the f1-score was 0.95 for myopia images. There is a significant improvement, as shown in the precision, recall, and f1-score values from the hybrid-based classification report. The macro-average shows how all categories equally contributed to the final averaged metrics, the weighted-average shows how each class appears to contribute to the average as weighted by its size, and the micro-average clearly demonstrates how all samples equitably make a contribution to the final averaged matrix.

- **Confusion matrix**: displays a table structure of the existing and proposed system prediction results of a binary-classification task to aid in visualizing its results. This is used to show the predicted and actual values of a classification model. Cell values above and below the main diagonal or off-diagonal elements showing the incorrectly predicted values, show the total number of correctly predicted values that are equal to the actual or true values. The greater the diagonal value, the more accurate the predicted target values.

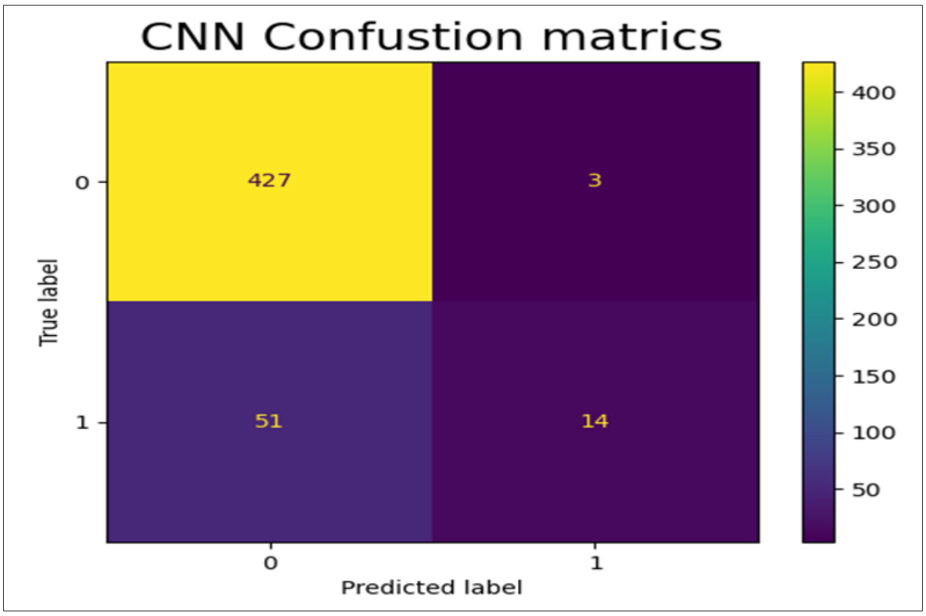

Figure 6 Confusion matrix of CNN

Figure 6 shows the confusion matrix, which displays a table structure of the CNN prediction results and outcomes of a binary-classification task to aid in visualizing its results. This is used to show the predicted and actual values of a classification model. Cell values above and below the main diagonal or off-diagonal elements showing the incorrectly predicted values, show the total number of correctly predicted values that are equal to the actual or true values. The greater the diagonal value, the more accurate the predicted normal and myopia images. According to the confusion matrix, normal cases had 3 incorrectly predicted cases with 427 correct predictions. While myopia cases provided 14 correctly predicted values with 51 wrongly classified predictions.

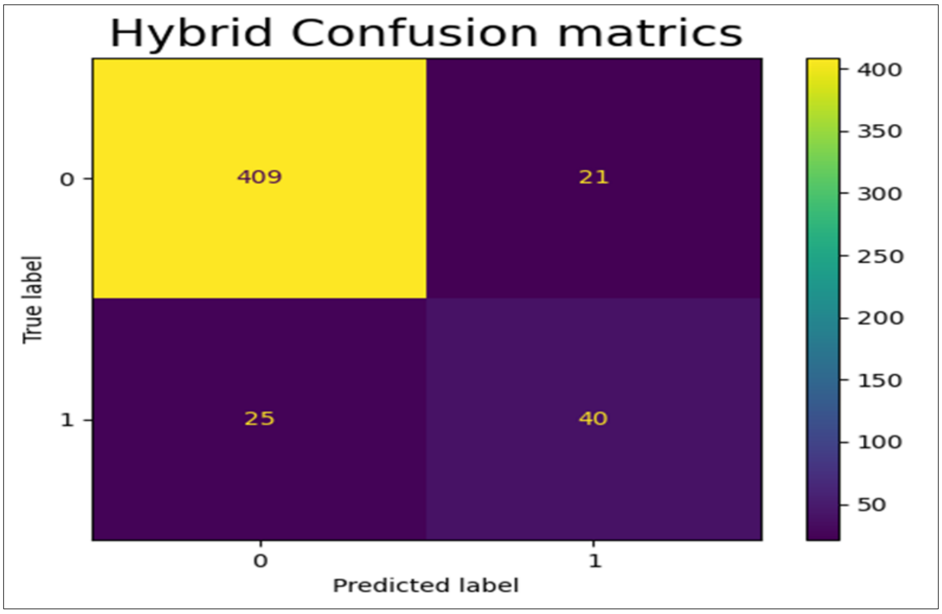

Figure 7 Confusion metrics of Hybrid

A confusion matrix is shown in Figure 7 as been used to evaluate the performance of the hybrid classification system using the validation set. It displays the type of errors made by the classifier. The suggested model confusion matrix was created, with accurate predictions displayed at the secondary diagonal and inaccurate predictions noted above and below the main diagonal, or "off-diagonal elements," in that order using the testing dataset. The overall number of correctly predicted values recorded to be 409+40 which is 449 instances while incorrectly predicted cases is 21+25 or 46 cases

6. Conclusion

This study developed a hybrid predictive model for the early detection of myopia in hospital patients. Given Nigeria's rapidly growing population and the rising incidence of myopia among both adults and children, there is an urgent need for affordable and high-quality healthcare solutions. Traditional CNN models, when applied to real-world datasets, exhibited significant challenges including overfitting, misclassification errors, and limited generalization capacity. These inefficiencies reduced their suitability for practical healthcare deployment.

The proposed hybrid framework, which combined multiple CNN architectures and employed feature engineering techniques to better align data features, effectively addressed these limitations. Comparative evaluation demonstrated that the hybrid model consistently outperformed a single CNN model in terms of detection accuracy and robustness. This research offers a promising step forward in utilizing machine learning for early diagnosis and management of myopia, providing a more reliable and scalable solution for clinical use.

To further enhance the transparency and clinical trust of predictive models for myopia detection, future research can explore the integration of Explainable AI (XAI) techniques. Approaches such as Grad-CAM for visualizing important regions in medical images, and SHAP or LIME for explaining the impact of clinical features, could offer valuable insights into model decisions and foster greater trust among healthcare practitioners

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Abayomi-Alli, O., Ojo, J., and Akinsola, J. (2021). Ensemble learning for COVID-19 prediction: A comparative study. Journal of Healthcare Engineering, 2-40.

[2] Abayomi-Alli, O. O., Damaevičius, R., Maskeliūnas, R., and Misra, S. (2022). An ensemble learning model for COVID-19 detection from blood test samples. Sensors, 22(6), 22-24. https://doi.org/10.3390/s22062224

[3] Ardabili, S. F., Mosavi, A., Ghamisi, P., Ferdinand, F., Varkonyi-Koczy, A. R., Reuter, U., Rabczuk, T., & Atkinson, P. M. (2020). COVID-19 outbreak prediction with machine learning. Algorithms, 13(10), 249. https://doi.org/10.3390/a13100249

[4] Ak, M. F. (2020). A comparative analysis of breast cancer detection and diagnosis using data visualization and machine learning applications. Healthcare, 8(2), 111. https://doi.org/10.3390/healthcare8020111

[5] Alfaiz, M. S., and Fati, A. (2021). Credit card fraud detection models using ensemble methods. Journal of King Saud University - Computer and Information Sciences. https://doi.org/10.1016/j.jksuci.2021.03.007

[6] Aljame, A., Khedher, L., & Jarraya, H. (2020). XGBoost for COVID-19 detection: A predictive analysis approach. International Journal of Biomedical Engineering and Technology, 33(1), 1-14.

[7] An, N., Ding, H., Yang, J., Au, R., and Ang, T. F. A. (2020). Deep ensemble learning for Alzheimer's disease classification. Journal of Biomedical Informatics, 105, Article 103411. https://doi.org/10.1016/j.jbi.2020.103411

[8] Chen, J., Liu, Q., and Tan, T. (2020). Advanced text classification techniques using machine learning. Journal of Computational Linguistics, 25(1), 50-66.

[9] Chen, T., Kornblith, S., Norouzi, M. and Hinton, C.(2020) A simple framework for contrastive learning of visual representations. In International Conference on Machine Learning. https://doi.org/10.48550/arXiv.2002.05709

[10] Chia, A., Ohno-Matsui, K., and Saw, S.-M. (2020). The Complications of Myopia: A Review and Meta-Analysis. Investigative Ophthalmology & Visual Science (IOVS). Retrieved from https://iovs.arvojournals.org/article.aspx?articleid=2765517

[11] Chicco, D., and Jurman, G. (2021). An ensemble learning approach for enhanced classification of patients with hepatitis and cirrhosis. IEEE Access, 9, 24485-24498. https://doi.org/10.1109/ACCESS.2021.3057196

[12] Da Conceição, L.R., da Costa, C.E., Rocha, G.N.d., Pereira-Filho, E.R., Zamian, J.R., (2015). Ethanolysis optimisation of jupati (raphia taedigera mart.) oil to biodiesel using response surface methodology. J. Brazil. Chem. Soc.26, 1321–1330

[13] Da Conceição, L.R.V., Carneiro, L.M., Rivaldi, J.D., de Castro, H.F., (2016). Solid acid as catalyst for biodiesel production via simultaneous esterification and transesterification of macaw palm oil. Ind. Crops Prod., 89, 416–424.

[14] Ebiaredoh-Mienye, S. A., Swart, T. G., Esenogho, E., and Mienye, I. D. (2022). A machine learning method with filter-based feature selection for improved prediction of chronic kidney disease. Bioengineering, 9(8), 350. https://doi.org/10.3390/bioengineering9080350

[15] Esenogho, E., Mienye, I. D., and Sun, Y. (2021). An ensemble learning-based credit card fraud detection method. Artificial Intelligence Review, 54(2), 1233-1245. https://doi.org/10.1007/s10462-020-09875-6

[16] Fitriyani, N. L., Syafrudin, M., Alan, G., and Rhee, J. (2019). Development of disease prediction model based on ensemble learning approach for diabetes and hypertension. IEEE Access, 7, 144777-144789. https://doi.org/10.1109/ACCESS.2019.2945129

[17] Gao, X. Y., Ali, A. A., Hassan, H. S., & Anwar, E. M. (2021). Improving the accuracy for analyzing heart disease prediction based on the ensemble method. Complexity, 2021, 1-10. https://doi.org/10.1155/2021/6663455

[18] Holden, B. A., Fricke, T. R., Wilson, D. A., Jong, M., Naidoo, K. S., Sankaridurg, P., Wong, T. Y., Naduvilath, T., Resnikoff, S., & (The International Myopia Institute) (2020). Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050: a systematic review and meta-analysis. Ophthalmology, 127(2), 202–210.

[19] Kazemi, M., and Mirroshandel, S. A. (2021). Novel ensemble method for early detection of kidney stones. Health Information Science and Systems, 9(1), 1-11.

[20] Lanca, C., and Saw, S. M. (2020). The association between digital screen time and myopia: A systematic review. Ophthalmic & Physiological Optics, 40(2), 216-229.

[21] Li, S. M., Kang, M. T., Chen, Q., and Zhao, X. (2021). Predicting the risk of developing myopia based on visual habits using an ensemble learning approach. International Journal of Ophthalmology, 13(7), 1000-1007.

[22] Mohammed, A. and Kora, R.(2023), A comprehensive review on ensemble deep learning: Opportunities and challenges, Journal of King Saud University – Computer and Information Sciences, 35, 757–774.

[23] Ragab, D., Eldesoky, M., Sakr, M., & Mohammed, M. (2021). A deep-learning ensemble for COVID-19 detection from chest CT scans using hybrid CNNs and RNNs. Computers in Biology and Medicine, 135, 104572

[24] Tedja, M. S., Guggenheim, J. A., & Woldehawariat, G. (2021). The Rotterdam Study—and beyond: insights into the etiology of myopia. Progress in Retinal and Eye Research, 80, 100878.

[25] Tran, L. T., and Wong, K. (2021). An ensemble deep learning framework for detecting COVID-19 from chest X-ray images. Expert Systems with Applications, 175, Article 114835. https://doi.org/10.1016/j.eswa.2021.114835

[26] Velusamy, R., & Ramasamy, R. (2021). Ensemble of heterogeneous classifiers for diagnosis and prediction of coronary artery disease with reduced feature subset. Computer Methods and Programs in Biomedicine, 202, 105866. https://doi.org/10.1016/j.cmpb.2020.105866

[27] Wu, S. Y., and Lin, C. Y. (2021). An ensemble learning approach for predicting the effectiveness of COVID-19 vaccines. Computer Methods and Programs in Biomedicine, 198, Article 105741.