2016 23rd International Conference on Pattern Recognition (ICPR)
Cancún Center, Cancún, México, December 4-8, 2016

Tattoo Detection and Localization using Region-based Deep Learning

Zhaohui H. Sun, Jeff Baumes, Paul Tunison, Matt Turek, Anthony Hoogs
Kitware Inc., Clifton Park, NY 12065

***Abstract*—Tattoos have been increasingly used as a discriminative soft biometric for people identification, such as criminal and victim identification in forensics investigation and law enforcement. However, automatic detection of tattoo images and accurate localization of the regions of interest are challenged by the large variations in artistic composition, color, shape, texture, location on the body, local geometric shape (e.g. neck and finger), imaging conditions, and image quality. In this paper, we train a tattoo detector from the Tatt-C and PASCAL VOC 2007 image datasets using region-based deep learning. The detector can effectively determine if an image contains tattoos and the locations of tattoo regions. We carry out a comprehensive evaluation of our tattoo image classification and detection localization. The detector improves upon the state-of-the-art algorithms in the Tatt-C challenge, achieving a better detection error trade-off curve. It yields low confidence scores on randomly sampled non-tattoo images from 397 scene categories in the MIT-SUN dataset. In addition, the same detector is also validated on the NTU Tattoo Image Dataset with 10000 images.**

I. Introduction

Tattoos are a form of permanent skin pigment change by inserting indelible ink into the skin, for unique identification, decoration, and association. It was reported in [1] that about 36 percent of Americans 18 to 29 years old have at least one tattoo. Due to their prevalence, tattoos are increasingly being used as a soft biometric for person identification, especially in forensics analysis and law enforcement. It provides discriminative information that is complementary to the other biometric modalities, such as fingerprint, face, and iris, and helps identify association, group affiliation, membership, gangs, criminals and victims.

Automatic tattoo image analysis has long been studied in the literature. The tasks include detection of tattoo images from an unorganized collection, localization of tattoo regions in an image, similarity matching of tattoo patterns (e.g. same tattoo over time, or images containing the same tattoo design), tattoo image search and indexing [2], [3], and tattoo image search from sketch [4]. Recently, the Tatt-C challenge was organized by NIST and a common dataset, including ground truth, was shared to foster research and development [5], [6], [7]. In addition, the NTU Tattoo Image Dataset V1 has been released to the public [8].

In this paper, we focus on the tasks of tattoo image detection and localization. We develop a tattoo detector, TATT-RBDL, to classify whether an image contains one or more tattoos. In addition, it provides localization information, such as a bounding box of [x,y,width,height], for each detected tattoo. See examples in Fig. 5. We customize the region-based deep learning method, Faster R-CNN [9], [10], [11], to the domain specific data, and train a tattoo detector using 3839 Tatt-C background images [5], and 9963 non-tattoo images from 20 object categories in the PASCAL Visual Object Classes (VOC) 2007 dataset [12]. Following the methodology in [7], we carry out extensive validation on the Tatt-C tattoo detection dataset [5], the MIT-SUN dataset [13], and the NTU Tattoo Image Dataset [8], [14].

Our tattoo detector improves upon the state-of-the-art tattoo detection algorithms reported in the Tatt-C challenge [7]. On the same evaluation dataset, TATT-RBDL yields consistently better detection error trade-off (DET) curve than all the other algorithms, as shown in Fig.2. It improves the overall detection accuracy to 98.25% from the best prior result of 96.3% in Table II. TATT-RBDL can detect most of the challenging images studied in the Tatt-C challenge, as shown in Table IV. It identifies a few images labeled as non-tattoo images but actually containing tattoos (Table V). The confidence scores on non-tattoo images are also quite robust to the other scene categories (Fig. 1(b)). In addition to image classification, the localization performance is evaluated in Fig. 3 with varying degree of overlap between the detection boxes and the groundtruth boxes. Finally, the same tattoo detector is applied to 10000 Flickr(10K) images from the NTU dataset, and detection ROC is plotted in Fig. 4.

After the review of region-based deep learning algorithm in Section II, the convolution network and region proposal network are customized and trained using a large number of tattoo images in Tatt-C dataset and non-tattoo images in PASCAL VOC 2007 dataset in Section III. Both qualitative and quantitative validations are carried out and the results are compared against the winning algorithms on the Tatt-C, MIT-SUN, and NTU datasets. The paper concludes with Section V.

II. Region-based Deep Learning

There are a number of challenges for fast tattoo detection and accurate localization. First, it is usually very hard to give a generic tattoo characterization, due to the large variation in tattoo composition and design, from images, graphics to text and symbols. The shape, color, and texture changes from case to case. The size and location vary from a small region on the finger to a full back tattoo. The skin tone also varies from race, age, and gender. Second, the imaging process could lead to different levels of contrast, geometric distortion, and image quality, which are impacted by the sensor, illumination, and 3D tattoo surface. Lastly, an exhaustive search of all possible

978-1-5090-4846-5/16/$31.00 ©2016 IEEE

spatial locations, scales, and aspect ratios in an image yields a prohibitively large number of region candidates.

To overcome these difficulties, we adopt the deep learning approach and customize the convolutional neural network (CNN) trained on generic object types to the tattoo patterns, using the tattoos previously seen to predict the class labels of the regions under study. It has been shown (e.g. in [15]) that CNNs are very effective for image classification. The shallow layers learn the representative image features directly from the training samples, and the deep layers specialize to specific object types. For generic object types, such as face, car, and dog, a number of models have been trained for various network architectures and datasets using Caffe [16]. Taking a similar approach, the tattoos are characterized by low-level image features trained from generic object types, with parameters fine-tuned from the training tattoo images, and classified by the image features at the deep layers.

Recently, Faster RCNN [9], [10] has been proposed to combine an image classification CNN network and a region proposal network, with both sharing the same full-image convolutional features. The object detection CNN network effectively predicts the class label through a number of convolution and pooling, followed by softmax classification. Meanwhile, feature vectors extracted from the shared convolution feature map are used to generate region proposals, i.e., the regions at different scales, aspect ratios, and locations where objects are likely to appear. During testing, the region proposal network can significantly reduce the number of candidate regions, usually down to the order of thousands, at a marginal overhead. Therefore it can greatly speed up the detection. Such a region-based deep learning approach has shown very competitive detection and recognition performance on several large-scale image datasets.

In addition to the well designed algorithm, carefully crafted data and annotation are equally important to achieve good performance. In the following section, we present the details on how to customize Faster RCNN to the domain specific tattoo data, and train a discriminative model.

III. Model Training and Testing

We train a tattoo detector using the Tatt-C dataset [5] and PASCAL VOC 2007 dataset [12], and fine tune the convolution neural network and the region proposal network [9], [10] from the pre-trained VGG_CNN_M_1024 network [17].

The networks are trained from images with and without tattoos and the corresponding location bounding boxes. The training samples come from two sources. The tattoo images are from the Tatt-C background image set with a total of 3839 tattoo images, each containing one or more tattoo regions. The ground truth of the location bounding boxes is available from the dataset (metadata.txt). In addition, 9963 images from VOC 2007 are selected as images without tattoos. These images contain 20 object types of airplane, bicycle, bird, boat, bottle, bus, car, cat, chair, cow, dining table, dog, horse, motor bike, person, potted plant, sheep, sofa, train, and TV/monitor. To this end, a total of 13802 training images and 28479 regions/labels are used for model training. Each training sample contains an image (a collections of pixels), one or more rectangular sub-regions, such as [x,y,w,h], and object labels (e.g. "tattoo" and "car") indicating the object types inside the regions.

The model is trained for 10,000 iterations with a learning rate of 0.001. On a Dell workstation with Xeon processors and a GPU of Nvidia GTX 980Ti, the training took 4 hours 22 minutes.

During testing, the tattoo detector scans through region proposals at different spatial locations, scales, and aspect ratios, and returns the rectangular regions [x,y,w,h] and the corresponding tattoo confidence scores. The detection bounding box is specified by the x/y coordinate of the upper left corner and the width and height of the region. The detection score is a real number between 0 and 1, indicating low and high detection confidence. A region is labeled as "Tattoo" when the confidence score of class "Tattoo" is higher than the other object classes and over a pre-specified detection threshold. When no tattoo regions are detected, an image is labeled as non-tattoo image.

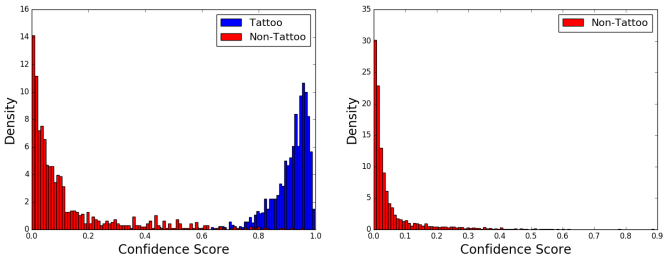
The source code and trained model have been released to the public, and can be found at github (https://github.com/z-harry-sun/TattDL.git).

IV. Tattoo Detection and Localization Results

We evaluate the tattoo detector performance on the Tatt-C dataset, the MIT-SUN dataset, and the NTU tattoo image dataset V1. The evaluation consists of two parts. The first is on binary image classification of Tattoo image and Non-Tattoo image. The performance is measured by classification accuracy (Table II), tattoo detection score histogram distribution (Fig. 1), detection error trade-off (DET) of false positive rate vs false negative rate (Fig. 2), and ROC of false alarm rate vs probability of detection (Fig. 4). The second part of evaluation is on the localization performance, i.e. how well the tattoo detection bounding boxes overlap with the groundtruth bounding boxes. It is measured by the ROC of false alarm count vs probability of detection (Fig. 3).

We first evaluate the performance on the Tatt-C dataset, by adopting the methodology outlined in [5] and comparing with the detection and localization results reported in [7], [6]. The proposed method (TATT-RBDL) compares favorably against the state-of-art methods in [7]. The overall accuracy improves to 98.25% from the best prior result of 96.3% in Table II. The detection error trade-off (DET) plot in Fig. 2(a) consistently outperforms the prior arts in Fig. 2(b). On average, the detection takes 0.083 second per image using a GPU, or 12 frames per second. An additional robustness test is carried out on 397 scene categories in the MIT-SUN dataset, and almost all detection scores fall below 0.5 as shown in Fig 1(b). In the following, the details of both qualitative and quantitative results are presented.

The Tatt-C detection test dataset consists of $N_{tattoo} = 1349$ tattoo images, $N_{non-tattoo} = 1000$ face images, and the ground truth of region type (Tattoo or Non-Tattoo) and locations of [x,y,width,height]. By comparing the detection at a



(a) (b)

Fig. 1. Tattoo score histogram of Tattoo images (blue) and Non-Tattoo images (red). (a) 2349 Tatt-C images with 1349 tattoo images and 1000 face images [5]. (b) 3000 randomly selected images from 397 scene categories in the MIT-SUN dataset [13].

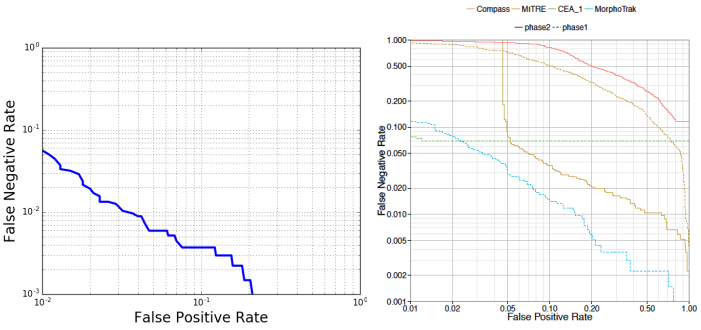
Operating points	FPR	FNR	P_d
1	0.3%	24.54%	75.45%
2	0.5%	12.60%	87.39%
3	1.0%	5.634%	94.37%
4	1.2%	4.448%	95.55%
5	1.5%	3.188%	96.81%
6	1.8%	2.372%	97.63%
7	2.0%	1.927%	98.07%
8	2.3%	1.334%	98.67%
9	3.0%	1.186%	98.81%
10	5.4%	0.593%	99.41%
11	10.2%	0.371%	99.63%
12	20.4%	0.148%	99.85%

TABLE I
TWELVE OPERATING POINTS ON THE DETECTION ERROR TRADE-OFF (DET) PLOT BETWEEN FNR AND FPR IN FIG. 2(A).

certain threshold with the groundtruth, the numbers of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) are counted, and summarized as overall accuracy, $accuracy = (TP + TN)/(N_{tattoo} + N_{non-tattoo})$, false positive rate, $FPR = FP/(FP + TN)$, and false negative rate, $FNR = (FN/(FN + TP)$. A false positive occurs when a Non-Tattoo image is mis-classified as a Tattoo image, and a false negative occurs when a Tattoo image is detected as a Non-tattoo image.

The histogram of the detection confidence scores is shown in Fig. 1. The X axis is the score value from 0 to 1, indicating low to high confidence. The Y axis is the occurrence of detected tattoo regions with a specific confidence score. The red and blue colors indicate the histogram distribution of Non-Tattoo and Tattoo classes. It is clear these two classes are well separated, and the threshold of 0.7 is a good trade-off of detection and false alarm.

Next we sweep the detection threshold from 0 to 1. At each operating point, the false positive rate (FPR), false negative rate (FNR), and probability of detection (P_d) are computed. The values at 12 operating points are summarized in Table I. For example, it reaches a very low FPR of 2% at FNR of 1.9%. The Detection Error Trade-off (DET) curve is plotted in Fig. 2(a), where the X and Y axes are the FPR and FNR on log scale, respectively. As a comparison, the DET plot by the Tatt-C challenge is shown in Fig. 2(b) [7]. The blue curve



(a) (b)

Fig. 2. The detection error trade-off (DET) plot between False Positive Rate (FPR) and False Negative Rate (FNR). (a) The proposed TATT-RBDL method. (b) Four algorithms reported in [7]. The blue curve in (a) consistently has lower FNR at the same FPR than all the curves in (b).

Algorithm	Overall Accuracy
TATT-RBDL	**98.25%**
MorphoTrak-Phase1	96.3%
CEA_1-phase1	95.6%
MITRE-phase2	93.4%

TABLE II
COMPARISON OF OVERALL ACCURACY BETWEEN TATT-RBDL AND THREE WINNING ALGORITHMS IN TATT-C CHALLENGE [7].

in 2(a) is closer to the origin point of (0,0), and consistently has lower FNR at the same FPR, or lower FPR at the same FNR, than the curves in 2(b). In addition, the overall accuracy of TATT-RBDL is compared to the three winning algorithms in Table II. It is improved to 98.25% from the best prior accuracy of 96.3%. Therefore the proposed Tatt-RBDL method has better tattoo detection performance than the state-of-the-art detection methods studied in the Tatt-C challenge. Recently 98.8% accuracy of Tattoo/Non-Tattoo image classification was reported in [8] using AlexNet. However it did not provide localization information of the detected regions.

Having presented the statistical results, we now turn to the representative examples, tough cases, and potential false alarms. Due to the restriction on the Tatt-C images to avoid person identification, no visual images or pixels can be shown. Instead we present the detection scores, detection bounding boxes, and the corresponding groundtruth boxes in Tables III,IV,V. The detection scores, bounding boxes, and groundtruth boxes on ten test images are shown in Table III.

Table IV lists detection results on a number of images deemed to be challenging to the algorithms in the Tatt-C challenge [7]. The first four images contain tattoos with human faces, which could easily defeat algorithms using face detection as most Non-Tattoo test images contain faces. The rest images have low contrast and poor image quality. Seven out of the eight challenging images are correctly detected. The first image with a tattoo on the neck (img_0085.jpg) is missed, as it has a low confidence score of 0.1.

We also dig into the false detection and find some face images labeled as Non-Tattoo actually contain tattoos, as shown

Tatt-C Image Name	Detection		Groundtruth
	Score	Region $[x, y, w, h]$	Region $[x, y, w, h]$
img_1417.jpg	0.952	[240,618,304,252]	[219,601,329,239]
	0.930	[71,258,626,259]	[189,288,202,174]
			[393,232,319,272]
	0.754	[0,809,100,148]	[0,778,140,183]
	0.667	[24,100,177,206]	[28,82,145,219]
	0.503	[675,769,91,164]	[682,761,70,121]
	0.442	[0,313,160,190]	[3,320,122,291]
img_1009.jpg	0.656	[110,69,609,695]	[90,68,769,753]
img_1009.jpg	0.964	[62,66,312,480]	[111,131,251,162]
			[64,290,240,311]
img_0012.jpg	0.753	[130,546,493,262]	[85,569,545,223]
	0.318	[538,0,228,244]	[542,0,212,249]
img_022.jpg	0.906	[132,37,305,353]	[140,0,259,396]
	0.707	[250,454,262,366]	[244,510,301,278]
	0.690	[398,882,150,76]	[368,843,209,167]
img_052.jpg	0.973	[23,75,410,399]	[35,89,383,366]
	0.945	[441,610,295,342]	[450,634,304,288]
	0.944	[509,222,247,290]	[512,203,339,304]
	0.876	[114,569,132,389]	[95,585,209,347]
img_0221.jpg	0.774	[57,106,672,820]	[18,33,749,879]
img_0898.jpg	0.973	[176,383,423,532]	[217,484,294,180]
			[254,608,281,340]
img_399.jpg	0.711	[5,215,87,158]	[0,213,100,152]
	0.594	[277,256,93,137]	[296,264,63,103]
	0.519	[85,234,77,112]	[101,254,52,80]
	0.262	[29,406,92,123]	[44,425,69,82]
img_415.jpg	0.680	[199,272,250,274]	[195,239,226,327]

TABLE III
TEN EXAMPLES OF TATTOO DETECTION AND ANNOTATED GROUND TRUTH IN TATT-C DATASET. EACH ROW CONTAINS THE TEST IMAGE NAME, DETECTION SCORE, DETECTION ROI AND GROUNDTRUTH ROI.

in Table V. The ground truth is questionable in these cases. If marked as true detection, the TATT-RBDL performance could become even better. In addition, some potential false detections are shown in Table VI. It appears the tattoo detector also fires up on text logos and scars at certain scales and locations.

The next question is whether the tattoo detector is suited for screening tattoo images in a massive image collection. Is the tattoo detector sensitive to other object types and is it going to generate a large number of false alarms? This is a valid question as most of the Non-Tattoo images in the test dataset are human faces. To answer this question, we randomly select 3000 images from 397 scene categories in the MIT-SUN dataset [13]. The object types are much more than the 20 categories in the VOC 2007 dataset. The histogram of the tattoo confidence scores are plotted in Fig. 1(b). Almost all confidence scores are below 0.5 (red in (b)), and well separated from the true tattoo scores (blue in (a)). Therefore the tattoo detector can be used as a screening filter, to quickly narrow down the tattoo images from a large Internet collection.

We now turn to the performance evaluation of localization, i.e., how well the detection bounding boxes align with the groundtruth boxes. The localization ROC's between 2435 detection boxes and 2217 groundtruth boxes on Tatt-C dataset are shown in Fig. 3. In (a), a detection box is classified as a true detection if it overlaps with a truth box and the overlap region covers at least 10% (purple), 25% (green), 50% (cyan), and 75% (yellow) of the truth box area. For example, given 10% overlap, the purple curve shows 95.26% P_d with 100 false

Tatt-C Image Name	Detection		Groundtruth
	Score	Region $[x, y, w, h]$	Region $[x, y, w, h]$
img_0085.jpg			[59,33,115,183]
img_0230.jpg	0.770	[117,151,297,444]	[160,184,171,290]
img_0125.jpg	0.665	[17,370,193,192]	[16,370,230,174]
img_0138.jpg	0.204	[181,426,405,159]	[198,429,209,156]
img_1381.jpg	0.949	[122,81,356,477]	[210,139,204,354]
img_175.jpg	0.708	[114,80,355,519]	[114,76,370,511]
img_0555.jpg	0.987	[58,566,617,187]	[192,565,435,187]
img_1287.jpg	0.887	[219,245,102,154]	[213,260,120,147]
	0.575	[95,247,100,175]	[81,223,129,215]
	0.442	[69,224,309,206]	

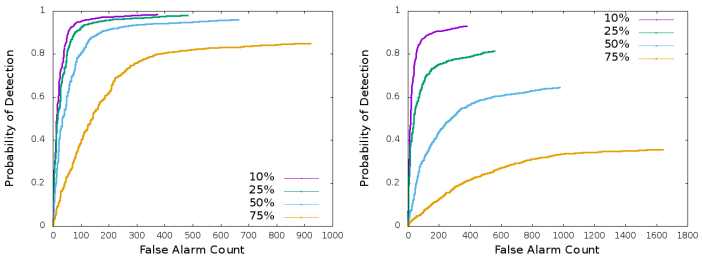
TABLE IV
TATTOO DETECTIONS ON CHALLENGING IMAGES REPORTED IN TATT-C CHALLENGE [7]. THE FIRST FOUR IMAGES CONTAIN TATTOOS OF FACES, AND THE REST HAVE LOW CONTRAST. EXCEPT FOR THE FIRST IMAGE, WITH A LOW SCORE OF 0.107, THE REST ARE ALL CORRECTLY DETECTED.

Tatt-C Image Name	Detection	
	Score	Region $[x, y, w, h]$
img_332.jpg	0.727	[72,547,184,52]
img_762.jpg	0.821	[160,514,187,85]
img_858.jpg	0.873	[100,466,282,133]
img_687.jpg	0.381	[431,534,48,65]
	0.243	[299,423,96,147]

TABLE V
TRUE DETECTIONS ON IMAGES MARKED AS NON-TATTOO IN TATT-C DATASET. THE GROUNDTRUTH IS QUESTIONABLE IN THESE CASES.

Tatt-C Image Name	Detection	
	Score	Region $[x, y, w, h]$
img_0984.jpg	0.619	[90,203,65,32]
img_1117.jpg	0.889	[7,187,94,49]
img_391.jpg	0.788	[111,94,82,78]
img_649.jpg	0.659	[76,480,342,119]

TABLE VI
POTENTIAL DETECTION FALSE ALARMS IN TATT-C DATASET, SUCH AS SCAR AND TEXT LOGOS.



(a) (b)

Fig. 3. Tattoo localization ROC between 2435 detection boxes and 2217 groundtruth boxes on Tatt-C dataset. (a) A true detection is counted when the overlap pixels between the boxes are at least 10% (purple), 25% (green), 50% (cyan), and 75% (yellow) of the pixels in groundtruth box. (b) A tighter condition when the overlap is at least certain percentage of both the groundtruth box and the detection box.

detection boxes. The criterion is more forgiving and favors large detection boxes. In (b), a detection box is called as a true detection only when the overlap region has the same percentage coverage on both the detection box AND the truth box. In other words, unless two boxes are similar in size and align well, they are called as a false detection.

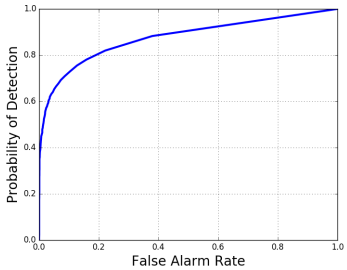


Fig. 4. Tattoo detection ROC on Flickr(10K) in the NTU Tattoo Image Dataset, containing 5740 Tattoo images and 4260 Non-Tattoo images. It uses the same model trained on Tatt-C dataset and PASCAL VOC 2007 dataset.

At last we run the same tattoo detector on the NTU tattoo image dataset V1 [8], [14]. It uses the same model trained on the Tatt-C and PASCAL VOC 2007 images. No NTU images are used in training. On the Flickr(10K) set, with 5740 Tattoo images and 4260 Non-Tattoo images, the overall classification accuracy is 80.66%. The detection ROC is depicted in Fig. 4. In addition, local rectangular regions containing the detected tattoos and the detection confidence are returned. Fig. 5 shows the representative detections on 30 test images, referenced by row (R) and column (C). For example, R1C5 refers to the image with a head tattoo. The blue color indicates the detection regions and the corresponding tattoo confidence score. The tattoo detections vary in scales and aspect ratios. The bounding boxes are pretty tight around the regions of interest. In general, the detections are fairly strong with scores close to 1. The detections include a single region (R1C5), two regions (R2C1), and three regions (R4C4). The tattoo can be a small one on the finger (R7C1) or as large as the whole back (R1C2). The tattoos appeared at different body parts, such as head (R1C5), neck (R3C4), shoulder (R1C4), chest (R4C3), arm (R1C1,R4C2), legs (R5C3), waist (R6C4), back (R5C2), hand (R6C3), foot (R3C1), and fingers (R7C1,R7C4).

V. Conclusions

We have developed a tattoo detector using region-based deep learning. The detector is trained using tattoo images from the Tatt-C background images and non-tattoo images from the VOC 2007 dataset containing 20 object types. The detector outperforms the state-of-the-art algorithms in the Tatt-C challenge, achieving better DET curve and overall precision, and the performance has been validated on 10000 Flickr(10K) images. Based on the test on randomly sampled non-tattoo images in the MIT-SUN dataset, the confidence score is also very robust to other scene categories. In addition, the localization performance has been validated by varying degrees of overlap between the detection bounding boxes and the groundtruth boxes. The detector runs at the speed of 12 images per second on a GPU. Both the source code and trained model have been released to the public. The developed tattoo detector can be used to filter tattoo images from a large image collection (e.g. mined from the Internet). The local regions of interest can also be used for detailed analysis, such as similarity matching, mixed media matching, and tattoo validation.

Acknowledgment

This work was partially supported by DARPA MEMEX program. The opinions and findings are those of the authors only. The Tatt-C dataset used in this study was made available by National Institute of Standards and Technology (NIST). Portions of the research in this paper use the Nanyang Technological University (NTU) Tattoo Image Database Version 1. Credit is hereby given to the Biometrics & Forensics Laboratory at the School of Computer Science and Engineering, NTU, Singapore for providing the database.

References

[1] A. E. Laumann and A. J. Derick, "Tattoos and body piercings in the united states: a national data set," *Journal of the American Academy of Dermatology*, vol. 55, no. 3, pp. 413–421, 2006.

[2] A. K. Jain, J.-E. Lee, and R. Jin, "Tattoo-ID: automatic tattoo image retrieval for suspect and victim identification," in *PCM*, 2007.

[3] J.-E. Lee, R. Jin, A. K. Jain, and W. Tong, "Image retrieval in forensics: tattoo image database application," *IEEE MultiMedia*, vol. 19, no. 1, pp. 40–49, 2012.

[4] H. Han and A. K. Jain, "Tattoo based identification: Sketch to image matching," in *The 6th IAPR International Conference on Biometrics (ICB)*, Madrid, Spain, June 2013, pp. 4–7.

[5] M. Ngan, P. Grother, M. Garris, and E. Phillips, "NIST tattoo recognition technology - challenge (Tatt-C) dataset, concept, and evaluation plan," http://www.nist.gov/itl/iad/ig/tatt-c.cfm, April 2015.

[6] NIST, "Tatt-C workshop online proceedings," http://www.nist.gov/itl/iad/ig/tatt-c_workshop_proceedings.cfm, June 2015.

[7] M. Ngan, G. W. Quinn, and P. Grother, "Tattoo recognition technology challenge (tatt-c) outcomes and recommendations," National Institute of Standards and Technology, Tech. Rep. NISTIR 8078, Sept. 2015.

[8] Q. X. Xu, S. Ghosh, X. Xu, H. Yi, and A. Kong, "Tattoo detection based on CNN and remarks on the NIST database," in *International Conference on Biometrics*, 2016.

[9] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 1, pp. 142–158, 2016.

[10] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing Systems*, 2015, pp. 91–99.

[11] R. Girshick, "Fast R-CNN," in *International Conference on Computer Vision*, 2015, pp. 1440–1448.

[12] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, 2010.

[13] J. Xiao, J. Hays, K. A. Ehinger, A. Oliva, and A. Torralba, "Sun database: Large-scale scene recognition from abbey to zoo," in *IEEE International Conference on Computer Vision and Pattern Recognition*, 2010, pp. 3485–3492.

[14] H. Yi, P. Yu, X. Xu, and A. Kong, "The impact of tattoo segmentation on the performance of tattoo matching," in *2015 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, 2015, pp. 43–46.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[16] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the ACM International Conference on Multimedia*. ACM, 2014, pp. 675–678.

[17] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, "Return of the devil in the details: Delving deep into convolutional nets," in *BMVC*, 2014.

Fig. 5. Thirty examples of tattoo detections, including both detection scores and bounding boxes, on the NTU Flickr(10K) dataset.