A Computer Vision System for Iris Recognition Based on Deep Learning

Shefali Arora
Division of Computer Engineering
Netaji Subhas Institute of Technology
Delhi,India
arorashef@gmail.com

M.P.S Bhatia
Division of Computer Engineering
Netaji Subhas Institute of Technology
Delhi,India
bhatia.mps@gmail.com

***Abstract*—Biometric systems are playing an important role in identifying a person, thus contributing to global security. There are many possible biometrics, for example height, DNA, handwriting etc., but computer vision based biometrics have found an important place in the domain of human identification. Computer vision based biometrics include identification of face, fingerprints, iris etc. and using their abilities to create efficient authentication systems. In this paper, we work on a dataset [1] of iris images and make use of deep learning to identify and verify the iris of a person. Hyperparameter tuning for deep networks and optimization techniques have been taken into account in this system. The proposed system is trained using a combination of Convolutional Neural Networks and Softmax classifier to extract features from localized regions of the input iris images. This is followed by classification into one out of 224 classes of the dataset. From the results, we conclude that the choice of hyperparameters and optimizers affects the efficiency of our proposed system. Our proposed approach outperforms existing approaches by attaining a high accuracy of 98 percent.**

Keywords— Computer Vision, Biometrics, Deep learning, Hyperparameter tuning

I. INTRODUCTION

Biometric systems are becoming promising technologies in identification and verification of a user. The user does not need to carry any physical proofs like passwords, IDs etc. [2]. Iris recognition is being used these days in many applications such as database access, Aadhar cards, financial services etc. It is considered as a reliable source of biometric identification. The advantage of using iris recognition over other traits is that is more reliable and accurate as compared to biometric systems making use of other traits. Iris represents the annular area between pupil and sclera. It is believed that the texture of iris is unique and there is less possibility of two iris patterns being the same. The richness in texture of iris makes it a reliable source of human identification, i.e. presence of crypts, ridges, furrows etc. [3]. Also, iris pattern remains stable for an individual until his death. Therefore, this method of biometric identification is secure and less prone to spoofing attacks. [4]. However, implementing iris recognition systems is a challenging task as it acquires eyelids, eyelashes and reflections which may tamper the efficiency of recognition in a system.

In this paper, deep learning has been used to recognize iris patterns of various users [5]. The technique is explored by using Convolutional neural networks, which make use of filters and learning of weights via backpropagation. An important aspect, known as hyperparameter optimization of network architecture [6] is considered, along with the use of an appropriate filter weights to provide an efficient framework for iris recognition. It has been observed that architecture optimization of Convolutional Neural Networks goes a long way in improving the performance of the biometric system [7].

While hyperparameters like number of filters, strides and padding are given importance in our proposed architecture, filter optimization using backpropagation is analyzed by using different optimization techniques. Understanding of these parameters is effective in improving the system performance. Section 2 describes the work done in iris recognition until now and the use of Convolutional Neural Networks. Section 3 describes our proposed architecture. Section 4 explains the results of experiments and the last section concludes the paper.

II. RELATED WORK

Computer vision is playing a crucial role in a number of applications these days. These include applications relating to surveillance, protection and security from various threats. Computer vision includes acquiring, processing and analyzing images from real world to take decisions. Biometrics is one such domain in which persons are recognized using facial images, images of fingerprints or iris. Thus it combines Computer vision with knowledge of human physiology. Daugman proposed an iris recognition system in 1993[8]. This method required images of high resolution and accuracy decreased in non-ideal conditions.

In [9], authors have worked on four approaches to identify iris patterns. These are Haar Wavelet, Gabor filter, Discrete Cosine Transform and Fast Fourier Transform. Identification using Log Gabor Filter gives the best performance out of these approaches. A statistical feature extraction technique based on Hamming distance and pixel correlation has been devised for efficient iris recognition. [10]

In most of the subsequent iris recognition systems developed, accuracy decreases in non-ideal conditions. Thus performance improvement over existing methods is still an open challenge. The efficiency of such systems is highly dependent on robustness of feature extraction and stages of classification. Use of MLP and neural networks help to attain high accuracy and reliability in such recognition problems [11]. Neural networks are considered to be the most powerful classifiers [12]

978-1-5386-6678-4/18/$31.00 ©2018 IEEE

Yet there are various hurdles in the application of neural networks to such a system. Input images have to go through a number of processing stages like enhancement, segmentation and extraction of features to get a good performance. Deformations in image like translations and rotations cannot be handled well by these networks and a good amount of domain knowledge is required [13]. Hyperparameter tuning is also an important factor to avoid overfitting problems.

Thus deep learning techniques are being used to get rid of such limitations and drawbacks. Using deep learning, a number of levels help to learn features automatically and provide good understanding of images, text etc.

Deep learning has made its mark in various domains like NLP, speech recognition, face recognition [14] etc.

III. CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Networks (CNN) combine various locally connected layers for automated feature recognition. These fully connected layers are followed by classification. The architecture of CNN consists of various layers like pooling layers.

The architecture has made CNN successful in fields like such as image processing and Natural Language Processing. Each neuron obtains input from the region of previous layer equal in size of convolutional filters used. This ensures that strong responses are captured to extract features from the object in question. The CNN layer also applies weights to reduce complexity of the model [15].

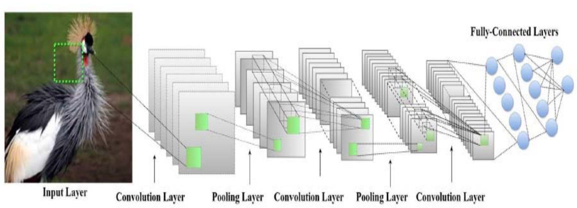

Fig. 1. Convolutional neural network

A. *CNN in Computer Vision*

These days, the use of Convolutional Neural Networks in various applications is justified by the availability of a huge amount of training data and efficient computer hardware. CNNs help to extract features by recognizing visual patterns from pixels of images. CNNs help to learn relationship between inputs and outputs, storing the learned experiences in filter weights. The use of nonlinear activation functions in CNNs is important as they help to clip the value of output attained eg. Sigmoid activation function gives an output between 0 and 1 whereas tanh gives a result between -1 and +1. ReLu or Rectified Linear Unit is the most commonly used activation function.

In every convolutional layer, convolution is performed by using filter weights, which are updated iteratively during backpropagation and minimization of cost function. Thus the operation of convolution followed by nonlinear activation functions modelled as 'Rectified Correlations on a Sphere' [15]. This model has been used in the paper, where ReLu is used to rectify all negative correlations to zero. The RECOS model is explained as followed.

Consider $x=(x_1 \ldots x_n)^T$ as an arbitrary vector on a unit sphere in N dimensional space given as:

$$S = \{x| \; ||x|| = (\textstyle\sum_{n=1}^{N} x_n^2)^{\frac{1}{2}} = 1 \quad (1)$$

Given two points x_i and x_j, the geodesic distance on sphere S can be calculated by finding cosine of angle between them.

$$\theta(x_i, x_j) = cos^{-1}(x_i^T x_j) \quad (2)$$

If correlation is a negative value, then it is not a good measure for calculating geodesic distance between points.

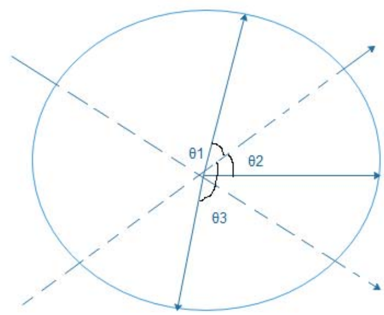

Fig. 2. Correlation rectification in a unit sphere in RECOS model

Given input x and three filter weights f_k (k = 1,2,3) and θ_i as the respective angles between them, a correlation between them can be viewed as a projection from filter weights to input. The calculated value is smaller for larger correlation between the points. $f_1^T x$, $f_2^T x$ will be positive if θ_1 and θ_2 are less than 90 degrees. Whereas correlation is negative if θ value is greater than 90 degrees.

Given k filter weights in N dimensional space, given by $f_k \in R_N$, k=1,2…K. For given input x and K rectified correlations with f_k, a nonlinear mapping from x to output vector is given as:

$$y = (y_1 \ldots y_k \ldots y_K)^T \quad (3)$$

where $y_k(x, f_k) = \max(0, f_k x) = Rec(f_k x)$

The RECOS model can be generalized as follows:

$$S = \{x| \; ||x - \mu|| = (\textstyle\sum_{n=1}^{N}(x_n - \mu)^2)^{\frac{1}{2}} = 1 \quad (4)$$

Here $\mu = \frac{1}{N}\sum_{n=1}^{N} x_n$ which is the mean of all input data points i.e x denotes the N pixel values of an image and μ denotes the mean of all the pixels. Given the input image is large in size, then image is processed in smaller patches and μ represents the local mean.

$$y_k(x - \mu, f_k) = Rec(f_k x + \mu f_{k,0}) \quad (5)$$

and $f_{k,0} = \sum_{n=1}^{N} f_{k,n}$

Consider two filter weight matrices:

$F = [f_1 \ldots. f_k \ldots.. f_K]$ and $K=[k_1 \ldots k_l \ldots k_L]$

Whose columns are filter weights f_k and k_l of two given RECOS units in a CNN architecture. $F \in R^{N \times K}$ and $K \in R^{K \times L}$. For analysis of correlation between them, we assume:

$y = F^T x$ and $z = K^T y$. Then we have:

$$z = K^T F^T x = N^T x \quad \text{where } N = KF \tag{6}$$

In a CNN with many layers, the use of cascaded layers has an advantage since deeper networks perform better in research domains involving pattern recognition.

The application of this CNN model has been implemented in various computer vision problems. Iris recognition is one such research problem, which can be solved using a CNN architecture involving convolutions and Rectified Linear Units. The next section explains the proposed architecture as well as its application in the area of iris recognition.

IV. PROPOSED SYSTEM ARCHITECTURE

Firstly, a preprocessing method is implemented by performing iris localization to detect iris and other features like pupil, eyelids, sclera etc. The main reason of doing this is that defining these areas instead of giving whole eye image as input to the CNN would help to reduce computational overheads. After detection of these regions, iris regions are normalized to allow comparison with other iris images.

These features are employed by CNN for further feature extraction. The feature vectors are further connected to a Softmax classifier, which matches scores of a given user and gives a final probability of recognition of human iris from among the available classes in the dataset. During the training phase, various configurations of CNN are trained on the training set. These are checked on the validation set in order to achieve the best configuration giving the minimum error.

Further, the methodology shows the tuning of filter hyperparameters and optimizers in such a manner that efficiency obtained is the maximum. The proposed architecture is shown in Fig.3.

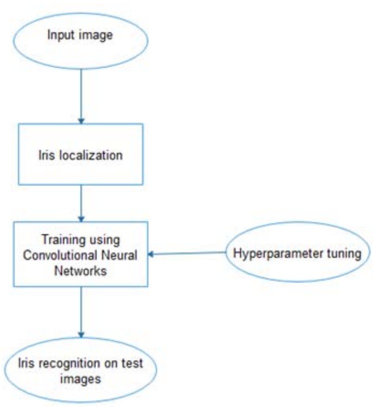

Fig. 3. Overview of proposed Computer Vision System for Iris Recognition

A. Iris Localization

Localization of the iris region plays a significant role to improve the efficiency of the iris recognition system. The iris localization is done to detect the boundary regions of pupil-iris as well as iris-sclera. Presence of eyelids, eyelashes and reflections make it a daunting task though. In this step, inner as well as outer boundaries have been determined by employing Circular Hough Transform, which helps to find center coordinates along with radius of the pupil as well as iris [19]. Eyelid boundaries are detected by eyelid detection algorithm [16].

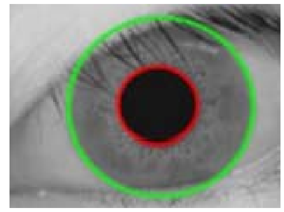

Fig. 4. Localization of iris

The process of circle detection in Hough transform works on the process of voting, in which circles are extracted as features from the image. This has been explained in the algorithm ahead:

Voting method in Circular Hough Transform
For given pixel (a ,b) For theta Ө =0 to 360 for every possible radius a = x – r * cos(Ө* π/180) b = x – r * cos(Ө* π/180) A[a,b,r]+=1 End End

B. CNN for Iris Recognition

The use of iris patterns for recognition was proposed long back by researchers in computer vision. Iris is very small to locate in an image therefore precise focus is needed to extract its details. The iris image might get obscured due to presence of eyelids etc. With emerging trends in technology, computer vision based biometrics has found a strong market today.

In this paper, Convolutional neural networks are trained after localization and normalization of iris images. These images are obtained from IITD iris dataset. The iris images in this database have been taken from students and staff of IIT Delhi under different environmental conditions affecting the iris.

TABLE I. SUMMARY OF IRIS DATASET

Number of classes	224
Number of images	2240(10 for each user class)
Image size	320x240
Image format	bmp

The architecture of the proposed CNN has convolutional layers and max pooling layers. The output from the final fully connected layer is given to Softmax classifier. A cross entropy function is used to define the loss between the predicted class labels and actual class labels.

The main parameters kept in mind while training are:

i) Input image size
ii) Hyperparameter tuning
iii) Data augmentation
iv) Difference between optimizers used for backpropagation

In the given work, 60% of the samples were used for training, 20 % for validation and the remaining 20% for the purpose of testing. Data augmentation is used so that more training data is fed to the network. Validation set is used to generalize ability of the network and storage of configuration of weights that perform best with least validation error.

Training procedure is performed and compared using three optimizers:

i) AdaDelta: It is an extension of AdaGrad and it helps to deal with decaying learning rate problem. This optimization helps to limit the previously accumulated gradients by a size w.

ii) SGD: The traditional batch descent algorithm is very slow. Therefore, SGD optimizer is used to update each training example and thus performs faster. It helps to attain a better local minimum.

iii) Adam: Adam or Adaptive Moment Estimation computes adaptive learning rates for each parameter and thus, give sthe best accuracy out of the three on our dataset.

iv) RMSProp: RMSProp or Root Mean Square Optimizer works on exponential decaying average and not sum of gradients.

In this paper, we focus on batch training. Therefore, RMSProp gives the best results for iris recognition on our dataset.

A four layered network is used and an overall set of all hyperparameters is included in a search space. These hyperparameters help to optimize the architecture of the network, thus increasing the accuracy.

The size of input images is one of the hyper-parameters which is very significant in CNN to affect the accuracy of the proposed network. In this paper, the size of input images is (320x 240) pixels.

Our four layered network has tuned hyperparameters. The three convolutional layers are succeeded by a fourth fully connected layer. These are as shown in the table below:

TABLE II. HYPERPARAMETER TUNING IN CNN

Operation	Hyperparameter	Values
Convolution	Size of filters	3,5,7
	Number of filters	32,64,128
	Fully connected layers	256
Activation	Applied	ReLu
Pooling size	Max Pooling	2,2,2

Using these hyperparameters and Adam optimization, the following steps are followed to implement iris recognition:

The main steps followed in the proposed training architecture are:

1. Division into training, validation and test set.
2. Use the described hyperparameters to train the CNN.
3. Evaluate the configurations using validation.
4. The number of epochs have been set to 100.
5. Repeat these steps through 100 epochs.
6. Choose the configuration giving minimal error using validation.
7. Evaluate this configuration using the test set.

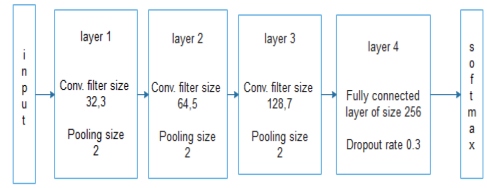

Fig. 5. Hyperparameter tuning in CNN architecture in RECOS model

Dropout is a method which helps to prevent overfitting in training set for a given neural network. The dropout technique has been implemented by ignoring the nodes having probability of 0.3, including their connections [17].

V. RESULTS

The performance of this proposed architecture has encouraging results i.e. an accuracy of 98 % is achieved. The iris localization is considered accurate if the inner and outer boundaries are localized in a correct manner. Accuracy of the proposed system is calculated as follows:

$$\text{Accuracy} = \frac{\text{Correctly localized images}}{\text{Total number of images}} \text{ x } 100$$

The accuracy of our proposed system is 98 percent.

The comparison of results from our proposed technique have been compared with existing state-of-the-art approaches applied on the given iris dataset in the following table:

TABLE III. COMPARISON WITH EXISTING METHODS

Method	Accuracy achieved
Haar Wavelet Kumar and Passi[9]	96.6 %
Log Gabor filter Kumar and Passi[9]	97.2%
Statistical feature extraction Bansal et al.[10]	97.86%
Proposed deep learning approach	98 %

VI. CONCLUSION AND FUTURE WORK

In this paper, we have proposed an iris recognition system using deep learning approach, which authenticates a person's identity. The proposed system by applying localization to iris using Hough transform, followed by automatic feature extraction using CNN.

of the same person. The overall accuracy is increased by hyperparameter tuning and the processing time is reduced.

As a part of future work, we would be working on more real world datasets and compare our achieved results with them. We would also work on advanced techniques like multimodal biometric identification and information flow to increase the accuracy of the proposed system.

ACKNOWLEDGEMENTS

We would like to thank the Biometrics Research laboratory, IIT Delhi for providing access to this database, which is being worked on since the year 2007.

REFERENCES

[1] IIT Delhi iris database, Biometrics Research laboratory,IIT Delhi

[2] B. Connor and K. Roy, "Iris Recognition Using Level Set and Local Binary Pattern", International Journal of Computer Theory and Engineering, vol. 6, no. 5, pp. 416-420, 2014.

[3] "Authentication System For IRIS Biometric Recognition Using Texture Analysis", International Journal of Modern Trends in Engineering & Research, vol. 4, no. 5, pp. 157-162, 2017.

[4] C. Ananth, "Iris Recognition Using Active Contours", SSRN Electronic Journal, 2017.

[5] A. Krizhevsky, I. Sutskever and G. Hinton, "ImageNet classification with deep convolutional neural networks", Communications of the ACM, vol. 60, no. 6, pp. 84-90, 2017.

[6] N. Pinto, D. Doukhan, J. DiCarlo and D. Cox, "A High-Throughput Screening Approach to Discovering Good Forms of Biologically Inspired Visual Representation", PLoS Computational Biology, vol. 5, no. 11, p. e1000579, 2009.

[7] A. Ali and F. Yangyu, "Unsupervised feature learning and automatic modulation classification using deep learning model", Physical Communication, vol. 25, pp. 75-84, 2017.

[8] J. Daugman, "High confidence visual recognition of persons by a test of statistical independence", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 11, pp. 1148-1161, 1993.

[9] "Personal Authentication Based on Iris Recognition", International Journal of Science and Research (IJSR), vol. 5, no. 2, pp. 446-450, 2016.

[10] A. Bansal, R. Agrawal and R. Sharma, "Statistical feature extraction based iris recognition system", Sadhana, vol. 41, pp. 507-518, 2016.

[11] S. Ding, H. Zhu, W. Jia and C. Su, "A survey on feature extraction for pattern recognition", Artificial Intelligence Review, vol. 37, no. 3, pp. 169-180, 2011.

[12] L. Blanger and A. Panisson, "A Face Recognition Library using Convolutional Neural Networks", International Journal of Engineering Research and Science, vol. 3, no. 8, pp. 84-92, 2017.

[13] M. Zeng, L. Nguyen, B. Yu, O. Mengshoel, J. Zhu, P. Wu and J. Zhang, "Convolutional neural networks for human activity recognition using mobile sensors", in 6th international conference mobile computing, applications and services, 2014, pp. 197-205.

[14] H. El Khiyari and H. Wechsler, "Face Recognition across Time Lapse Using Convolutional Neural Networks", Journal of Information Security, vol. 07, no. 03, pp. 141-151, 2016.

[15] C. Kuo, "Understanding convolutional neural networks with a mathematical model", J. Vis. Commun. Image R. ,2016.

[16] A. Al-Waisy, R. Qahwaji, S. Ipson and S. Al-Fahdawi, "IEEE international conference computer and information technology; ubiquitous computing and communications; dependable, autonomic and secure computing; pervasive intelligence and computing", 2015, pp. 1028-1034.

[17] A. Al-Waisy, R. Qahwaji, S. Ipson, S. Al-Fahdawi and T. Nagem, "A multi-biometric iris recognition system based on a deep learning approach", Pattern Analysis and Applications, 2017.

[18] D. Menotti, G. Chiachia, A. Pinto, W. Robson Schwartz, H. Pedrini, A. Xavier Falcao and A. Rocha, "Deep Representations for Iris, Face, and Fingerprint Spoofing Detection", IEEE Transactions on Information Forensics and Security, vol. 10, pp. 864-879, 2015.

[19] S. Srihari and V. Govindaraju, "Analysis of textual images using the Hough transform", Machine Vision and Applications, vol. 2, no. 3, pp. 141-153, 1989.

[20] Y. Imai, K. Kaneko and M. Nakagawa, "Application of a Visual Computer Simulator into Collaborative Learning", Journal of Computing and Information Technology, vol. 14, no. 4, p. 267, 2006.