Identifying Illicit Drug Abuse

1st Aishik Biswas
Computer Science
Lakehead University
abiswas3@lakeheadu.ca
1158823

2nd Bhuwan Dutt
Computer Science
Lakehead University
bdutt@lakeheadu.ca
1150522

***Abstract*—In this paper we propose a framework for identifying illicit drug use using Computer Vision. We extracted features from multiple images of people who abuse drugs and who do not and let the machine learning algorithm give a prediction if the person in the image is a drug user or not.**

***Index Terms*—Drug Abuse, Classification, Social Media**

I. Introduction

Drugs have existed in our society since the start of humanity. The chemicals that can change our neuro-chemistry and leave euphoric effects has always fascinated humans. But as these drugs have evolved,so have our use. More people are dying of prolonged and wrong use of drugs every year. World Drug Report from 2019 suggests that over 35 million people worldwide suffer from drug use disorders. With stigma attached to drug abuse and people not been able to reach for help because of societal judgement and shame, this problem becomes more complex. Another report from United Nations suggests that due to pandemic the opiod use have increased by 73 percent and opioids may not be the most abused drug but it is responsible for most deaths in the world [2]. With this alarming rise in substance abuse and deaths, it becomes very necessary to identify and educate people on the safe and controlled use of drugs and the harm of prolonged and wrong use of them.
Drug abuse is obviously harmful and determent to health of a person but it can be hard to find if a person is using drugs or not. The person abusing drugs mostly try to hide it as much as they can because of the shame attached to it. It is difficult to tell if person is using drugs but there are some signs that are visible on the face of a user. Figure 1 shows the effect some drugs have on facial features of it's users. These images are suggestive and it depends on the amount of drugs a person use and for how long. Our aim is to extract those physical features and give a prediction if a person is using illicit drugs or not. Previous approaches have shown that [3] features like left cheek, right cheek, forehead, binocular region and full face are the regions where the drug abuse can have a distinctive effect. These features were extracted using approach like SIFT and HOG and then we applied machine learning models, SVM and KNN, to classify if the person has been using illicit drugs or not

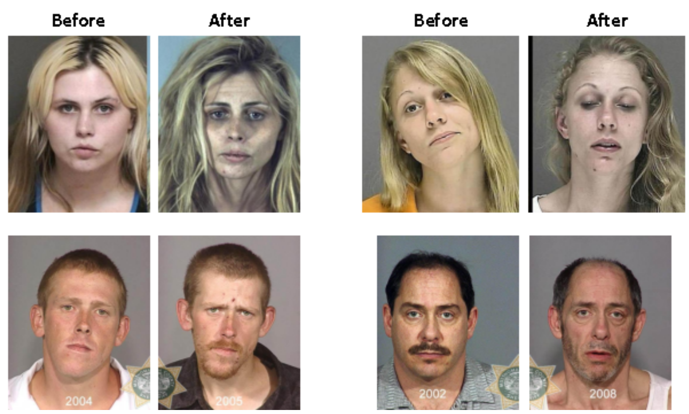

Fig. 1. Before and after use of drug

II. Dataset

Acquiring these types of confidential dataset was pretty hard. Because none of the people would like to give their details and their photographs of their before an after use of drug abuse due to the stigma of the society as well as the fear of being caught and harassed in various ways. As they want to lead a normal life without being judged. And some of the people are still using it and would not like to come in the light. So thanks to initiatives such as "Face the Meth" which is a project started in 2008 by Oregon Police contains mugshots of people abusing methamphetamine and Rehab.com. Therefore thanks to these organisations we were able to curate the dataset for our experimental use.

III. Methodology

A. Data Augmentation And Pre-Processing

1) Pre-Processing: Data pre-processing is the first and foremost step that is very essential and paves the way for better feature extraction and more accurate predictions. Data pre-processing was required to remove noise and extract better features to give accurate predictions. We firstly manually cropped the images so only the face is left in the image and other parts are cropped out. We also resized our images and converted them to gray-scale for feature extraction.

B. Augmentation

Data Augmentation is a technique used to add new data to our dataset using the existing data. This. new data can be modified data of existing data or can be new synthetic data created from existing data [4]. Data Augmentation was needed as our dataset was very scant due to reasons foretold. In Computer Vision, data Augmentation can be done by making slight changes to the images like rotating, tilting, cropping, resizing, etc. We Augmented the data using the library Albumentation. Albumentation has pre-written functions to do various Augmentation tasks using single command, we just need to pass the required parameters. We increased the volume of our training dataset using Albumentation by increasing the Drug Addicts Images by approximately 33

Fig. 2. Image cretaed using Albumentation

C. Feature Extraction Techniques

1) SIFT (Scale-invariant Feature Transform): SIFT, or Scale Invariant Feature Transform, is a computer vision feature detection algorithm. SIFT assists in locating the image's local features, often known as the image's 'keypoints.' These scale and rotation invariant keypoints can be utilised for picture matching, object detection, scene detection, and other computer vision applications. During model training, we can use the keypoints created by SIFT as features for the image. SIFT features have a significant benefit over edge or hog features in that they are unaffected by image size or orientation. Here's another shot of the Eiffel Tower, this time with a smaller version. The object's keypoints in the first image correspond to the keypoints in the second image. The same is true when the object in one image is slightly rotated in the other.

a) Gaussian Blur and Difference of Gaussian: We use the Gaussian blur to reduce the noise of the image. So, for every pixel in an image, the Gaussian Blur calculates a value based on its neighboring pixels. Below is an example of image before and after applying the Gaussian Blur. As is it seen that all the texture and the minor details of the image is removed and only the distinguishable information like shape and edges remains. So we need to increase the amount of blur scale and make four octaves to it. Next we perform Difference of Gaussian which is an enhancement algorithm which subtracts the blurred version of the image from another to make a less blurred version of the original. It creates another set of images for each octave by subtraction.

b) Keypoints: As a result, we identify the image's important points. The image's local maxima and minima are the keypoints. The points on the photos represent the image's unique points, allowing us to compare them to similar images with different transformations. We'll remove keypoints with low contrast or that are too close to the edge. So, we run a check to see if there are any keypoints that aren't in the right place. These are keypoints that are close to the edge and have a high edge reaction, but are susceptible to minor noise. To find such keypoints, a second-order Hessian matrix is used.

2) Bag Of Visual Words: Bag of Visual Words is a technique that takes inspiration from Bag of Words technique in Natural Language Processing (NLP). The basic concept behind a bag of visual words (BOVW) is to portray an image as a collection of attributes. Keypoints and descriptors are two types of features. Keypoints are the "stand out" points in an image, so they will remain the same no matter how the image is rotated, shrunk, or expanded. The descriptor is the keypoint's description. To build vocabularies, we employ keypoints and descriptors, and each image is represented as a frequency histogram of characteristics in the image. We may later use the frequency histogram to locate additional similar photos or predict the image's category. We detect features, extract descriptors from each image in the dataset, and build a visual dictionary. Detecting

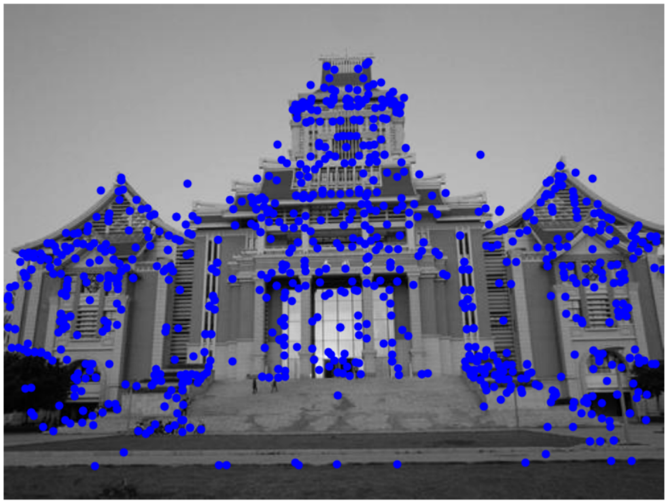

Fig. 3. Example of SIFT Feature Detection.

features and extracting descriptors was done using SIFT and then we made clusters using K-Means. After which we created a frequency histogram from the vocabulary in each image, as well as the frequency of the vocabularies in the image. Those histograms are our visual dictionary (BOVW).

3) HOG: HOG, or Histogram of Oriented Gradients, is a feature descriptor for extracting features from image data. It is commonly utilised in object detection tasks in computer vision. The extraction of features through HOG has below components

a) Preprocessing of Images: The image must be preprocessed to reduce the width to height ratio to 1:2. The picture should ideally be 64 × 128 pixels in size. This is because the features will be extracted by splitting the image into 8*8 and 16*16 patches. Having the stated size (64 x 128) simplifies all of our computations.

b) Calculating gradients of x and y direction: The gradient for each pixel in the picture must then be calculated. Gradients are minor changes in x and y coordinates. I'm going to pick a tiny section of the image and compute the gradients on it:

For this fix, we'll get the pixel values. The pixel value 85 has been highlighted. Now subtract the value on the left from the pixel value on the right to get the gradient (or change) in the x-direction. To compute the gradient in the y-direction, subtract the pixel value below the chosen pixel from the pixel value above it. As a consequence, for this pixel, the resulting gradients in the x and y directions are:

- X-direction change (Gx) = 89 – 78 = 11
- Y-direction change (Gy) = 68 – 56 = 8

121	10	78	96	125
48	152	68	125	111
145	78	85	89	65
154	214	56	200	66
214	87	45	102	45

Fig. 4. Pixel matrix

c) Magnitude and Orientation: Using the gradients we calculated in the last step, we will now determine the magnitude and direction for each pixel value. For this step, we will be using the Pythagoras theorem.The gradients are basically the base and perpendicular here. So, for the previous example, we had Gx and Gy as 11 and 8. Let's apply the Pythagoras theorem to calculate the total gradient magnitude:

Total Gradient Magnitude = $\sqrt{(G_x)^2 + (G_y)^2}$

Total Gradient Magnitude = $\sqrt{[(11)^2 + (8)^2} = 13.6$

Then, for the same pixel, determine the orientation (or direction). We already know how to write the tan for the angles:

$tan(\phi) = G_y/G_x$

As a result, the angle's value is: $tan(G_y/G_x)$

When we punch in the data, the orientation comes out to be 36. So now we get the entire gradient (magnitude) and the orientation for each pixel value (direction). Using these gradients and orientations, we must create the histogram.

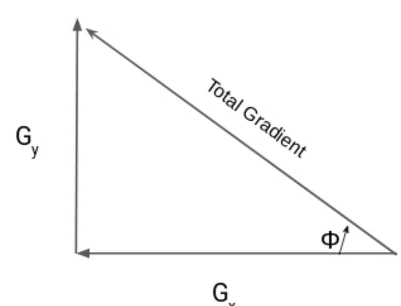

Fig. 5. Pythagoras Theorem

Fig. 6. Example of HOG Feature Detection.

D. Classification Techniques

1) KNN: KNN is one of the simplest classification algorithms. Even with such simplicity, it can give highly competitive result. KNN can also be used for regression problems. KNN captures the idea of similarity (sometimes called distance, proximity, or closeness). For calculating distance different methods like Euclidean distance, Minkowski distance, hamming distance are used. In KNN classification, after finding nearest neighbors majority voting is done among nearest neighbors i.e., target variable from majority class is chosen as output for test instances [9]. The main task of KNN algorithm is to find best K value. When K=1 then target variable of first nearest neighbor is directly assign to test instance. In this case, there are high chances for over-fitting problem. When K=n, then it chooses majority class among all instances available for training. In this scenario, our model becomes more simpler So there are chances for under-fitting problem. We can choose value of K manually putting K values and also, we can run hyper parameter tuning to find best value of K. Distance formula for minkowski distance is as follows:

$$(\sum_{i=1}^{n} |X_i - Y_i|^p)^{1/p}$$

2) SVM: A Support Vector Machine (SVM) is a machine learning method that examines data and divides it into two categories. The goal of the SVM method is to determine the best line or decision boundary for categorising n-dimensional space so that fresh data points can be placed in the correct category easily in the future [8]. The extreme points/vectors that contribute in the creation of the hyper-plane are chosen using SVM.

IV. Results

We were able to achieve the highest accuracy using SVM with Bag of Visual Words. We observed that even KNN showed lower accuracy, we were able to get a better accuracy in BOVW every time.

The results are reported below.

Results		
Classifier	Features	Accuracy
KNN (k=3)	HOG BOVW	72% 75%
KNN (k=5)	HOG BOVW	65% 67%
KNN (k=7)	HOG BOVW	60% 63%
SVM	HOG BOVW	82% 92%

V. Discussion

As we were able to achieve better accuracy using BOVW it is safe to say that BOVW can be used a good feature extraction technique in classification of Drug Abuse. The reason for KNN performing poorly can be the high dimensionality of the data. As dimension becomes huge, KNN is unable to identify thw nearest neighbour. This phenomenon is known as Curse of dimensionality. As the SVM uses hyper-planes to draw the decision boundary, it is able to achieve high accuracy. In the previous study [7], only HOG and SIFT were used as feature extractors but the algorithms used were more advanced (AlexNet, VGG), which we could not test due to computational limitations.

VI. Conclusion

Illicit Drug Abuse is a grave issue that is seeing a steep rise and we believe that this issue needs urgent attention. This study focuses to take help of Computer Vision to identify the damage done to a person's facial features by prolonged illicit drug use. But, this kind of technology has a lot of potential to be misused as it can be used to marginalise people and increase the discrimination against people who need help with their addictions as it can equip someone to identify the drug addict just by feeding a picture to a computer program. We believe that any such study should be proceeded with caution.

References

[1] "Distinctive Image Features from Scale-Invariant Keypoints" David G. Lowe International Journal of Computer Vision volume 60(2004)

[2] "Face recognition using HOG" EBGM A Albiol, D Monzo, A Martin, J Sastre, A Albiol - Pattern Recognition Letters, 2008 - Elsevier

[3] "The Faces of Meth" Updated: Jan. 11, 2019, 9:07 a.m. — Published: Dec. 28, 2004

[4] P. Mishra, “Data Augmentation – Towards Data Science,” Data Augmentation – Towards Data Science, Dec. 2021. https://towardsdatascience.com/tagged/data-augmentation.
[5] United Nations, “World Drug Report 2019: 35 million people worldwide suffer from drug use disorders while only 1 in 7 people receive treatment,” Unodc.org, 2019. https://www.unodc.org/unodc/en/frontpage/2019/June/world-drug- report-2019-35-million-people-worldwide-suffer-from-drug-use- disorders-while-only-1-in-7-people-receive-treatment.html.
[6] United Nations Office on Drugs and Crime, “UNODC World Drug Report 2020: Global drug use rising; while COVID-19 has far reaching impact on global drug markets,” United Nations : Office on Drugs and Crime, 2020. https://www.unodc.org/unodc/press/releases/2020/June/media-advisory—global-launch-of-the-2020-world-drug-report.html.
[7] D. Yadav, N. Kohli, P. Pandey, R. Singh, M. Vatsa, and A. Noore, “Effect of illicit drug abuse on face recognition,” 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, Mar. 2016. doi: 10.1109/wacv.2016.7477556.
[8] J. Brownlee, “Support Vector Machines for Machine Learning,” Machine Learning Mastery, Aug. 12, 2019. https://machinelearningmastery.com/support-vector-machines-for-machine-learning/.
[9] [3]T. Srivastava, “Introduction to KNN, K-Nearest Neighbors : Simplified,” Analytics Vidhya, Mar. 07, 2019. https://www.analyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering/.