

© 2012 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Pre-print of article that will appear at BTAS 2012.

# Detecting and Classifying Scars, Marks, and Tattoos Found in the Wild

Brian Heflin\*

bheflin@securics.com

Walter Scheirer

wscheirer@securics.com

T.E. Boult \*

tboult@vast.uccs.edu

Securics Inc and University of Colorado Colorado Springs

## Abstract

Within the forensics community, there is a growing interest in automatic biometric-based approaches for describing subjects in an image. By labeling scars, marks and tattoos, a collection of these discriminative attributes can be assigned to images and used to assist in large-scale person search and identification. Typically, the imagery considered in a forensics context consists to some degree of uncontrolled, unprofessionally generated photographs. Recent work has shown that it is quite feasible to detect scars and marks, as well as categorize tattoos, presuming that the source imagery is controlled in some manner. In this work, we introduce a new methodology for detecting and classifying scars, marks and tattoos found in unconstrained imagery typical of forensics scenarios. Novel approaches for initial feature detection and automatic segmentation are described. We also consider the “open set” nature of the classification problem, and describe an appropriate machine learning methodology that addresses it. An extensive series of experiments for representative unconstrained data is presented, highlighting the effectiveness of our approach for images found “in the wild”.

## 1. Introduction

Digital image forensics encompasses a wide range of applications, including person search and identification, where suspects, victims and even general scenes must be considered by visual appearance. While traditional biometric matching is helpful for specific person identification, often times a search for a broader spectrum of potential candidates is what is desired during an investigation. Further, when only uncontrolled, unprofessionally generated photographs are available, traditional biometric matching might not be feasible. Cases like these can still benefit from automatic biometric-based approaches for describing subjects in an image by significant dermatological features such as scars, marks and tattoos.

\*This work was supported by Army SBIR W15P7T-12-C-A210



Figure 1. Constrained vs. unconstrained imagery for forensics scenarios. In many cases, constrained imagery that is controlled for lighting, pose or feature of interest is simply not available, leaving us with a more challenging problem when trying to label scars, marks and tattoos. In this work, we look at extending support for dermatological feature detection and classification to unconstrained imagery. Further, we consider the problem of open set recognition, where every candidate image is not assumed to contain a face, skin region or tattoo.

Very promising recent work has shown the feasibility of detecting scars and marks, as well as categorizing tattoos. Jain et al. [10] note that facial mark detection has now reached the levels of accuracy required for image retrieval and augmented face matching applications that are tailored to forensics. Similarly for tattoos, Lee et al. [15] demonstrated good accuracy for forensics oriented image retrieval for tattoo categories over a database of many thousands of images. The approaches commonly deployed for these tasks include a collection of well-known image modeling methods and feature descriptors that have typically produced good results for related image categorization tasks. Given a frontal image of a person, the face can be pre-processed using 3D morphable models for alignment, followed by the application of a Laplacian of Gaussian (LoG) operator for mark detection [10]. Given a tattoo image, SIFT features can be directly computed and used in a distance-based comparison to other images stored in a database [15].

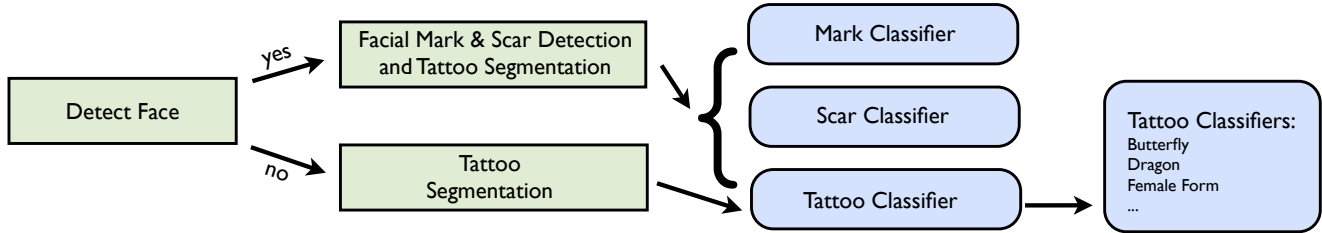


Figure 2. The components of our pipeline methodology integrating scars, marks and tattoos. The detection components (green squares) are specifically designed to isolate candidate regions that might contain a dermatological feature of interest. The classification components (blue ovals) allow us to label particular dermatological features, and in an open set context, eliminate objects that are not scars, marks or tattoos. Since a wide variety of tattoos exists, we also support identifying particular tattoo categories.

While the results found in the literature are promising for reasonably constrained imagery, they do not perform as well on unconstrained imagery. For example, consider the individual images in Fig. 1. The tattoo work described by Lee et al. [15], the facial mark work of Park and Jain [18], and the skin mole work of Cho et al. [5] all constrain their input imagery to pre-cropped regions where the feature of interest can be isolated without considering a larger scene. In this work, we are interested in extending support for dermatological feature detection and classification to unconstrained images, such as those shown on the right of Fig. 1.

Another important aspect of unconstrained recognition is that of the “open set” nature of the problem. Prior work in this area has always made the assumption that a face, skin region or tattoo is present in the image. For instance, in content-based image retrieval (CBIR) [15], a SIFT-based distance comparison to known tattoo images will always return the closest matching candidates, regardless of whether or not the input image is actually of a tattoo. We consider the open set model, where every image is not assumed to contain a face, skin region or tattoo, to be correct for this problem, and structure our learning approach around it.

In this paper, we introduce a new methodology for detecting and classifying scars, marks and tattoos found in unconstrained imagery typical of forensics scenarios. Our specific contributions include:

1. **Detection and Segmentation for Unconstrained Imagery:** We introduce a novel algorithm for detecting dermatological features on the face, and a automatic variant of the GrabCut [23] segmentation algorithm coupled with a quasi connected components approach for tattoo detection.
2. **Open Set Classification:** When considering unconstrained imagery, it is possible that images that do not contain what we are looking for will be submitted for classification. We introduce an open set classification approach to account for the “unknown” class.
3. **A Pipeline Methodology Integrating Scars, Marks**

**and Tattoos:** In prior work, smaller features like scars and marks were treated distinctly from more complicated tattoo features. Here we present an integrated methodology (Fig. 2) that can detect and classify all dermatological features as one process.

## 2. Related Work

The medical imaging community first considered the problem of detecting marks on the skin for diagnostic analysis. Cho et al. [5] describe a reliable skin mole localization scheme that utilizes skin detection to isolate candidate regions in an image, a Difference of Gaussian (DoG) filter to detect specific mole candidates, and a support vector machine (SVM) classifier. In biometrics, marks such as freckles, moles and scars are considered “soft biometric” features – too weak to stand alone as discriminative features for identity purposes, but useful features to improve matching accuracy when combined with more traditional biometric features.

Jain et al. [10] provide a comprehensive overview of the work at Michigan State University and elsewhere for incorporating facial marks into forensics oriented face applications. Park and Jain [13, 18] developed an automatic facial mark detection method that utilizes an active appearance model for locating primary facial features that can be eliminated from consideration (eyes, nose and mouth), a LoG feature detector, and morphological operators to enhance accuracy. Experimental results show that this approach is able to enhance the matching performance of a competitive face recognition algorithm. In similar work, Ramesha et al. [21] describe a template based mole detection approach that utilizes normalized cross correlation, complement of Gaussian templates and skin segmentation. Park et al. [19] showed that facial marks are also able reduce image retrieval time for very large scale face databases. When considering challenging unconstrained data, the accuracy of this existing dermatological mark work is significantly impacted by false positives caused by several constraints we discuss in Sec. 3.

Tattoos represent another discriminative feature that can be used in a broader biometric context for person identification. A good survey treatment of tattoo classification is also provided by MSU [15]. Jain et al. [11] first looked at color, shape and texture features to describe tattoos, with a simple histogram bin distance metric for comparisons. Subsequent work at MSU [14, 12, 15] moved towards SIFT-based feature extraction and distance comparison for CBIR. For image retrieval, results were noted to vary as a function of image quality. When text labels outside of the vision system were incorporated, accuracy increased to acceptable levels when poor quality images were considered. Noting weaknesses in the edge-based segmentation incorporating morphology in [11], Acton and Rossi [1] proposed active contour-based segmentation and a global-local feature for tattoo specific CBIR.

While evaluating the above tattoo approaches for unconstrained imagery, we found two significant shortcomings. First, the prior work makes use of pre-cropped imagery [11, 1, 14, 12, 15], and does not consider a broader scene as input. Second, in all cases, distance metrics are used for matching to find the images closest to an exemplar image, and must make use of an empirical threshold to reject non-tattoo candidates. In our work here, we consider full scenes where a person and tattoo might be present, as well as a flexible learning approach that can reject features from detected objects that are not of interest.

Full scenes require a segmentation approach that is stronger than edge detection & morphology [11] or active contours [1]. Some of the best recent approaches to the general problem formulate segmentation as a pixel-based energy function that can be optimized using a graph cut for energy minimization. Connectivity priors, shape priors, and random walker-based algorithms are recent innovations that show promise. However, in practice, segmentation algorithms that require a manual bootstrapping stage provide better results. The GrabCut algorithm [23] iteratively re-estimates region statistics, which are modeled as mixtures of Gaussians in color space, based on minimal user input. In this work, we look at replacing the need for user input by automatically estimating a bootstrap region.

For the open set problem, the 1-class SVM has received some attention in the computer vision literature – mostly in the areas of image retrieval and face recognition. The application of 1-class SVMs to problems in computer vision was first made by Chen et al. [3] a decade ago. For binary classification, equal treatment is usually given to positive and negative training examples. However, Chen et al. argue that while it is reasonable to assume that positive training examples cluster in a certain way, the same cannot be said about negative examples, since they can belong to *any* class. Thus, for an open set problem, it seems natural to consider a 1-class SVM, which is trained using only positive examples

for a target class.

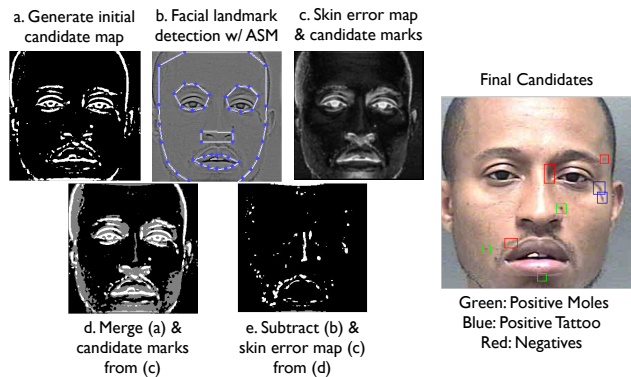


Figure 3. The automatic facial mark detection process.

### 3. Facial Mark Detection with Refinement

Like the prior work in this area [18], we make use of a combination of common image processing and facial modeling techniques (Fig. 3). However, our set is tailored to the specific refinement we need to handle a broader range of imagery. Our first stage facial mark detection approach begins with a LoG filter for pre-processing, since the facial landmarks we are trying to segment and classify all appear as salient regions on the face. We perform filtering at two different scales. The first scale utilizes a LoG filter with a  $15 \times 15$  kernel and  $\sigma = \sqrt{1.6}$ . This filter allows us to detect small facial marks such as moles and freckles. The second scale utilizes a LoG filter with a  $25 \times 25$  kernel and  $\sigma = \sqrt{11.5}$ . This filter is used to detect larger structures such as scars. The filtered images are normalized with respect to lighting using the SQI algorithm [26].

After SQI normalization, a histogram is computed over the filtered image to segment the skin pixels from the facial mark pixels. The histogram will bin the elements of the image into six equally spaced containers ranging from 0.0 to the maximum pixel value in the filtered image. Since the value of skin in a LoG filtered and SQI normalized image should be approximately equal to 0.0, the first histogram bin will contain a large majority of the skin pixels. Therefore, we threshold the image at approximately the edge of the largest bin in the histogram to separate skin pixels from the candidate facial mark pixels. If there are too many pixels in the first bin our algorithm will automatically adjust the threshold based on the overall distribution. Approximately 85% to 90% of the pixels are eliminated using this technique. Finally, the two thresholded images are combined using a binary OR operation to produce our initial candidate map image (Fig. 3(a)).

While the use of the LoG filter is very effective at segmenting skin and non-skin regions, the preliminary facial mark candidate map can still contain the primary features

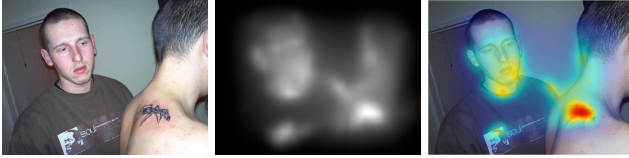


Figure 4. An example of a graph-based visual saliency algorithm (GBVS). Left: Original image. Center: Raw saliency map. Right: Deep saliency map, in which red regions are the most salient.

of the face including the eyes, eyebrows, nose, mouth, and hair. Therefore, our next step in the pipeline is to eliminate primary feature areas before we move to classification. To delineate the primary facial features we use an Active Shape Model (ASM) [17]. Once the facial landmarks are located, we construct a mask (Fig. 3(b)) from the results of the ASM algorithm to suppress false positives caused by primary facial features. The ASM based mask does not eliminate user specific facial features such as beards, mustaches, or wrinkles around eyes that can also increase the false positive rate during classification, so the next step in our pipeline is to build a second mask based on skin detection results.

This second user specific mask is constructed using a skin detection algorithm based on the work of Pierrard and Vetter [20], but adapted for color images. The first step in our skin detection process is color SQI normalization of the original image. Based on the results of the ASM algorithm, we take a sample patch of skin from the left and right cheek of the subject. We then use the sample patches in conjunction with the skin detection algorithm to form a skin error map (Fig. 3(c)). The brightness of a pixel in the map, ranging from 0.0 to 1.0, correlates directly to the probability of a pixel being a non-skin area. Once the skin error map has been obtained we again threshold the image using the histogram technique described above.

Prior work in this area [18] also builds a user specific mask based on an edge image that is obtained by using the conventional Sobel operator. However, in our own evaluation, we observed that it also eliminates a number of areas with true facial marks. In our algorithm, small structures are eliminated from the skin error map after thresholding, since these structures are potential facial marks. Subsequently, these small structures are merged with the preliminary facial mark candidate image as potential facial marks for classification (Fig. 3(d)) and the remaining structures in the thresholded skin error map are combined with the ASM based mask to form a third user specific mask.

The values of active pixels in the facial mark candidate image at this point range from 1 to 2, where a value of 2 corresponds to active points in both the original facial mark candidate map image AND the small structures that were added to it. Values of 1 correspond to an active point in either the original facial mark candidate image OR the added small structures. Finally, the facial mark candidate image is filtered with the third user specific mask (Fig. 3(e)). We fil-

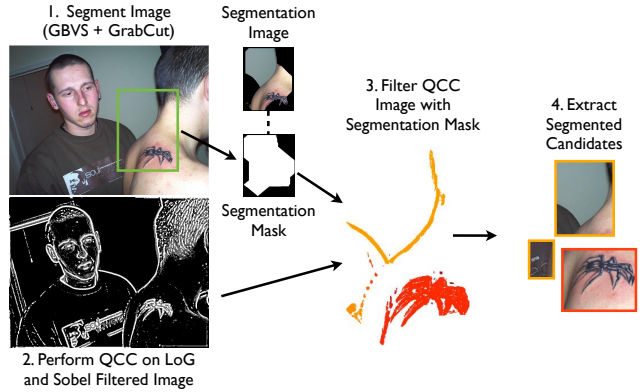


Figure 5. The automatic tattoo segmentation process.

ter the final facial mark candidate image at both the 1 and 2 value levels to produce two final candidate facial mark images. The final candidate facial mark images are then sent to the facial mark classifiers (described in Sec. 5) to determine what they are. Before classification, the blobs are sorted for saliency based on their value in the LoG and skin detection images. This allows us to return the top  $N$  blobs per image.

#### 4. Tattoo Segmentation: Automatic GrabCut + Quasi Connected Components

Tattoos are a more complicated feature than marks, thus, we handle them using a separate approach. Our approach consists of two parts. First, we segment the image using an iterative image segmentation based on the GrabCut algorithm of Rother et al. [23]. We then apply a variant of connected components to the segmented image to extract coherent tattoo objects.

To initially segment the raw image we use an iterative image segmentation technique based upon the GrabCut algorithm. We chose the GrabCut algorithm because it consistently segments and returns a full tattoo image while only requiring a simple bounding box input. The GrabCut segmentation approach is based on optimization by graph-cut and utilizes both texture information and edge/contrast information. User interaction is simplified to drawing a rectangle around the desired foreground, followed by a small amount of corrective editing. In practice, this manual bootstrapping yields excellent results, but has the drawback of human intervention.

Similar to the work in [4], we have automated the GrabCut algorithm through the use of the saliency, specifically the graph-based visual saliency model (GBVS) [7]. The idea of saliency maps is that the sight or gaze of people will direct to areas which, in some way, stand out from the background. For our particular application these areas of interest are tattoos within the image. The GBVS algorithm consists of two steps. First, the algorithm creates feature maps using the technique of Itti et al. [9], and then performs

normalization using a graph-based approach. The GBVS model is simple and powerfully predicts human fixations; it has achieved 98% of the ROC area on 749 variations of 108 natural images [7], whereas the visual saliency algorithms of Itti et al. only achieve 84%. Fig. 4 shows a sample input image and the saliency maps obtained by the GBVS algorithm. Since the GBVS saliency map does not output a full resolution saliency map, we upsample the map to the size of the input image. Once the saliency map is computed it is thresholded at 60% of the maximum value, where we classify all points above the threshold as foreground (the targets) and all points below as background. The thresholded saliency map defines our region of interest input into the GrabCut segmentation algorithm described below. If more than one region exists after thresholding, we subsequently process each region separately. After the bounding box is computed, the GrabCut algorithm is executed on the image to produce a segmentation mask and segmented image (Step 1 in Fig. 5).

We then perform LoG filtering on the original image and automatically threshold the results using the technique described in Sec. 3 to produce an initial candidate tattoo image map. Performing LoG filtering on the segmented image does not work due to the sharp contrast around the boundary of the image, which affects the automatic thresholding operation. A second filtering operation is then performed on the image using a Sobel kernel with a lower threshold than the LoG filter to produce a second candidate tattoo image map. We chose to use a Sobel kernel instead of a LoG with a smaller scale to facilitate the detection of blobs that are not detected by LoG filtering. The results of the filtering operations are combined to produce the initial candidate tattoo map (Step 2 of Fig. 5).

Subsequently, to group the pixels of the individual components in the image, we perform a variant of image connected components called quasi-connected components (QCC) [2]. While this technique was originally designed for target detection and tracking applications, we have adapted it for the extraction of tattoos. QCC is designed to fill in gaps in a thresholded image while eliminating noise. The quasi-connected components idea can be viewed as a direct extension of the idea of thresholding with hysteresis, where we allow the connections to jump over small gaps within the “parent” or high threshold pixel. We perform QCC for our LoG and Sobel filtered images using 8-neighbor connectivity, with a threshold value ranging from 0 to 1 for both images, as the primary pixel inclusion test. The high threshold image produced from the LoG filter allows us to filter out components that may have been noise not filtered out by the lower threshold Sobel filtered image.

Our modified QCC algorithm can be summarized as follows. If a low threshold pixel is within a user defined distance from a high threshold pixel then it is labeled as part



Figure 6. Comparison of tattoo segmentation algorithms. Left: Original image. Center: Segmentation results using algorithm of Lee et al. [15]. Right: Segmentation results using our algorithm.

of the foreground, otherwise the pixel is labeled as background and subsequently discarded. A region size parameter controls how aggressively pixel groups are combined to form single components. The output image contains all of the high threshold pixels and the low threshold pixels labeled as foreground. The QCC image is then filtered with the mask produced from the GrabCut algorithm to produce our final candidate tattoo image (Step 3 in Fig. 5). Finally, 8-way connected components is performed on the image to produce  $N$  blobs that will be individually extracted and sent to the tattoo classifiers (Step 4 in Fig. 5).

Extraction methods similar to Lee et al. [15] that use standard connected components with the morphological operations of closing and opening were also evaluated. However, these techniques frequently return most, if not all, of the entire image due to the opening and closing. As part of our research we identified and corrected this problem using the QCC algorithm. Fig. 6 shows a visual comparison between our tattoo extraction algorithm and the technique of Lee et al. [15].

## 5. Skin Feature Classification Approach

From the candidate image regions that are determined by the detection techniques described in Sec. 3 & 4, we compute features for our open set machine learning. The underlying features used for classification of tattoos (first and second level classifiers) are generated by extracting points of interest (POIs) from the image regions using Difference of Gaussians as proposed in [16], and then computing an LBP-like [24] feature descriptor in a window around each detected PoI. Feature vectors are composed of histogram bins that summarize the feature descriptor information for each sample image.

Scars and marks are significantly smaller, and do not have enough POIs for the above algorithm to work effectively. Thus, for the first level classifiers for marks and scars, we make use of Histograms of Oriented Gradients (HOG) [6] as low-level features. HOG features are accurate for specific object detection, and capture a large amount of information from our small candidate image regions. They are, however, not as accurate as the PoI + LBP-like descriptor approach for tattoos, due to the impact of the orientation information during matching (tattoos tend to be found

in more diverse geometrical configurations than scars or marks). Both approaches produce low-level features that are suitable for SVM-style learning.

The 1-class SVM introduced by Schölkopf et al. [25] adapts the familiar SVM methodology to the open set recognition problem. With the absence of a second class in the training data, the origin defined by the kernel function serves as the only member of a “second class”. The goal then becomes to find the best margin with respect to the origin. The resulting function  $f$  after training takes the value +1 in a region capturing most of the training data points, and -1 elsewhere.

Let  $p(x)$  be the probability density function estimated from the training data  $\{x_1, x_2, \dots, x_m \mid x_i \in X\}$ , where  $X$  is a single class. A kernel function  $\Psi : X \rightarrow H$  transforms the training data into a different space. To separate the training data from the origin, the algorithm solves a quadratic programming problem for  $w$  and  $\rho$  to learn  $f$ :

$$\min \frac{1}{2} \|w\|^2 + \frac{1}{\nu m} \sum_{i=1}^l \xi_i - \rho \quad (1)$$

subject to

$$(w \cdot \Psi(x_i)) \geq \rho - \xi_i \quad i = 1, 2, \dots, m \quad \xi_i \geq 0 \quad (2)$$

In the 1-class SVM,  $p(x)$  is cut by the margin plane minimizing Eq. 1 and satisfying Eq. 2. Regions of  $p(x)$  above the margin plane define positive classification and capture most of the training data. The kernel function  $\Psi$  impacts density estimation and smoothness. The regularization parameter  $\nu \in (0, 1]$  controls the trade-off between training classification accuracy and the smoothness term  $\|w\|$ , and also impacts the choice and number of support vectors.

We use the 1-class SVM formulation to train our first stage scar, mark, and tattoo classifiers, as well as our second stage tattoo category classifiers. The 1-class SVM gives us the flexibility to handle any “unknowns” that might be submitted to a classifier. Since our processing pipeline is completely automated, the possibility exists that image regions that are not drawn from the classes of interest will be submitted for classification. Further, like Lee et al. [15], we can also support multiple labels for a tattoo, by considering multiple positives decisions from a set of classifiers. This is necessary when class overlap occurs (for instance, our classes “female” and “artistic human rendering”).

## 6. Experimental Evaluation

To evaluate our overall pipeline, we consider final “open set” classification results as a metric of success. This takes into account the complete flow of information from initial detection, through segmentation to the 1-class SVM (as would be typical in an operational system). We collected a

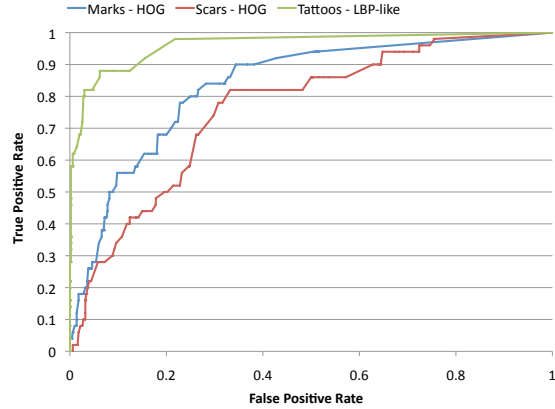


Figure 7. “Open Set” first level classifier accuracies for the dermatological features. For the smaller mole and scar objects, we make use of HOG [6] low-level features, while for tattoos, we make use of LBP-like [24] low-level features. Negative test data is composed of features from other classes, as well as false positive data from the dermatological feature detector.

significant amount of unconstrained data by crawling popular tattoo and dermatology forums found on the web, as well as sampling from the unconstrained face set Labeled Faces in the Wild [8], and the representative mugshot data set MORPH [22]. To generate ground truth labels for training and evaluation, we turned to Amazon’s Mechanical Turk service. In total, we evaluated 6,322 images (selected by the labels assigned by the Mechanical Turk workers) for detection and 12,600 subsequent candidate region images + negative images (objects other than scars, marks and tattoos) for classification.

The classification component of each experiment utilized 150 images for training, 50 positives testing samples and 500 negative testing examples. This emphasizes the open set nature of the problem, with negatives representing an amount of data that is an order of magnitude larger than the positives. Each 1-class SVM is trained using a linear kernel. The ROC curves found in Figs. 7, 9 and 10 are generated by varying the 1-class SVM parameter  $\nu$ , which changes the positive and negative classification rates (as described in Sec. 5). Points approaching the upper left of each plot indicate higher levels of accuracy.

In our first experiment, we assess the accuracy of our “open set” first level classifiers. Negative test data is composed of features from other classes, as well as false positive data from the dermatological feature detector (for example, the negative data for tattoos consists of marks, scars and other non-dermatological objects). While feature rich tattoos provide plenty of information for the machine learning, yielding very good accuracy, the smaller marks and scars are more difficult to discern from the negatives that are often mark- or scar-like in appearance (Fig. 8). Considering the difficult nature of our data, these accuracies are a step in

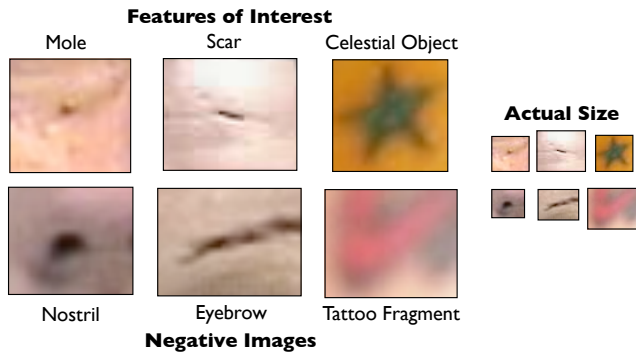


Figure 8. A selection of image regions containing different types of small features of interest and similar looking negative images. The image regions at their actual sizes are on the right, while enlarged versions showing detail are on the left.

the right direction towards an operationally viable solution.

Moving on to specific tattoo classes, when we began this work, we considered the approach of Lee et al. [15], which consists of a SIFT-based distance comparison between images. That approach is very well suited to problems such as CBIR, where we have an exemplar image, and want to find the closest visual matches in a closed set context. However, in an open set classification scenario, a distance comparison to known tattoo images will always return the closest matching candidates, regardless of whether or not the image is actually of a tattoo. We adapted the approach of Lee et al. (reimplemented based on the description in their article) to open set classification by applying a series of threshold tests at set intervals over the distance scores to reject “non-matches” when comparing against a gallery composed of images of the tattoo class of interest. In the summary plot of Fig. 9, we can see a comparison between the work of Lee et al. with thresholds and our proposed approach for all 15 tattoo classes we evaluate in this paper. Our 1-class SVM approach, which was designed specifically for the open set problem, shows superior accuracy.

With respect to the 15 individual tattoo classes, we examined a broad range of classes related to animal and human figures, as well as other object forms. Fig. 10 highlights our results, with very good accuracies (note that the y-axis begins at 0.5) for all classes. Similar to what we observed with moles and scars, classes representative of smaller objects (butterflies, wild birds, celestial objects) are more difficult to classify. In general, small objects are a significant challenge area for this problem (also noted in [15]), which we are continuing to tackle.

## 7. Discussion

In this paper, we took a look at the next challenge for the detection and classification of scars, marks and tattoos: unconstrained imagery for forensics applications. While very

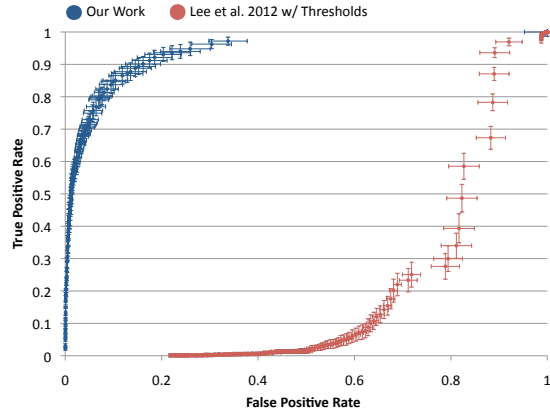


Figure 9. A comparison between our work and that of Lee et al. [15] for “open set” tattoo classification for 15 classes. The SIFT-based distance comparison method of Lee et al. is not well suited to problems where non-tattoo data is present. A more flexible 1-class SVM approach shows superior accuracy. Negative test data is composed of features from other tattoo classes, as well as false positive data from the tattoo detector. Vertical error bars indicate the standard error of the true positive rate, while horizontal error bars indicate the standard error of the false positive rate.

promising recent work has demonstrated that these dermatological features can be detected and classified, there is much work yet to be done to accurately process images found in the wild. During our study of this topic, we concluded that automated facial mark detection and tattoo segmentation that can flexibly filter candidate regions are essential for a good forensics solution. We also discovered that approaches designed for closed set evaluation do not readily apply to open set problems where we don’t have complete control over the input images. Our current work is assessing more descriptive feature sets and visual attributes for dermatological object labeling (especially for very small objects), as well as new machine learning algorithms that are specific to the open set recognition problem.

## References

- [1] S. Acton and A. Rossi. Matching and Retrieval of Tattoo Images: Active Contour CBIR and Glocal Image Features. In *IEEE SSIAI*, March 2008.
- [2] T. E. Boult, R. J. Micheals, X. G. X. Gao, and M. Eckmann. Into the Woods: Visual Surveillance of Noncooperative and Camouflaged Targets in Complex Outdoor Settings. *Proceedings of the IEEE*, 89(10):1382–1402, 2001.
- [3] Y. Chen, X. Zhou, and T. Huang. One-class SVM For Learning in Image Retrieval. In *IEEE ICIP*, pages 34–37, 2001.
- [4] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu. Global contrast based salient region detection. In *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2011.
- [5] T. Cho, W. Freeman, and H. Tsao. A Reliable Skin Mole Localization Scheme. In *IEEE MMBIA*, October 2007.



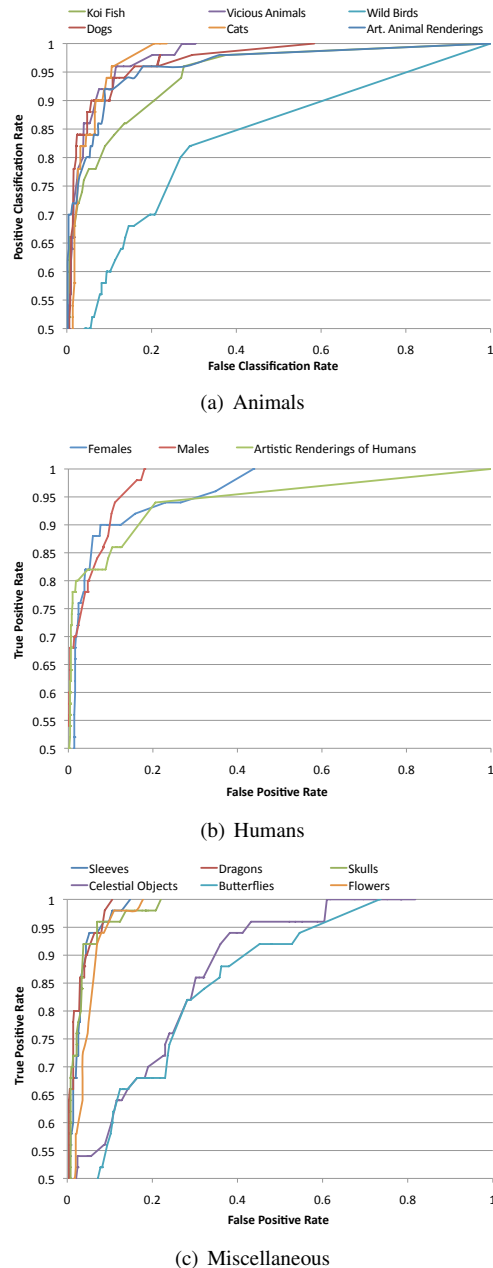


Figure 10. “Open Set” classifier accuracies for a selection of different tattoo classes. Note that classes representative of smaller objects (butterflies, wild birds, celestial objects) are more difficult.

[6] N. Dalal and B. Triggs. Histograms of Oriented Gradients for Human Detection. In *IEEE CVPR*, June 2005.

[7] J. Harel, C. Koch, and P. Perona. Graph-Based Visual Saliency. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Advances in Neural Information Processing Systems 19*, pages 545–552. MIT Press, Cambridge, MA, 2007.

[8] 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 Re-

port 07-49, University of Massachusetts, Amherst, October 2007.

[9] L. Itti, C. Koch, and E. Niebur. A Model of Saliency-based Visual Attention for Rapid Scene Analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(11):1254 – 1259, November 1998.

[10] A. Jain, B. Klare, and U. Park. Face Matching and Retrieval in Forensics Applications. *IEEE Multimedia*, 19(1):20–28, January 2012.

[11] A. Jain, J.-E. Lee, and R. Jin. Tattoo-ID: Automatic Tattoo Image Retrieval for Suspect and Victim Identification. In *PCM*, December 2007.

[12] A. Jain, J.-E. Lee, R. Jin, and N. Gregg. Content-based Image Retrieval: an Application to Tattoo Images. In *IEEE ICIP*, November 2009.

[13] A. Jain and U. Park. Facial Marks: Soft Biometric for Face Recognition. In *IEEE ICIP*, November 2009.

[14] J.-E. Lee, A. Jain, and R. Jin. Scars, Marks and Tattoos (SMT): Soft Biometric for Suspect and Victim Identification. In *Biometrics Symposium*, September 2008.

[15] J.-E. Lee, R. Jin, A. Jain, and W. Tong. Image Retrieval in Forensics: Tattoo Image Database Application. *IEEE Multimedia*, 19(1):2–11, January 2012.

[16] D. Lowe. Distinctive Image Features From Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[17] S. Milborrow and F. Nicolls. Locating Facial Features with an Extended Active Shape Model. *ECCV*, 2008.

[18] U. Park and A. Jain. Face Matching and Retrieval Using Soft Biometrics. *IEEE Trans. on Information Forensics and Security*, 5(3):406–415, September 2010.

[19] U. Park, S. Liao, B. Klare, J. Voss, and A. Jain. Face Finder: Filtering a Large Database Using Scars, Marks and Tattoos. Technical Report MSU-CSE-11-15, Michigan State University, 2011.

[20] J.-S. Pierrard and T. Vetter. Skin Detail Analysis for Face Recognition. In *IEEE CVPR*, pages 1–8, 2007.

[21] K. Ramesha, K. Raja, K. Venugopal, and L. Patnaik. Template Based Mole Detection for Face Recognition. *International Journal of Computer Theory and Engineering*, 2(5):1793–8201, October 2010.

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

[23] C. Rother, V. Kolmogorov, and A. Blake. GrabCut: Interactive Foreground Extraction Using Iterated Graph Cuts. *ACM Trans. on Graphics*, 23(3):309–314, August 2004.

[24] A. Sapkota, B. Parks, W. J. Scheirer, and T. E. Boult. FACEGRAB: Face Recognition with General Region Assigned to Binary Operator. In *The IEEE Computer Society Workshop on Biometrics*, June 2010.

[25] B. Schölkopf, J. Platt, J. Shawe-Taylor, A. Smola, and R. Williamson. Estimating the Support of a High-dimensional Distribution. Technical Report MSR-TR-99-87, Microsoft Research, 1999.

[26] H. Wang and J. Chen. Improving Self-Quotient Image Method of NPR. *Int. Conf. on Computer Science and Software Engineering*, December 2008.