FPSensorNet: A Deep CNN for Fingerprint Sensors Classification

Abstract—**Voluminous biometric projects acquire data from multiple sensors. Like in UIDAI project of India various fingerprint sensors of different technology and model are used for data acquisition. This heterogeneity within sensors give rise to: sensor inter-operability issues and problems in regulating sequence of commands for law-enforcement. Both of these potential problems can be addressed by automatically identifying the input fingerprint sensor type. In this paper we present FPSensorNet, for fingerprint sensor identification using Deep-Convolutional Neural Network(D-CNN). The proposed approach is computationally efficient and yields high accuracy.**

I. Introduction

Massive biometric projects (like Aadhaar[1]) foment the tremendous growth in sensor technology and design. Heterogenous sensors have been employed in voluminous projects for data acquisition in order to prevent sensor monopoly. In such cases, sensor inter-operability is an issue where images from different sensors are matched using various matching algorithm. When a biometric modality is captured using different sensors, the acquired image although belonging to the same modality, differs in resolution, distortion and size. For example, Fig.1 shows fingerprint images captured from different types of sensors. It is evident from Fig.1, that image quality largely depends upon underlying sensor employed.

Motivation : Fingerprint sensors can be classified into various categories *e.g.* (i) basis of imaging technology they are classified as optical, capacitive and thermal; (ii) basis of user interaction they are classified as press, sweep and non-contacted ones. In such a heterogenous sensor environment, it is crucial to identify the source sensor by which the acquired image is captured. This is essentially required to handle sensor inter-operability issues and further in identifying various attacks on biometric systems, where biometric templates can be modified or mis-used. Another interesting application of sensor identification is in establishing the sequence of commands for law enforcement for identifying spurious activity in online systems. An image can be altered or fabricated during the acquisition phase, transmission or during storage. In order to understand whether the image has been fabricated or not it is necessary to know the source that generates the image.

Related Work : Automatically recording the command of actions for forensic and law reinforcement is a less explored research field while a lot of work has been done for solving fingerprint sensor inter-operability issues. The existing sensor identification techniques can be grouped into three main categories based on : (i) hand crafted features, (ii) sensor pattern noise and (iii) colour filter.

Fig. 1. Example of fingerprint images taken from different sensors (a) Futonic, (b) Lumidigm, (c) SecuGen

Author	Significant Contribution
Ross & Jain [8]	Optical versus Solid State
Bartlow [4]	Photo Response Non Uniformity
Modi [9]	False non match rate, minutia count
Jia [10]	Cross Database(Fingerpass)
Lungini [11]	Optical fingerprint sensor interoperability
Agarwal [5]	Combining handcrafted features
Debiasi [12]	Multiple PRNU enhancements

TABLE I
Summarized sensor classification literature review

Bayrem et al. [2] proposed a method for sensor identification based on measuring the interpolation artifacts occurred in image using color filter arrays. Lukas et al. [3] proposed a technique in which sensor is identified by measuring the pixel non uniformity (PNU) noise of each image using wavelet based denoising. Further Barlow et al. [4] used a variant of PNU technique known as photo response non-uniformity ($PRNU$) for fingerprint sensor identification. Agarwal et al. [5] used handcrafted features that includes features based on entropy, texture, image quality and statistics for sensor recognition. Recently, Sudipta et al. [6] identified sensors from NIR iris images. In their work they have reported that enhanced Sensor Pattern Noise Scheme (SPN), works better for detecting image sensor than maximum likelihood and phase based SPN methods. Uhl and Holler [7] have also used $PRNU$, to identify NIR iris sensor from their images. Table I, summarizes the related work done in fingerprint sensor classification.

Contribution : In this paper, we have proposed a Deep-Convolutional Neural Network (D-CNN) architecture for fingerprint sensor classification inspired from ResNet50 [13]. The main contribution of this paper is as follows:

1) An architecture based on Deep Convolutional Neural Network is proposed that is capable of detecting input fingerprint sensor by systematically pruning and training

ResNet-50.

2) Proposed network has been evaluated on three publicly available databases *viz.* : FVC 2002 [14], FVC 2004 [15] and FVC 2006[16]. It has also been rigorously test over IITK dataset which is largest available multi-sensor, single and 4-Slap fingerprint dataset with more than $40,000$ images.
3) We have trained faster RCNN [17] (from scratch) over IITK data to localize exact finger print regions from 4-slap images.

Rest of the paper has been organized as follows: Section 2, explains in detail the proposed Deep CNN architecture. Section 3, discuss the databases and testing protocol. Section 4, discuss four slap fingerprint ROI extraction and sensor detection. Section 5, provides the experimental results. Finally, layer specific feature analysis has been presented in Section 6 while Last Section concludes our paper.

II. PROPOSED FPSENSORNET

Our model is inspired by ResNet50 [13] architecture. We took pre-trained ResNet50 model on ImageNet images and performed extensive experimentation over it using multi-sensor fingerprint in order to fine tune it. The existing ResNet50 model consists of five main branches (out of which we have dropped fifth branch) as described below.

- Branch-1 is the initial one that convolves the input image of size $224 * 224$ with 64 kernels and gives an output image of size $112*112$. After that max-pooling has been done to avoid over-fitting and to reduce the amount of parameters and in turn computation. The max-pooling layer gives an output of size $55 * 55$.
- Branch-2 consists of three sub-branches $branch\_2a$, $branch\_2b$ and $branch\_2c$. At the end of branch2 an output of size $55 * 55$ has been obtained.
- Branch-3 of ResNet50 model consists of four further branches branch_3a, branch_3b, branch_3c and branch_3d. At the end of branch3 an output of size $28*28$ is generated.
- Branch-4 of ResNet50 model consists of five further branches namely branch_4a, branch_4b, branch_4c, branch_4d, branch_4e and branch_4f and the end output of size $14 * 14$ is generated.

ResNet Pruning : ResNet50 is a very deep network with around 170 layers and trained over 1000 ImageNet classes. Since we have less number of classes and that too with few thousands of images, we have decide to prune the network systematically in order to avoid over-fitting as well as reducing training time. It is also famous for its notorious training hence extensive experimentation has been done to $fine-tune$ as well as systematically prune existing ResNet50 model. During experimentation we found that $Branch-2$ and $Branch-4$ of ResNet50 are extremely important for learning discriminative information, which is necessary for differentiating between images acquired from different fingerprint sensors. We also have observed that $Branch-5$ and $Branch-3$ have "similar" contribution in final classification. Hence one can drop anyone layer, but dropping both have caused drastic performance deterioration. Hence we have dropped $Branch-5$ because generated feature map size was only $7*7$. Fig. 2 shows the proposed FPSensorNet network architecture, which has been designed in a manner so that it can classify commonly used fingerprint sensors. In this model, we have dropped $Branch-5$, since in our case this branch was not learning much discriminative information. By doing so, we have decreased the computation time while retaining the performance.

ResNet Description and Parameterization [13] : The first layer of proposed architecture is an input layer which takes an input image of size $224*224$. Given the fingerprint image and its corresponding label, the first convolutional layer $CONV1$, filters the input image of size $224*224$ using 64 kernels and pool it to an output of size $112*112$. The filters of $CONV1$ layer generally detects the edges and colors of the input image (as shown in Fig 6). The edges are detected at various angles. The output of the $CONV1$ is connected to max-pooling layer. After that $Branch-2$, $Branch-3$ and $Branch_4$ blocks of ResNet50 architecture has been utilized. At the end a fully connected layer with 2048 neuron has been added along with a dropout layer in the model with a probability of 0.4, in order to avoid over-fitting as well as to force the network to learn only robust features. Finally $softmax$ function has been used to generate the probability distribution by minimizing the categorical cross-entropy loss-function. While fine tuning this network we have used $adam$ optimizer with a mini-batch size of 64, initial learning rate has been set as 0.001 for 30 epoches. All the parameters that are used in this work are calculated empirically over a small validation set.

III. DATABASE AND TESTING PROTOCOL

In this section, we provide the details about the database and the testing protocol used in this work. We have tested our proposed architecture on publicly available databases as well as over largest available IITK database which has more than $40,000$ images. We have used single fingerprint images as well as four slap fingerprint images. The IITK dataset of single fingerprints consists of $41,129$ images collected using three different types of sensors *viz.*, (i) Futronic (FS88H), (ii) Lumidigm V310 (V31X) and (iii) SecuGen Hamster I. All of them are of same 500 DIP, but the light source and the image size generated is different. The IITK dataset of four slap fingerprints consists of $26,215$ images collected using three different types of sensors *viz.*, (i) Crossmatch, (ii) Sagem and (iii) Morpho. Proposed architecture has also been tested upon three publicly available single fingerprint databases *viz.*, (i) FVC2002 [14], (ii) FVC2004[15] and (iii) FVC2006 [16]. Few sample images of these datasets are shown in Fig.3. All FVC dataset consists of images collected from four different types of sensors. The detail description about the dataset considered has been given in Table II.

For our work we have considered only the $set-A$ of FVC datasets. More information about these databases can be obtained from the reference papers [16], [14], [15]. We have

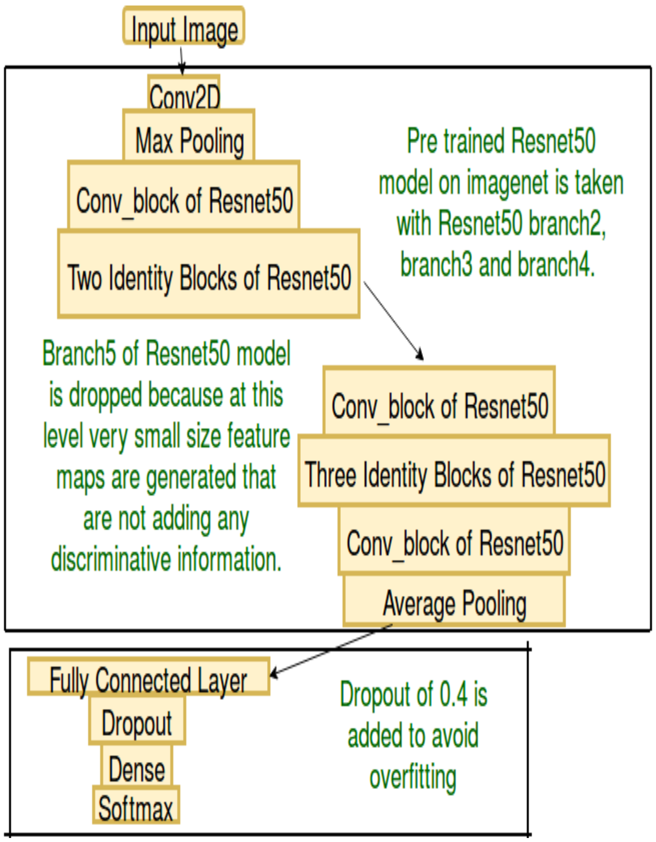

Fig. 2. Block diagram of the proposed FPSensorNet Architecture.

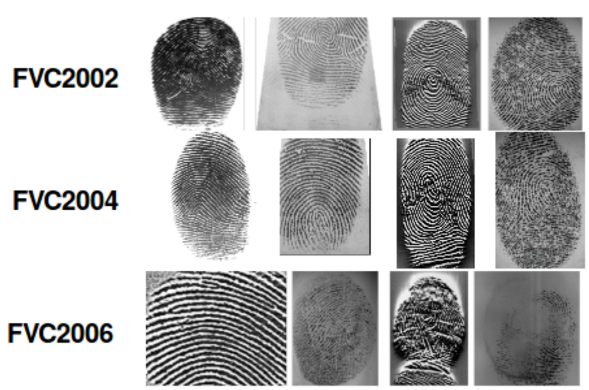

Fig. 3. Sample images from FVC dataset.

trained the proposed model by taking only 10% of the single finger print images as training set and rest 90% as testing set, thus adopting a very difficult protocol.

IV. Four Slap Fingerprint Experimentation :

Re-Training Faster R-CNN [17], (FPSegNet) : We have used one state of the art localization network *viz.*, Faster R-CNN [17] that is based on region proposals for extracting the ROI from four slap fingerprints. The ground truth for finger bounding boxes have been generated using method suggested

Database	Sensor Model	Images
FVC 2002 [14] $(Single Finger)$	Optical Sensor "TouchView" Optical Sensor "FX2000" Capacitive Sensor "100SC" Synthetic fingerprint generation	3,200
FVC 2004 [15] $(Single Finger)$	Optical Sensor "V300" Optical Sensor "U4000" Thermal Sweeping Sensor Synthetic fingerprint generation	3,143
FVC 2006 [16] $(Single Finger)$	Electric Field Sensor Optical Sensor Thermal sweeping Sensor SFinGe V3.0	6,720
IITK-S Dataset $(Single Finger)$	Futronic FS88H Lumidigm V310(V31X) SecuGen Hamster I	41,129
IITK-F Dataset **(Four Slap Finger)**	Crossmatch Sagem Morpho	26,215

TABLE II
Detail Database Specifications

in [18] which was around 92% accurate. Fig. 5 shows the ROI extracted on Four Slap fingerprint dataset. Extensive experimentation has been done in order to select the appropriate threshold and other network parameters for extracting the correct ROI. We have trained Faster R-CNN from scratch, for four slap fingerprint as well as over around $20,000$ iris images and named the resultant network $FPSegNet$. Fig. 4, shows the graph that can summarize the performance of the the trained localization network where x axis is threshold applied over IOU (Intersection over Union) computed between predicted and actual bounding box, while y axis represents the accuracy at some threshold.

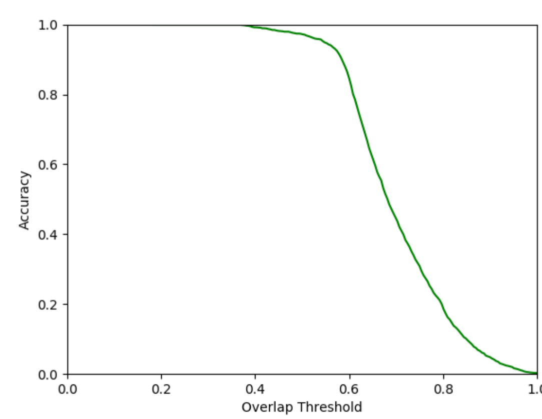

Fig. 4. Accuracy Graph

Sensor Detection : For detecting the sensor of four slap fingerprints we have used the same proposed architecture (as discussed previously). We fed a four slap fingerprint image (sized 1672×1572) to our trained model as well four single fingerprint images (extracted using above trained

$FPSegNet$). These four single fingerprint images has now been fed to the network. Finally $softmax$ function has been used to generate the probability distribution. Each single fingerprint have probability distribution corresponding to its sensors. In order to predict the corresponding sensor for four slap fingerprint the aggregated sum of all probabilities has been calculated and maximum among them is assigned as its label class.

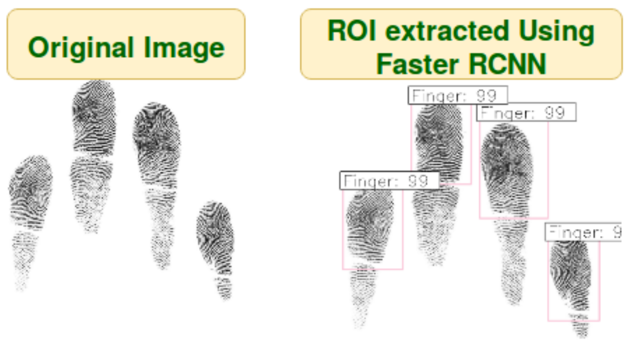

Fig. 5. Four Slap Fingerprint (a) Original image, (b) Corresponding ROI extracted using FPSegNet.

V. Experimental Results and Discussion

In this section we present the classification results of the proposed $FPSensorNet$. The results are computed for two scenarios namely (a) Intra sensor classification (b) Multi sensor classification [19]. The Correct Classification Rate (CCR%) is computed for performance evaluation, higher the value better is the result.

A. Intra-sensor Classification

In this type of classification training and testing is performed on images acquired from same type of sensor model. Fig. 6, visualize the proposed architecture weights from obtained at different convolutional layer when trained on intra-sensor images. One can observe from Fig.6, that the activation output of the first convolutional layer looks like a dense blob, the main reason behind this is that the first convolutional layer looks directly at the raw pixels which makes it more interpretative. As the network goes on deeper and deeper the activation maps obtained at higher convolutional layers get localized and sparsed which makes it less interpretative. It can be seen from activation maps of various convolutional layers that different blobs and different orientations are learned for each sensor which serves as distinct feature required to identify a particular sensor.

Table III, indicates the computed performance of the proposed $FPSensorNet$ on different datasets. Based on the obtained results, following are the main observations. The proposed architecture yields quite good results. The CCR(%) rate is quite high for all the datasets considered except FVC 2004-DB3.

Database	Classification Type	FingerNet
FVC 2002 [14]	DB1-DB1	100
	DB2-DB2	96.79
	DB3-DB3	99.72
	DB4-DB4	99.45
	Aggregate	**98.99**
FVC 2004[15]	DB1-DB1	100
	DB2-DB2	100
	DB3-DB3	93.78
	DB4-DB4	99.17
	Aggregate	**98.23**
FVC 2006[16]	DB1-DB1	99.87
	DB2-DB2	99.28
	DB3-DB3	99.80
	DB4-DB4	100
	Aggregate	**99.73**
IITK S Dataset	Futronic-Futronic	99.34
	Lumidigm-Lumidigm	100
	SecuGen-SecuGen	100
	Aggregate	**99.78**

TABLE III
INTRA-SENSOR QUALITATIVE PERFORMANCE IN CCR(%) ON PROPOSED ARCHITECTURE (WHERE IITK S DATASET IS SINGLE FINGERPRINT DATASET)

B. Multi-sensor Classification

In multi-sensor classification, data is fused together from various fingerprint sensors. We have merged data from all FVC datasets *i.e.* FVC 2002[14], FVC 2004[15] and FVC 2006[16]. The combined dataset consists of $13,063$ images resulting from 12 different sensors and trained our network with 12 output neurons. We have trained our proposed model on 1306 images and tested on remaining $11,757$ images (*i.e.* 10% training and 90% testing). Table IV indicates the computed performance in terms of CCR(%). We have observed that the obtained results are quite remarkable.

Database	Classification Type	CCR (%)
FVC Combined	FVC2_DB1-FVC2_DB1	99.29
	FVC2_DB2-FVC2_DB2	98.37
	FVC2_DB3-FVC2_DB3	99.57
	FVC2_DB4-FVC2_DB4	99.72
	FVC4_DB1-FVC4_DB1	99.16
	FVC4_DB2-FVC4_DB2	100
	FVC4_DB3-FVC4_DB3	99.40
	FVC4_DB4-FVC4_DB4	97.47
	FVC6_DB1-FVC6_DB1	100
	FVC6_DB2-FVC6_DB2	100
	FVC6_DB3-FVC6_DB3	98.56
	FVC6_DB4-FVC6_DB4	100
	Aggregate	**99.29**

TABLE IV
MULTI-SENSOR PERFORMANCE OF $FPSensorNet$ (WHERE FVC2 IS FVC2002, FVC4 IS FVC2004, FVC6 IS FVC2006)

Four Slap Fingerprint Results : We have used the same network for predicting the sensor of the Four Slap Fingerprint image. In total we have $26,215$ images of four slap fingerprints, out of these, when we have used $15,729$ images as whole for training the network, obtained CCR(%) are as follows Crossmatch-Crossmatch 99.98, Sagem-Sagem 96.82, Morpho-Morpho 80.9. In the second set of experiment,

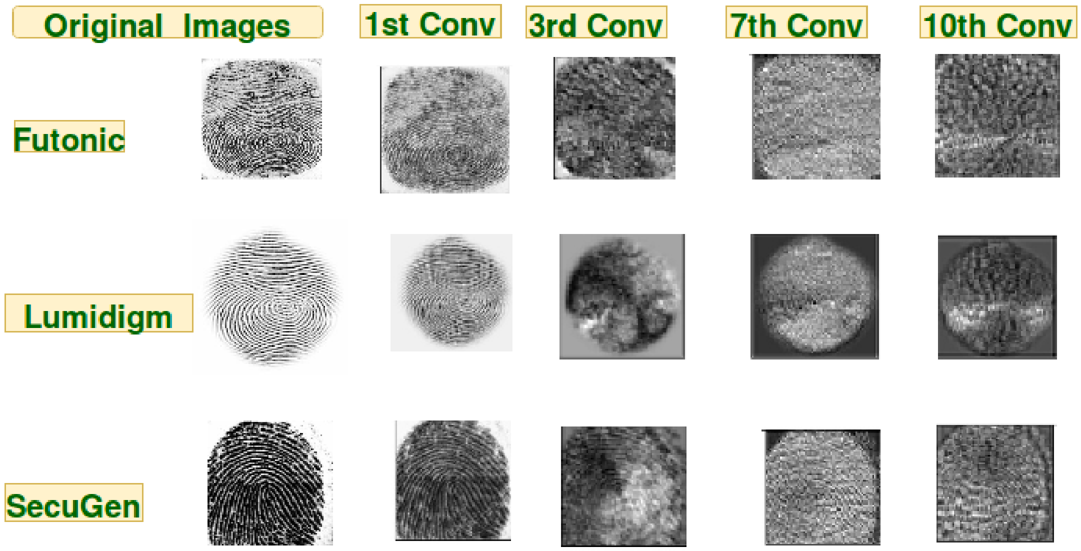

Fig. 6. Visualization of different convolutional layers on proposed architecture for fingerprints acquired using (a) Futonic (b) Lumidigm (c) Secugen sensors

we have used only 4080 images for training the network. In this case we have only shown the four single fingerprint images (extracted using Faster $FPSegNet$), corresponding to the single four slap finger print for training the network. In this case the CCR(%) are as follows, Crossmatch-Crossmatch 100, Sagem-Sagem 94.76, Morpho-Morpho 65.2. In the first set of experiments we are getting low CCR(%) for Morpho due to big re-sizing, one have to do before feeding it to our network (from $1672 * 1572$ to $224 * 224$), hence a lot of discriminative information got lost during size reduction. In order to address this we have consider the second set of experimentation but due to the restricted computational resources at the dispose we have to train our network only over 1360 images per sensor.

Comparative Analysis : In [5], fingerprint sensor classification has been done *via.* combination of handcrafted features. They have combined four databases namely FVC 2002, FVC 2006, IIIT-D MOLF and CASIA Cross Sensor Fingerprint dataset. In total they have $29,320$ images out of that they have used 3000 images for training and remaining for testing. They have achieved an accuracy of about 96.52%. In our experimentation we have also considered training on 10% dataset and testing on remaining 90% dataset. We have also considered FVC 2004 dataset which is quite challenging, and four- slap finger-print dataset of IIT-K . In all the cases we have achieved an aggregated accuracy of 98% and above which is more than the available state-of the art techniques. To the best of our knowledge this is the first attempt in which Deep Convolutional Neural Network is used for identifying the sensor of the underlying fingerprint image.

VI. Layer Specific Feature Analysis

Fig 7 shows the layer specific feature analysis of Lumidigm fingerprint sensor. It is clearly evident from the Fig 7, that initial convolutional layers are learning general specific features, while as we go deeper and deeper more sensor specific features, localized and sparse features are learned. Features learned by $Conv - 2$ layer are very basic features, but as we move deeper in the network more specific learning has been performed like in $Conv - 9$ layer its has learned sector significance information.

Interestingly we have observed that our network automatically learned the state-of-the-art $compcode$ [20], and its closely related features at $Conv - 31$ layer. It can be observed that oriented gabor filter like features are learned at different orientations [21]. The lower layers of the network are learning high level aggregated discriminative features, as shown in Fig 7 for $Conv - 83$ and $Conv - 169$ features. At the lower layers of the network, the resolution and features becomes mostly an encoding of few discriminative intrinsic information. It has been observed that apart from few generic layers that are learning some general features there are few sensor specific layers that use to learn very clear and highly discriminative sensor specific data representation.

VII. Conclusion

The main contribution of this paper is the proposed novel deep architecture for fingerprint sensor classification termed as $FPSensorNet$. To the best of our knowledge this is the first attempt in which deep convolutinal neural network architecture has been used for fingerprint sensor classification. Extensive experiments have been performed on IITK dataset and over three publicly available datasets namely - FVC 2002, FVC 2004 and FVC 2006. The proposed architecture yields correct classification rate above 98% and requires less images to train as well as very less computation time which makes it ideal for real-time applications.

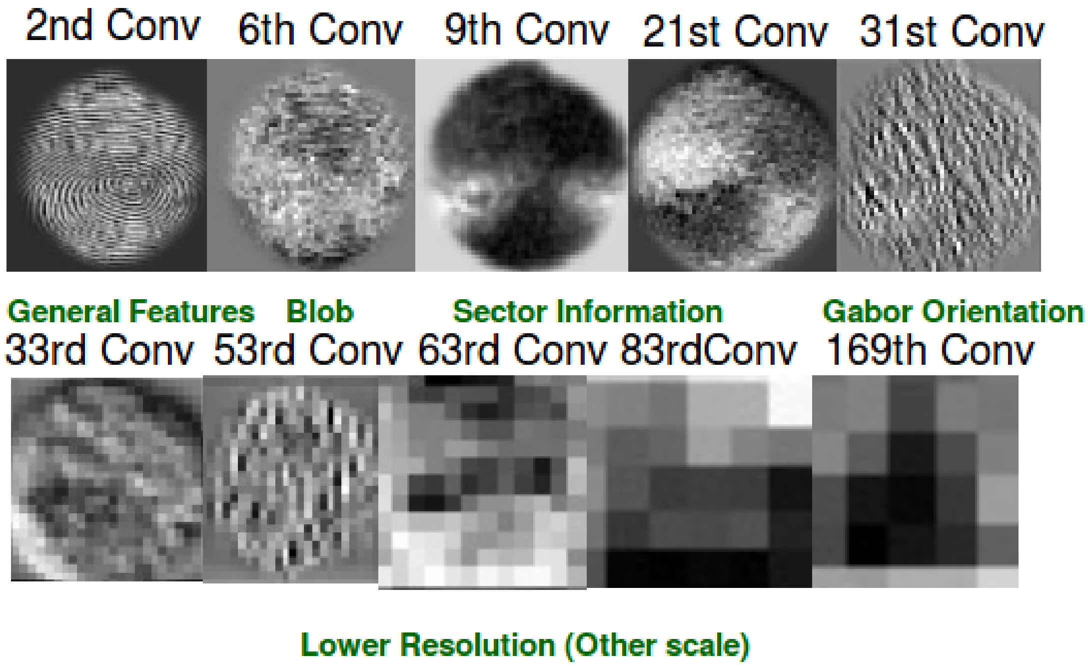

Fig. 7. Features Learned by different Convolutional Layers of Lumidigm Sensor, the source image of Lumidigm Sensor Figure 6 column 2

REFERENCES

[1] "Unique identification authority of india."

[2] S. Bayram, H. Sencar, N. Memon, and I. Avcibas, "Source camera identification based on cfa interpolation," in *IEEE International Conference on Image Processing*, vol. 3, 2005, pp. 69–72.

[3] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, pp. 205–214, 2006.

[4] N. Bartlow, N. D. Kalka, B. Cukic, and A. Ross, "Identifying sensors from fingerprint images," in *IEEE Conference on Computer Vision and Pattern Recognition, CVPR Workshops , Miami, FL*, 2009, pp. 78–84.

[5] A. Agarwal, R. Singh, and M. Vatsa, "Fingerprint sensor classification via mélange of handcrafted features," in *23rd International Conference on Pattern Recognition, ICPR*, 2016, pp. 3001–3006.

[6] S. Banerjee and A. Ross, "From image to sensor: Comparative evaluation of multiple prnu estimation schemes for identifying sensors from nir iris images," in *5th International Workshop on Biometrics and Forensics (IWBF), (Coventry, UK),2017*.

[7] A. Uhl and Y. Hller, "Iris-sensor authentication using camera prnu fingerprints," in *5th IAPR International Conference on Biometrics (ICB)*, 2012, pp. 230–237.

[8] A. Ross and A. K. Jain, "Biometric sensor interoperability: A case study in fingerprints," pp. 134–145, 2004.

[9] S. K. Modi, S. J. Elliott, and H. Kim., "Statistical analysis of fingerprint sensor interoperability performance," *In IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, 2009.

[10] X. Jia, X. Yang, Y. Zang, N. Zhang, and J. Tian, "A cross-device matching fingerprint database from multi-type sensors," in *21st International Conference on Pattern Recognition*, 2012, pp. 3001–3004.

[11] L. Lugini, E. Marasco, B. Cukic, and I. Gashi, "Interoperability in fingerprint recognition: A large-scale empirical study," in *43rd Annual IEEE/IFIP Conference on Dependable Systems and Networks Workshop*, 2013.

[12] L. Debiasi and A. Uhl, "Blind biometric source sensor recognition using advanced prnu fingerprints," in *23rd European Signal Processing Conference (EUSIPCO),2015*.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition, CVPR , Las Vegas, NV, USA*, 2016, pp. 770–778.

[14] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain, "Fvc2002: Second fingerprint verification competition," in *16th International Conference on Pattern Recognition,2002*, vol. 3.

[15] D.Maio, D.Maltoni, R.Cappelli, J. Wayman, and A.K.Jain, "Combining multiple matchers for fingerprint verification: A case study in FVC2004," in *13th International Conference, Image Analysis and Processing ICIAP, Cagliari, Italy,2005*.

[16] R. Cappelli, M. Ferrara, A. Franco, and D. Maltoni, "Fingerprint verification competition 2006," in *Biometric Technology Today,2007*, vol. 15.

[17] S. Ren, K. He, R. B. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1137–1149, 2017.

[18] N. Singh, A. Nigam, P. Gupta, and P. Gupta, "Four slap fingerprint segmentation," in *8th International Conference, Intelligent Computing Theories and Applications ICIC, Huangshan, China*, 2012, pp. 664–671.

[19] R. Raghavendra, K. Raja, and C.Busch, "Contlensnet: Robust iris contact lens detection using deep convolutional neural networks," in *IEEE Winter Conference on Applications of Computer Vision, WACV*.

[20] Zhang, L. Zhang, D. Zhang, and H. Zhu, "Ensemble of local and global information for finger-knuckle-print recognition," *Pattern Recognition*, pp. 1990–1998, 2011.

[21] L. Zhang, L. Zhang, D. Zhang, and H. Zhu, "Online finger-knuckle-print verification for personal authentication," *Pattern Recognition*, pp. 2560–2571, 2010.