Identifying Prolonged Narcotics Users using Face Images

Prateekshit Pandey
Indraprastha Institute of
Information Technology, Delhi
2012078

Sarthak Ahuja
Indraprastha Institute of
Information Technology, Delhi
2012088

***Abstract*—World Drug Report by United Nations Office on Drugs and Crimes in 2014 clearly suggests that during the period 2003-2012 the increase in crime rates for possession for personal use worldwide was due to the increase in the total number of drug users, esp. cannabis and ATS (Amphetamine-Type Stimulants). Also with the recent improvements in the CCTV surveillance and the introduction of wearable video cameras for police officers in the United States and some other countries, a large amount of data is available for biometric analysis.**

We propose a system which can use the data of face images from such sources and identify faces possibly altered by prolonged narcotic drug usage. Experiments were conducted majorly on before-after drug mug-shot images made public by Multinomah Sheriff County Office. We use three different types of feature extraction techniques: HoG, Local Binary Patterns and Color Histogram, over which we apply a Support Vector Machine with different kernels to classify the face images.

I. Introduction

According to the World Drug Report by United Nations Office on Drugs and Crimes in 2014, the number of persons arrested for possession of drugs increased by 31% in the period of 2003-2012. While other kinds of crime like burglary, motor vehicle theft, robbery etc. declined, there were significant increases in drug-related crimes like drug trafficking and drug possession.

Crime related to drug trafficking varies depending on the type of drug and the supply patterns involved in different regions. In the Americas, cocaine follows cannabis as the second most prominent drug with respect to possession related to personal use, and was almost at par with cannabis (in first place) with respect to trafficking. In Asia, illicit opioids offer some competition to cannabis as the most prominent drugs for possession related to personal use, and illicit ATS emerge as the most prominent for trafficking offences. In Europe, illicit ATS ranked last among these four drug classes in terms of trafficking offences, despite being in second place (after cannabis) in terms of personal drug use offences.

The recent BSRIA study of the world electronic security market, which covers both established and important emerging (BRIC) markets, reveals that the world electronic security systems market is now worth more than US$58 billion, has continued to grow on aggregate throughout the recession, and is forecast to achieve double-digit compound annual growth in the period up to 2015. To add to that, Body Worn Video Cameras (BWVCs) have become available and police departments in the some areas of the United States and Hong Kong. They offer a potentially more practical choice for policing purposes given their size, convenience and ease of use.

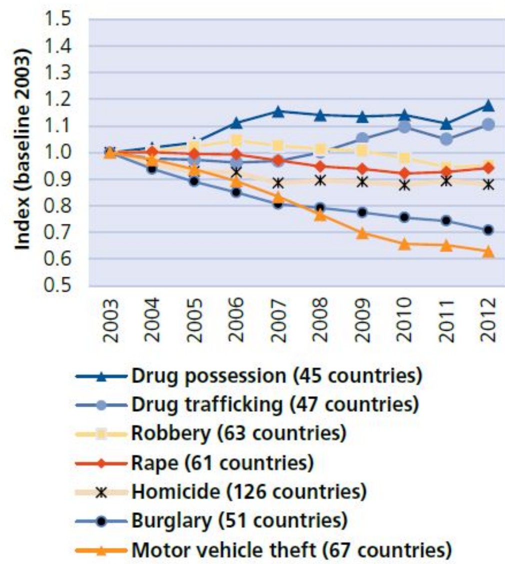

Fig. 1. Global trend in crime rates per population for selected types of crime, 2003-2012 (Source: UNODC)

Given the amount of data captured by these video cameras, there relevance to face biometrics is tremendous. Face images detected from these sources can help in criminal identification and investigation. As a step towards the larger goal of creating an autonomous system capable of identifying criminals, we targeted the fraction of criminals involved with narcotic drugs. The advantages of this selection is two-fold:

1) Prolonged use of narcotic drugs have an adverse effect on the face, which is very prominently visible
2) Given the increasing rates of drug related crimes, identification of prolonged drug users will dramatically reduce the size of gallery for recognition purpose.

We propose a system which uses standard feature extraction techniques like HoG, LBP and color Histogram and an SVM to classify face images into two classes: drug user, and non drug user. Section 2 describes the database we have used. Section 3 describes the feature extraction methods. Section 4 explains the results and throws light on the challenges faced by this problem. Section 5 discusses some potential future work.

II. Dataset: BA-Drug Dataset

The BA-Drug dataset (BA for "before-after") was compiled using 75 before-after drug usage images from the Internet. These images comprise of mug-shot images made public by the Multinomah Sheriff County's office. Each before-after image typically has one face image before doing drugs, and one face image for after doing drugs. Some of the before-after images did have more than one "after drugs" face image.

Fig. 2. Example of a before-after methamphetamine usage face image from the database (Source: Multinomah Sheriff County)

Separating the before and after images into two separate bins ("Normal" and "Drugged"), there were a total of 81 images in the "Drugged" bin and 75 images in the "Normal" bin. We ran Viola-Jones face detection algorithm, on which we normalized the eye-coordinates to the same position in order to align their face orientations. All these were normalized to a size of $200x200$.

III. Feature Extraction

We applied three techniques for feature extraction, namely: Color Histogram, Local Binary Pattern and Histogram of Oriented Histograms.

A. Color Histogram

A distinctive feature of the "Drugged" face images is the facial scars and changes in the color at certain regions of the face. This could be captured well using a color histogram technique, which will encode the changes in the color distribution.

Color histograms for all extracted (un-normalized) RGB face images were taken and fed into an SVM for classification. Each histogram was composed using 20 bins for each red, blue, and green color. Hence, the length of feature vector is $20^3 = 8000$.

B. Local Binary Pattern

The differences in the "Normal" and "Drugged" face images was also visible in the texture of the faces. Textures are very well encoded using Local Binary Patterns (LBP).

LBP encodes the local texture for a pixel using a binary pattern. A standard library (VLFeat) was used to compute the LBP patterns (of length 58) for all the face images.

C. Histogram of Oriented Gradients

As a standard object recognition technique, HoG works well to encode the shape of a given object. HoG computed the histogram of orientation of absolute gradients for each uniformally spread block of pixel around the given key point.

Here, for each face image, we took 64 equally spaced key-points all over the image, and computed Histogram of Oriented Gradients over them. Since, each HoG feature is of length 128, the total length of the feature vector was 64*128 = 8192.

Over these features, we applied a Support Vector Machine, with different kernels.

IV. Template based matching

Another approach we used involved making a template for each of the two classes (before drugs and after drugs), and classifying the images based on their similarity with the given template.

A naive approach to making a template for face image would be to average all the images in a defined training set. The results shown in the following figure explains why this approach will be invalid.

Fig. 3. Template of before and after drug images made simply averaging all the images in each category

From the figure, one cannot make out the difference in the two types of faces. The difference in texture and skin patterns gets averages out and hence, does not show up in the bigger picture. This gave us the intuition that the *edge maps* of the images must be used to incorporate the difference in facial features.

From the images of the dataset, it is apparent that there are a lot of irregularities in the face images in "after" category. These irregularities have been captured successfully in the following figure, which is able to record the aberrations in contrast to the *smoothness* of the "before" face images. These templates was computed in two steps:

1) Dividing each class in the dataset into 50% training and 50% testing sets
2) On the training set, canny edge detector with threshold 0.22 and sigma=2.2 was applied over each image

3) For each category, the edge maps were averaged out. Since this averaged image had a lot of spurious edges, these edges were removed by re-running canny edge detector (with threshold = 0.22 and sigma = 2.2) over the resultant average images, which gave us the templates.

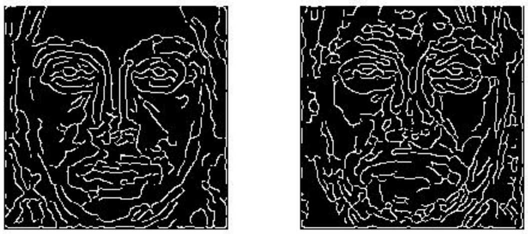

Fig. 4. Template of before and after drug images made using (canny) edge maps

The broad protocol we designed for this template based matching was to (1) compute distance (or similarity measure) of the edge map of a probe image with the template of both of the classes, and then (2) compute true positive rate and false positive rate based on which template it is nearer to (according to the distance measure). We used two ways to execute this protocol:

1) Use L1 distance over raw images
2) Use Chi-squared distance over HoG features of the edge map

The first approach fell flat, given that the L1 distance of any image (be it "before" or "after" face image) from the "before" template was lower as compared to "after" template. This is not surprising, since a simple modulus distance from a smoother edge map is usually lesser, given the less amount of irregularities.

We then applied Histogram of Oriented Gradients (HOG) over 64 uniformly spread key-points over the templates. For each probe image, canny edge detector was run with threshold = 0.22 and sigma = 2.2, and HOG features were computed over this edge map the exact same way we did for the template images. These HOG features were then compared using the Chi-squared distance metric, and based on that, we arrived at True Positive Rate = 100% and False Positive Rate = 100%. Though the results for True Positive Rate is elating, the results for False Positive Rate is disheartening, given that False Positive Rate must be very low in face recognition systems.

V. Patch-wise Features

Since these holistic features were not working very well, we decided to go for patchwise features. That too not all features: only those which we thought counted towards discriminability of drugged and non-drugged faces. We considered three broad patches: forehead, left cheek and right cheek. HOG features were computed over them, with 30, 20 and 15 uniformly distributed keypoints in the three regions respectively.

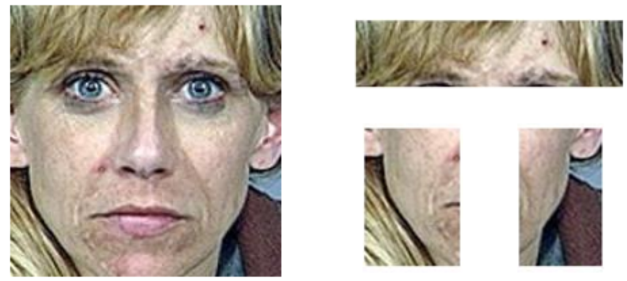

Fig. 5. Patches extracted from face images

The best accuracy fetched over classifying using SVM with polynomial was equal to 50%, with 98 support vectors (which is a lot, since the total number of data points in testing was 100). Thus, selective patchwise features weren't the best way to go forward with the problem.

VI. Results

We applied a Support Vector Machine over the given feature extractions of the face images and tried to classify them into "Normal" and "Drugged". Of the 154 images we had in total (74 Normal + 80 Drugged), we took a selection of 100 random images for training and the remaining 54 for testing. We use the following SVM kernels to obtain results:

SVM Kernels	Accuracy			Number of SVs		
	HoG	LBP	Color	HoG	LBP	Color
Linear Kernel	66.6667%	68.5185%	64.8148%	100	99	81
Radial Basis Function Kernel	55.5556%	68.5185%	57.4074%	100	100	98
Sigmoid Kernel	55.5556%	59.2593%	51.8519%	100	100	96
Polynomial Kernel	57.4074%	53.7037%	59.2593%	100	98	99

As it is apparent from the table, for most of these SVMs, the number of support vectors (nSV) is coming out to be very near to 100 (which is also the number of training samples). This means that the support vector machine is considering each data point as a support vector, which implies that it is overfitting the data. Hence, these results are not too reliable.

VII. Conclusions and possible future work

Most of the accuracies obtained during the course of this project have been erroneous, given the inadequate amount of data we have (which is the primary curse of the problem), and the lack of discriminability of our feature spaces. We look forward to implementing the following to get better results:

1) Transfer learning: Run autoencoder over a large set of good face images, learn the representation, and then use it to discriminate between drugged and non-drugged face images.

2) Using deeper edge features: one way to go about it is using scattering wavelet transform, which we would like to explore.
3) Cascading classifiers: we could use multiple weak classifiers in cascade so that each weak classifier is able to contribute towards decreasing the false positives.

We conclude by saying that this is an extremely non-trivial problem, with many constraints. But at the same time, it is of high relevance in context of the increasing crime rates related to drug abuse.

VIII. Work distribution

Prateekshit: Data collection, formatting data, Coding (HoG, LBP, Patchwise, Template Matching)

Sarthak: Data collection, formatting data, documenting face images, Coding (Color histogram, CrAngleA, Average Face, SVM)