Tattoo Detection Based on CNN and Remarks on the NIST Database

[1,2]Qingyong Xu, [1]Soham Ghosh, [1]Xingpeng Xu, [1]Yi Huang, and [1]Adams Wai Kin Kong
[1]School of Computer Science and Engineering, Nanyang Technological University, Singapore
[2]Department of Computer, Nanchang University, China
{n1509179e, soham002}@e.ntu.edu.sg, {xpxu, yhuang, adamskong}@ntu.edu.sg

Abstract

Detecting tattoo images stored in information technology (IT) devices of suspects is an important but challenging task for law enforcement agencies. Recently, the U.S. National Institute of Standards and Technology (NIST) held a challenge and released a tattoo database for the commercial and academic community in advancing research and development into automated image-based tattoo recognition technology. The best tattoo detection result in the NIST challenge was achieved by MorphoTrak with accuracy of 96.3%. This paper aims to answer three questions. 1) Is the NIST database suitable for training algorithms to detect tattoo images stored in IT devices of suspects? 2) Can convolutional neural networks (CNNs) outperform the MorphoTrak's algorithm? 3) How do training databases impact on tattoo detection performance? The NIST tattoo detection database containing 2,349 images and a database containing 10,000 collected from Flickr are utilized to answer these questions. The Flickr images taken in diverse environments and poses are used to simulate images stored in the IT devices. A CNN is trained on the NIST and Flickr images for this study. The experimental results demonstrate that the CNN outperforms the MorphoTrak's algorithm by 2.5%, achieving accuracy of 98.8% on the NIST database. When the CNN is trained on the NIST database to detect Flickr images, the accuracy drops to 65.8%. It implies that the NIST database is not an ideal database for training algorithms to detect tattoo images in IT devices of suspects. However, when the training database size increases, the detection performance improves.

1. Introduction

Searching evidence in IT devices of suspects, e.g., servers, laptops, tablets and smartphones, is an essential but challenging task for law enforcement agencies because of the importance of the evidence and the huge and increasing storage capacity of the devices. Even in one single case, data seized by law enforcement agencies to be investigated can be over 120 terabytes [1]. It is impossible to check all the data manually. Face and pornographic material detectors have been a part of many computer forensic tools for processing image and video data. They detect face and pornographic images (videos) from other images (videos). Tattoos, being widely used by law enforcement agencies, are a vital forensic trace for criminal and victim identification. It is estimated that 45 million Americans have tattoos [16]. However, research on tattoo detection was neglected by the academic community.

Text based searching methods are widely used by many law enforcement agencies to retrieve tattoo records based on witness' descriptions. To apply these methods, law enforcement officers manually annotate registered tattoo images, which can be collected from prisoners, gangsters or offenders. Some standards have been established for this annotation, but some countries do not employ any standard. To match tattoo images collected from crime scenes with tattoo images in a given database, researchers applied content-based image retrieval (CBIR) techniques. Jain et al. used CBIR techniques based on low-level image features e.g., color, shape and texture, and a histogram intersection method for representation and matching [2]. Acton and Ross utilized active contour for tattoo segmentation and proposed a global image feature approach to improve matching performance [3]. Lee et al. adopted the concept of visual similarity for image retrieval and examined the use of SIFT features for matching [4]. Lee and Jain pinpointed that though the local descriptors like SIFT can provide good matching performance, they are difficult to be applied to large scale retrieval problems directly [5]. Therefore, they proposed an ensemble ranking method to achieve more accurate retrieval results for large scale databases. It combines the ranks from multiple bag-of-words models. Another application of content-based tattoo image retrieval is sketch-to-image matching. It is for the cases where query images are not available, but witnesses can remember and sketch the tattoos of criminals. Han and Jain [8] employed local invariant features to match tattoo sketches with tattoo images. In addition to tattoo retrieval, researchers studied tattoo segmentation. Allen et al. developed a method which splits each tattoo image into clusters through a bottom-up process [6]. The clusters containing the skin were merged through learning and the tattoo patterns were distinguished from other skin tissues via a top-down prior in the image itself. To segment and classify tattoos in images collected in uncontrolled environments, Heflin et al. [7] introduced a new methodology. They first used a saliency model to find regions of interest and then applied Grabcut and quasi-connected components to perform final segmentation. Wilber et al. proposed exemplar codes for tattoo detection and classification [17]. Huynh et al. [9] noted that prison and police departments do not have an effective way to collect tattoo images. Many processes in this collection are still manual and very time consuming. Huynh et al. developed a full-body imaging system which has an automatic and systematic routine for collecting and processing tattoo images and other biometric traits.

Heflin et al. and Wilber et al.'s works are the closest ones to this study. For each classification component, including tattoo detection, Heflin et al. used 150 images for training, 50 positive samples and 500 negative samples for testing. Their non-tattoo images were collected from dermatology forums and face image databases. Because face detectors were used to remove negative images; how many of these training and testing images without faces were not clearly given and the database is private, the actual tattoo detection accuracy is not clear. In Wilber et al.'s study, 50 segmented butterfly tattoos were used as positive training samples and 100 non-segmented images were used for testing samples. The size of negative samples was 800. In their detection evaluation, only butterfly tattoo images were tested. In other words, these two studies imposed some constraints on either positive or negative images in their evaluations. In IT devices of suspects, images can be very diverse, e.g., pets, scenery, buildings, and tattoo images with different classes on different parts of the body.

Recently, NIST held a challenge for the commercial and academic community in advancing research and development into automated image-based tattoo recognition technology. The challenge aimed to evaluate automatic methods for identifying tattoos, detecting region of interest, matching visually similar or related tattoos using different types of non-tattoo imagery (e.g., scanned print and sketch), matching similar tattoos from different subjects and detecting tattoos from images [10]. For tattoo detection, NIST aimed to evaluate methods classifying whether an image contains a tattoo or not. Four organizations, French Alternative Energies and Atomic Energy Commission, Compass Technical Consulting, MITRE Corporation and Morpho/MorphoTrak, joined the evaluation of tattoo detection. Their short names used in the NIST report [10] and this paper are respectively CEA, Compress, MITRE and MorphoTrak. No academic institute joined the tattoo detection evaluation. The NIST tattoo detection database contains 1,349 tattoo and 1,000 non-tattoo images. The evaluation was "open book", meaning that the participants were provided with the dataset and the ground-truth data and ran their algorithms by themselves. Therefore, results reported in this paper can be compared with the results given by the four organizations directly. The four organizations did not disclose their algorithms used in the evaluation. Their results range from 62.2% to 96.3%. The best result was obtained by MorphoTrak. This paper attempts to answer three questions. 1) Is the NIST database suitable for training algorithms to detect tattoo images in IT devices of suspects? 2) Can convolutional neural networks (CNNs) outperform the MorphoTrak's algorithm? 3) How do training databases impact on detection performance? In addition to the NIST database, 10,000 images are collected from Flickr for this study. Note that in addition to detecting tattoo images in suspects' IT devices, tattoo detection algorithms can be used for database construction and maintenance, which were mentioned in the NIST report [10]. However, due to the lack of public databases in this area, the NIST database is considered for training algorithms to detect tattoos in suspects' IT devices.

The rest of this paper is organized as follows. Section 2 describes the CNN used in this study. Section 3 presents the databases and evaluation protocols. Section 4 reports and discusses the experimental results. Section 5 gives conclusive remarks.

2. The CNN for this study

Convolutional neural networks have a feed-forward network architecture with multiple interconnected layers which may be of any of the following types: convolution, normalization, pooling and fully connected layers. CNNs are chosen as a detector for this study because they outperform other traditional methods in many image classification challenges, such as ImageNet [11] and many other image-based recognition problems, e.g., face recognition and digital recognition [12-13]. Comparing with traditional methods which rely on feature engineering, CNNs are able to learn feature representation through the backpropagation algorithm without the need for much intervention and also achieve much higher accuracy.

The aim of this study is not to propose another CNN but use a CNN to answer the questions given before. Fig. 1 illustrates the architecture of the CNN employed in this study. It is modified from the CNN used by Krizhevsky et al [11] in the ImageNet Challenge 2012. The network includes five convolutional layers and three fully connected layers. The first five convolutional layers have 64, 192, 384, 256 and 256 kernels and their sizes are respectively $11\times11\times3$, $5\times5\times64$, $3\times3\times192$, $3\times3\times384$ and $3\times3\times256$. The first fully connected layer has 4,096 neurons and the second fully connected layer has 2,048 neurons. In the last fully connected layer, there are two neurons, one for tattoo image output and the other for non-tattoo image output. The response normalization processes follow the first and second convolutional layers. The max-pooling with a size of 3 and a stride of 2 follows the response normalization processes and the fifth convolutional layer. Dropouts are

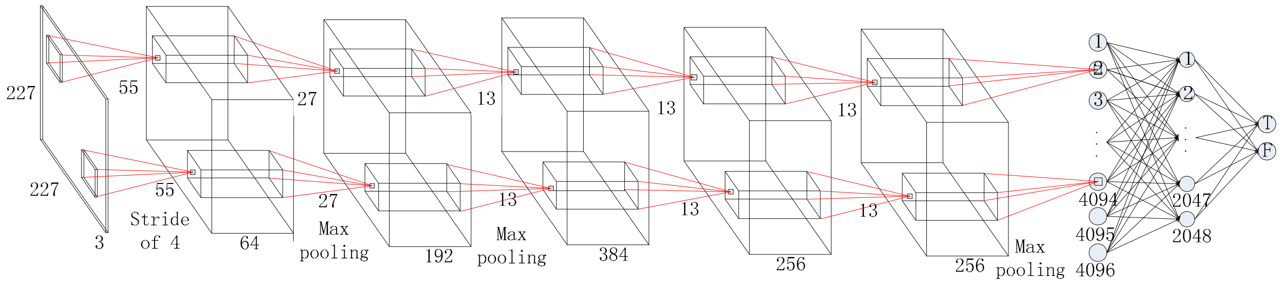


Fig. 1 The architecture of the CNN employed in this study.

done in the first and second fully connected layers and the dropout rate is 0.5. As with Krizhevsky et al.'s [11] network, our network maximizes the multinomial logistic regression objective. The shorter side of the training and testing images is rescaled to 256 pixels and then, the central patch with a size of 227×227 pixels is cropped. The mean of training images is subtracted as a preprocessing step. 128 training images form a batch to optimize the CNN through stochastic gradient descent (SDG). The initialized learning rate is 0.01; the momentum is 0.9 and the weight decay is 0.0005. The number of iteration in each experiment is 1,000. The weights in the CNN are initialized with a zero-mean Gaussian distribution with standard deviation of 0.01. The bias is set to zero in the first and third layers, but it is set to one in the other layers.

3. Databases and evaluation protocols

The NIST tattoo detection database in the challenge consists of 2,349 images, with 1,349 tattoo images and 1,000 non-tattoo images. Their raw sizes range from 271 by 291 pixels to 1,944 by 2,592 pixels. Fig. 2 shows samples of the NIST images. The tattoo images were taken from different body sites of different subjects, while the non-tattoo images were taken from face and upper body. All the NIST images were collected from indoor environments. Many of them are similar to custodies. These images were extracted from the Multiple Encounter Database 2 (MEDSII), which stores images of deceased persons [14].

In addition to the NIST database, 10,000 images with 5,740 tattoo images and 4,260 non-tattoo images were collected from Flickr. The ratio of the tattoo images to the non-tattoo images is same as that of the NIST database. The Flickr images were taken from diverse viewpoints, poses and environments with complex backgrounds, including indoor and outdoor. The raw image sizes range from 72 by 96 pixels to 500 by 500 pixels. Fig. 3 shows samples of the Flickr images. The Flickr images were used to construct four datasets. They are named Flickr(2349), Flickr(3.5K), Flickr(5K) and Flickr(10K). Table 1 lists the details of these datasets. The ratios of the tattoo images to the non-tattoo images in these datasets are the same as that of the NIST database. These Flickr images will be publicly available [15]. The NIST evaluation protocol is also used in this study, so that our results can be directly compared with the results given by the four participants in the NIST challenge. More precisely, a five-fold cross-validation scheme is used and detection accuracy is employed as a performance index.

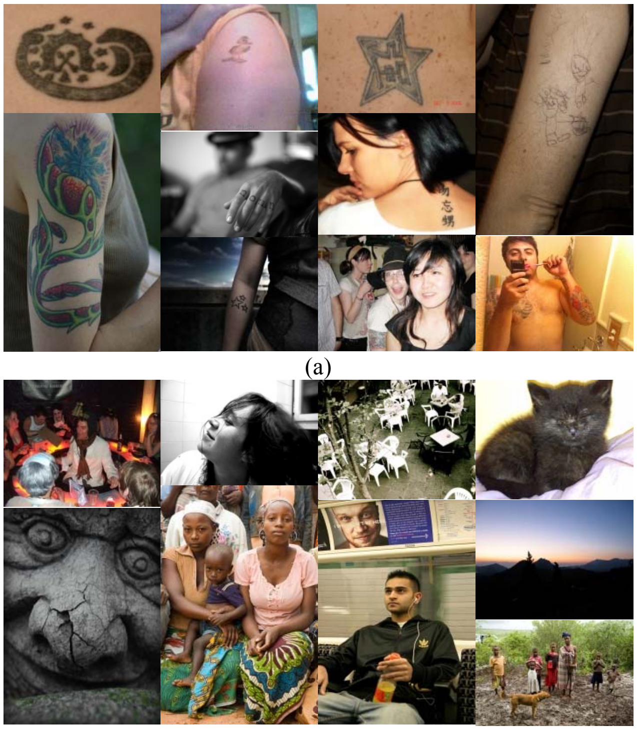
Table 1 Details of the four Flickr datasets.

	Flickr(2349)	Flickr(3.5K)	Flickr(5K)	Flickr(10K)
Tattoo	1349	2010	2870	5740
Non-tattoo	1000	1490	2130	4260
Total	2349	3500	5000	10000

(a)

(b)

Fig. 2 Samples of the NIST tattoo images. (a) Tattoo images and (b) non-tattoo images



(a)

(b)

Fig. 3 Samples of the Flickr images. (a) Tattoo images and (b) non-tattoo images.

4. Experimental results and discussion

Seven sets of experiments were performed to study the performance of the [1]CNN and the impacts of the NIST database and Flickr datasets. The aim of the first experiment was to study the performance of the CNN on the NIST database. The CNN achieved accuracy of 98.81%. It wrongly classified 15 tattoo images as non-tattoo images and 13 non-tattoo images as tattoo images. Fig. 4 shows some misclassified images. Table 2 compares our CNN result with the results given by French Alternative Energies and Atomic Energy Commission, Compass Technical Consulting, MITRE Corporation and Morpho/MorphoTrak. The non-tattoo (tattoo) detection accuracy listed in Table 2 is the percentage of testing non-tattoo (tattoo) images being classified as non-tattoo (tattoo) images. The CNN employed in this study outperforms all the other algorithms in the NIST challenge by at least 2.5%. The NIST challenge was an "open-book" evaluation. Thus, the CNN results can be directly compared with the results given by the four participants. These results demonstrate the performance of CNN for tattoo detection.

In the first experiment, the CNN achieved high accuracy. Does it mean that the NIST images are suitable for training CNNs to search tattoo images in IT devices of suspects? To answer this question, in the second experiment, the Flickr(2349) dataset was used as a testing dataset, while the NIST database was used as a training dataset. In each fold, 80% of the NIST images were used for training, while 20% of the Flickr(2349) images were used for testing. Note that the NIST database and the Flickr(2349) dataset have the same number of images. In this setting, the detection accuracy drops to 65.77%. The Flickr(2349) images are more similar to images stored in suspects' devices because they were taken from diverse viewpoints, poses and environments. This experimental result implies that the NIST database is not the most ideal database to train algorithms to detect tattoo images for the target application.

The second experimental result raises another question. Are the Flickr images more challenging than the NIST images? To answer this question, in the third experiment, the CNN was trained and tested on the Flickr(2349) dataset. The CNN achieved accuracy of 78.20%. Comparing with the results in the first experiment, the accuracy of the CNN dropped over 20%. The performance drop pinpoints clearly that the Flickr images are much more challenging. In other words, the experimental results listed in Table 2, including the results given by the four participants and the CNN, may not reflect the performance of the algorithms used to search tattoo images in IT devices of suspects. The Flickr images are more difficult to be recognized, because they are more diverse. They were collected from different lighting environments and viewpoints. The subjects in the images have very different poses and their backgrounds are also very complex. Thus, the CNN is harder to learn common features from the Flickr images.

For the sake of completeness, in the fourth experiment, the Flickr(2349) dataset was used as a training dataset, while the NIST database was used as a testing dataset. In each fold, 80% of the Flickr(2349) images were used for training, while 20% of the NIST images were used for testing. The CNN achieved accuracy of 83.31%, which is even higher than the accuracy of 78.20% obtained in the third experiment. Note that in the third experiment, the Flickr(2349) dataset was used for both training and testing. This experimental result once again indicates that the NIST images are not as difficult as the Flickr images. Table 3(a) summarizes the results from the first four experiments and Table 3(b) lists their non-tattoo detection accuracy and tattoo detection accuracy. The non-tattoo detection accuracy from the second experiment is significantly lower than that of the other experiments. This result matches our observation that the non-tattoo images in these two datasets are very different. The NIST non-tattoo images cannot represent the Flickr non-tattoo images well. However, using the Flickr(2349) dataset to train the CNN, but using the NIST database to test it, the tattoo detection accuracy is 93.18%, which is much higher than the tattoo detection accuracy obtained from the second and third experiments.

[1] The final trained CNN will be publicly available in [16].

It implies that the Flickr tattoo images can represent the NIST tattoo images.

Table 2. Performance comparison of the CNN and the algorithms of the four participants in the NIST challenge.

Algorithm	Non-tattoo detection accuracy	Tattoo detection accuracy	Overall accuracy
CEA_1	98.8%	93.2%	95.6%
Compass	38.6%	79.8%	62.2%
MITRE 1	75.0%	73.4%	74.1%
MITRE 2	94.8%	92.4%	93.4%
MorphoTrak	95.0%	97.2%	96.3%
CNN	**98.9%**	**98.7%**	**98.8%**

Table 3. The results of the first four experiments. (a) Overall accuracy and (b) non-tattoo detection accuracy and tattoo detection accuracy.

(a)

Experim ents	Training database	Testing database	Accuracy
1	NIST	NIST	98.81%
2	NIST	Flickr(2349)	65.77%
3	Flickr(2349)	Flickr(2349)	78.20%
4	Flickr(2349)	NIST	83.31%

(b)

Experiment s	Non-tattoo detection accuracy	Tattoo detection accuracy	#Accurac y difference
1	98.70%	98.89%	-0.19%
2	43.40%	82.36%	-38.96%
3	74.40%	81.02%	-6.62%
4	70.10%	93.18%	-23.08%

#Non-tattoo detection accuracy subtracts tattoo detection accuracy.

The first four experiments show that when the Flickr(2349) dataset was used for testing, the detection accuracy is not over 78.20%, which is lower than that of the first experiment by 20%. In the fifth, sixth, and seventh experiments, Flickr(3.5K), Flickr(5K) and Flickr(10K) were used as training datasets to evaluate the performance change when the sizes of training datasets increased. The 5 fold cross-validation scheme employed in the NIST challenge and the previous experiments were kept using in these experiments. The Flickr(2349) dataset was a subset of the Flickr(3.5K), Flickr(5K) and Flickr(10K) datasets. The sizes of the training datasets in each fold were respectively 2,800, 4,000 and 8,000 images. The Flickr(2349) dataset and the NIST database were used as the testing datasets. When the Flickr(2349) dataset was used for testing, 20% of its images were employed as testing images in each fold. These images did not overlap with the training images in that fold. The experimental results are given in Table 4 and Fig. 5. Clearly, when the size of training datasets increased, the performance on the both testing datasets improved. However, the NIST images gained much more improvement. Even when the Flickr(10K) dataset was used for training, the performance on the Flickr(2349) dataset was still below 85%. Fig. 6 shows some misclassified images. This result pinpoints that even using 8,000 images to train the CNN is not enough to achieve high accuracy for detecting tattoo images in real operation. Thus, a large training dataset is demanded.

We did not re-implement Heflin et al. and Wilber et al.'s detection methods [7, 17] for comparison because some implementation details in their papers are not very clear. Table 5 summarizes the databases and techniques used in different studies.

Fig. 4 NIST images that are misclassified by the CNN in the first experiment.

Table 4. The performance of the CNN on training datasets with different sizes.

Training datasets	Testing datasets: NIST	Testing datasets: Flickr(2349)	*Performance difference
Flickr(2349)	83.31%	78.20%	5.29%
Flickr(3.5K)	88.04%	80.76%	7.28%
Flickr(5K)	91.06%	82.42%	8.64%
Flickr(10K)	93.78%	84.76%	9.02%

*The accuracy of the NIST testing dataset subtracts the accuracy of the Flickr(2349) testing dataset.

Table 5 A summary of the training and testing databases in different studies

	NIST [10]	*Heflin et al. [7]	*Wilber et al. [17]	This study
Total number of training samples	^Pos: 1349, Neg: 1000	Total: 150	Pos: 50, Neg: 800	Pos: 5,740, Neg: 4,260
Total number of testing samples	Pos: 1349, Neg: 1000	Pos:50, Neg: 500	Total: 100	Pos: 5,740, Neg: 4,260
Remarks	5-fold cross-validation. Images were collected from inner environments. Negative images are faces	Negative images were collected from dermatology forums and face databases [18-19].	All positive images are butterfly.	5-fold cross-validation. No limit on positive and negative samples. Images were collected from Flickr
Major technique	NIL	One class SVM	Examplar Codes	CNN
Public or private	Public	Private	Private	Public

* Only the tattoo detection schemes are considered.
^Pos standards for positive samples and Neg standards for negative samples.

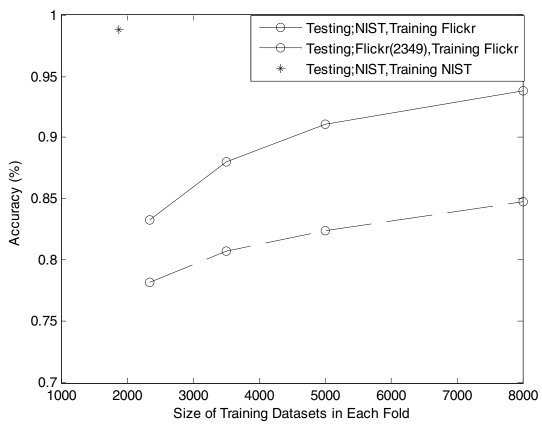


Fig. 5 The performance of the CNN on training datasets with different sizes.

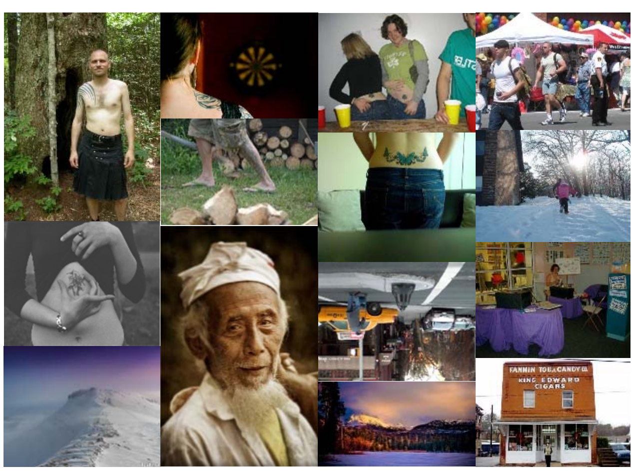


Fig. 6 Some misclassified images, when the Flickr(10K) dataset was used for training.

5. Conclusion and further work

Detecting tattoo images in IT devices of suspects is an important step for searching criminals and victims. However, this research direction was neglected by the academic community. Recently, the U.S. National Institute of Standards and Technology held a tattoo recognition challenge and released databases for the commercial and academic community in advancing research and development into automated image-based tattoo recognition technology. One governmental and three commercial organizations joined the tattoo detection challenge and achieved accuracy in a range from 62.2% to 96.3%. In this study, a CNN is first trained and evaluated on the NIST tattoo detection database and achieves accuracy of 98.8%, 2.5% higher the best result given by MorphoTrak in the challenge. Ten thousand tattoo and non-tattoo images are downloaded from Flickr to further analyze the NIST database and the tattoo detection performance based on CNN. When the CNN is trained on the Flick images with the same number of images as the NIST database, the performance of the CNN drops to 78.20%. It indicates that the Flickr images are more challenging. Using the Flickr images as a training set and the NIST images as a testing set and vice versa, the non-tattoo detection accuracy and the tattoo detection accuracy show that the NIST images, especially the non-tattoo images, are not the most ideal training images for developing algorithms to detect tattoo images taken from diverse backgrounds, viewpoints and poses, as those stored in IT devices of suspects. The NIST images are more suitable for training detection algorithms for database construction and maintenance. In this application, tattoo images are collected in controlled environments, e.g., prisons, with user cooperation [9]. When the size of training datasets increases, the performance of the CNN improves for the both NIST and Flickr testing datasets. However, the rate of improvement of the NIST testing images is much higher than that of the Flickr testing images. It pinpoints that the NIST images are much easier than the Flickr images. When the number of the training images is 400% more than the NIST training images, the CNN achieves accuracy of 84.76% for the Flickr testing images. This study shows that though tattoo detection algorithms can be used for database construction and maintenance and search of tattoo images in IT devices of suspects, they require different training databases to develop effective algorithms. It also indicates that more research effort should be put into this direction. We are going to collect more challenging images to train the CNN, improve its architecture and analyze the relationship between detection accuracy and properties of misclassified tattoo images, e.g., sizes of tattoos in the images

Acknowledgments

The authors would like to thank the U.S. National Institute of Standards and Technology for sharing their databases. This work is partially supported by the Ministry of Education, Singapore through Academic Research Fund Tier 1, RG14/14. The first author is supported by the China Scholarship Council under Grant 201506820025. All his work was done in the Nanyang Technological University, Singapore.

References

[1] T. Frieden, "72 charged in online global child porn", CNN, August 3, 2011 http://edition.cnn.com/2011/CRIME/08/03/us.child.porn.ring/ (access on 26 Dec 2015)

[2] A.K. Jain, J.E. Lee, and R. Jin, "Tattoo-ID: Automatic tattoo image retrieval for suspect and victim identification", *Proceedings of the multimedia 8th Pacific Rim conference on Advances in multimedia information processing*, pp. 256-265, 2007.

[3] S.T. Acton, and A. Ross, "Matching and retrieval of tattoo images: Active contour CBIR and glocal image features", *Proc. of the IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 21-24, 2008.

[4] J.E. Lee, A.K. Jain, and R. Jin, Scars, "marks and tattoos (SMT): Soft biometric for suspect and victim identification", *2008 Biometrics Symposium*, BSYM, pp. 1-8, 2008.

[5] J.E. Lee, R. Jin, and A.K. Jain, "Unsupervised ensemble ranking: Application to large-scale image retrieval", *ICPR*, pp. 3902-3906, 2010.

[6] J.D. Allen, N. Zhao, J. Yuan, and X. Liu, "Unsupervised tattoo segmentation combining bottom-up and top-down cues", *Proc. of SPIE - The International Society for Optical Engineering*, 2011.

[7] B. Heflin, W. Scheirer, and T.E. Boult, "Detecting and classifying scars, marks, and tattoos found in the wild", *BTAS*, pp. 31-38, 2012.

[8] H. Han, and A.K. Jain, "Tattoo based identification: Sketch to image matching", *ICB*, 2013.

[9] N.Q. Huynh, X. Xu, A.W.K. Kong, and S. Subbiah, "A preliminary report on a full-body imaging system for effectively collecting and processing biometric traits of prisoners", *IEEE Workshop on Computational Intelligence in Biometrics and Identity Management*, pp. 167-174, 2015.

[10] M. Ngan, G.W. Quinn and P. Grother, "Tattoo recognition technology — Challenge (Tatt-C) Outcomes and Recommendations", *National Institute of Standards and Technology, U.S. Department of Commerce*, Sept 2015 http://dx.doi.org/10.6028/NIST.IR.8078

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks,", NIPS, pp. 1097–1105, 2012

[12] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning appliedto document recognition", *Proceeding of the IEEE*, pp. 1-46, 1998.

[13] F. Schroff, D. Kalenichenko and J. Philin, "FaceNet: a unified embedding for face recognition and clustering", *CVP*R, pp. 815-823, 2015.

[14] NIST Special Database 32 - Multiple Encounter Dataset 2 (MEDS-II), *NISTIR 7807*. http://www.nist.gov/itl/iad/ig/sd32.cfm

[15] NTU Forensic Skin Database http://forensics.sce.ntu.edu.sg/.

[16] Tattoo statistics, Statistic Brain Research Intitute, http://www.statisticbrain.com/tattoo-statistics/

[17] M.J. Wilber, E. Rudd, B. Heflin, Y.M. Lui and T.E. Boult, "Exemplar codes for facial attributes and tattoo recognition", *IEEE Winter Conference on Applications of Computer Vision*, pp. 205-212, 2014

[18] K. Ricanek and T. Tesafaye. "MORPH: A Longitudinal Image Database of Normal Adult Age-Progression". In *IEEE AFGR*, April 2006.

[19] G. Huang, M. Ramesh, T. Berg, and E. Learned-Miller. "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments", *Technical Report 07-49, University of Massachusetts*, Amherst, October, 2007