Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

Hybrid Transfer Learning Model Combining Xception and NASNetMobile for Diabetic Retinopathy Detection

Hiba Kahdum Dishar [1]
[1]Dhi Qar Education Directorate (Almotfuqat Secondary School, Assistant Lecturer, Postal Code, Iraq).
Email: hibakd@gmail.com

Article's Information

Received: 30.09.2025
Accepted: 10.01.2026
Published: 31.03.2026

Keywords:
Convolutional Neural Network
Image Classification
Transfer Learning
Xception
NASNetMobile

Abstract

Convolutional Neural Networks (CNNs) are sole of the various forms of deep learning that have become effective tools in computer vision. Machine learning has been transformed by Deep Neural Networks (DNNs) with many parameters; their impact is particularly evident in network architecture. CNNs are very good at image classification problems because of their ability to concentrate on objects in images and extract information using spatial relationships. In this work, two robust architectures Xception and NASNetMobile are integrated into a unique CNN model for image categorization based on transfer learning. Using Xception and NASNetMobile, the model uses photographs with specified dimensions, often known as "best windowing of images," as input to categorize the images into two groups. To avert overfitting issues in CNN, a dropout layer is introduced after the outputs of these designs are concatenated using a concatenate layer. This research evaluates the proposed model for an exceptionally challenging dataset associated with diabetic retinal disease. The "Diabetic Retinopathy 224*224 Grayscale images" dataset, which is part of the "APTOS 2019 Blindness Detection" dataset on Kaggle, has 3662 images, of which 1875 show abnormal cases and the residual 1805 show normal instances. With an accuracy of 97.50%, precision of 96.39%, recall of 98.64%, and F1-score of 97.36%, the model fared extremely well in this test.

https://doi.org/10.46649/fjiece.v5.1.8a.31.3.2026
*Corresponding author: hibakd@gmail.com

1. INTRODUCTION

CNN is a particular genre of deep neural network that was developed primarily to process and evaluate visual data, including images and videos. Since CNNs are modelled after the human visual system, they constitute the base of many advanced machine learning models in computer sight issues [1]. CNNs utilize a specialized layer called a convolutional layer to preserve the data's spatial structure, in contrast to traditional neural networks that process input data as an irregular vector. By applying convolutional filters or kernels to the input data, this layer aids in the local pattern or feature detection of forms, edges, and textures [2]. The best efforts in the domain of artificial intelligence have long been hampered by issues that deep learning is trying to resolve. Its exceptional ability to uncover intricate patterns in high-dimensional data can be beneficial in many scientific, industrial, and political realms [3]. One of CNNs' key benefits is that it can automatically learn feature structures from raw pixel values, eliminating the need for human recognition of features. CNNs have been highly effective and successful for a wide range of computer vision applications, such as object identification, image segmentation, and image classification. Recent advances in deep learning and the accessibility of massive datasets have made CNNs incredibly successful in an assortment of applying, altering sectors inclusive healthcare, automotive, entertainment, and beyond [2].

Diabetic retinopathy (DR) is a dangerous sickness that may lead to blindness without any warning indicators. The lead of the illness from its early to its severe phases must thus be regularly screened for and audited. The felled condition known as diabetic retinopathy (DR) can rice blindness without alarms signs. The progression of the disease from its early to its severe phases must thus be regularly screened for and audited [4]. Deep learning approaches are not the only image processing technologies that have been developed. In image processing methods, intricate characteristics are manually recognized Deep learning approaches are not the only image processing technologies that have been developed. In image processing methods, intricate characteristics are manually detected [5]. Previous studies on diabetic retinopathy (DR) detection have primarily focused on using individual deep learning or transfer learning models like VGG, ResNet, Inception, or DenseNet architectures. Despite their remarkable accuracy, these models frequently have drawbacks, including overfitting on short medical datasets, poor generalization across various imaging situations, and a lack of feature variety for intricate patterns of retinal lesions. Also, the majority of current methods only use one pre-trained network, which limits their capacity to recognize complementary spatial and hierarchical patterns found in retinal images. The synergistic potential of hybrid deep transfer learning models that integrate the superiority of several architectures has not received much attention. To label these gaps, to improve feature representation and classification performance, this study suggests a unique hybrid transfer learning framework that combines Xception and NASNetMobile. It does this by utilizing Xception effective depthwise separable convolutions and NASNetMobile architecture search-based optimization. In comparison to conventional single-model or ensemble approaches, the combined technique seeks to boost detection accuracy, resilience, and computational efficiency, offering a more dependable and scalable automated diabetic retinopathy screening solution.

2. RELATED WORKS

Eman Abdel Maksoud et al. [4] institute a groundbreaking hybrid deep learning mechanism named the E-DenseNet model, prepared for verify various Diabetic Retinopathy (DR) grades. This innovative method melds the powers of EyeNet and DenseNet through transfer learning, achieving remarkable advantages by customizing EyeNet and incorporating dense bricks. The suggest CAD system exhibits exceptional performance, needs the least amount of memory and couching time. Notably, the E-DenseNet-based CAD system showcases promising results on two benchmark datasets, the EyePACS dataset and APTOS dataset. It establishes an impressive average accuracy (ACC) of 91.6%, a Dice identity point of 92.45%, and a Kappa score of 0.883, demonstrating its efficiency in precise DR-grade diagnosis.

Md. Robiul Islam et al. [5] conduct a deep learning model that employ the VGG16 model's transfer learning. Next, it uses a cutting-edge color version preprocessing technique that boosts the realism of the photos by softening them. appoint image smoothing (blurring) techniques proved beneficial in reducing noise, and the Gaussian filter was chosen as the method of choice for image smoothing. To achieve this, they applied the Gaussian filter to the images, creating a smoothed version. Then, they obtained a mask by deduct the polished image from the native image, thereby preserving the elevated-frequency element that were subsequently abandon out by the smoothing filter. Involve this technique to the new Kaggle dataset "APTOS 2019 Blindness Detection" produced an average accuracy of 91.32% and successfully reduced the training time. In addition, to address the long-term issue of overfitting, they employed Stratified K-fold cross-validation to ensure a more robust and generalized model performance. Niloy Sikder et al. [6] took a holistic strategy by combining several different kinds of image processing with two different kinds of feature extraction and a feature selection strategy. Utilizing the APTOS 2019 Blindness Detection (BD) dataset, their method achieved impressive classification accuracy values of 94.20% with a marginal mistake of 0.32%, as well as an F-measure of 93.51% with a marginal error of 0.5%. These outstanding outcomes highlight the high execution and dependability of the recommended technique. Furthermore, the method's

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

robustness and dependability were illustrated through the evaluation of its additional parameters. As a noble tool for mass retinal inspection, this strategy has the potential to dramatically lower the prevalence of diabetic retinopathy (DR)-related blurred vision. The recommended approach helps to improve care and decrease vision-related consequences by facilitating early and accurate detection of DR. Rahman et al. [7] There are 3662 fundus images from the Asia Pacific Tele-Ophthalmology Society (APTOS). They carefully selected features using transfer learning, first classifying DR into five levels, Afterward, it was streamlined to a binary format using three promising machine learning models: Histogram Gradient Boosting (HGB**),** k-nearest neighbors (KNN), and support vector machine (SVM) Class 1–4: No DR and DR, With an astounding accuracy of 96.9%, the SVM model beat the other methods in the writings using the same dataset, according to the results, while the KNN and HGB models both achieved 95.6%. Wan et al [8] apply (CNNs) power to DR detection, including three main difficult tasks: detection, segmentation, and classification. use AlexNet, VggNet, GoogleNet, and Res Net in conjunction with transfer learning and hyper-parameter tuning, and estimate how competently these models implement in the DR image classification. Use the Kaggle platform, which is openly accessible, to train these models. The outcome view that CNNs and transfer learning are more accurate at classifying DR images, with the best classification accuracy being 95.68%. Skouta et al [9], based on the ophthalmologist's experience, they have compiled representative Diabetic Retinopathy (DR) photos into five categories. A variety of deep Convolutional Neural Network techniques have been used to classify DR stages. The precision results of Inception Net V3, AlexNet 37.43%, VGG16 50.03%, and Inception Net V3 63.23% are state-of-the-art.

Sanjana et al [10], in their study, 1115 retinal fundus images from two public databases are used. A binary categorization of DR was suggested in their study using five transfer learning models: Xception, InceptionResNetV2, MobileNetV2, DenseNet121, and NASNetMobile. These models reach the elevated validation precision, with respective values of 86.25%, 96.25%, 93.75%, 81.25%, and 80.00%.

Table. 1: summary of related work

References	Year	Technology	Dataset	Accuracy
Eman AbdelMaksoud et al.[4]	2020	The suggested method is a transfer learning hybrid model of EyeNet and DenseNet.	APTOS 2019 Blindness Detection and EyePACS dataset	86.5%,96.8%
Robiul Islam et al. [5]	2020	Developed a unique color version pretreatment technique and then used model-based transfer learning to train on the VGG16 model.	APTOS 2019 Blindness Detection	0.9132%
Niloy Sikder et al.[6]	2021	Utilized method uses a plethora of image processing methods, including two feature extraction strategies and a single feature selection method.	APTOS 2019 BD	94.20%
Atta Rahman et al. [7]	2024	Use three promising three machine learning models: (HGB), KNN, and (SVM) after carefully choosing features with transfer learning.	APTOS 2019 BD	96.9%, 95.6%

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

3. MATERIALS AND METHODS

3.1Transfer Learning

Transfer learning is a key downstream application of acquired image classification systems. In some real-world machine learning scenarios, this assumption does not hold [11]. In certain situations, collecting training data can be difficult or costly. Therefore, it is crucial to create high-achieving students who are instructed utilizing more readily available facts from multiple domains; this approach is known as transfer learning [2]. Utilizing prior expertise of related tasks to improve execution on a new task is the aim of transfer learning (TL) with convolutional neural networks. In addition to saving time and hardware resources, it has resolved the problem of data shortages, which has greatly advanced medical image analysis. But in most instances, transfer learning has been intentionally confused [10]. Text classification, picture classification, and sensor-network-based localization are just a few examples of the small-scale, narrowly varied applications that have mostly used transfer learning approaches. Techniques for transfer learning will be widely employed in the future to address other difficult applications, like logical reasoning, social network analysis, and video categorization [11][12]. An abundant supply of provenance data but a restricted quantity of target domain data can remarkably progress the model's rendering through transfer learning. Using transferred weights from a less significant network is preferable to using reckless weights [13].

3.2. Nasnet Mobile Model

Three phases of CNN development have occurred since AlexNet received international recognition. These phases are based on the ideas of profound is superior, Architecture Engineering, and Auto ML. Nasnet, an acronym for Neural Search Architecture (NAS) Network, is an expandable CNN model [14]. There are 12 cells, 5.3 million parameters, and 564 million multiplied collect in the mobile version of Nasnet called Nasnet Mobile [15]. NASNetMobile autonomously designs a CNN architecture optimized for mobile devices. This architecture comprises double convolutional layers with different filter sizes, put up with by max-pooling layers to down sample the feature maps. NASNetMobile achieves high accuracy, making it suitable for deployment on appliance with finite resources such as smartphones and embedded systems [16].

3.3. Xception Model

In 2017, Xception (Chollet, 2017) was put forth as a CNN architecture. As extreme inception, it was first presented. There are 36 layers of convolutions in Xception. Three flows compose this system. The initial flow, which is the first flow, has pooling layers, separable convolution, and convolution. The middle flow, which has separable convolution layers, is the second flow. Additionally, the middle proceed occurs eight times. The last flow comes in at number three. It is the final flow and produces the dense layer as a result [17]. Xception's depth-wise separable convolution divides the learning of space-wise and channel-wise characteristics. Moreover, by establishing a shortcut within the sequential network, the residual relevance helps to address the issues of representational bottlenecks and disappearing gradients [18]. With fewer overfitting issues, the Xception has demonstrated increased resilience and generalization abilities [19].

3.4. MATERIALS

The illness famed as diabetic retinopathy (DR) develops as a consequence of diabetes. Because it is frequently overlooked and can issue blindness if not identified soon, it is especially harmful. Although the obvious significance and severity of this sickness, there is still no accurate method for DR early identification. Convolutional neural networks (CNNs), which gained popularity in the medical imaging

Al-Furat Journal of Innovations in Electronics and Computer
Engineering (FJIECE)
ISSN -2708-3985

industry because they can be successfully incorporated into different systems in a way that greatly enhances performance, are one example of deep learning that can be used to create such a system [20][21].

The objectives of this work are to use transfer learning to push the limits of severity categorization from photographs, specifically focusing on medical images of diabetic retinopathy. The ultimate goal is to achieve remarkable accuracy while mitigating the issue of overfitting. The early disclosure of diseases associated with diabetic retinopathy is crucial to improving the lives of individuals. Therefore, we present a new model that blend the advantages of the Xception and NASNETMobile architectures in order to achieve these significant aims. This innovative fusion improves the model's performance by transfer learning using pre-trained deep CNNs and precisely calibrated hyperparameters. Contrast to models trained with randomly initialized weights, these techniques are more effective and are well-established in the field of medical image processing. Overall, our suggest classification model outperforms conventional deep CNN models in terms of accuracy and effectiveness, evidence its superiority with confidence. Proposed approach shown below in fig. 1.

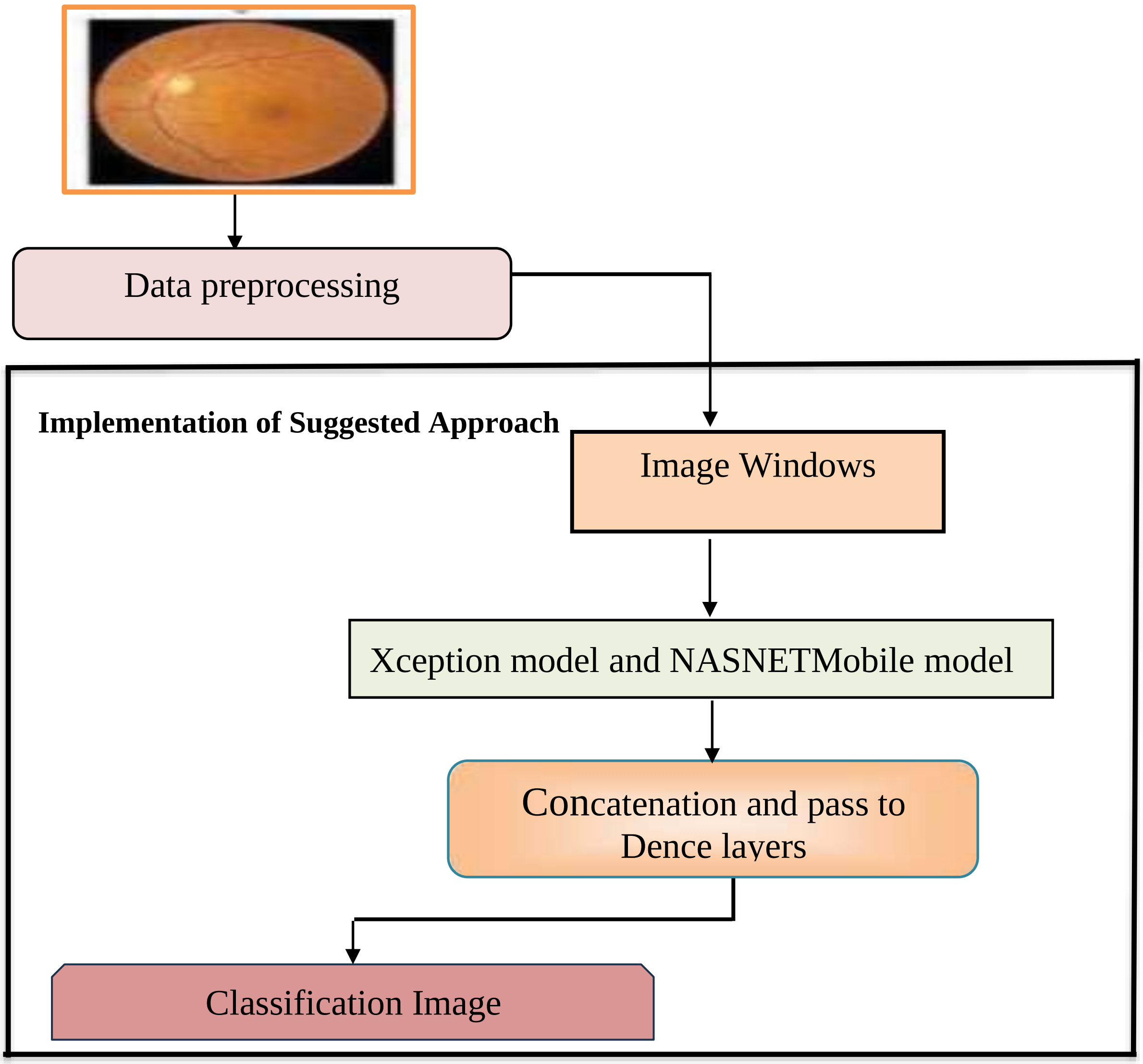

Figure. 1: Overview of suggested framework

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

3.5. METHODS

In this research, an approach to enhance CNN for prediction more precisely was suggested. The proposed CNN model combines two pretrained models, namely the Xception model and the NASNetMobile model, to achieve improved classification precision. To begin, assemble the data with the Diabetic Retinopathy dataset, ensuring a comprehensive representation of medical images. This dataset was carefully chosen to consist of a variety of medical conditions and image quality alteration. Next, resize the images to a standardized dimension of (224, 224, 3) to facilitate compatibility with the pre-trained models. Resizing the images ensures that they are uniform in size and can be efficiently processed by the proposed model. additionally, introduces a set of window sizes for image analysis, including (214, 214), (210,210), (200, 200), and (185, 185). By experimenting with different window sizes, the aim is to determine the optimal window that captures the generality pertinent and informative features for input to the proposed model, in this dataset found that the (200*200) best window achieved the highest accuracy, because this window contains the best view of the patient’s image. Provide a visual depiction to better illustrate the workflow and all the phases involved. 60% of the training set is split up, 20% is used for validation, and 20% is used for testing.

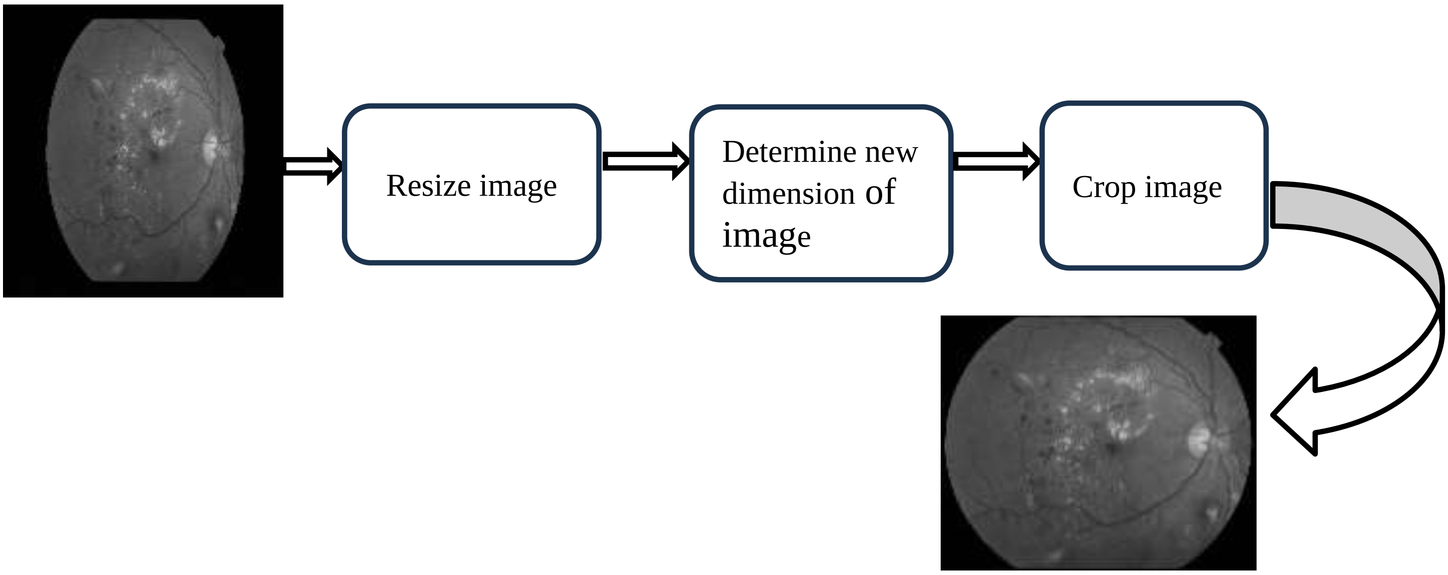

Figure. 2: Explain phases of image windows

4. RESULTS AND DISCUSSION

4.1. RESULTS

This part is related to the testing combination model with different datasets. The testing included image windowing with different states, first, the original image was tested, then different size windows were applied. With each state the model was experimented. Table (1) for (Diabetic Retinopathy 224*224 Grayscale images) dataset. Also, these details explain in Fig. 3 and table.1 consequently.

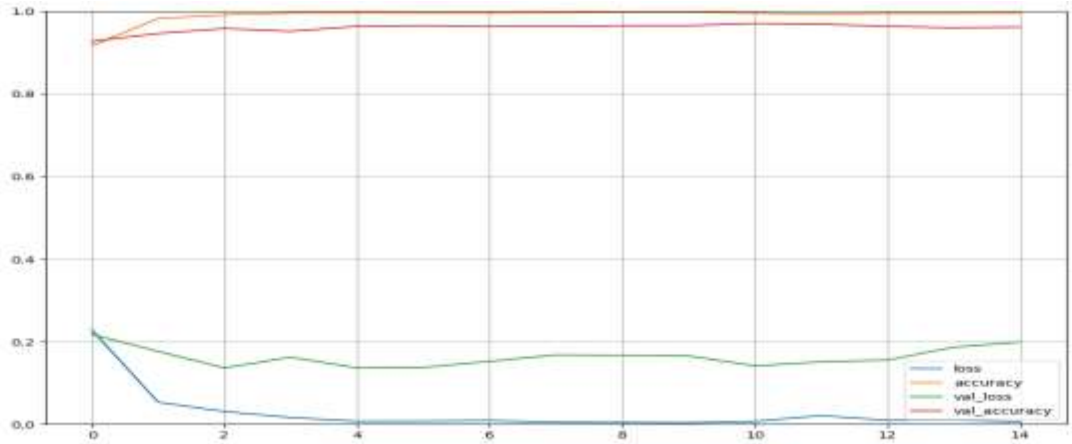

Figure. 3: Accuracy and loss function of the proposed model in Diabetic Retinopathy 224*224 Grayscale images data window (200,200)

Table. 2: Performance a proposed model on Diabetic Retinopathy Grayscale images dataset

Windows of image	Accuracy%	Recall%	F1-score%	Precision%
Original window (224,224)	95.83	98.08	95.64	93.61
Window (214,214)	96.67	97.73	96.50	95.56
Window (210,210)	96.81	96.68	96.56	96.64
Window (200,200)	**97.50**	**98.64**	**97.36**	**96.39**
Window (185,185)	97.08	98.57	96.66	95.56

4.2. DISCUSSION

Retinal pathology is often evaluated and detected by the examination of various imaging modalities that enable structural analysis [22]. Patients will benefit from early diabetes research and fewer serious health consequences like vision impairment [23]. In this work, suggested a novel CNN architecture for diabetic retinal image categorization that is based on transfer learning. To establish a neural network using parallel models, two pre-trained models Xception and NASNetMobile were combined to build a model. An input layer is then generated to receive the picture data after the input image has been cropped and shrunk in various dimensions to reach a suitable area of the image to boost the categorization method. The input layer is made available across each model when Xception and NASNetMobile instances are created without their top layers.

5. CONCLUSIONS

To prevent overfitting, using a comprehensive average pooling layer, the outcomes of jointly models are concatenated and then added as an input to a dropout layer. The accuracy of the final model is enhanced by combining the two pre-trained models. Deep learning technicality has the capability to develop the precision and dependability of medical image interpretation, as demonstrated by the model's effectiveness. Even though my study has shown how successful the suggested CNN architecture is, more investigation is required to completely grasp the possibilities of deep learning techniques for processing medical images. Bigger dataset sizes, various imaging modalities, the application of certain optimization strategies to improve picture classification, and the influence of various high-level variables on model performance should all be the subject of future research. The hybrid model (Xception and NASNetMobile) performs better than individual models and current cutting-edge techniques, according to the experimental data achieving an accuracy of 97.50%, a precision of 96.39%, a recall of 98.64%, and an F1-score of 97.36%. Our aim in the future is to train our models on a large sample. Applying these models to various datasets related to diabetic retinopathy is what we will try to do. Additionally, we will attempt to boost the accuracy of the model. The study will be expanded to include several classifications of diabetic retinopathy, and latest elements can be compiled to improve categorization and accuracy.

REFERENCES

[1] Taye, Mohammad Mustafa. "Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions." *Computation* 11.3 (2023): 52.
[2] Kim, Hee E., et al. "Transfer learning for medical image classification: a literature review." *BMC medical imaging* 22.1 (2022): 69.
[3] T. He, Z. Zhang, H. Zhang, Z. Zhang, J. Xie, and M. Li, “Bag of tricks for image classification with convolutional neural networks,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 558–567.
[4] E. AbdelMaksoud, S. Barakat, and M. Elmogy, “Diabetic retinopathy grading based on a hybrid deep learning model,” in *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*, IEEE, 2020, pp. 1–6.
[5] M. R. Islam, M. A. M. Hasan, and A. Sayeed, “Transfer learning based diabetic retinopathy detection with a novel preprocessed layer,” in *2020 IEEE Region 10 Symposium (TENSYMP)*, IEEE, 2020, pp. 888–891.
[6] N. Sikder, M. Masud, A. K. Bairagi, A. S. M. Arif, A.-A. Nahid, and H. A. Alhumyani, “Severity classification of diabetic retinopathy using an ensemble learning algorithm through analyzing retinal images,” *Symmetry (Basel)*, vol. 13, no. 4, p. 670, 2021.
[7] Rahman, Atta, et al. "Diabetic Retinopathy Detection: A Hybrid Intelligent Approach." Computers, Materials & Continua 80.3 (2024).

[8] Wan, Shaohua, Yan Liang, and Yin Zhang. "Deep convolutional neural networks for diabetic retinopathy detection by image classification." *Computers & Electrical Engineering* 72 (2018): 274-282
[9] Skouta, Ayoub, et al. "Automated binary classification of diabetic retinopathy by convolutional neural networks." *Advances on Smart and Soft Computing: Proceedings of ICACIn 2020*. Singapore: Springer Singapore, 2020. 177-187.
[10] Sanjana, Silvia, Nazmus Shakib Shadin, and Mayisha Farzana. "Automated diabetic retinopathy detection using transfer learning models." *2021 5th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*. IEEE, 2021.
[11] Weiss, Karl, Taghi M. Khoshgoftaar, and Ding Ding Wang. "A survey of transfer learning." *Journal of Big data* 3.1 (2016): 9.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

[12] S. Niu, Y. Liu, J. Wang, and H. Song, “A decade survey of transfer learning (2010–2020),” *IEEE Transactions on Artificial Intelligence*, vol. 1, no. 2, pp. 151–166, 2020.

[13] Yosinski, Jason, et al. "How transferable are features in deep neural networks." *Advances in neural information processing systems* 27 (2014).

[14] Ghazal, Mohammed, et al. "Accurate detection of non-proliferative diabetic retinopathy in optical coherence tomography images using convolutional neural networks." *IEEE Access* 8 (2020): 34387-34397.

[15] Reddy, Narsi, Ajita Rattani, and Reza Derakhshani. "Comparison of deep learning models for biometric-based mobile user authentication." *2018 IEEE 9th international conference on biometrics theory, applications and systems (BTAS)*. IEEE, 2018.

[16] Ameer, Ashwaq Waleed Abdul, Pedram Salehpour, and Mohammad Asadpour. "Deep transfer learning for lip reading based on NASNetMobile pretrained model in wild dataset." *IEEE Access* (2024).

[17] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[18] He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[19] Wu, Xizhi, et al. "An xception based convolutional neural network for scene image classification with transfer learning." *2020 2nd international conference on information technology and computer application (ITCA)*. IEEE, 2020.

[20] Doshi, Darshit, et al. "Diabetic retinopathy detection using deep convolutional neural networks." *2016 international conference on computing, analytics and security trends (CAST)*. IEEE, 2016.

[21] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[22] Kumar, Athota Manoj, Atchukola Sai Gopi Kanna, and Ramya G. Franklin. "Diabetic retinopathy detection." *International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy*. Singapore: Springer Nature Singapore, 2020.

[23] Martinez-Murcia, Francisco J., et al. "Deep residual transfer learning for automatic diagnosis and grading of diabetic retinopathy." Neurocomputing 452 (2021): 424-434.

[24] https://www.kaggle.com/datasets/sovitrath/diabetic-retinopathy-224x224-grayscale-images