Deep Learning Framework for Detection of an Illicit Drug Abuser Using Facial Image

Tanu Gupta, Meghna Goyal, Goutam Kumar and Balasubramanian Raman

Abstract The detection of an illicit drug abuser by analyzing the subject's facial image has been an active topic in the field of machine learning research. The big question here is up to what extent and with what accuracy can a computer model help us to identify if a person is an illicit drug abuser only by analyzing the subject's facial image. The main objective of this paper is to propose a framework which can identify an illicit drug abuser just by giving an image of the subject as an input. The paper proposes a framework which relies on Deep Convolutional Neural Network (**DCNN**) in combination with Support Vector Machine (SVM) classifier for detecting an illicit drug abuser's face. We have created dataset consisting of 221 illicit drug abusers' facial images which present various expressions, aging effects, and orientations. We have taken random 221 non-abusers' facial images from available dataset named as, Labeled Faces in the Wild (LFW). The experiments are performed using both datasets to attain the objective. The proposed model can predict if the person in an image is an illicit drug abuser or not with an accuracy of 98.5%. The final results show the importance of the proposed model by comparing the accuracies obtained in the experiments performed.

Keywords Deep learning · Convolutional neural network · Face recognition

1 Introduction

The number of illicit drug abusers has been increasing rapidly on a global scale. As per the World Drug Report 2017 released by United Nations Office of Drugs and Crime [1], 255 million people were using illicit drugs in 2015 whereas there were 247 million people using the illicit drugs in 2014. Considering such a huge number and

T. Gupta (✉) · M. Goyal · G. Kumar · B. Raman
Indian Institute of Technology Roorkee, Roorkee, India
e-mail: tgupta@dm.iitr.ac.in

B. Raman
e-mail: balarfma@iitr.ac.in

B. B. Chaudhuri et al. (eds.), *Proceedings of 3rd International Conference on Computer Vision and Image Processing*, Advances in Intelligent Systems and Computing 1022,
https://doi.org/10.1007/978-981-32-9088-4_2

Fig. 1 Sample images from our dataset representing effects of illicit drugs on human face

the growth at which the problem of illicit drug abuse is growing, it is becoming more apparent and challenging for humanity. Globally the most commonly used illicit drugs are cocaine, methamphetamine (meth), heroin, marijuana, and crack cocaine. This list is not exhaustive but will be used for the scope of this research.

As per the extensive research done by an organization named Rehab [2], the repeated and prolonged use of illicit drugs causes various side effects. A noticeable change in the face is the most prominent side effect of almost all illicit drugs as shown in Fig. 1. For instance, the physiological side effects of meth like acne, dry mouth, teeth clenching, dull skin, and self-inflicted wounds on the face and body are documented by the project Faces of Meth [3]. The regular consumption of certain illicit drugs causes physiological changes in the skin which make it possible to detect illicit drug abusers only by observing the facial changes. The detection of illicit drug abusers has its application in the areas like a job or army recruitment where the background verification of a person is required.

The detection of illicit drug abusers based on facial analysis is an interesting and useful subject, but has been little explored. The research in [4] has proposed a dictionary learning framework which will help one to recognize the face of a subject even after he has been abusing illicit drugs for a long time. In order to add value to that research, we have built a framework which can tell if a person is an illicit drug abuser or not just by analyzing his current image. In this research work, various experiments have done in order to detect illicit drug abusers from facial images. Initially, features have been extracted using handcrafted representations like HOG, GIST, BISF, LBP, BagOfWords (BoW), and **DCNN** and after that classifier algorithms like support vector machine (SVM), Random Forests (RF), and K-Nearest Neighbor (KNN) have been used to classify the images. In proposed method, SVM classifier is used with **DCNN**. To showcase the final results, comparison between accuracies obtained from the various handcrafted descriptors used with various classifiers and the proposed method has been done.

Our contribution through this paper is summarized below:

1. We have created an **Illicit Drug-Abused Face dataset** containing images of 221 illicit drug abusers which are collected from various sources from Internet.
2. We have proposed a **DCNN** framework with SVM classifier to detect an illicit drug abuser's face and a non-abuser's face in a combined dataset of images.

The rest of the paper has been organized as follows. A description of related work is presented in Sect. 2. The proposed method is described in Sect. 3. The dataset and scenarios that have been used for experimentation have been described in Sect. 4. The experimental results have been presented in Sect. 5. Section 6 concludes the paper.

2 Related Work

Since drug abuse is a serious concern in today's society, several types of research have been conducted on various impacts of illicit drugs on human health [5]. In order to alert people about the adverse effects of meth, the images of haggard and sunken-cheeked faces of meth addicts were released [3]. Further, images of cocaine, crack, and heroin addicts representing weight loss and fast aging effect were collected and disseminated, so that people could be made aware of the effects of these drugs.

An ample amount of research has been completed in order to use face recognition for various applications. In [6], face recognition is done using the LBP descriptor. The images in which the faces could be recognized may have different facial expressions, illumination, aging effects, and other external variations. These faces could be recognized with good accuracy. A deep learning framework was proposed in [7] for recognizing the developmental disorders. A fine-tuned DCCN was used for the feature extraction along with the SVM classifier. The model is tested on different experimental scenarios which include differentiation of the normal faces from those with cognitive disabilities. The dataset consisted of 1126 images of normal subjects and disabled subjects each. The accuracy obtained is 98.80%.

In spite of all the researches regarding illicit drug abuse, there is a little work done on the detection of illicit drug abusers. In 2017, [4] research was conducted which objectified the facial changes of a subject due to the consistent consumption of illicit drugs and the deterioration in the performance of two commercial face recognition systems. Yadav et al. [4] proposed a dictionary learning framework to recognize the effect of illicit drug abuse on face recognition. The accuracy of the proposed model was 88.88%.

3 Proposed Method

The objective of the paper is to efficiently detect an illicit drug abuser's face and a non-abuser's face. Machine learning based classification techniques can give good accuracy only if the features are extracted properly from the images. We have used basic handcrafted representations such as HOG [8], GIST [9–11], LBP [12–14], BISF [15], and BOWs [16] to get the best possible facial features. The classifiers RF, SVM, and KNN are trained with all the representations, respectively. Further improvement in the accuracy is done by proposing a deep learning framework CNN for representations and again train SVM classifier to classify the images. As shown in the block diagram in Fig. 2, two CNN architectures have been used in the framework. The facial image is given as an input to the first CNN and the salient features extracted from the facial images are given as an input to the second CNN. The features obtained from the above two branches are merged and then fed to the classifier.

3.1 Saliency-Based Feature Extraction

The cluster based co-saliency method has used to get the salient features of the face. Figure 3 shows the salient images corresponding to the input drug abused facial images. The highly salient parts and the lower salient parts represent the unique parts of the images and the background of the images, respectively.

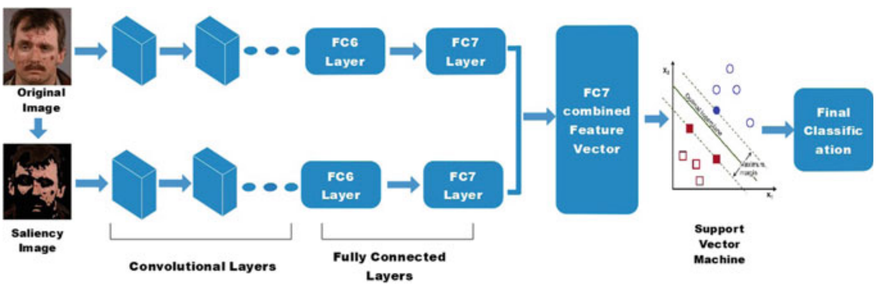

Fig. 2 A block diagram representing proposed architecture

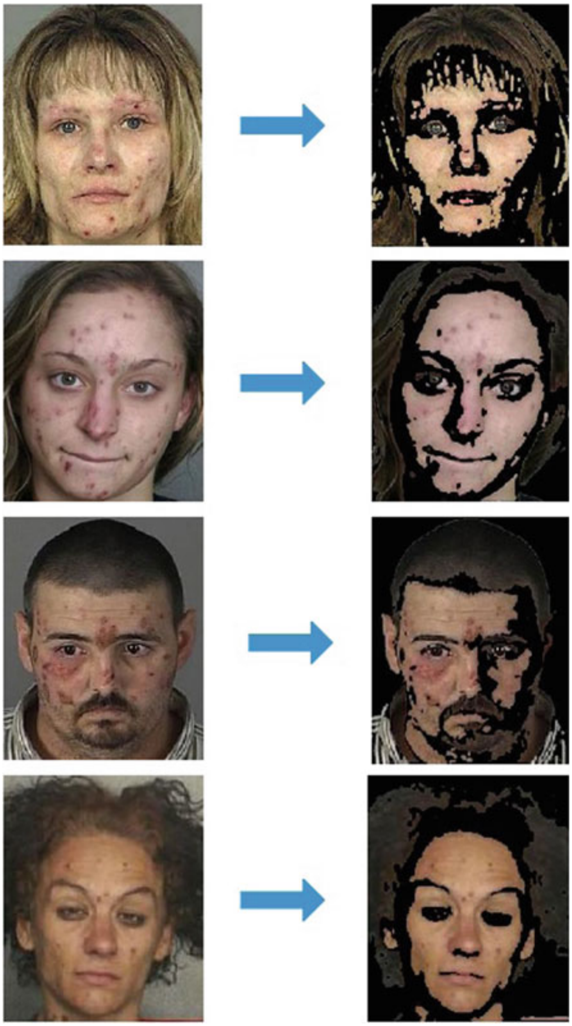

Fig. 3 Salient images corresponding to the drug abused images

3.2 Using Deep CNNs for Extracting Features

We have used deep representations over other handcrafted representations mainly because of the superiority of **DCNN** over traditional descriptors for complex image recognition tasks. Also, it requires minimal processing [17]. Face recognition of drug abusers is done using deep CNNs fc6 and fc7 layers. The Alex-Net architecture [18] is used for the fine-tuning of all different CNNs. The input layer is of $227 \times 227 \times 3$, conv of $96 \times 11 \times 11$, max pooling of 3×3, conv of $128 \times 5 \times 5$, max pooling of 3×3, conv of $256 \times 3 \times 3$, conv of $192 \times 3 \times 3$, conv of $192 \times 3 \times 3$, 3×3 pooling 4096×1, fc $- 4096 \times 1$ fc 1000×1 fc layers. Before being fed into the respective

CNNs, facial images are resized to 227×227. The initial learning rate is set to 0.001 for the final fc layer and 0.0001 for the remaining layers while fine-tuning. The dropout is 0.5 and the momentum is 0.8 for a total 10,000 iterations. For capturing the different aspects of the facial images, two CNNs are used and 4096 representations are extracted overall from them. These representations are concatenated together to form a feature vector comprising of 67,840 elements.

3.3 Classifier

The representations extracted from the above methods are trained and classified with the help of SVM classifier. The SVM is trained on the linear kernel and the value of kernel offset is set to 0.

4 Experimental Setup

A description of the experimental scenarios on which the model is tested along with the details of the dataset has been elaborated below. The experiments are conducted using the MATLAB 2017a.

4.1 Description of Dataset

Illicit drug abuse is a shameful and confidential issue of an addict. As a result, dataset of limited size is available for the images of illicit drug abusers. However, few websites [2] have released images of drug abusers in order to put light on the devastating physiological side effects of illicit drugs. Moreover, few images and stories of drug abusers who have quit the drugs are available to motivate other drug abusers. We have collected the number of images in the database using facial images of Internet subjects. Illicit Drug abuse Database consists of 221 facial images of drug abusers. The database includes the addicts of drugs such as cocaine, heroin, meth, alcohol, crack, and coke. The images available on the web are of different facial orientations and few of them are blurred. So we have cropped them in order to get the better accuracy.

We have divided the dataset into the two categories, facial images of illicit drug abusers and images of the non-drug abusers (LFW). LFW dataset comprises facial images of various orientations. The images are different in many aspects such as face orientation and clarity of images. The dataset is divided in the ratio of 70:30 for training and testing purpose.

4.2 Experimental Scenario

In order to check the performance of the proposed model, we have performed experiments using handcrafted representations on the given dataset. The purpose of the experiments is to differentiate between an illicit drug abuser's face and non-abuser's face using the proposed model for the improved accuracy.

For the comparison purpose, we have performed the experiments in which the features extracted using HOG, BISF, GIST, BoW, and LBP are fed to the classifiers (KNN, SVM, RF). Then we have taken a record of accuracy obtained on the given dataset. Afterward, the same experiment is performed using the proposed model.

5 Experimental Results

The experimental results are recorded in the order described below. To begin with the experiments, we have extracted features using the HOG descriptor. The images in the database are classified using the extracted features with SVM, RF, and KNN classifier algorithms. The accuracies as recorded in Table 1 are 89.57%, 91.30%, and 84.35%, respectively. Next we have used GIST for the purpose of feature extraction. Again the images are classified using the extracted features with all the three earlier mentioned algorithms. This time the accuracy as recorded are 94.78, 93.91, and

Table 1 Comparison of the proposed model and handcrafted representation accuracies

Technique	Accuracy (%)
HOG+SVM	89.57
GIST+SVM	94.78
BSIF+SVM	72.17
LBP+SVM	91.30
BagOfWords+SVM	87.96
HOG+Random Forests	91.30
GIST+Random Forests	93.91
BSIF+Random Forests	89.57
LBP+Random Forests	93.91
BagOfWords+Random Forests	92.59
HOG+KNN	84.35
GIST+KNN	94.78
BSIF+KNN	74.78
LBP+KNN	93.91
BagOfWords+KNN	83.33
Proposed approach	98.5

Table 2 Confusion matrix for drug-abused and non-abused dataset

	Drug abused	Non-abused
Drug abused	0.925	0.075
Non-abused	0	1

Fig. 4 Samples of correctly classified images from the used dataset

Fig. 5 Samples of wrongly classified images from the used dataset

94.78%. In the next experiment, LBP is used to extract the features which are then used with SVM, RF, and KNN classifiers. The accuracies obtained this time are 91.30%, 93.91%, and 93.91%, respectively. The performance of BSIF descriptor for feature extraction with each of the abovementioned classifier is measured as 72.17, 89.57, and 74.78%. Moreover, the accuracies obtained for image classification using BagOfWords for feature extraction with all the three classifiers are 87.96, 92.59, and 83.33%. In order to compare the accuracy of different traditional methods with that of the proposed method, the classification of abuser's and non-abuser's faces is performed using DCNN method with SVM classifier. The accuracy obtained in this case is 98.5%.

There are 60 images of illicit drug abusers which have been used for the testing the framework. The confusion matrix represented in Table 2 gives the accuracy up to which the model is able to do the recognition. Figure 4 represents the few examples of correctly detected images whereas Fig. 5 consists of some wrongly detected images.

6 Conclusions

Nowadays, the impact of illicit drug abuse is becoming a serious concern in the world. The ability to differentiate between an illicit drug abuser's face and a non-abuser's face with better accuracy is one of the challenges in the field of machine learning. Our work would facilitate in solving many real-life practical problems such as: it can be used by law enforcement to identify if any suspect is an addict, it can be used at airport terminals to be able to spot drug traffickers. As the facial changes are drastic when images from before and after the use of drugs are compared, a combination of several experiments is performed to differentiate between an illicit drug abuser's face and a non-abuser's face in a combined dataset. This paper presents a deep learning framework for face recognition, which will identify if the subject in a given image is a drug addict or not. The proposed DCNN framework is able to achieve an overall accuracy of 98.5% with SVM. Also, in this paper, we have presented a database consisting of 221 images of illicit drug abusers. These images cover variations in the features such as expressions, aging effects, and orientations. This database can further be used in future works for validating similar approaches.

References

1. United Nations Office on Drugs and Crime: World drug report 2017, Vienna. https://www.unodc.org/wdr2017/index.html (2017). Accessed 20 Jan 2018
2. Rehabs: Faces of addiction: image collection of drug abusers. https://www.rehabs.com/explore/faces-of-addiction. Accessed 20 Jan 2018
3. Multnomah County Sheriff Office: Faces of meth. http://www.mcso.us/facesofmeth/. Accessed 25 Jan 2018
4. Yadav, D., Kohli, N., Pandey, P., Singh, R., Vatsa, M., Noore, A.: Effect of illicit drug abuse on face recognition. In: 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1–7. IEEE (2016)
5. Devlin, R.J., Henry, J.A.: Clinical review: major consequences of illicit drug consumption. Crit. Care **12**(1), 202 (2008)
6. Ahonen, T., Hadid, A., Pietikainen, M.: Face recognition with local binary patterns. In: Computer Vision—ECCV 2004. ECCV. Lecture Notes in Computer Science, vol. 3021. Springer, Berlin, Heidelberg (2004)
7. Saini, A., Singh, P., Shukla, P., Gupta, T., Raman, B.: A deep learning framework for recognizing developmental disorder. In: IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA, pp. 1–10. IEEE (2017)
8. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005, vol. 1, pp. 886–893. IEEE (2005)
9. Hays, J., Efros, A.A.: Scene completion using millions of photographs. Commun. ACM **51**(10), 87–94 (2008)
10. Douze, M., Jégou, H., Sandhawalia, H., Amsaleg, L., Schmid, C.: Evaluation of gist descriptors for web-scale image search. In: Proceedings of the ACM International Conference on Image and Video Retrieval, p. 19. ACM (2009)
11. Oliva, A., Torralba, A.: Modeling the shape of the scene: a holistic representation of the spatial envelope. Int. J. Comput. Vis. **42**(3), 145–175 (2001)
12. He, D.-C., Wang, L.: Texture unit, texture spectrum, and texture analysis. IEEE Trans. Geosci. Remote Sens. **28**(4), 509–512 (1990)
13. Wang, L., He, D.-C.: Texture classification using texture spectrum. Pattern Recognit. **23**(8), 905–910 (1990)
14. Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: application to face recognition. IEEE Trans. Pattern Anal. Mach. Intell. **28**(12), 2037–2041 (2006)
15. Kannala, J., Rahtu, E.: BSIF: binarized statistical image features. In: 2012 21st International Conference on Pattern Recognition (ICPR), pp. 1363–1366. IEEE (2012)

16. Zhang, Y., Jin, R., Zhou, Z.-H.: Understanding bag-of-words model: a statistical framework. Int. J. Mach. Learn. Cybern. **1**(1–4), 43–52 (2010)
17. Shamov, I.A., Shelest, P.S.: Application of the convolutional neural network to design an algorithm for recognition of tower lighthouses. In: 2017 24th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), pp. 1–2. IEEE (2017)
18. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)