Performance Comparison of CNN and DNN Algorithms for Automation of Diabetic Retinopathy Disease

B. Sankara Babu[1*], *Bhavita* Mandapati[2], *Bhavana* Mandapati[2], *Harshith* Nallapu[2], *Prasenjeet Samanta*[3], K.Maithil[4]
[1]Department of CSE, Gokaraju Rangaraju Institute of Engineering and Technology (GRIET), Hyderabad, Telangana State, India
[2]UG Student, Department of CSE, GRIET, Hyderabad, Telangana State, India.
[3]Uttaranchal Institute of Management, Uttaranchal University, Dehradun, 248007, India.
[4]KG Reddy College of Engineering & Technology, Hyderabad.

Abstract. Automation of medical image analysis helps medical practitioners to ensure early detection of certain diseases. Diabetic Retinopathy (DR) is a widespread condition of diabetes mellitus and a main global cause of vision impairment. The manual diagnosis of diabetic retinopathy by ophthalmologists requires a significant amount of time, causing inconvenience and discomfort for patients. However, the use of automated technology makes it possible to quickly identify diabetic retinopathy, permitting the continuation of therapy without interruption and averting future ocular damage. This paper presents a comprehensive comparative analysis of six Convolutional Neural Networks and Deep Neural Networks based machine learning models, including simple CNN, VGG16, DenseNet121, ResNet50, InceptionV3, and EfficientNetB3, for the recognition of diabetic retinopathy using fundus photographs. The accuracy of various models is evaluated using the Cohen Kappa metric. The results of this study add a contribution to the research on the application of machine learning models for diagnosing diabetic retinopathy.

1 Introduction

Diabetic retinopathy is a serious eye condition that affects people with diabetes. It is one of the major causes of blindness in the world, timely treatment is highly dependent on early and precise detection. The evolution in machine learning techniques has led to the development of automated systems for diabetic retinopathy detection by analyzing fundus photographs. This research report compares various machine learning models for detecting diabetic retinopathy in fundus photographs. We seek to enhance the precision and effectiveness of diabetic retinopathy diagnosis through the use of ML algorithms, ultimately enabling early intervention and averting vision loss.

*Corresponding Author: bsankarababu81@gmail.com

The comparative analysis involves evaluating different ML models, particularly those based on Convolutional Neural Networks (CNNs). These models will be trained and tested on a large dataset of fundus photographs to identify diabetic retinopathy. The evaluation will be based on Cohen's Kappa Score, a metric that measures the agreement between two raters assessing the same object. This study aims to provide insights into the strengths and weaknesses of each model, helping the authors determine the most effective strategy for diabetic retinopathy detection in fundus photographs. This study aims to improve diabetic retinopathy diagnoses by evaluating different ML models.

Findings offer insightful comparisons, facilitating the development of reliable automated systems for precise and timely diagnosis. By integrating these tools into clinical practice, early identification, personalized treatment strategies, and improved patient outcomes can be achieved.This study compares Machine Learning (ML) algorithms for detecting diabetic retinopathy in fundus images. We want to enhance automated methods in identifying diabetic retinopathy and advance ophthalmic healthcare by evaluating its effectiveness and examining its strengths and limits.

2 Existing Methods

Diabetic Retinopathy is a severe complication of diabetes mellitus that poses a great risk to patients' vision. However, the current practice of annual DR screening for diabetic patients faces challenges such as resource limitations, high patient volumes, and financial load. Machine learning classifiers offer potential remedies by assisting in the diagnosis process. Existing models have shown promising results but lack external validation and may not generalize well to different populations. This research aims to classify DR using ML techniques on Saudi diabetic data. We aim to identify the most accurate method and extract interpretable features from socio-demographic and clinical information. For automated DR detection, this study compares and contrasts support vector machines, random forests, and convolutional neural networks ML models. A big dataset of fundus pictures is used for the evaluation. The accuracy, precision, recall, and F1 score are among the metrics the authors use to evaluate the models' performance. We seek to identify the most successful ML model for DR recognition by comparing the outcomes [1].

This study investigates the use of Deep Learning Convolutional Neural Networks for automated diabetic retinopathy detection. The authors propose a comprehensive model that incorporates 26 state-of-the-art DL networks for deep feature extraction and image classification of DR fundus photographs. The performance of these networks is evaluated, with a particular focus on overfitting. Among the tested models, ResNet50 exhibits the highest overfitting, while Inception V3 demonstrates the lowest overfitting when trained on Kaggle's EyePACS fundus image dataset. EfficientNetB4 becomes the most effective, optimal, and reliable DL method for DR detection. This research paper centers around the application of CNNs for detecting diabetic retinopathy. The authors compare the performance of various CNN architectures, including AlexNet, VGGNet, and ResNet, in terms of accuracy and sensitivity. They discuss the strengths and limitations of each model and provide insights into the potential of deep-learning approaches for diabetic retinopathy recognition [2][10].

Untreated diabetic retinopathy, a complication of uncontrolled chronic diabetes, can lead to complete vision loss. Early detection and medical intervention are crucial to mitigate the severe consequences of this condition. Manual diagnosis of diabetic retinopathy is time-consuming for ophthalmologists and can be discomforting for patients. Automating the diagnosis process using technology allows for prompt identification of diabetic retinopathy, facilitating timely treatment to prevent further eye damage. This study proposes the utilization of machine learning techniques to extract three key features: exudates, hemorrhages, and microaneurysms. These features play a vital role in diagnosing diabetic retinopathy[11]. A hybrid classifier, combining support vector machines, k-nearest neighbors, and random forests, is employed to categorize these extracted features and facilitate accurate classification [3][9].

Diabetic retinopathy is a leading cause of blindness worldwide, associated with diabetes. Detecting and treating it early is crucial. Researchers have explored artificial intelligence methods, particularly deep learning techniques, using fundus retina images to identify and classify diabetic retinopathy. This review thoroughly examines the application of deep learning at different stages of the detection pipeline. It covers datasets, preprocessing methods, and the development of deep learning models for diagnosis, grading, and lesion localization. Real-world applications and future research areas are discussed. The paper evaluates the effectiveness of deep learning models, specifically CNNs and DBNs, in automating diabetic retinopathy detection. It emphasizes their potential for high accuracy and efficiency [4][8].

Diabetic retinopathy, a common complication of diabetes affecting the retina, causes vision impairment and blindness in working-age adults. Automated systems for diagnosing this condition present new challenges, including the need for resource-efficient architectures. This paper compares EfficiencyNet, an optimized architecture, with DenseNet and ResNet. The study utilizes a dataset from the APTOS Symposium, classifying retinal images into five cases. Various machine learning techniques, including SVM, decision trees, and naive Bayes, are analyzed for detecting diabetic retinopathy. Performance is evaluated using publicly available datasets and metrics like accuracy, precision, recall, and F1 score. The authors provide insights into the strengths and limitations of each algorithm [5][7].

3 Proposed Method

a. Proposed Method

By utilizing image processing, the proposed automated system is capable of discerning between retinal images showing diabetic retinopathy and those depicting a normal condition. The preprocessing of Fundus images includes color space conversion, cropping, resizing, and a combination of weighted addition and Gaussian blur. These steps are commonly used in image preprocessing to enhance features, reduce noise, and prepare the image for further analysis or machine learning tasks.

The proposed method uses the Fundus images dataset and preprocesses the images using various OpenCV functions to convert color space from BGR to RGB, cropping the images, resizing them, and adding Gaussian blur to make sure that no data item in the dataset becomes an anomaly. The dataset is split into train, test, and validation datasets. This system uses Cohen's Kappa metric to show the agreement between predicted and true labels. A custom callback is used to monitor the validation of Cohen's kappa metric during training and save the model whenever an improvement in the metric is observed.

This research paper seeks to construct an automated system for diagnosing and classifying diabetic retinopathy from fundus photos. The goal is to design a CNN architecture that is capable of accurately identifying and classifying various stages of the disease, thus allowing for more timely diagnosis and treatment decisions. The main objectives of the study are to optimize the CNN model's performance by investigating preprocessing techniques such as image enhancements and normalization to enhance feature extraction from the fundus photographs, as well as various data augmentation techniques to improve the model's overall coverage and reduce overfitting. Furthermore, the proposed model will be tested against a variety of patient populations and imaging conditions to determine its reliability and applicability to real-world conditions. The results of this study will be beneficial to the medical image analysis field by providing a dependable and automated method for diagnosing and grading diabetic retinopathy with CNNs.

3.2 Architecture Diagram

The input stage involves capturing retinal images from diabetic patients using specialized imaging devices. These images serve as the primary input for the system. In the image detection phase, computer vision techniques are applied to detect the presence of retinal images in the input data. This step ensures that only relevant images are processed further. Then the image processing techniques are employed to enhance the quality and clarity of the retinal images. The goal is to prepare the images for feature extraction and analysis.

Feature selection methods are applied to extract relevant features from the preprocessed retinal images. These features could include blood vessel abnormalities, exudates, microaneurysms, or other indicators of diabetic retinopathy. The selected features are used to classify and determine the severity of diabetic retinopathy in the retinal images. The model learns from the extracted features to make predictions and classify the retinal images. Then the output stage provides the results of the diabetic retinopathy analysis. It may include information such as the severity level of diabetic retinopathy, the presence of specific abnormalities, or a risk assessment. Figure 1 represents the manuscript flow.

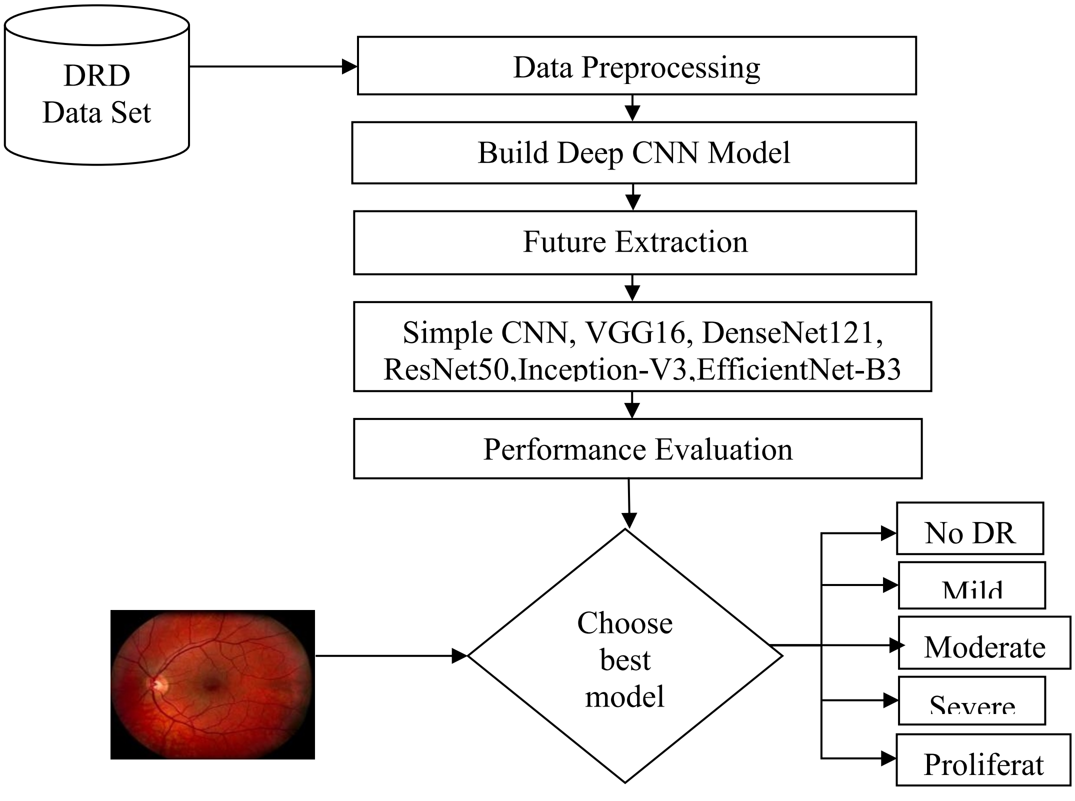


Fig. 1. Architecture diagram and flow overview of the suggested system

https://doi.org/10.1051/e3sconf/202343001075

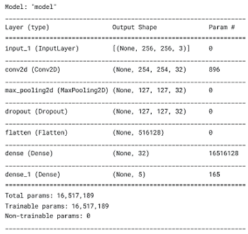
4. Results and Discussions

The CNN-based machine learning techniques that are performed for classification in this system are:

4.1 Simple Convolutional Neural Network (CNN)

Simple CNN is a basic CNN architecture used for pattern recognition and image classification, often with fewer layers and simpler design.

Figure 2 (a) illustrates the summary of the Conventional CNN parameters, 2 (b) illustrates the Kappa Score and 2 (c) depicts the Confusion Matrix.



Model: "model"

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 256, 256, 3)]	0
conv2d (Conv2D)	(None, 254, 254, 32)	896
max_pooling2d (MaxPooling2D)	(None, 127, 127, 32)	0
dropout (Dropout)	(None, 127, 127, 32)	0
flatten (Flatten)	(None, 516128)	0
dense (Dense)	(None, 32)	16516128
dense_1 (Dense)	(None, 5)	165

Total params: 16,517,189
Trainable params: 16,517,189
Non-trainable params: 0

Fig.2(a) CNN Model Summary

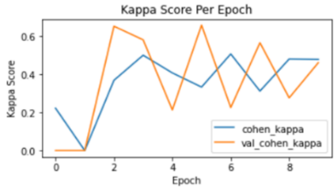


Fig.2(b) Kappa Score Per Epoch of Simple CNN

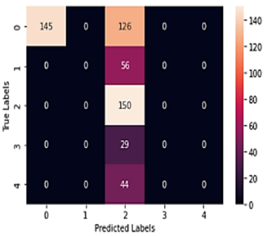


Fig.2(c) Confusion Matrix of Simple CNN

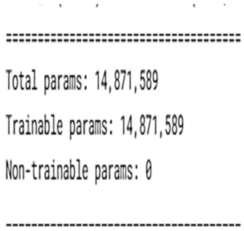
Fig. 2 Results of Simple CNN model

We can observe the Kappa Scores of Simple CNN are fluctuating for each epoch. This shows that simple CNN might be experiencing difficulties inlearning certain patterns or features present in the data, leading to lower efficiency. By observing the confusion matrix of Simple CNN, as the total instances are 550 and correctly predicted instances are 295, the authors got an accuracy of 53.63%.

4.2 VGG 16

VGG16 is a deep learning algorithm that is 16-layer deep and it is loaded with pre-trained weights for the dataset.

Figure 3(a) illustrates the summary of VGG16 model parameters, 3(b) illustrates the kappa score and 3(c) depicts the Confusion Matrix.



Total params: 14,871,589
Trainable params: 14,871,589
Non-trainable params: 0

Fig.3(a) VGG16 Model Summary

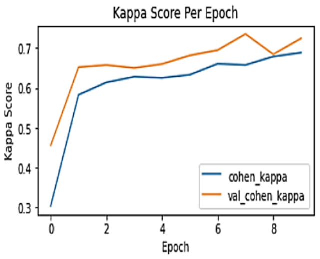


Fig.3(b) Kappa Score Per Epoch of VGG16

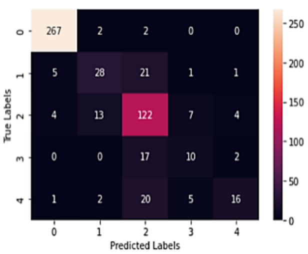


Fig.3(c) Confusion Matrix of VGG16

Fig. 3 Results of using VGG16 model

We can observe the Kappa Scores of DenseNet121 in validation are slightly fluctuating for each epoch, but Kappa Scores in training are significantly high. This shows that DenseNet121 is less efficient for validation datasets, and the model suggests overfitting.

By observing the confusion matrix of VGG16, as the total instances are 550 and correctly predicted instances are 408, the authors got an accuracy of 74.18%.

4.3 DenseNet121

DenseNet121 is a CNN algorithm with dense connections between layers. It consists of four dense blocks and 121 layers.

Figure 4(a) illustrates the summary of DenseNet121 model parameters ,4(b) illustrates the kappa score and 4(c) depicts the Confusion Matrix.

```
Total params: 7,042,629
Trainable params: 6,958,981
Non-trainable params: 83,648
```

Fig.4(a) DenseNet121 model Summary

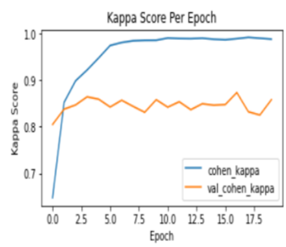


Fig.4(b) Kappa Score Per Epoch of DenseNet121

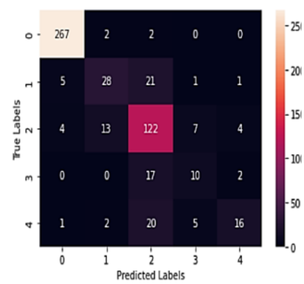


Fig.4(c) Confusion Matrix of DenseNet121

Fig.4 Results of using DenseNet121 model

We can observe the Kappa Scores of DenseNet121 in validation are slightly fluctuating for each epoch, but Kappa Scores in training are significantly high. This shows that DenseNet121 is less efficient for validation datasets, and the model suggests overfitting. By observing the confusion matrix of DenseNet121, as the total instances are 550 and correctly predicted instances are 375, the authors got an accuracy of 68.18%.

4.4 ResNet-50

ResNet-50 is a deep learning algorithm that works with 50 neural network layers.

Figure 5(a) illustrates the summary of ResNet-50 model parameters,5(b) illustrates the kappa score and 5(c) depicts the Confusion Matrix.

Total params: 23,597,957

Trainable params: 23,544,837

Non-trainable params: 53,120

Fig.5(a) ResNet-50 model Summary

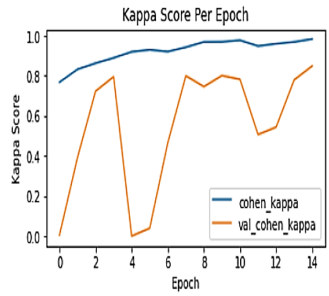


Fig.5(b) Kappa Score Per Epoch of ResNet50

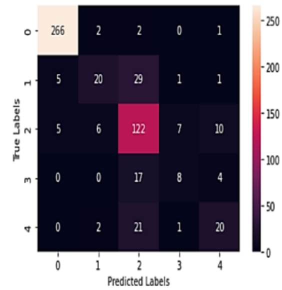


Fig.5(c) Confusion Matrix of ResNet50

Fig.5 Results of using ResNet-50 model

We can observe high Kappa Score values of ResNet50 in training are high, but in validation, Kappa Scores are greatly fluctuating and are highly unstable. This shows that ResNet50 is less efficient than DenseNet121, and suggests overfitting. By observing the confusion matrix of ResNet50, as the total instances are 550 and correctly predicted instances are 436, the authors got an accuracy of 79.27%.

4.5 Inception-V3

Convolutional filters with various receptive field sizes are used in the Inception-V3 CNN design, allowing the network to efficiently collect both local and global data. It is 48 layers deep.

Figure 6(a) illustrates the Inception-V3 model parameters, 6(b) illustrates the kappa score and 6(c) depicts the Confusion Matrix.

Total params: 21,813,029

Trainable params: 21,778,597

Non-trainable params: 34,432

Fig.6(a) Inception-V3 model Summary

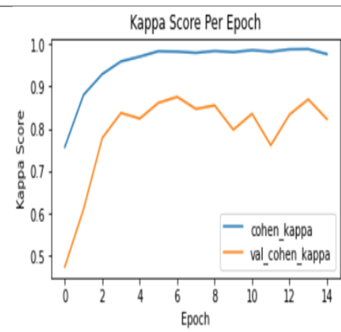


Fig.6(b) Kappa Score Per Epoch of Inception-V3

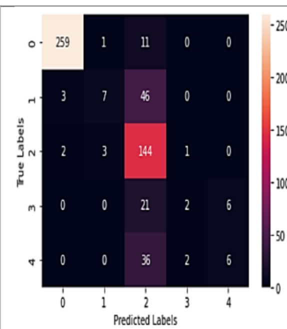


Fig.6(c) Confusion Matrix of Inception-V3

Fig.6 Results on using Inception-V3 model

We can observe that Kappa Scores of Inception-V3 in validation are slightly fluctuating for each epoch, but are higher than DenseNet121. This shows that Inception-V3 is more efficient than DenseNet121. By observing the confusion matrix of Inception-V3, as the total instances are 550 and correctly predicted instances are 418, the authors got an accuracy of 76.00%.

4.6 EfficientNet B3

EfficientNet B3 is a scalable CNN architecture that achieves cutting-edge performance while maintaining computational efficiency. Model depth, width, and resolution are balanced using a compound scaling process, producing extremely accurate and precise models.
Figure 7(a) illustrates the EfficientNet-B3 model parameters,7(b) illustrates the kappa score and 7(c) depicts the Confusion Matrix.

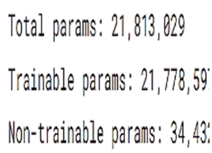


Fig.7(a) EfficientNet-B3 model Summary

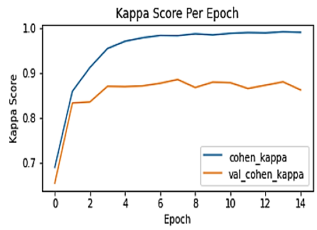


Fig.7(b) Kappa Score Per Epoch of EfficientNet-B3

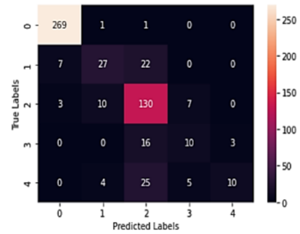


Fig.7(c) Confusion Matrix of EfficientNet-B3

Fig.7 Results of using EfficientNet-B3 model

We can observe that the Kappa Scores of EfficientNet-B3 are significantly high in the training dataset. In the validation dataset, Kappa scores are high and increasing for each epoch, with minimal fluctuations. This shows that EfficientNet-B3 is the most efficient of all the compared algorithms.

By observing the confusion matrix of EfficientNet-B3, as the total instances are 550 and correctly predicted instances are 456, the authors got an accuracy of 82.90%.

For all the algorithms, the authors use the activation function “softmax” to the output layer, to obtain the probability distribution over multiple classes.

Table 1. Comparison of all Models

S.No	Model	Validation loss	Validation Kappa Score	Accuracy
1	Simple CNN	1.0031180381774902	0.6567116975784302	53.63%.
2	VGG16	0.7033660411834717	0.735814094543457	74.18%
3	DenseNet121	0.5636863708496094	0.8736634850502014	68.18%
4	ResNet50	0.7095967531204224	0.8481181859970093	79.27%.
5	Inception-V3	0.848162055015564	0.875112771987915	76.00%
6	EfficientNet-B3	0.538705050945282	0.8848318457603455	82.90%.

From the tabulated values in Table 1, the authors and readers can observe that EfficientNet-B3 has the highest validation Kappa Score and lowest validation loss.

5 Conclusion

This paper presented a comprehensive analysis of various machine learning models for the detection of diabetic retinopathy. The models were evaluated based on their performance in terms of cross-entropy loss and Cohen's kappa score. Among the evaluated models, the EfficientNet-B3 model emerged as the most efficient and effective approach.The EfficientNet-B3 model consistently exhibited low cross-entropy loss, indicating its capability to minimize errors during training. Additionally, it achieved a high Cohen's kappa score, which reflects its ability to accurately classify diabetic retinopathy cases. The superior performance of the EfficientNet-B3 model suggests its suitability for automated diabetic retinopathy detection, offering promising prospects for timely intervention and treatment.The findings of this study underscore the significance of leveraging advanced machine learning techniques, such as EfficientNet-B3, for diabetic retinopathy recognition. The model's exceptional performance in terms of low loss and high kappa score demonstrates its potential to improve diagnostic accuracy and facilitate efficient screening processes. Further research and real-world implementation of the EfficientNet-B3 model is recommended to fully realize its benefits in clinical settings and contribute to the effective management of diabetic retinopathy.

References

1. Alabdulwahhab KM, Sami W, Mehmood T, Meo SA, Alasbali TA, Alwadani FA. *Automated detection of diabetic retinopathy using machine learning classifiers*. Eur Rev Med Pharmacol Sci. 2021 Jan;25(2):583-590, (2021).
2. Das D, Biswas SK, Bandyopadhyay S. . Multimed Tools Appl. **29,**1-59, (2022) doi: 10.1007/s11042-022-14165-4. Epub ahead of print. PMID: 36467440; PMCID: PMC9708148.
3. Moe Myint Thu, Set SetSwe and Tin ThandarZaw, Security System with PIR Sensor(Iconic Research and Engineering Journals, Volume 4, Issue 6, 2020)
4. Matla. Elimarteena, V Siva Nagaraju, *IOT Based Home Security System with PIR and Ultrasonic Detection Using Arduino*, in Proceedings of the 5th International Conference on Electronics, Communication and Aerospace Technology, ICECA, 2-4 December 2021, Coimbatore, India (2021).
5. Muhammad Zeeshan Saeed, Raja Raheel Ahmed, Omar Bin Samin, and Nusrat Ali, *IoT-based Smart Security System using PIR and Microwave Sensors*, in Proceedings of the 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics, MACS, 14-15 December 2019, Karachi, Pakistan (2019).
6. S. Mohammadian, A. Karsaz and Y. M. Roshan, "*A comparative analysis of classification algorithms in diabetic retinopathy screening,*" 2017 7th International Conference on Computer and Knowledge Engineering (ICCKE), Mashhad, Iran,84-89, (2017). doi: 10.1109/ICCKE.2017.8167934.
7. K Madhavi , "*Algorithm Selection and Model Evaluation in Application Design Using Machine Learning*", Lecture Notes in Computer Science. 12081 LNCS, pp. 175-195, 2020.
8. Karanam Madhavi et.al, "*A Survey on Cardivascular Prediction using Variant Machine learning Solutions*." E3S Web of Conferences 309, 01042, ICMED 2021, (2021)..
9. Ramesh, G., Reddy, K.S.S., Ramu, G., Reddy, Y.C.A.P., Somasekar, J. *An Empirical Study on Discovering Software Bugs Using Machine Learning Techniques.* In: Buyya, R., Hernandez, S.M., Kovvur, R.M.R., Sarma, T.H. (eds) Computational Intelligence and Data Analytics. Lecture Notes on Data Engineering and Communications Technologies, vol 142. Springer, Singapore, (2023).

10. V. Tejaswini Priyanka, Y. Reshma Reddy, D. Vajja, G. Ramesh and S. Gomathy (2023). *A Novel Emotion based Music Recommendation System using CNN*. 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 592-596, (2023). doi: 10.1109/ICICCS56967.2023.10142330.
11. Chandrika Lingala, and Karanam Madhavi, "*A Hybrid Framework for Heart Disease Prediction Using Machine Learning Algorithms* ", E3S Web of Conferences 309, 01043 , ICMED 2021, (2021). https://doi.org/10.1051/) e3sconf/202130901043.