DOI: 10.21917/ijivp.2021.0370

IRIS DETECTION FOR BIOMETRIC PATTERN IDENTIFICATION USING DEEP LEARNING

M. Ramkumar[1], V. Amirtha Preeya[2], R. Manikandan[3], T. Karthikeyan[4]
[1]*Department of Computer Science and Engineering, HKBK College of Engineering, India*
[2]*Department of Computer Science and Engineering, Presidency University, India*
[3]*Department of Computer Science, The Quaide Milleth College for Men, India*
[4]*Department of Electronics and Telecommunications Engineering, University of Technology and Applied Sciences, Oman*

Abstract

In this paper, we develop a liveness detection of iris present in the study to reduce various spoofing attacks using gray-level co-occurrence matrix (GLCM) and Deep Learning (DL). The input images of iris are been given to this technique for the extraction of texture and colour features with fine details. The details are fused finally and given to a DL classifier for the classification of liveness detection. The simulation is conducted to test the efficacy of the model and the results of simulation shows that the proposed method achieves higher level of accuracy than existing methods.

Keywords:

Iris Detection, Pattern Identification, Liveness Detection, Biometric, Deep Learning

1. INTRODUCTION

When it comes to enhancing user comfort and security, biometric technology has attracted a lot of attention over the last few decades [1]- [9] and is widely employed in numerous applications. In contrast, recent research has shown that biometric recognition systems can be hacked by attackers who give bogus samples to data collectors [1]-[5].

Direct or indirect attacks on a biometric recognition system can be used to identify an unauthorised person using appropriate artificial biometric traits [6]. Therefore, presentation assault detection technologies are needed to defend a biometric recognition system from hackers and increase its security level.

With its high level of security and dependability [7]-[9], the iris pattern has become a popular biometric feature in recent years. Much research has shown that it is possible to create an imitation of an iris pattern by either image or printing an iris pattern onto the lens of a contact lens.

To address this issue, we suggest a new presentation attack detection method for an iris recognition system based on hybrid image characteristics and offer a classification method to overcome the constraints of earlier research. In comparison to past studies, our method is new in five ways.

In this paper, we develop a liveness detection of iris present in the study to reduce various spoofing attacks using gray-level co-occurrence matrix (GLCM) and Deep Learning (DL). The input images of iris are been given to this technique for the extraction of texture and colour features with fine details. The details are fused finally and given to a DL classifier for the classification of liveness detection.

2. RELATED WORK

For iris recognition systems, a number of strategies have been developed for detecting presentation attack images. Expert-knowledge-based (handcrafted) image features are used in some of these investigations, such as the iPAD techniques, whereas learning-based image features are used in other studies.

There were various feature extraction approaches developed by the authors in the first group, all of which were based on their expertise in the field. To identify genuine and presentation attack images, they used classification methods such as support vector machines [10]. Iris images were detected using a variety of local descriptors in this study.

Presentation images can be detected using local descriptors such as the local binary pattern (LBP) and its derivatives, the local phase quantization (LPQ), the binarized statistical image features (BSIF), and the shift-invariant descriptors (SID). It was shown that the detection accuracy varied depending on the feature extraction methods and datasets used, which lowered the detection system overall reliability.

For an iris identification system, the BSIF feature extraction method [11] to detect the textured contact lenses. According to this study, proper segmentation of the iris region is not necessary for accurate detection of findings. It has been shown that eye movement information can be used to assess the liveness of an iris.

Imposters [12] with extensive knowledge can, however, mimic eye movements. In [13] used colour information from distinct channels instead of a gray-textured image to detect a presenting attack ocular image. According to these investigations, the handcrafted image features were useful in detecting presentation attack iris images.

Secondly, the authors use a learning-based strategy on a huge amount of training data to develop a detection model that hides the details of feature extraction and classification. A convolutional neural network (CNN) is a spoofnet framework [13] to identify textured cosmetic contact lenses. Experiments on the Notre Dame Contact Lens (NDCL-2013) dataset indicated that the CNN approach was capable of detecting objects at the highest level of accuracy possible.

This strategy gave findings that were less than optimal when applied to a dataset generated by IIIT-Delhi. This study also employed a very shallow spoofnet (two convolution layers and one fully connected layer). The detection accuracy may be affected by this issue. Structure and filter optimization were

applied to the CNN network [4] in a similar fashion to our research.

Face, fingerprint, and iris biometric traits were tested to see how well they could be recognized. The fingerprint test proved to be a good test for their proposed way of integrating the architectural and filter optimizations. However, their face and iris detection findings were just as good as current state-of-the-art detection results. The CNN networks employed in this study had two convolution layers and a dense layer that was fully coupled. For biometric recognition systems, this research shows that a deep convolutional neural network is good at detecting assault images. This research may be hindered by the lack of training data as well as the adoption of a shallow network architecture.

3. PROPOSED IRIS DETECTION USING GLCM

There are numerous threats to the security of an Iris Recognition System. For highly secure applications, these flaws make a system less trustworthy. Feature-level DL is attempted in this article using the GLCM and DL features of iris images, which can tell if an iris is real or false, respectively. The proposed method removes any preprocessing, such as segmentation and normalisation, commonly used in the literature, making the proposed approach faster and comparably easier. In the proposed method, the only preprocessing step is to resize the iris image to square size. The DL system block diagram is shown in Fig.1. It is proposed that the proposed method be broken down into four phases: iris image scaling, feature generation, classification, and DL classification.

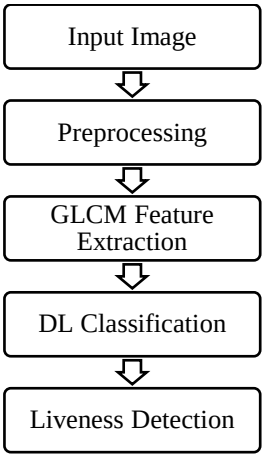

Fig.1. Block diagram of the proposed DL using feature fusion with GLCM features

3.1 RESIZING

In DL, iris pre-processing is critical. There are two iris preparation methods in the proposed technique. Four standard datasets are used to obtain the image, each of which requires a different number of images to be kept. To ensure the integrity of the experiment, the original images are standardised to 128 x 128 pixels during pre-processing. The RGB format is used by some sensors, whereas the grayscale format is used by others. Images capture different datasets using different sensors. In order to preserve the originality of each image, it is converted to grayscale.

3.2 FEATURE FORMATION AND FUSION

GLCM is used on iris images in the proposed method to try feature fusion.

3.2.1 GLCM:

With the use of GLCM, you can obtain statistics on an image gray-level distribution. A scaled iris image is processed using GLCM. In contrast, energy, entropy, and correlation were all calculated using GLCM. Iris image feature extraction using GLCM.

- ***Energy***: Local gray-level consistency, is a measure of energy, which is high in pixels that are in close proximity.
- ***Entropy***: The unpredictability of a image can be described using the image entropy equation. More entropy means it is harder to draw conclusions from your data.
- ***Contrast***: Using contrast, you can tell if a pixel is brighter or darker than its next-door neighbour. Because of the lack of contrast, the GLCM intensity value is quite low.
- ***Correlation***: Gray-level values in the co-occurrence matrix are linearly related to each other.

Four aspects of an image are taken into account: The 10 cross-validation method is used to accurately estimate accuracy.

3.3 CLASSIFICATION AND DETECTION

An ensemble of ML classifiers is used in the proposed DL technique. These classifiers for DL are trained using a tenfold cross-validation approach. Cross-validation tenfold is an excellent method for training machine learning classifiers. As a result of this method, the trained classifier is less likely to be biased because it has access to all of the dataset samples. SVM, naive Bayes, random forest, random tree, and J48 are the ML classifiers used here. Ensembles of a few ML classifiers are also used. A large portion of this is devoted to the creation of ML classifier ensembles by employing voting logic.

Low-level vision uses the iris detection approach to break the system down into smaller subproblems at various scales, each of which is responsible for the residual solution between a coarser and a finer scale. Instead of relying on iris detection, hierarchical basis preconditioning makes use of residual vectors between two scales as an alternative to this. Compared to ordinary solvers, which are oblivious of the residual nature of the solutions, these solvers have been demonstrated to converge substantially more quickly. These approaches imply that the optimization process can be made simpler through the use of preconditioning or reformulation.

An underlying mapping, $H(x)$, can be fit by several layers, with the first layer designating the inputs to $H(x)$. It is analogous to hypothesising that the residual functions, i.e., $H(x)$-x, can be asymptotically approximated if one hypothesises that numerous nonlinear layers can do so. This means that instead of expecting stacked layers to approximate $H(x)$, we expressly allow them to approximate the residual function $F(x):=H(x)-x$. As a result, $F(x)+x$ replaces the original function. Although both forms should be able to asymptotically approximate the necessary functions, the ease of learning may differ.

The degradation problem has prompted this rethinking, which is based on some seemingly counterintuitive observations (Fig. 1,

left). As we explained in the introduction, a deeper model should have no more training error than its shallower counterpart if the additional layers can be created as identity mappings. Identity mappings may be difficult to approximate by several nonlinear layers because of the degradation problem. Solvers can get close to identity mappings by reducing the weights of the various nonlinear layers to zero using the residual learning reformulation if identity mappings are the best option.

Although optimum identity mappings are unlikely to be found in real-world situations, we believe that our reformulation can help to alleviate some of the underlying causes. As long as the optimal function is nearer to an identity mapping than to a zero mapping, finding perturbations should be simpler for the solver than learning the optimal function from scratch. An experiment shows that the learnt residual functions in general have tiny responses, suggesting that identity mappings provide adequate preconditioning.

3.4 IDENTITY MAPPING

Every few stacked layers of the model are used to incorporate residual learning. A building block is defined in this text as:

$$y = F(x,\{W_i\})+ x. \quad (1)$$

where, x and y are the input and output vectors of the layers that are being studied. The residual mapping is represented by the function $F(x,W_i)$. $F= W_{2\sigma}(W_{1x})$ in which σ represents ReLU and the biases are omitted for convenience. A shortcut connection and element-wise addition are used to achieve the operation F+x is adopted as the second nonlinearity after the addition.

In Eq.(1), connections do not add any extra parameters or complexity to the computation. When we compare ordinary and residual networks in practise, this is not only appealing, but also critical. Networks with similar number of parameters, depth, width and computing cost can be compared honestly and objectively.

In Eq.(2), x and F must have the same dimensions as in Eq.(1). W_s can be linearly projected by the shortcut connections to fit the dimensions if this is not the case.

$$y = F(x,\{W_i\})+W_{sx}.$$

A square matrix W_s can also be used in Eq.(1). Identity mapping is sufficient for fixing the degradation problem, and it's more cost-effective than W_s, therefore it's only employed when the dimensions match.

The residual function F can take any shape. A function F with two or three levels is used in the experiments presented in this paper, while additional layers are feasible. In this case, Eq.(1) is analogous to a linear layer, which we have not seen any advantages for. The study should also point out that, for the sake of simplicity, the above notations refer to fully-connected layers, although the same principles apply to convolutional layers as well ($y = W_{1x} + x$). There are numerous convolutional layers that can be represented by the function $F(x,\{W_i\})$. The element-wise addition is performed on two feature maps, channel by channel.

3.4.1 Ensemble Method:

Using more than one model simultaneously on the same set is always preferable to using only one model for classification. Ensemble learning is the name given to this approach. Classifiers are used to train a model, and the ultimate result is an ensemble of such classifiers. This technique uses an ensemble of ML classifiers with a majority voting mechanism.

4. EXPERIMENTAL SET-UP

The studies employed an Intel (R) Core (TM) i3-6006U CPU @ 2.0 GHz, 12 GB of RAM, and a 64-bit operating system with MATLAB platform. Clarkson LiveDet2013, Clarkson LiveDet2015, IIITD Contact Lens, and IIITD Combined Spoofing are the datasets used for experimental explorations of the proposed DL technique.

4.1 DESCRIPTION OF THE DATASET

The four standard and publicly available datasets are Clarkson LivDet2013, Clarkson LivDet2015, IIITD Combined Spoofing Database, and IIITD Contact Lens.

4.2 PERFORMANCE MEASURES

In order to compare the proposed DL method performance measures, accuracy, recall, F-measure, and precision are employed.

$$Accuracy=(TP + TN)/(FP + FN + TP + TN) \quad (1)$$

$$Precision=TP/(FP + TP) \quad (2)$$

$$Recall=TP/(TP + TN) \quad (3)$$

$$F\text{-measures}=2*(Precision* Recall)/(Precision + Recall) \quad (4)$$

4.3 RESULTS

The DL experiments with four benchmark datasets have been presented as a method. Performance criteria such as accuracy, recall, precision, and F-measure are utilised to evaluate the proposed DL approach variants. Classifiers and ensembles of classifiers are trained using these extracted features.

Table.1. Accuracy

Images	VGG 16	VGG 19	DenseNet 121	DenseNet 169	DenseNet 201	ResNet 50
150	0.830	0.823	0.846	0.844	0.851	0.833
300	0.845	0.839	0.862	0.860	0.867	0.849
450	0.869	0.862	0.886	0.884	0.891	0.872
600	0.892	0.886	0.910	0.908	0.915	0.896
750	0.921	0.915	0.939	0.937	0.945	0.925
900	0.954	0.947	0.972	0.970	0.978	0.958

Table.2. Recall

Images	VGG 16	VGG 19	DenseNet 121	DenseNet 169	DenseNet 201	ResNet 50
150	0.818	0.781	0.805	0.830	0.831	0.852
300	0.834	0.796	0.821	0.845	0.847	0.868
450	0.857	0.818	0.843	0.869	0.871	0.892
600	0.880	0.841	0.867	0.892	0.894	0.916
750	0.909	0.869	0.895	0.921	0.923	0.946
900	0.941	0.900	0.927	0.954	0.956	0.979

Table.3. Precision

Images	VGG 16	VGG 19	DenseNet 121	DenseNet 169	DenseNet 201	ResNet 50
150	0.954	0.947	0.972	0.970	0.978	0.958
300	0.921	0.915	0.939	0.937	0.945	0.925
450	0.892	0.886	0.910	0.908	0.915	0.896
600	0.869	0.862	0.886	0.884	0.891	0.872
750	0.845	0.839	0.862	0.860	0.867	0.849
900	0.830	0.823	0.846	0.844	0.851	0.833

Table.4. F-measure

Images	VGG 16	VGG 19	DenseNet 121	DenseNet 169	DenseNet 201	ResNet 50
150	0.823	0.802	0.825	0.838	0.841	0.842
300	0.839	0.817	0.841	0.854	0.857	0.858
450	0.862	0.840	0.864	0.877	0.881	0.882
600	0.886	0.863	0.888	0.901	0.905	0.906
750	0.915	0.891	0.917	0.930	0.934	0.935
900	0.947	0.923	0.949	0.963	0.967	0.968

According to the data, the DL utilising the RF classifier achieves the maximum DL accuracy of roughly 94.16%, followed closely by an ensemble classifier. The highest recognition rates were shown by the emphasised figures.

For the Clarkson 2013 dataset, the DL accuracy with the maximum fusion of DL highest performance with GLCM features and the RF classifier is found to be around 93.78% on average. An accuracy of roughly 95.57% is achieved by combining DL top performance with GLCM features, and doing so with the RF classifier used for the Clarkson 2015 dataset.

When applied to the IIITD Contact dataset, the best DL performance combined with the best GLCM features produced an accuracy of 78.88% using the RF classifier. For the IIITD combined spoofing dataset, the greatest recorded DL accuracy is about 99.68%, combining DL best performance with GLCM features utilising RF and an ensemble of J48 + RF + MLP classifiers.

In the current studies, GLCM and DL have been found to be the most effective approaches for extracting image features. There are a number of image classification applications. Iris presentation attacks have never been assessed before in this way. Feature-level fusion is achieved by combining DL and GLCM features. Although DL has shown promise in the classification of coloured images for many purposes, such as land use identification, gender classification, and so on, it has also shown promising results for the detection of iris presentation attacks. In the experiments, the proposed method was found to be quite effective in the detection of iris spoofing attacks.

In comparison to the most up-to-date state-of-the-art techniques, the feature-level fusion of local GLCM and global DL can distinguish between real and false artefacts and provide better outcomes. Results show that our proposed strategy for detecting presentation attacks in an iris detection system reduces classification errors and improves accuracy compared to previous approaches.

5. CONCLUSIONS

This paper developed a unique technique that relies on a textured lens and print attacks. Attacks are recognised by the proposed method, as well as a variety of iris spoofing attempts utilising different sensors. It has been extensively used in preprocessing, including iris segmentation and normalisation, as well as localisation. With the use of direct extractions from the images, this limitation is addressed. A global DL and a local GLCM feature are used to perform feature-level fusions. These fusion properties of iris images are used to train several ML algorithms and their ensemble combinations. There are four benchmark datasets used to test the proposed liveness detection technique.

REFERENCES

[1] Z. Zhao and A. Kumar, "A Deep Learning based Unified Framework to Detect, Segment and Recognize Irises using Spatially Corresponding Features", *Pattern Recognition*, Vol. 93, pp. 546-557, 2019.

[2] S. Karthick and P.A. Rajakumari, "Ensemble Similarity Clustering Frame work for Categorical Dataset Clustering Using Swarm Intelligence", *Proceedings of International Conference on Intelligent Computing and Applications*, pp. 549-557, 2021.

[3] A. Khadidos, A.O. Khadidos and S. Kannan, "Analysis of COVID-19 Infections on a CT Image using Deep Sense Model", *Frontiers in Public Health*, Vol. 8, pp. 1-18, 2020.

[4] K. Srihari, G. Dhiman and S. Chandragandhi, "An IoT and Machine Learning-based Routing Protocol for Reconfigurable Engineering Application", *IET Communications*, Vol. 23, No. 2, pp. 1-15, 2021.

[5] S.B. Sangeetha, R. Sabitha and B. Dhiyanesh, "Resource Management Framework using Deep Neural Networks in Multi-Cloud Environment", *Proceedings of International Conference on Operationalizing Multi-Cloud Environments*, pp. 89-104, 2021.

[6] H. Proenca and J.C. Neves, "Deep-Prwis: Periocular Recognition without the Iris and Sclera using Deep Learning Frameworks", *IEEE Transactions on Information Forensics and Security*, Vol. 13, No. 4, pp. 888-896, 2017.

[7] H. Proenca and J.C. Neves, "Segmentation-Less and Non-Holistic Deep-Learning Frameworks for Iris Recognition", *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1-8, 2019.

[8] N.V. Kousik and M. Saravanan, "A Review of Various Reversible Embedding Mechanisms", *International Journal of Intelligence and Sustainable Computing*, Vol. 1, No. 3, pp. 233-266, 2021.

[9] I.J. Jacob, "Capsule Network based Biometric Recognition System", *Journal of Artificial Intelligence*, Vol. 1, No. 2, pp. 83-94, 2019.

[10] M. Vatsa, R. Singh and A. Majumdar, "*Deep Learning in Biometrics*", CRC Press, 2018.

[11] V. Maheshwari, M.R. Mahmood, S. Sravanthi and N. Arivazhagan, "Nanotechnology-Based Sensitive Biosensors

for COVID-19 Prediction Using Fuzzy Logic Control", *Journal of Nanomaterials*, Vol. 2021, pp. 1-14, 2021.

[12] S. Umer, A. Sardar and B.C. Dhara, "Person Identification using Fusion of Iris and Periocular Deep Features", *Neural Networks*, Vol. 122, pp. 407-419, 2020.

[13] S. Arora and M.P.S. Bhatia, "Presentation Attack Detection for Iris Recognition using Deep Learning", *International Journal of System Assurance Engineering and Management*, Vol. 8, No. 2, pp. 1-7, 2020.