Semantic Segmentation of Iris using U-Net in Deep Learning

Dugimpudi Abhishek Reddy, Deepak Yadav, Nishi Yadav, Devendra Kumar Singh

Abstract: In the field of medicine, iris segmentation has become a great field of interest from the past few years. Iris segmentation is also largely used in iris recognition systems [3] which are extensively used in security control [1][2]. Here iris segmentation is done using semantic segmentation which is based on the U-Net architecture. The typical U-net architecture contains two paths- contracting path containing convolutional and pooling layers and the expanding path consists of transposed convolutional operations. The UBIRIS dataset is trained on the traditional U-Net model with some modifications according to the size of the images present in the UBIRIS dataset. The results obtained were very close to the ground truths and accuracy obtained is also appreciable.

Keywords: Iris Segmentation, Iris recognition system, Semantic segmentation, U-Net architecture, Convolutional layers, Pooling layers, Transposed Convolution layers

I. INTRODUCTION

Segmentation is a process that partitions an image into regions. Segmentation is mainly used in applications like remote sensing or tumour detection [9]. Segmentation is an image processing method that allows us to separate objects and textures in images. There are different kinds of segmentation like semantic segmentation, instance segmentation. But our topic of discussion is semantic segmentation.

A. Semantic segmentation

The input to the semantic segmentation is an image and the output is a category for each individual pixel. The important task in this technique is to classify each and every pixel into one of the different categories. In other words, all pixels belonging to aeroplane would be classified into one group and similarly the sky etc.
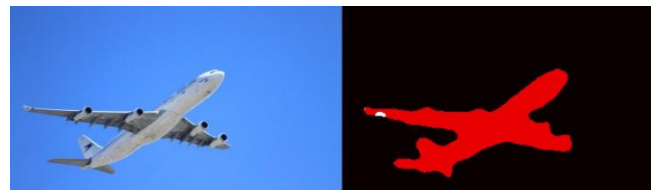
Dugimpudi Abhishek Reddy, Department of CSE, School of Studies in Engineering and Technology, Guru Ghasidas Vishwavidyalaya, Bilaspur, India.Email:abhishek.duggimpudi@gmail.com

Deepak Yadav, Department of CSE, School of Studies in Engineering and Technology, Guru Ghasidas Vishwavidyalaya, Bilaspur, India.Email:deepkyadav28031998@gmail.com

Nishi Yadav*, Department of CSE, School of Studies in Engineering and Technology, Guru Ghasidas Vishwavidyalaya, Bilaspur, India.Email:nishidv@gmail.com

Devendra Kumar Singh, Department of CSE, School of Studies in Engineering and Technology, Guru Ghasidas Vishwavidyalaya, Bilaspur, India.Email:devendra.singh170@gmail.com

Fig. 1. Example depicting semantic segmentation

Semantic segmentation is done based on the UNET architecture. The UNET was developed by Olaf Ronneberger for Bio Medical Image Segmentation[5]. The architecture consists of two parts [13].

The first part is the contraction path which is also known as Encoder and

1. The second part is the expanding path which is also known as Decoder.

The encoder path is where the context in the image is captured and it is made up of several convolutional and max pooling layers. Decoder is the second path that is used to find the localization with the help of transposed convolutions. Hence it is an end-to-end Fully Convolutional Network[18]. It can take image of any given size as input as it is not containing any dense layers.

B. Convolutional Layer:

A typical convolutional layer takes two inputs-

a) A 3D volume which is the input image of dimensions : H1 X W1 X D1
b) A set of k filters each of size p x p x channels where p is usually 3 or 5

The hyper parameters used are K (filters), J (size of filter), S (Stride) and P (zero padding)

The output from the convolutional layer is also a 3D volume of dimensions H2 X W2 X D2

where $H2 = (H1-J+2P)/S+1$

$W2 = (W1-J+2P)/S+1$

$D2 = K$

C. Max pooling Layer:

The main work of pooling is to decrease the size of the feature map so that network contains less amount of parameters.

The input to this layer is a volume of size H1 X W1 X D1

The hyper parameters are F (filter size) and S (stride)

The output from this layer is volume of dimensions H2 X W2 X D2

Where H2 = (H1-J)/S+1

W2 = (W1-J)/S+1

D2 = D1

Generally from every block of certain size of the input feature map the pixel with the maximum value is selected and get a feature map which is pooled. The important function is to keep only the main features from each region and get rid of the unnecessary information. In a convolutional network, the width and height of the image eventually decreases due to max pooling or down sampling which helps in focusing on the information in the inner layers and on a bigger context of the image. The number of channels eventually increases which enables to obtain many complex features from the image. By this process of pooling, the model loses the information of "WHERE" it is existing but it gets the information of "WHAT" is there in the image. But we need to keep both "WHERE" and "WHAT" information from the image. Hence we need to do up sampling along with down sampling to get "WHERE" the information is present in the image.

D. Transposed Convolution:

It is a method which is used to perform up sampling of an image with learnable parameters. It is also called as fractionally strided convolution or deconvolution. Transposed convolution is exactly the reverse procedure of a usual convolution which means it gives high resolution image by taking low resolution image[18]. The parameters are learnable through back propagation.

E. Importance of iris segmentation:

The present methods use deeply learned features in association with a conventional pipeline[3][14][15].Or they use a pre-trained model which is fine-tuned that is used for iris recognition [16].The iris has been a very important organ in the human body for biometric identification for many reasons:

It is the most protected internal organ which is least highly vulnerable and it is covered with sensitive and transparent membrane cornea. The iris is generally flat, and its geometric configuration is controlled by a pair of complementary muscles- sphincter pupillae and dilator pupillae which controls the pupil diameter[17][20]. The iris also has a fine texture which makes it very unique which is very tough to prove [17]. Iris segmentation is the most powerful iris recognition system because regions that are wrongly segmented as iris may corrupt biometric templates which result in very poor recognition. There are many methods for iris segmentation as mentioned in [7].There are also many factors that affect the recognition of iris like occlusion, image quality, light etc. [8].

II.   LITERATURE REVIEW

In [1], the importance of biometrics in daily life is explained in detail and they mainly discussed how the integration of biometrics into security revolutionised the field.

In [2], the biometrics recognition like fingerprints, iris recognition were discussed and their entry to the field of security and also how this helped in privacy control.

In [3], they discussed how iris recognition systems do function in less constrained environments and the acquisition of images for iris recognition.

In [4], the main discussion was on the analysis of error rates and their effect on the accuracy of the segmentation stages.

In [5], Olaf Ronneberger for the first time introduced U-Net and its use in biomedical image segmentation. The U-Net architecture had given good results in the field of biomedicine.

In [6], they have given a database of ground truth for segmentation of iris and a procedure for using that database to find an algorithm for iris segmentation.

In [7], they discussed the various problems with the process of segmentation, and proposed some important techniques for iris segmentation.

In [8], they evaluated some of the components which affect the image in iris segmentation and also evaluated the effect of quality components in the execution of iris matchers based on Log-Gabor wavelets and SIFT keypoints.

In [9], they leveraged the significance of DNNs for object detection. They had shown an easy method of detection as DNN-base object mask regression can give better output when it is applied along with a multi-scale course-to-fine procedure.

In [10], they gave a procedure for rapidly increasing the training of deep networks. It is based on the fact that covariate shift known to affect the machine learning systems' training, which also applies to sub-networks and layers, and erasing it from internal activations of the network may help in training.

In [11], they have studied and extracted the deep features from VGG-Net and used it in iris recognition. The proposed method gave good results with very good accuracy rate when tested on two popular iris databases.

In [12], they presented two iris segmentation models for iris images with noise obtained at some distance and in motion, namely hierarchical convolutional neural networks (HCNNs) and multi-scale fully convolutional network (MFCNs).

In [13], to segment the images with less quality, they gave a deep neural network and it was trained on Bath800 and CASIA1000 databases and later performed on UBIRIS and MobBio.

III.   PROPOSED METHODOLOGY

As described above, the layers present in the U-Net model, let us proceed to the training and testing of the UBIRIS dataset.

The sample of the UBIRIS dataset with the image and its corresponding ground truths [6] are shown below in the Fig.2.
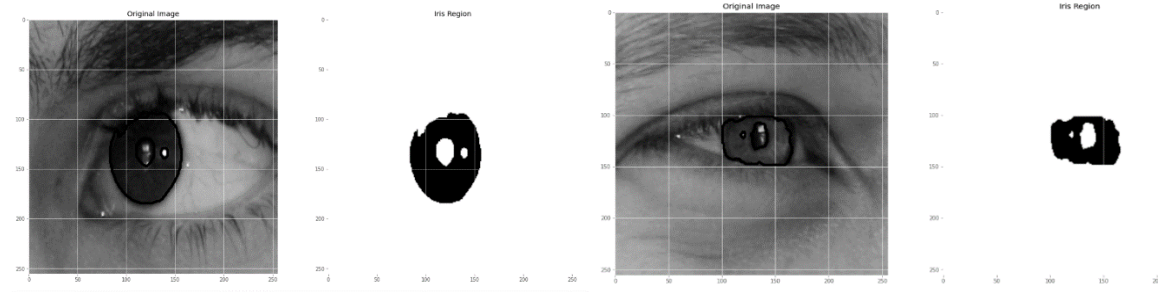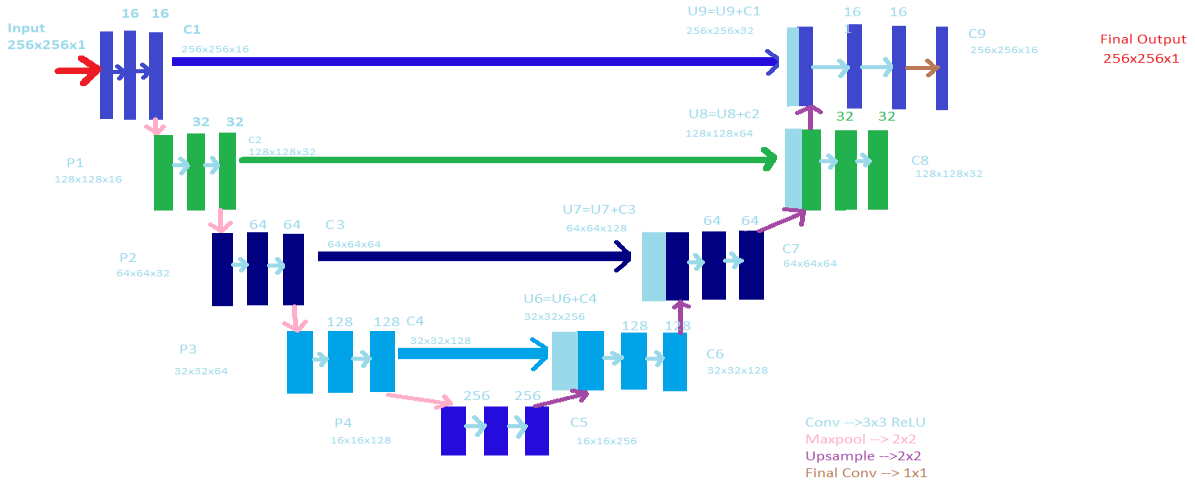
Fig. 2. Images from UBIRIS dataset

Fig. 3: Structure of U-Net architecture

A. DEFINING THE U-NET ARCHITECTURE

The U-Net model which is already available is taken and fine-tuned in such a way that it works well with our UBIRIS dataset. The input to this network is of dimensions 256x256x1 and the output is also of the same dimensions but with the segmented iris from the original input image The network as we know contains the convolution, max pooling, transposed convolution layers in it stacked in the shape of "U". Internally there are many other operations going on like batch normalisation [10], activation which usually happen in ConvNets.

IV.   IMPLEMENTATION AND RESULTS

A. Preprocessing

The UBIRIS dataset mainly contains two sets of images. The original eye images and their corresponding ground truths where each ground truth image is the segmented iris image of its eye. There are a total of 1000 images each in both parts. Initially the images are paired in such a way that each image of eye is linked with its segmented iris image for training.

B. Experimental setup

The simulation is done on Google Colaboratory environment. Google Colaboratory provides an environment to write and run python programs efficiently. It is mainly used for deep learning projects as it provides large amount of resources like RAM, GPU and CPU [Colab]. The details of hardware and software specifications of Google Colab are as shown in the Table 1:

Table 1: Hardware and software components used

Specification	Description
Graphical Processing Unit	Tesla P100-PCIE-16GB
Central Processing Unit	Intel Xenon CPU @2.30GHz
Random Access Memory	~12.72 GB
Size of disk	~68.4 GB
Programming Language	Python 3

The U-Net architecture as specified above is used to train the UBIRIS dataset which consists of 1000 images of eye and corresponding ground truth images. The experiment is done to achieve highly segmented iris images with great accuracy and minimal segmentation error.

C. Simulation and Results

The experiment is carried out with the U-Net architecture on the UBIRIS dataset on the Google Colab with a lot of hyper parameters[11].

Some of those include- Adam Optimizer, loss used is binary_crossentropy, batch size is 6, and the number of epochs used to train are 100. Although the number of epochs are high, callbacks are used so that the model may not get overfitted on the UBIRIS dataset. Callbacks help us to have control over the training process where training gets automatically stopped when it reaches certain accuracy/loss and saving the model as checkpoint after every successful epoch adjusting the learning rates over time.

The EarlyStopping function is used to stop the training before the model gits overfitted. The training was performed well and the best model was obtained at a point between 25th epoch and 30th epoch although training continued till 38th epoch.

Fig. 4. Accuracy curve during training and validation process

The training loss and validation losses are shown according to number of epochs and the best model is shown with x symbol. As shown, it is very clear that the loss obtained during the training process is very minimal which resulted in a good model for segmentation. The accuracies obtained during training and validation is also shown in the Fig 4 and Fig 5. As the dataset size used is small, the accuracy [4] is very high but the results were also close to the ground truths. The model is tested with certain images and the results are as shown:
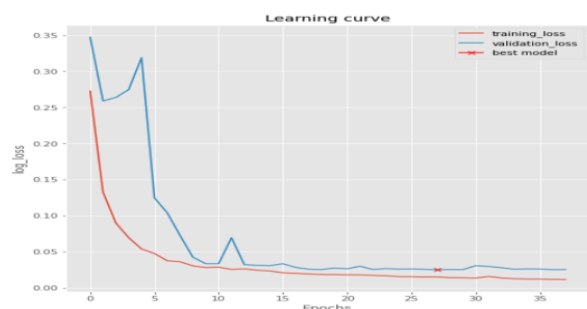

Fig. 5. Loss curve during training and validation process

The average segmentation error when evaluated got the value of 8.056640625e-06 and the jaccard index obtained is 0.9999999084472656 which shows that the training has been done in a very good manner.
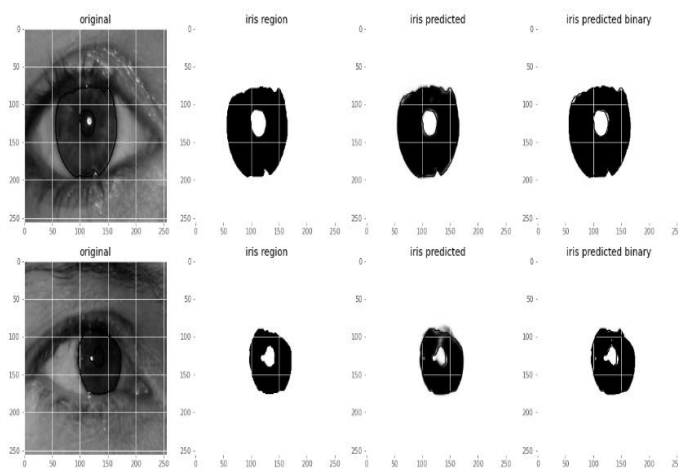

Fig. 6. Sample images of testing the model on UBIRIS dataset

The above figure shows four column of images where each column represents different images-first column represents the original image and the second shows its ground truth. The third column shows the predicted output from the model which is iris-segmented image of the actual image. The last column shows the iris predicted binary which is obtained from the predicted image by removing the greyish region and only shows the black and white regions of the image. It is also clear that the predicted images are very close approximations to the original ones.

V. CONCLUSION:

We described a U-Net based system for Iris segmentation. This CNN based Segmentation model shows very good segmentation of the Iris which are very close to the ground truth images. This Model required small amount of training data even though the results were better. U-Net performs very well in semantic segmentation. So in conclusion we can say that using Deep learning method for Iris segmentation provides better accuracy in the field of biometrics compared to conventional methods. The UBIRIS dataset if large may also give even better results were better. U- Net performs very well in semantic segmentation. So in conclusion we can say that using Deep learning method for Iris segmentation provides better performance in an area of biometrics compared to conventional methods .

REFERENCES:

1. A.Pando, "Beyond Security: Biometrics Integration Into Everyday Life," forbes. [Online].
2. 2. S. Prabhakar, S. Pankanti, and A. K. Jain, "Biometric recognition: security and privacy concerns,"IEEE Secur. Priv. Mag., vol. 1, no. 2, pp. 33–42, 2003.
3. 3. J. R. Matey et al., "Iris on the Move: Acquisition of Images for Iris Recognition in Less Constrained Environments," Proc. IEEE, vol. 94, 2006
4. 4. H. Proença and L. A. Alexandre, "Iris recognition: Analysis of the error rates regarding the accuracy of the segmentation stage," Image Vis. Comput., vol. 28, no. 1, pp. 202–206, 2010

Retrieval Number: A2614059120/2020©BEIESP
DOI:10.35940/ijrte.A2614.059120
Journal Website: www.ijrte.org

2027

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

5. U-Net: Convolutional Networks for Biomedical Image Segmentation , Olaf Ronneberger, Philipp Fischer, and Thomas Brox, Computer Science Department and BIOSS Centre for Biological Signalling Studies, University of Freiburg, Germany
6. H. Hofbauer, F. Alonso-Fernandez, P. Wild, J. Bigun, and A. Uhl, "A ground truth for Iris segmentation," in Proceedings - International Conference on Pattern Recognition, 2014
7. 7. R. Jillela and A. A. Ross, "Methods for Iris Segmentation," in Handbook of Iris Recognition, M. J.Burge and K. W. Bowyer, Eds. London: Springer London, 2013, pp. 239–279.
8. 8. F. Alonso-Fernandez and J. Bigun, "Quality factors affecting iris segmentation and matching," in Proceedings - 2013 International Conference on Biometrics, ICB 2013, 2013.
9. C. Szegedy, A. Toshev, and D. Erhan, "Deep neural networks for object detection," in Advances in neural information processing systems, 2013, pp. 2553–2561.
10. 10. S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in International Conference on Machine Learning, 2015, pp. 448–456.
11. S.Minaee, A Abdolrashidiy, and Y.Wang, "An experimental study of deep convolutional features for iris recognition," in 2016 IEEE Signal Processing in Medicine and Biology Symposium, SPMB 2016 - Proceedings, 2017.
12. N. Liu, H. Li, M. Zhang, J. Liu, Z. Sun, and T. Tan, "Accurate iris segmentation in non-cooperative environments using fully convolutional networks," in 2016 International Conference on Biometrics,ICB 2016, 2016.
13. E. Jalilian and A. Uhl, "Iris segmentation using fully convolutional encoder–decoder networks," Adv. Comput. Vis. Pattern Recognit., vol. PartF1, pp. 133–155, 2017.
14. Z. Zhao and A. Kumar. Towards more accurate iris recognition using deeply learned spatially corresponding features. In ICCV, pages 22–29, 2017.
15. X. Tang, J. Xie, and P.a Li. Deep convolutional features for iris recognition. In CCBR, pages 391–400. Springer, 2017.
16. M. Arsalan, H.G. Hong, R.A. Naqvi, M.B. Lee, M.C. Kim, D.S. Kim, C.S. Kim, and K.R. Park. Deep learning-based iris segmentation for iris recognition in visible light environment. Symmetry, 9(11):263, 2017
17. https://en.wikipedia.org/wiki/Iris_recognition
18. www.towardsdatasciene.com
19. Jus Lozej, Blaz Meden, Vitomir Struc, Peter Peer. "End-to-End Iris Segmentation Using UNet", 2018 IEEE International Work Conference on Bioinspired Intelligence (IWOBI), 2018
20. Al-Hawamd Kawther Hamadala. "Improvement of Iris Recognition System Using Fuzzy Segmentation and Neural Network", Al Al-Bayt University, 2010
21. Network", Al Al-Bayt University, 2010

AUTHORS PROFILE

Dugimpudi Abhishek Reddy, is currently pursuing Bachelor of Technology in Computer Sciene and Engineering from School of Studies in Engineering and Technology, Guru Ghasidas Vishwavidyalaya, Bilaspur, India. He is currently studying in the final year. His research interest fields are Machine Learning, Computer Vision and Network Security.

Deepak Yadav, is currently pursuing Bachelor of Technology in Computer Science and Engineering from School of Studies in Engineering and Technology, Guru Ghasidas Vishwavidyalaya, Bilaspur, India. He is currently studying in the final year. His research interest fields are Machine Learning, Computer Vision and Network Security.

Nishi Yadav, is currently working as an Assistant Professor in Computer Science and Engineering Department, School of Studies in Engineering and Technology Guru Ghasidas Vishwavidyalaya. A Central University, Bilaspur, Chhattisgarh,495009, India. She received Master of Technology (M.Tech) degree in 2010 from National Institute of Technology ,Rourkela (NITRKL),Odisha, India. She has published more than 20 papers in reputed journals and confrences like IEEE, Springer, Ugc approved, etc .Her research interest fields are Wireless Ad Hoc Network, Network Security, Deep learning, etc. She has more than 10 year of experience.

Devendra Kumar Singh, is currently working as an Assistant Professor in Department of Computer Science and Engineering, School of Studies in Engineering and Technology, Guru Ghasidas Vishwavidyalaya (Central University), Bilaspur, Chhattisgarh, 495009, India. He received Master of Technology (M.Tech) degree in 2006 from AAI-DU, Allahabad University, India. He has published more than 20 papers in reputed journals and conferences like IEEE, Springer etc. His research interest fields are Intrusion Detection System, Computer Security, Deep Learning, etc. He has more than 17 years of teaching experience.