JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

DENSE RESIDUAL BRIDGE UNET FOR RETINAL VESSEL EXTRACTION

[1]Marada Amrutha Reddy, [2]Dr. Anuradha Padala

[1]M.Tech Scholar, [2]Associate Professor

[1] Dept of Computer Science and Engineering,

[1] GITAM University, Visakhapatnam, India

Abstract : Digital photography provides prominent diagnostic information used to identify and treat diseases. Fundus imaging is performed to acquire the two-dimensional representation of an eye's optic nerve, retinal tissue, and retinal vessels. The imaging captures retinal vessels, which are used to diagnose ophthalmic diseases. Retinal vessel segmentation helps visualize and identify the abnormalities in vessels. Proper segmentation helps make an accurate diagnosis and early surgical decisions and avoids the risks of blindness. The tiny and curved lines of vessels make the process of segmentation challenging. Although different state-of-art segmentation models are built using a sequence of encoder-decoder deep learning techniques, they fuse semantically dissimilar feature maps. The skip connections in UNet provide feature maps. The thinner vessels show a low contrast against the background. This research aims to improve the contrast between retinal vessels for better segmentation. A modified architecture using UNet is used to precisely localize the vessels. Dense blocks replace the convolutional blocks. The encoder-decoder ladder is bridged using a residual block. It is observed that the proposed architecture improves sensitivity, reaching a value of 95.7 percent.

Keywords **- Image segmentation; Deep Learning; Retinal vessel segmentation; Unet; Dense layer; Residual Block**

I. INTRODUCTION

The task of segmenting vessels in retina images is called retina vascular segmentation. Blood vessel segmentation is an essential part of color retinal fundus image processing for monitoring and detecting diabetic retinopathy, hypertensive retinopathy, age-related macular degeneration, and arteriosclerotic retinopathy. The retinal vessels supply the retina with blood, oxygen, and nutrients. Retinal vascular patterns are as different and unique as fingerprints, so that that retina scans may be utilized as a high-tech identification technique. Abnormalities caused to the vessels within the retina obstruct circulation and result in light blockage and sudden vision loss. Qualitative and quantitative retinal observations can detect imminent retinal, systemic, neurologic, or cerebrovascular illness.

Changes in the retinal vasculature and retinopathy can be monitored via fundus photography. The method of taking serial images of the eye's interior through the pupil is fundus photography. The optic disc, retina, and lens are examined with a fundus camera. Known as a retinal camera, it is a specialized low-power microscope with an attached camera used to picture the retina, retinal vasculature, optic disc, macula, and posterior pole of the eye (i.e., the fundus). It allows comparison of the retinal vasculature in each eye side by side, which is a valuable predictor of disease worsening or resolving.

The pictures that arise might be photographic or digital, forming part of the member's medical file. Unless it is essential for image capture or clinically contraindicated, fundus photos usually are taken through a dilated pupil to improve the quality of the photographic record. Fundus photography can document abnormalities related to eye disease processes or track disease progression. It is considered medically necessary for conditions like macular degeneration, retinal neoplasms, choroid disturbances, and diabetic retinopathy to diagnose glaucoma, multiple sclerosis, and other central nervous system abnormalities.

Retinal vessel segmentation demonstrates the retinal vasculature of the human eye. Accurate segmentation of retinal vascular structure helps diagnose diabetic retinopathy, retinal vascular occlusion, atherosclerosis, and hypertension. The retinal vessels comprise arteries and veins. The arteries and veins keep branching out until they form tiny thin vessels called capillaries. The arteries and veins are more prominent than the capillaries. Localizing and visualizing capillaries is challenging because the network structure is constricted, and the vessels are closely connected. The vessels are too tiny and curved, which makes the segmentation of the vessels challenging. The variations present in the shape and structure of vessels are challenging to localize. Visualizing the morphology of the retinal vessels is difficult due to vessels being too closely structured. Segmenting vessels is challenging because of the complex, tiny and curved lines. This research aims to enhance the extraction of retinal vessels for accurate detection of any blood clots and abnormalities. The fundus images used are taken from the (Digital Retinal Images for Vessel Extraction) DRIVE dataset. The dataset is thoroughly processed using image preprocessing techniques. Data augmentation is used to increase the number of samples. A custom Unet architecture is built using Tensorflow and Keras framework. The encoder and decoder are built using dense blocks, and

layers are bridged using a residual block. The model is trained on the augmented train dataset and is evaluated on the test set based on accuracy, sensitivity, and specificity.

II. RELATED WORK

Chen C et al. [1] presented a detailed overview of deep learning-based retinal blood vessel segmentation. Retinal vascular geometry reflects a patient's health condition and aids in diagnosing disorders such as diabetes and hypertension. This study looked at recent research on deep learning-based retinal vascular segmentation. They examined these recommended methodologies, particularly network designs, to determine the model trend. The study looked at the challenges and crucial features of using deep learning for retinal vascular segmentation and suggested future research topics.

Sun X et [2] suggested two novel data augmentation modules: channel-wise random Gamma correction and channel-wise random vascular augmentation. *Gamma correction* is an operation used to encode and decode luminance or tristimulus values. It is extensively employed in automated vessel segmentation systems as an image preprocessing step. Gamma correction was directly applied in the RGB (Red, Green, Blue) color space.

Khan T. M et al. [3] proposed an encoder-decoder architecture with spatial pyramid pooling modules to produce semantic segmentation. Shah S. A. A et al. [4] proposed a technique based on the Gabor wavelet and multi-scale line detector. The Gabor wavelet transform provides high-frequency precision in low frequencies, while it provides high spatial accuracy in high frequencies. It is used for edge augmentation. Following that, a line detector was used to the augmented picture to overcome the central reflex problem and improve vessel recognition.

Dash J et al. [5] studied three methods for segmenting blood vessels. The first method focused on contrast differences in big and thin blood vessels. The second technique uses a 2-D Gabor wavelet to improve the vascular pattern. In contrast, the third method employs a Star Networked Pixel Tracking Algorithm to eliminate noise aligned in a vessel format. These techniques segment blood arteries, making it easier to identify and treat various eye problems.

III. DEEP LEARNING

Deep learning models try to mimic or replicate the working of the human brain. A perceptron, a single neuron in a Neural Network, is a deep neural network [6]. It permits computational models made of multiple processing layers to learn data representations with multiple levels of abstraction [7]. Put differently, deep learning employs artificial neural networks to tackle challenges. An input layer, many hidden layers, and an output layer are all present in every artificial neural network. Each layer has nodes, and the nodes between the layers are all connected and have weights. An illustration of an artificial neural network is shown in the diagram below:

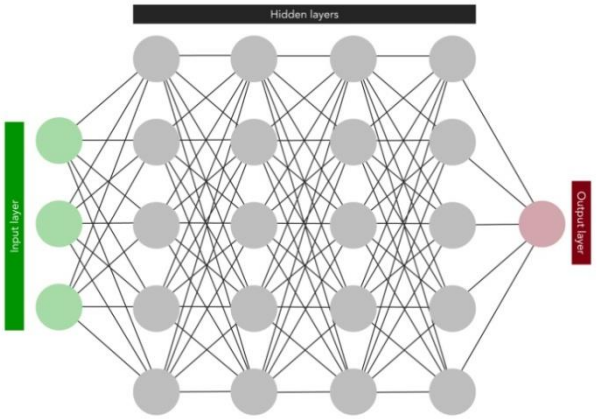

Fig. 1. Structure of an artificial neural network

The three most common neural networks are feed-forward, convolutional, and recurrent neural networks. Information moves in one direction along connected pathways in a feed-forward neural network, from the input layer to the final output layer through hidden layers.[8]. Image data is processed using convolutional neural networks. It was motivated by biological processes in which the connection network between neurons mirrors the animal visual cortex's architecture. [9]. All forms of learning, including supervised, unsupervised, and reinforcement learning, can benefit from deep learning models. Learning a mapping between a collection of input variables X and an output variable Y and using that mapping to predict the outputs for unseen data is referred to as supervised learning. [10]. Supervised learning includes regression and classification. Deep Learning is being used in various fields to extract characteristics from data and build intermediate representations. [11]. It may be used in almost any sector and helps get better outcomes.

IV. IMAGE SEGMENTATION

Splitting a digital image into several image segments, also known as image regions or image objects, is called image segmentation (sets of pixels). The purpose of segmentation is to make a picture more intelligible and straightforward to examine by simplifying and changing its representation. Objects and boundaries (lines, curves.) in pictures are often located via image segmentation. Image segmentation is giving a label to each pixel in an image so that pixels with the same label have similar features.

Image segmentation produces either a collection of segments that encompass the entire image or a set of contours taken from the image. Each pixel in an area is comparable in characteristics or calculated features, such as color, intensity, or texture.

4.1. Semantic Segmentation

Semantic image segmentation aims to assign a class to each picture pixel that represents anything. This assignment is known as a dense prediction since we predict every pixel in the image. The expected result is more than simply labels and bounding box

parameters in semantic segmentation. The output is a high-resolution picture (usually the same size as the input), with each pixel categorized into a different class. As a result, it is a pixel-by-pixel image categorization.

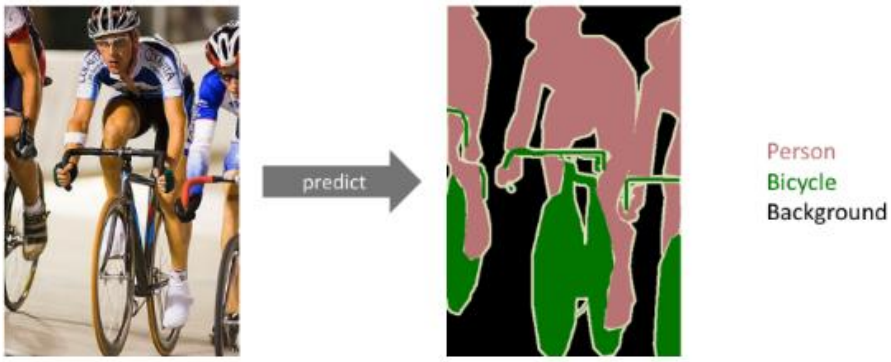

Fig. 2. Semantic Segmentation

V. RETINAL VESSEL SEGMENTATION

The study of vessels is vital for diagnosis, treatment planning, execution, and assessment of clinical results in various professions, including laryngology, neurosurgery, and ophthalmology; hence blood vessel segmentation is a hot issue in medical image analysis. The segmentation of the picture areas corresponding to the vasculature is the first stage in a retinal analysis pipeline. As some characteristics of retinal blood vessels need segmentation, it is the foundation of retinal fundus image analysis. Vessel segmentation is a binary classification issue in theory.

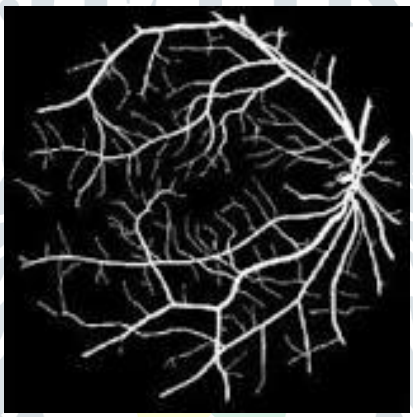

Fig. 3. Annotated Vasculature Provided by Experts

VI. DATASET

(Digital Retinal Images for Vessel Extraction) DRIVE Dataset is used. The images were used as part of a diabetic retinopathy screening program in the Netherlands. [12] It comprises 40 color fundus images, including seven abnormal pathology cases. The 40 photographs were divided evenly across the training and testing sets, yielding 20 images. A circular field of view (FOV) mask of roughly 540 pixels in diameter for each picture is present inside both sets. One manual segmentation by an ophthalmological specialist was applied to each picture inside the training set. Within the testing set, two distinct observers applied two different manual segmentations to each image, with the first observer segmentation being regarded as the ground truth for performance evaluation.

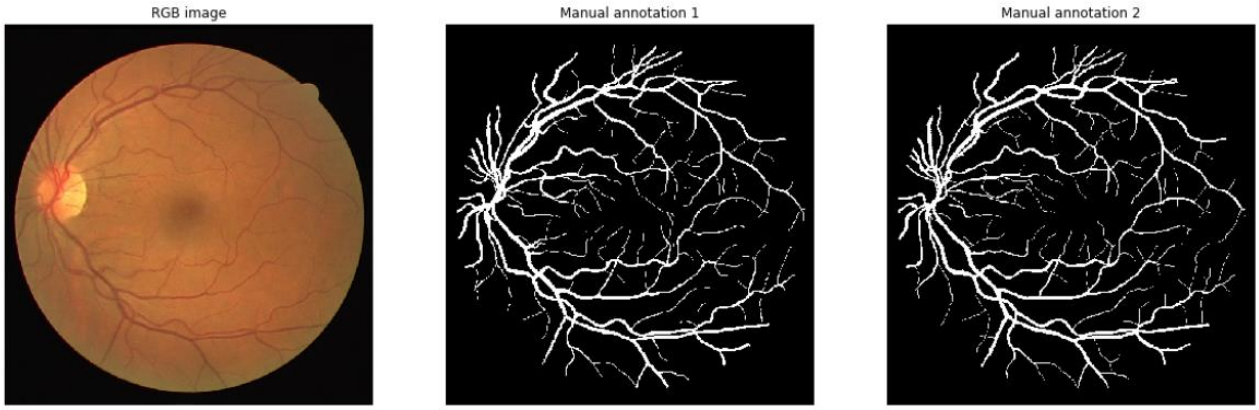

Fig. 4. DRIVE Dataset

VII. PROPOSED SYSTEM

There are 20 images in each of the dataset folders. i.e., training set and test set. Each dataset folder comprises RGB images and the mask images annotated by experts. The images of the training and test set are preprocessed. After preprocessing, the train images are augmented using the albumentations library, and an 80:20 train-validation split is performed. The test images are resized to 512x512 pixels. The model used for training is a custom architecture based on Unet.

7.1 UNet Architecture

UNet, which developed from the classic convolutional neural network, was built and used to analyze biological pictures for the first time in 2015. It can localize and discern boundaries because it performs classification on each pixel, resulting in the same input and output size. It is an encoder-decoder convolutional neural network created to segment biological pictures. Its architecture resembles the letter U when imagined, thus the name U-Net. Its architecture is divided into the contracting path on the left and the expanding way on the right. The contracting path's job is to capture context, while the expanding path's job is to help with exact localization.

Olaf Ronneberger et al. [13] created the UNET for Biomedical Image Segmentation. There are two paths in architecture. The contraction path (the left part, also known as the encoder) is the first path, and it is used to record the image's context. The encoder is simply a convolutional and maximum pooling layer stack. The symmetric expanding path (the right path, which is also known as the decoder) is the second way, and it is employed to achieve exact localization via transposed convolutions. As a result, it is an end-to-end, fully convolutional network (FCN). It only has Convolutional layers and no Dense layers, allowing it to accept images of any size.

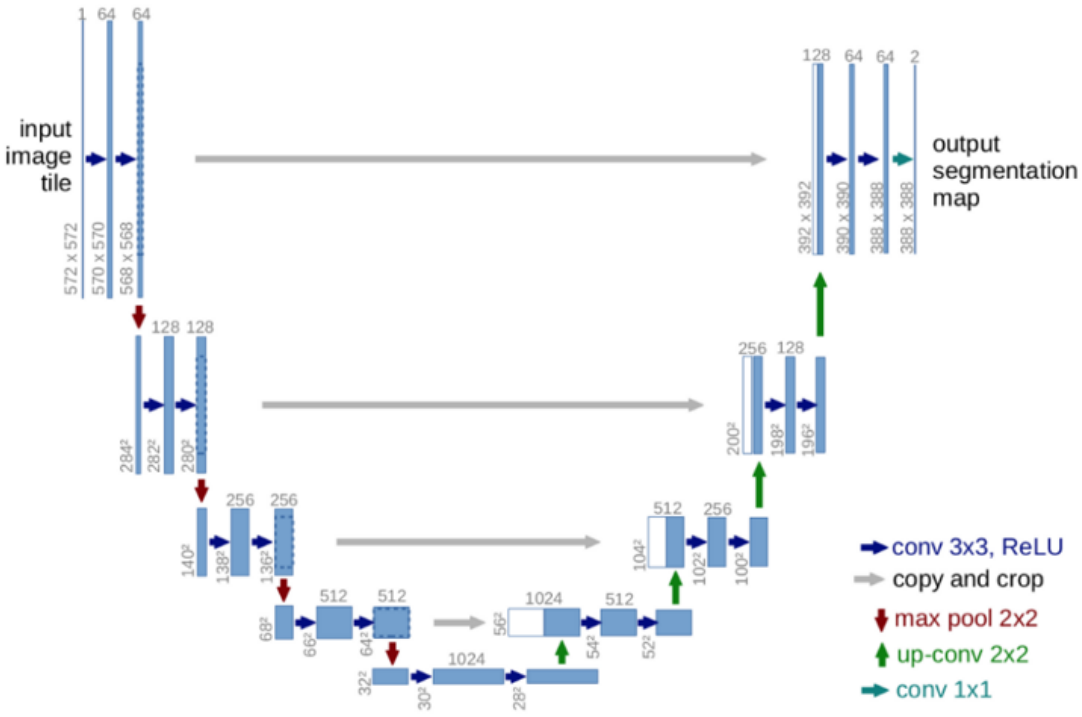

Fig. 5. UNet Architecture

7.2 Dense Residual Bridge Connected UNet

In the proposed system, the convolution layers in the contraction and the expansive path have been replaced by dense layers. A residual block replaces the convolutional bridge between the expansive and the contraction path.

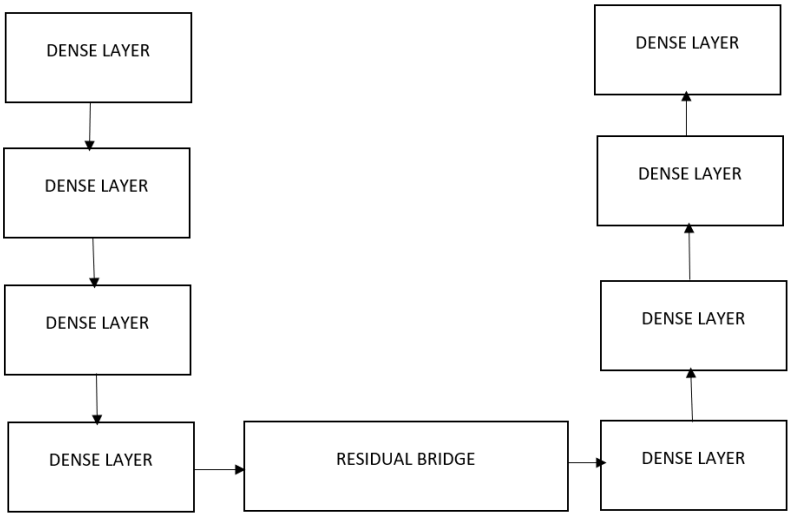

Fig. 5. Dense Residual Bridge Connected UNet

Unlike convolutional layers wherein the image matrix is multiplied with a kernel for feature extraction, dense layers use vector-matrix multiplication, leading to an efficiently retained feature extraction. A residual block is a collection of layers in which the output of one layer is taken and added to a layer more profound in the block. After that, the nonlinearity is applied by combining it with the output of the relevant layer in the main route. The shortcut or skip-connection is the name for this by-pass connection. The encoder-decoder ladder is bridged using a residual block to ensure proper mapping. Below is an illustration of the approach:

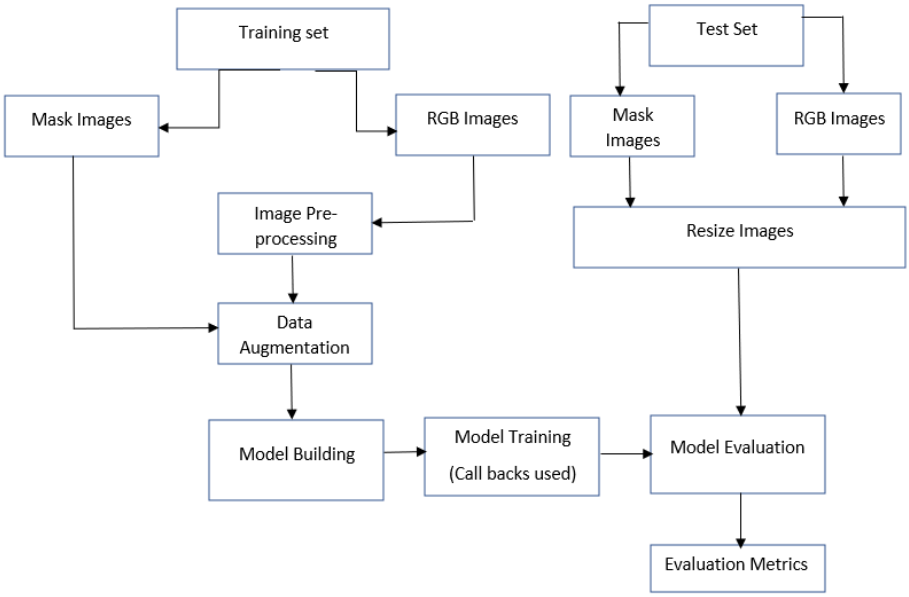

Fig. 6. Approach

VIII. EXPERIMENT

Google Drive is used to import the data into Google Colab. Keras and Tensorflow 2.8.0 were used to create the model's environment. The experiment was conducted on a Windows 11 PC with 16 GB RAM. Google Colab's hardware accelerator was employed to speed up the model training. GPU and TPU are the two types of hardware accelerators offered by Colab. GPU was used as a hardware accelerator to run the code.

8.1 Data Preparation

The dataset is mounted from google drive. Two folders are maintained, one for preprocessing the images and the other for storing the augmented images. The images go through five steps of preprocessing. Since there are only 20 images in the train set, the size of the data set is increased using data augmentation. The images for model training are retrieved from the folder of augmented images.

8.2 Data Pre-Processing

A 5-step process that involves the following steps:

1. Green Channel Extraction
2. CLAHE (Contrast Stretched Adaptive Histogram Equalization)
3. Gamma Correction
4. Fast cv2 Denoising
5. Normalization

8.2.1. Green Channel Extraction: The image's green channel is extracted first during preprocessing. The pictures of the retina are typically low contrast. The extreme contrast in the green channel makes microaneurysms readily apparent. To improve the contrast of the green channel, further preprocessing is used.

Fig.7. Green Channel Extraction

8.2.2. CLAHE: It cannot be performed on RGB images as it would lead to dramatic changes in the image's color balance. Hence, it is performed on the green channel extracted images.

Fig.8. CLAHE performed on green channel extracted images

8.2.3. Gamma Correction: Gamma correction of CLAHE images is performed using a lookup table.

Fig.9. Gamma Correction performed on CLAHE images

8.2.4. Fast cv2 Denoising: Denoising estimates the original image by suppressing noise from the image. Gamma corrected images are denoised using OpenCV. A function fastNlMeansDenoising() is used, which works with a single grayscale image.

Fig.10. Fast cv2 Denoising of gamma corrected images

8.2.5. Normalization: Normalization is an important step that ensures that each input parameter (pixel, in this case) has a similar data distribution.

Fig.11. Normalization of denoised images

8.3 Data Augmentation

Albumentations is a Python module for picture augmentations that is quick and versatile. It provides a succinct yet powerful image augmentation interface for various computer vision applications, such as object classification, segmentation, and detection, while effectively implementing various image transform operations geared for speed. It helps in classification, semantic segmentation, instance segmentation, object identification, and posture estimation, among other computer vision problems. The pre-processed train images are augmented, and test images are resized to dimension 512x512.

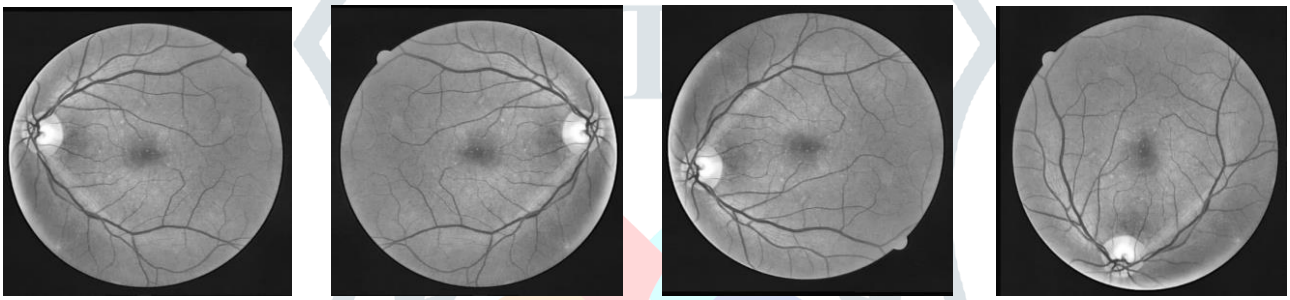

Fig.12. Augmented images

8.4 Model Training

The augmented train dataset is split into 80 percent training and 20 percent validation. With a batch size of four and a learning rate of 0.0001, each model is trained for 200 epochs. Callbacks are employed to avoid overfitting.

IX. RESULTS AND DISCUSSION

After training, the models are evaluated based on accuracy, sensitivity, and specificity. One parameter for assessing classification models is accuracy.

Informally, **accuracy** refers to the percentage of accurate predictions made by our model. The following is the formal definition of accuracy:

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \qquad (9.1)$$

For binary classification, accuracy can also be calculated in terms of positives and negatives as follows:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (9.2)$$

Where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

Sensitivity measures the proportion of actual positive cases that got predicted as positive (or true positive). The higher the sensitivity, the higher the real positive value and the lower the false negative value. The lower the sensitivity, the lower the real positive value and the larger the false negative value. Models with great sensitivity will be sought in the healthcare and finance domains.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \qquad (9.3)$$

The fraction of real negatives projected as negatives is specificity (or true negative). This means that a part of true negatives will be forecasted as positives, referred to as false positives.

$$\text{Specificity} = \frac{TN}{TN+FP} \qquad (9.4)$$

9.1. Evaluation Metrics

Model	Metrics		
	Accuracy	Sensitivity	Specificity
Unet	0.94660	0.90074	0.94824
Dense Residual Bridge Unet	0.96903	0.95732	0.94910

9.2. Output Visualization

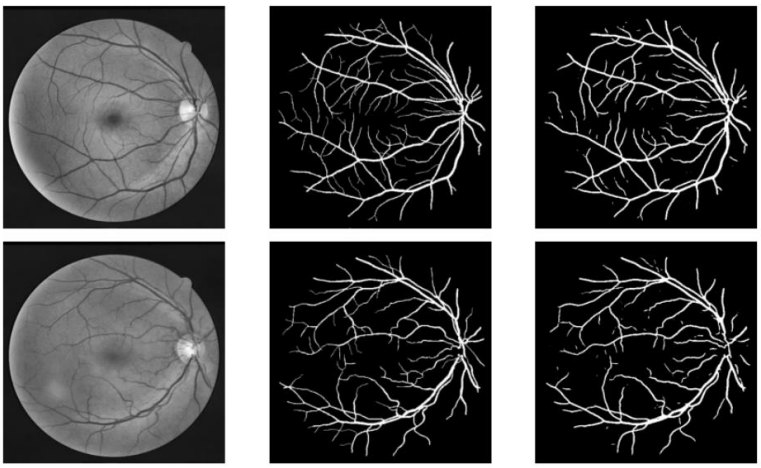

Fig. 13. Extracted Vessels using UNet

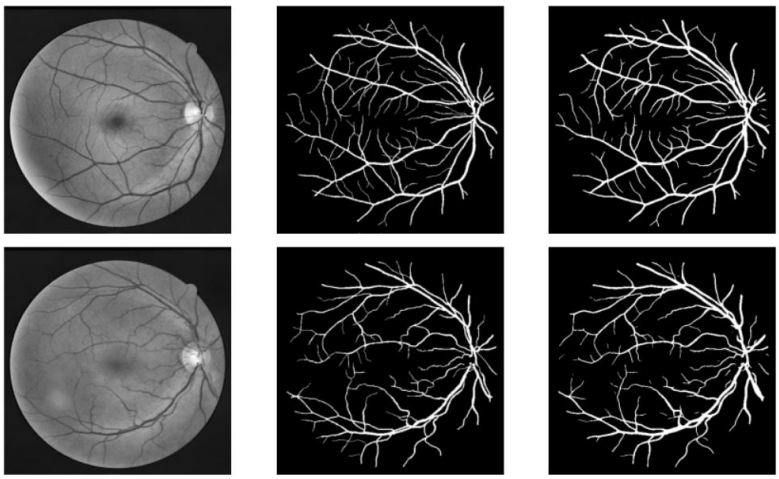

Fig. 14. Extracted Vessels using Dense Residual Bridge UNet

	Image	Acc	Specificity	Sensitivity
0	01_test	0.937878	0.938113	0.931113
1	02_test	0.94054	0.940918	0.931936
2	03_test	0.941071	0.944245	0.867173
3	04_test	0.9506	0.952536	0.900553
4	05_test	0.952271	0.954082	0.903712
5	06_test	0.951496	0.954591	0.870422
6	07_test	0.948513	0.950468	0.895084
7	08_test	0.955002	0.958659	0.850209
8	09_test	0.958622	0.961987	0.849746
9	10_test	0.950127	0.951421	0.909451
10	11_test	0.94416	0.944925	0.921387
11	12_test	0.946804	0.948896	0.886143
12	13_test	0.94767	0.949714	0.895096
13	14_test	0.942177	0.942255	0.939697
14	15_test	0.944653	0.944581	0.947263
15	16_test	0.945511	0.947685	0.885185
16	17_test	0.951214	0.954988	0.840689
17	18_test	0.942322	0.943431	0.907542
18	19_test	0.936245	0.935486	0.960559
19	20_test	0.945152	0.945813	0.921797

Fig. 15. Evaluation Scores of UNet

	Image	Acc	Specificity	Sensitivity
0	01_test	0.957635	0.936579	0.978054
1	02_test	0.959035	0.93718	0.991322
2	03_test	0.963949	0.945464	0.918666
3	04_test	0.974248	0.954108	0.96787
4	05_test	0.97846	0.95916	0.949692
5	06_test	0.971936	0.952845	0.938145
6	07_test	0.973268	0.953805	0.948579
7	08_test	0.980588	0.962934	0.903358
8	09_test	0.97994	0.960831	0.94112
9	10_test	0.974229	0.954151	0.966685
10	11_test	0.964673	0.94385	0.979224
11	12_test	0.970815	0.951213	0.949283
12	13_test	0.966745	0.946267	0.969026
13	14_test	0.967103	0.946227	0.984969
14	15_test	0.972371	0.952449	0.959524
15	16_test	0.964719	0.944611	0.957712
16	17_test	0.971372	0.952956	0.914986
17	18_test	0.9628	0.942291	0.968772
18	19_test	0.958368	0.936879	0.996058
19	20_test	0.968255	0.948113	0.963289

Fig. 16. Evaluation Scores of Dense Residual Bridge UNet

X. CONCLUSION

The ability to accurately segment retinal blood vessels in the fundus is instrumental in aiding doctors in diagnosing fundus disorders to avoid inadequate micro-vessel excision and severe mis - segmentation. A new approach to traditional retinal segmentation is presented. The dense residual bridge unet is a hybrid of the U-Net, dense layers, and residual networks. The suggested approach provided better segmentation accuracy, sensitivity, and specificity. The architecture performance can further be improved by training it on patches obtained after patch extraction of the given annotated mask images.

XI. ACKNOWLEDGEMENT

[1] Chen, C., Chuah, J. H., Raza, A., & Wang, Y. (2021). Retinal vessel segmentation using deep learning: a review. IEEE Access.

[2] Sun, X., Fang, H., Yang, Y., Zhu, D., Wang, L., Liu, J., & Xu, Y. (2021, September). Robust retinal vessel segmentation from a data augmentation perspective. In International Workshop on Ophthalmic Medical Image Analysis (pp. 189-198). Springer, Cham.

[3] Khan, T. M., Abdullah, F., Naqvi, S. S., Arsalan, M., & Khan, M. A. (2020, July). Shallow vessel segmentation network for automatic retinal vessel segmentation. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-7). IEEE.

[4] Dash, J., & Bhoi, N. (2015, January). A survey on blood vessel detection methodologies in retinal images. In 2015 International Conference on Computational Intelligence and Networks (pp. 166-171). IEEE.

[5] Shah, S. A. A., Shahzad, A., Khan, M. A., Lu, C. K., & Tang, T. B. (2019). Unsupervised method for retinal vessel segmentation based on Gabor wavelet and multi-scale line detector. IEEE Access, 7, 167221-167228.

[6] Pamina, J., & Raja, B. (2019). Survey on deep learning algorithms. *International Journal of Emerging Technology and Innovative Engineering*, *5*(1).

[7] İçöz, B., Karayagli, S., & Kanlitepe, S. DEEP LEARNING. Ankara Yildirim Beyazit University.

[8] Jha, G. K. (2007). Artificial neural networks and its applications. *IARI, New Delhi, girish_iasri@ rediffmail. com*.

[9] Deepthi, T. H., Gaayathri, R., Shanthosh, S., Gebin, A. S., & Nithya, R. A. (2019). Firearm Recognition Using Convolutional Neural Network.

[10] Cunningham, P., Cord, M., & Delany, S. J. (2008). Supervised learning. In Machine learning techniques for multimedia (pp. 21-49). Springer, Berlin, Heidelberg.

[11] Sk, S., Jabez, J., & Anu, V. M. (2017). The Power of Deep Learning Models: Applications. Networks, 33.

[12] Staal, Joes & Abràmoff, Michael & Niemeijer, Meindert & Viergever, Max & Ginneken, Bram. (2004). Ridge-Based Vessel Segmentation in Color Images of the Retina. IEEE transactions on medical imaging. 23. 501-9. 10.1109/TMI.2004.825627.

[13] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.