Variational Autoencoder based synthetic data generation method on Diabetic Retinopathy Images

Nunna Lakshmi Manasa[1], Mahesh Kumar Morampudi[2]

[1] III/IV B. Tech, Department of Computer Science and Engineering, SRM University – AP, Amaravati, Andhra Pradesh, India

[2] Assistant Professor, Department of Computer Science and Engineering, SRM University – AP, Amaravati, Andhra Pradesh, India

Email: lakshmimanasa_nunna@srmap.edu.in, maheshkumar.m@srmap.edu.in

Abstract: Diabetic Retinopathy (DR) is one of the complications of diabetes, in which high blood sugar levels cause damage to the eye by causing lesions on the retina. DR is mainly classified into 5 classes based on its severity of appearance of lesions on retina image. The essential challenge in this complication is early detection which is very important for treatment success. Manual detection of DR is time consuming and prone to misdiagnosis. So, recently many computer-aided diagnosis systems have been adopted using deep learning techniques for medical image analysis, in which convolutional neural networks are used for diagnosing diabetic retinopathy through analysing fundus images that have proven their superiority in detection and classification tasks. However, CNN relies on a largely diverse training dataset to learn which is difficult to collect, particularly for high severity levels. In, largest public DR dataset, EYEPacs images of DR, images of classes 0, 4 account for 73.67% and 2.16% respectively. Adopting such imbalanced data to train makes the model less sensitive to samples with higher DR levels and thus leads to over fitting. So, to solve this problem, we propose a variational autoencoder (VAE) based synthetic data generation method. Using the generative kind of VAE model, to capture the dimensional dependencies, and produce new samples. Variational Autoencoder samples some values of the latent variable and then the new samples are generated from the conditional distribution of the data given the latent variable. Here we aim to try and approximate the minority class distribution.

Keywords: Diabetic Retinopathy; Convolutional neural networks; Class-imbalance problem; Variational autoencoder.[i]

I. INTRODUCTION

Deep learning is a branch of machine learning in which artificial neural networks adapt to and learn from large amounts of data and are efficient in decision making. DL has many applications like image recognition, virtual assistants, language translations of which in health care medical image analysis [1], processing, segmentation, classification, registration has been most advancing and developing rapidly. Especially in health care field, using deep learning techniques to detect the diseases at early stage can be helpful in effective treatment. Out of many deep learning networks, Convolutional neural networks (CNN) is one of the widely used deep learning technique and has been consistently advancing in object detection, classification, segmentation. CNN's deep architecture provides high performance for training models by learning patterns from original images. Because of its ability to learn data representation, CNN has helped the medical field tremendously.

Diabetes [2] is a complication of diabetes mellitus that occurs when glucose level in blood high due to lack of insulin. 425 million adults are affected worldwide by causing damage to retina, heart, feet or kidneys. Diabetic retinopathy [3] is one of complication of diabetes mellitus a “disease of retina” where sugar levels in blood are high and cause damage to blood vessels of the retina by swelling and leaking fluids, blood that causes lesions on the retina. This damages the retina of an eye that leads to loss of vision in advanced stages. Therefore, patients with diabetes need to be screened regularly to diagnose and treat DR as soon as possible to reduce the risk of DR. DR is mainly detected by appearance of different types of lesions on fundus images of retina. They are microaneuryms (MA), Haemorrhages (HM), soft and hard exudates.

- MA - small red round dots on the retina due to the weakness of the vessel's walls.
- HM - large spots on retina, with its size greater than 125 μm and with an irregular margin.
- Hard exudates - bright exudates appear as bright-yellow spots on the retina caused by leakage of plasma.
- Soft exudates - white spots on retina triggered by the swelling of the nerve fiber.

According to the appearance and severity of different types of lesions on retina, DR is mainly divided into 5 stages [4] namely, No DR, Mild non-proliferative DR, Moderate non proliferative DR, Severe non-proliferative DR, and Proliferative DR which are briefly described in Table 1.

Table 1

DR Severity Level	Types of Lesions
No DR	No abnormalities
Mild non-proliferative DR	Presence of MA only
Moderate non-proliferative DR	Any of the following: -Micro aneurysms -Retinal dot, blood haemorrhages -Hard exudates , cotton wool spots

Severe non-proliferative DR	Any of the following: - >20 intra-retinal hemorrhages in each of 4 quadrants - venous beading in 2 or more quadrants - intra-retinal micro vascular abnormality in 1 or more quadrants
Proliferative DR	Neovascularization, Vitreous hemorrhage

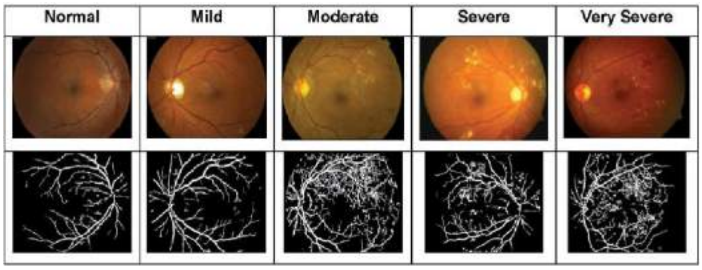


Fig. 1.Different stages of DR

The essential challenge in this complication is early detection which is very important for treatment success. Manual diagnosis for DR detection may be prone to misdiagnosis and cost, time, effort consuming unlike computer-aided diagnosis systems. Therefore, this article reviews the recent automated DR method [5] for classifying DR images using deep learning technology, and proposes a method to generate synthetic data from diabetic retinopathy images based on a variational autoencoder to avoid classifying images to avoid noise and class imbalance [6] issues.

II. DEEP LEARNING

Deep learning is a branch of machine learning technology that uses neural networks to automatically extract useful information from raw data to provide information for future decisions. It involves hierarchical layers for unsupervised feature learning and non-linear processing steps of classification patterns. In recent years DL has majorly contributed in medical image classification by computer-aided medical diagnosis methods.

For years, deep learning methods are being widely used in Diabetic retinopathy detection and classification. In this deep learning model, input data can be learned successfully even when heterogeneous sources are inputted. There are many deep learning methods like Convolutional neural networks, auto-encoders [7], recurrent neural networks, Boltzmann machines of which CNN is widely used in medical image classification tasks and have achieved significant performance till date. Also, the DL method does not require manual hand-crafted feature extraction, but unlike machine learning methods, it requires a large amount of training data.

The major difference [8] between Machine learning and Deep Leaning Model are briefly explained in Table 2.

Table 2

	ML	DL
Hand-crafted feature extraction	Required	Not required
Amount of training data	Not require large data	Require large data

The process to detect and classify DR images using DL is done by collecting the dataset and by applying preprocessing methods to improve and enhance images. Then, the images are fed in to the any DL method to extract the important features and classify the images as shown in figure 2.

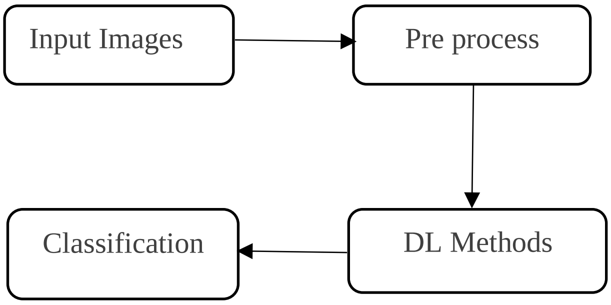


Fig. 2.Process of classifying DR using DL

A convolutional neural network [9] is a model of deep neural networks used in image recognition and classify particular features from images and processing, that is specifically designed to process pixel data and analyse visual images. The main layers are

- convolutional
- Normalization
- Pooling
- Fully Connected Layers

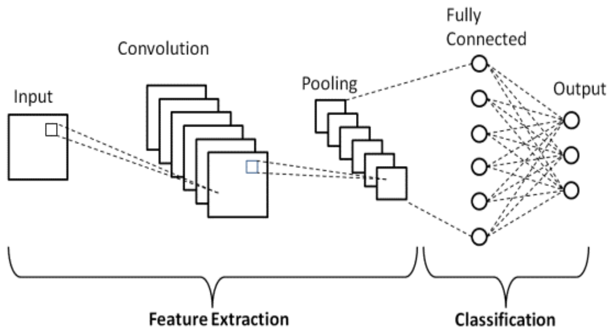


Fig. 3.Architecture of Convolutional Neural Networks

Various important features are difficult to identify in images. So, some filters are applied to the input images to extract those features from the input image in convolutional layer. The resulted images are linear. So, RELU non-linear activation function in normalization is used to change the linearity into non-linear.

Pooling layer to extract the dominant features and maintain the process of effectively training of the model. Last layers in the neural network are fully connected layers

The input to the fully connected layer is the flattened output from the final Convolutional Layer. After passing through the fully connected layers, softmax activation function is used in neural network to normalize the output of a network. This gives the probability distribution over predicted output classes.

$$softmax(z_i) = \frac{exp(z_i)}{\sum_j exp(z_j)} \quad (1)$$

However, CNN relies on a largely diverse training [10] dataset to learn features unlike machine learning methods which is difficult to collect, particularly for high severity levels. We come across one of the most extremely common problems when performing medical images tasks is a class imbalance problem which means the total number of observations of target classes of interest only appear in small portions in the entire dataset i.e., training dataset is not balanced as it contains relatively less images of one class compared to severe class. There are many publicly available datasets in kaggle, etc. But all those datasets are not readily available and have class imbalance problems. So, recently various synthetic sampling methods have been used of which one is the Variational Autoencoder (VA).

III. RELATED WORKS

- Gardner used a neural network on 200 images and divided them into blocks, achieving a sensitivity of 88.4% and a binary classification specificity of 83.5%.
- Dr. Nayakc classified diabetic retinopathy into three categories by identifying blood vessels and hard exudates from 140 images. The model classifies between normal, non-proliferative, and proliferative retinopathy. The results are verified by clinicians and show an accuracy of 93%, a specificity of 100%, and a sensitivity of 90% [11].
- Most of the research on automatic detection of more than 3 types of DR is done using support vector machines (SVM). Acharya *et al.* Detected retinopathy by entering the extracted features of the image into the SVM classifier [12] to capture changes in contour and shape using 5 categories. The characteristic regions of hard exudate, bleeding, blood vessels, and microaneurysms are calculated from the image and then entered into the SVM classifier. The results show that the precision and specificity are 82% and 88%, respectively.
- Dr.Adarsh uses image processing for automatic detection by detecting retinopathy-related textural features. Support vector machines are built using texture features and by running this method on the public image databases DIARETDB1 (130 images) and DIARETDB0 (89 images), precision rates of 94.6% and 96% were obtained.
- Harry Pratt proposed a method for diagnosing DR from digital fundus images using CNN [14], which consists of 80,000 training images, which are combined with 5,000 verification images belonging to 5 categories to achieve 75% each And 95% accuracy and sensitivity. This training is the hardware-intensive training required by NVIDIA K40c. Normalization is used to preprocess the images, and then feed them into your custom stacked convolutional layer network, followed by the fully connected layer.
- Xiaogang proposed a CNN-based transfer learning method in Alex Net and VGG16 [15]. Experiments were performed with 1200 fundus images from MESSIDOR and 1014 images from the DR1 data set. The model previously trained in Image Net was fit on the DR Dataset.
- Carson Lam is dedicated to classifying 4ary data into a data set containing 35,000 training images [16]. They used Otsu's method for systematic clipping, then normalized as a preprocess, and then additionally used limited-contrast adaptive histogram equalization for contrast adjustment. After removing the last dense layer with the help of the Tesla K80 GPU hardware, the model was trained on 22-layer Google Net through migration learning. Precision rates of 74.5%, 68.8%, and 57.2% were obtained in categories 2-ary, 3-ary, and 4-ary, respectively.
- Maithra Raghu explored transfer learning in a data set consisting of fundus photographs to diagnose 5 multiple class eye diseases and diabetic retinopathy. This method paves the way for a deeper understanding of the impact of learning transfer [17] from irrelevant data to medical data sets. They have observed that reuse of features at the lowest level and excessive parameterization of the model lead to biases from transfer learning.

A. *Generative Adversial Neural Networks [18] in Retinal Image Synthesis:* Recent works include exploiting GANs to synthesize retinal fundus images to data imbalance problem.

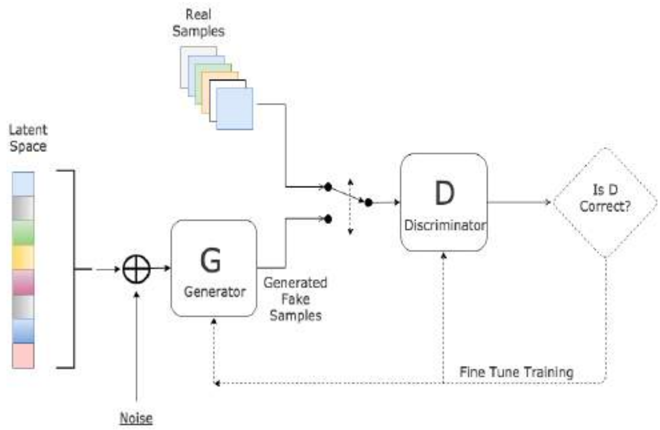


Fig. 4.Generative Adversial Neural Networks

- Initially, by Costa *et al*. [19] the UNet architecture and the vanilla GAN architecture are used to transfer the segmentation mask of blood vessels to the fundus image. However, the generated samples have block defects and no controllable grading sort information.
- Recently, Niu *et al*. [20] is committed to generating fundus images with pathological descriptors and vessel segmentation masks, although this method can only perform lesion operations, and does not consider the global representation to distinguish levels.

IV. PROPOSED WORK

Supervised learning methods in deep learning require large labelled training data to classify fundus images into a class or category. It is mainly where class imbalance occurs when one class, the minority group, contains particularly fewer samples than the other major classes. Therefore, we can observe that in the existing data set, the distribution of DR in different categories is very unbalanced, because the fundus abnormal (proliferative) image only accounts for a small part. Though there are many publicly available datasets they are many inaccessible datasets that have balanced data which can be helpful for training and research due to privacy-concerns, or lack of data-sharing incentives. In, largest public DR Dataset, EYEPacs [21] images of DR, levels 0, 4 account for 73.67% and 2.16% respectively. Training with this unbalanced data will make the model less sensitive to samples with a higher DR level, leading to overfitting.

A. Autoencoders:

Autoencoders are special implementation of artificial neural networks (ANNs) [22] and they are fully developed in an unsupervised manner. They learn compressed representations of data especially "coding" without requiring any labelled information. The input is compressed into a lower-dimensional code, and then the output is rebuilt according to the representation.

Basic architecture of autoencoders is illustrated in fig 5.

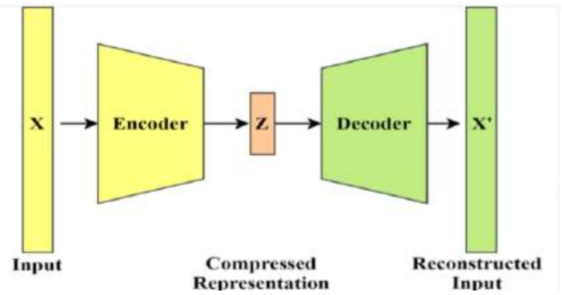


Fig. 5.Basic architecture of Autoencoders

B. Variational Autoencoders:

In general a VAE [23] model takes an image as an input runs it through the neural network to compress the data into smaller representations which passes through a hidden layer and then the decoder reconstructs the input by using CNN as shown in Figure1. VAE carries out the process described above using mainly 3 parts:

- Encoder Network
- Hidden Layer/Bottle Neck
- Decoder Network

Encoder takes input and produces some continuous representation (latent variable) of given sample. Decoder takes this representation as input and tries to reconstruct the input from original data.

Latent Space Representation: Encoder will produce which are two vectors which are mean and standard deviation. Using this mean and standard deviation vectors we can generate some amount of new samples and propagate them through decoder.

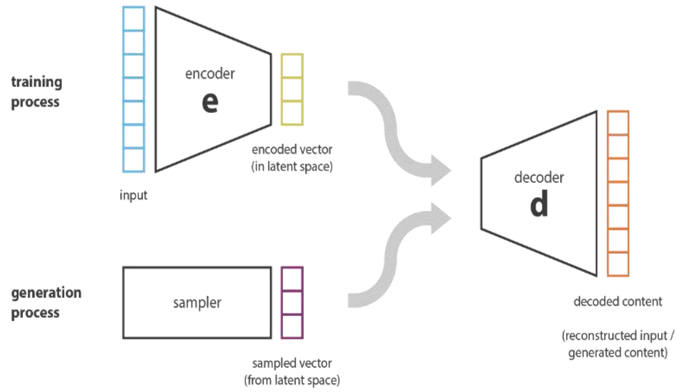


Fig. 6.Steps involved in Variational Autoencoder

C. Method and Structure:

The entire experiment is performed on APTOS Blind Detection Dataset. Labelled train and test images are in the dataset that can be helpful in classification before applying the model.

D. Preprocessing:

First, perform image preprocessing to remove noise in the image, improve image characteristics, and ensure image consistency. The image is converted, resized, cropped and rotated to obtain a clear and sharp enhanced image.

E. Data Augmentation:

Variational autoencoders one of the data augmentation techniques was performed on the imbalance data to increase the dataset size.

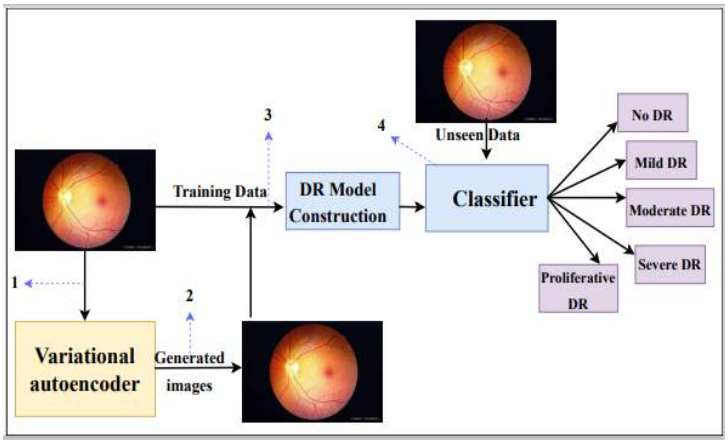


Fig. 7.Block diagram of VAE-CNN model for screening/identification of DR

In Figure7 there are 4 stages that represent:

1. Experimenting the input images into encoder of Variational autoencoder.

2. Synthetic generation of fundus images by reconstructing the output of encoder in decoder of VAE.

3. Training the deep learning model with synthetic images and actual fundus image dataset.

4. Classification of fundus images based on levels of severity with the help of classifiers.

Initially, VAE samples some latent variable values, and these values often produce raw data and generate new samples based on the conditional distribution of the data. Then they are inputted to latent variable. The generated samples are similar to original to dataset but not exactly the same.

V. RESULT ANALYSIS

The proposed model is experimented on APTOS (Asia Pacific Tele-Ophthalmology Society) Blindness dataset. We were able to demonstrate how a VAE-based approach could be used to generate synthetic data.

Image preprocessing is done by cropping the images and then resizing them. By this the important features were enhanced and then images were obtained with highest resolution are shown in Fig 8.

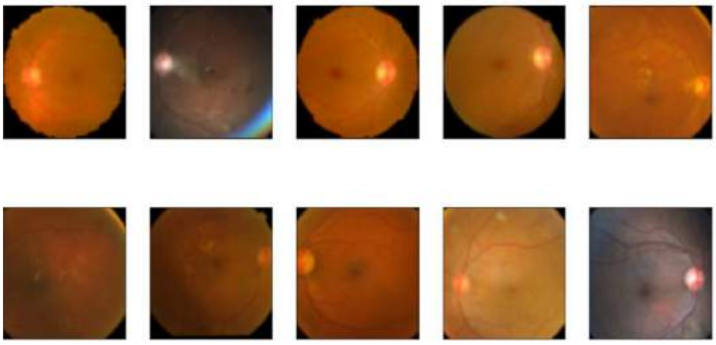
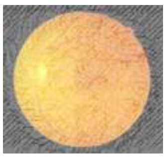
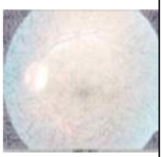
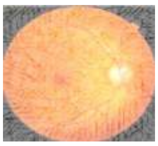
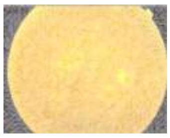


Fig. 8. Preprocessed Fundus Images

In Google colabaratory, an input layer to VAE is created and then it is connected to the encoder to encode the input and return the latent vector using tensor flow at the backend using Graphics processing Unit accelerator to reconstruct the input to generate sample images decoder is connected to output of encoder

100 images from each class are taken and fed to VAE. They are learned and were able to produce 100 synthetic images of each class with higher quality and enhanced features which are depicted in Table 3

Table 3

			Stage 0 Normal
			Stage 1 Mild
			Stage 2 Moderate
			Stage 3 Severe
			Stage 4 Very Severe

A. *Loss of VAE:*

Reconstruction Loss: It is the distance between original and generated data and help us to separate classes. And basically the loss in reconstructed images loss is pixel-wise Binary Cross Entropy of this network is some defined distance between original input and reconstructed output.

Kullback–Leibler divergence: It is the difference between two distributions and the distance between standard normal distribution and hidden variable of the variational autoencoder.

$$\text{KL Loss} = \sum_{i=1}^{n} \sigma_i^2 + \mu_i^2 - \log(\sigma_i) - 1 \quad (2)$$

Where σ = standard deviation μ = mean

Final Loss = KL + Reconstruction Loss. (3)

After validating the dataset final validation loss of **0.1745** and value accuracy of **0.9143** are obtained.

VI. CONCLUSION

We had described how fundus images in diabetic retinopathy are classified into different stages and the imbalance in the training data available to train the model. So, in this study we aimed to propose a variational autoencoder for synthetic data generation and tried to produce some samples. Some values of latent variable are sampled in variational autoencoder and from the conditional distribution of the data given the latent variable new samples are generated. The research needs to be still continued for producing sharp and clear sample images that can used for grading purpose so that there would not need of large datasets for training the model.

REFERENCES

1) Justin Ker, Lipo Wang, “Deep Learning Applications in Medical Images Analysis”, IEEE.
2) Stephanie Watson (2020), “Everything You Need to Know about Diabetes”, healthline.
3) Stanley S.Schwartz, “An Underdiagnosed and Undertreated Inflammatory, Neuro-Vascular Complication of Diabetes”, frontiers.
4) https://www.sciencedirect.com/science/article/abs/pii/S00200255070036 (accessed on 6 June 2021)

5) Valverde, Carmen, Garcia, Maria, “Automated detection of diabetic retinopathy in retinal images”, Indian Journal of Ophthalmology.
6) Justin Johnson, “Survey on deep learning with class imbalance problem”, Journal of Big Data (2009).
7) Jeremy Jordan, “Introduction to Autoencoders”, Jeremy Jordan.
8) https://www.zendesk.com/blog/machine-learning-and-deep-learning/
(Accessed on 20 June 2021)
9) https://www.ibm.com/cloud/learn/convolutional-neural-networks
(Accessed on 7 June 2021)
10) Laurence Moroney, “Use convolutional neural networks (CNNs) with large datasets to avoid overfitting”, Codelab.
11) J.Nayak, P.S. Bhat, C.Lim, M.Kagathi, “Automated identification of diabetic retinopathy stages using digital fundus images”, J. Med. Syst. 32 (2) (2008) 107-115.
12) R.Acharya, C.K. Chaua, “Application of higher order spectra for the identification of diabetes retinopathy stages”, J. Med. Syst. 32 (6) (2008) 481-488.
13) P.Adarsh, D.Jeyakumari, Multiclass svm-based automated diagnosis of diabetic retinopathy”, in: Communications and signal Processing (ICCSP), 2013 International Conference on, IEEE, 2013, pp. 206-210.
14) Harry Pratt, Frans Coenen, Deborah M.Broadbent, Simon P.Harding, Yalin Zheng, “Convolutional neural networks for diabetic retinopathy”,2016.
15) Xiaogang Li, Tiantian Pang, Weixiang Liu, Tianfu Wang, Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification, 2017.
16) Carson Lam, Darvin Yi, Margaret Guo, Tony Lindsay, Automated detection of diabetic retinopathy using deep learning, 2018.
17) Maithra Raghu, Chiyuan Zhang, Jon Kleinberg, Samy Bengio, “Transfusion: understanding transfer learning for medical imaging”, 2019.
18) Ian J. Goodfellow, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, “Generative Adversial Netwoks”, arXiv, Cornell University.
19) C.P, G. A, M. M I, N. M, A. M, M. A M, “End-to-end adversarial retinal image synthesis,” IEEE Transactions on MedicalImaging, vol. 37, no. 3, pp. 781–791, 2018.
20) Y.Niu, L.Gu, F.Lu, F.Lv, I. Sato, Z. Zhang, Y. Xiao and T. Cheng, “Pathological evidence exploration in deep retinal image diagnosis,” in AAAI, 2019.
21) https://www.kaggle.com/c/diabetic-retinopathy-detection/data
(Accessed on 23 June 2021).
22) Jake Frankenfield, “Artificial Neural Network (ANN)”, Canvs.
23) https://towardsdatascience.com/intuitively-understanding-variational-autoencoders-1bfe67eb5daf
(Accessed on 7 June 2021)

AUTHORS PROFILE

Manasa Nunna is currently an undergraduate student from the stream, Computer Science engineering at SRM University AP, Amravati. Research interests include Artificial Intelligence, Deep Learning and Data science.

Dr. Mahesh Kumar. M is currently working as Assistant Professor in the stream of Computer Science engineering at SRM University AP, Amaravati. Research interests include Privacy-preserving Biometric Authentication, Homomorphic Encryption, Privacy in Machine Learning.