SAPIENZA
UNIVERSITÀ DI ROMA

Lie Detection Thesis

Facoltà di Ingegneria dell'informazione, Informatica e Statistica

Corso di Laurea Magistrale in Informatica

Candidate

Emanuele Orfanelli
ID number 1383726

Thesis Advisor

Prof. Luigi Cinque

Co-Advisors

Dr. Danilo Avola
Dr. Daniele Pannone

Academic Year 2018/2019

Lie Detection Thesis
Master thesis. Sapienza – University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: emanueleorfanelli@gmail.com

Dedicated to
my Family and Friends

Acknowledgments

fill

Contents

Chapter 1

Introduction

In this chapter we give an overview of the work (Par. 1.1). We start with giving some information about how people perform at detecting lies (Par. 1.2). We then present a taxonomy of the current state of the art (Par. 1.3) concerning lie detection with particular emphasis on computer vision. The last section is about the structure of the rest of this work, and our contribution to it (Par 1.4).

1.1 Overview of the work

Deception detection has always been a very interesting topic since it has numerous social implications in many different fields. In the past 40 years there has been a steady increase in researchers interested in studying deceptive behavior, first from a sociological and psychological point of view, and in more recent years from a technological standpoint, aided by the advance of machine learning techniques.
Our work aims to recognize whether a person is lying or telling the truth by performing a frame by frame analysis on a database of videos, through the OpenFace framework [65], and extracting a subset of the different facial movements performed by the person in the video. After finishing the extraction, the data is submitted to analysis by using different machine learning techniques.
The dataset we are using was provided by Perez-Rosas et al [50] and is composed by 121 videos taken from real life trials. These 121 videos have been labeled by the authors with the help of experts to indicate the facial movement occurring in each frame. Since those videos are taken from real life trials, the illumination, pose, audio and video features are not homogeneous and often substandard. Substantial work was done to eliminate the not relevant parts, and many video were trimmed or deleted as result.
From the remaining videos we have 58 different subjects, meaning that there are more videos with the same subject. This means that it was important to divide the training and test set not to just make the classifier "remember" the person. This was done by never having the same subject appear in the training and test set.
After trimming and dividing the videos we extracted all the facial movements and split the data into training and test set based on the subject ID. Then fed them through a carefully tuned SVM for classification, and achieved x% accuracy on the test set.
The result of this work, especially when improved by having a better and bigger training dataset, can be useful in different situations, even though it is important to remember that privacy is a real concern.
We think the result of this work can be used to aid in airport security, work interview, many kind of social interactions. It could be eventually used by the public to review a speech of a political candidate or by the police force as an aid to interrogation. Another use can be in trials where people life are at stake and discerning a true or false testimony might be vital. We hope this works proves to be socially useful.

1.2 People Lie Detection

It's very rare for people to be able to consistently discern between lies and truths, even though we hear lies regularly in the course of our lives. In fact most untrained people perform like chance (50%) when tasked with detecting lies [52], and considering that the ordinary person lies at least twice a day on average [42], and that in this digital age the amount of lies told daily are increasing [31] due to on-line communication making us feel more protected and confident in our deceptions, the problem of detecting lies is an important one and we think it deserves a good deal of research effort.
The problem with people's lie detection is based on the difficulty of being objective. We are biased by so many factors and skilled deceivers can take advantage of that. An important reason is that people generally think it is easy to spot liars, underestimating the effort it takes by having too much confidence on their own judgment and capabilities, and by believing in gut feelings instead of empirical clues [54].
In a study by Baker et al. [6] 116 participants were asked to judge 20 videos of people pleading for the safe return of missing family members, half of which were lies where the pleader was the one responsible for the disappearance of the victim. The partecipants provided confidence ratings, the cues they utilized to make their judgment, and their emotional response to each video.
Weirdly, emotionally intelligent people perform worse at deception detection. This is due to their greater sympathetic feelings towards others (enhanced gullibility) that can often cloud their judgment. In [15] De Paulo et al. analyze the accuracy of deception judgments from a collection of 206 documents and 24.483 judges. They found that people can differentiate lies and truths with an accuracy of 54%, with a lie detection accuracy of 47% and a truth detection accuracy of 61%. Their findings reveal that is easier for people to discriminate lies from the audio cues rather than the visual ones.
In another study [33] 192 students obtained an accuracy of 55.2% on lie detection, with 61% accuracy for guilty suspects and 49% for innocent ones. Police officers and other trained officials seems to perform better at lie detection obtaining around 70% accuracy at detecting both lies and truths correctly [67].

Lie Catchers	*Number of Lie Catchers*	*Overall Accuracy*	*Liar Spotting Accuracy*	*Truth-teller Spotting Accuracy*
Diverse (206 documents synthesized)	N/A	54%	47%	61%
Undergraduate students	192	55.2%	61.5%	49%
Police officers, CIA and FBI agents, lawyers, college students, therapists, judges, etc.	N/A	50%	N/A	N/A
Secret Service agents	N/A	70%	N/A	N/A
Police officer	37	**72%**	73%	70%
College students	20	50%	N/A	N/A

Figure 1.1. Performance on deceit detection by human observers [63]

1.3 State of the Art

How are lies detected? At the moment there are a lot of different instruments and technologies to detect lies, ranging from the good old polygraph and cameras to MRI machines. In computer vision, lie detection is done using an array of different techniques, employing not only RGB cameras but also physical sensors and thermal cameras, with machine learning techniques, and often combining many of them to achieve better results. We now proceed to describe the state of the art, based on the latest researches done in the field:

Speech

Speech is one of the many methods that can be used to recognize if a person is lying, in fact the speech signal contains linguistic, expressive, organic and biological data. [48]
One of the most used indicator of deception in various studies has been the response latency [66], since inventing a lie requires additional cognitive load as opposed to remembering the truth. The authors also notice that habitual lying makes it easier, and conversely often being spontaneous and telling the truth makes lying harder.
Another indicator of lying is the speech rate, especially when it's different from the average rate of a specific person [13]. Other verbal cues like grammar usage and word frequency have been used and have achieved high accuracy in psychological researches [53].
Speech analysis can reveal changes that affect behavior, such as stress, emotion, deception etc. by analyzing the pitch and the stress level. When a stressful situation arises, the hormonal levels of the body change, and this causes an increase in blood pressure and heart rate. This in turn affects the muscle in the respiratory system, and so speech is affected [48].
In sound processing, the mel-frequency cepstrum (MFC) is a representation of the short-term power spectrum of a sound. Mel-frequency cepstral coefficients (MFCCs) are coefficients that collectively make up the MFC [70].

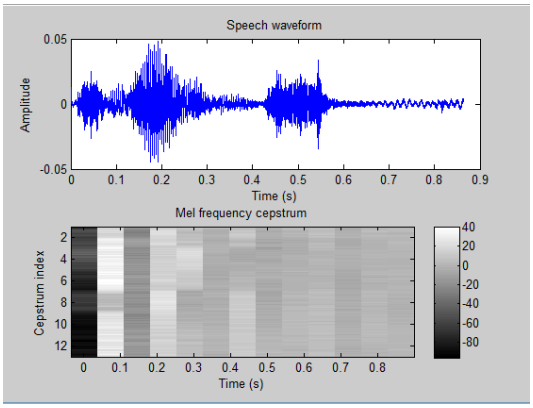

Figure 1.2. Topside is the spoken word, bottom is the MFCC derived from the word

In [46] the authors created a new database by making 40 candidates try to

deceive them while telling truthful or deceptive statements for about one to two minutes each. From this experiment they extracted MFCC and pitch, so that they could process them through Matlab's Voice Box. After acquiring the data, an SVM classifier was trained to classify new data, obtaining an accuracy of Lie and Truth detection from speech audio respectively of 88.23% and 84.52%.
In [50] [45] Perez et al. utilize real life trial data to identify deception, achieving 60-75% accuracy employing a model that extracts features from both linguistic and gesture modalities.

Eyes

Using the eyes to detect lies is one of the most studied approaches, as the eyes hold a significant amount of information regarding our thinking and emotional state [27]. Moreover is possible to generate a non invasive approach while analyzing the eyes, meaning that subjects do not need to willingly participate or even know if they are being examined or not (the moral matter should be considered in another setting), and there could be no need to have big and expensive machinery, like for example a polygraph or fMRI machine. Cognitive load, which is set to increase while lying, is one of the most significant factor for deception detection using the eyes. Important are also the blink rate, gaze aversion and pupil dilation.
In [29] the authors utilize high speed cameras to record and analyze blink count and blink duration of 50 subjects while asking 10 control questions, to see if there is a variation in them while the subject is being questioned. The authors analyzed the resulting images frame by frame and based on the facial landmarks around the eye they recognize AU45, the action unit for blinking. The results shows that both blink duration and count are increased while lying.
In another study, Leal and Vrij [38] asked 26 people, 13 liars and 13 truth tellers, to lie or tell the truth in a target period, while having a baseline from two preceding periods. The eye blinks during the target and baseline periods and directly after the target period (target offset period) were recorded. Compared to the baseline periods, lying subjects show a decrease in eye blinks during the target period and an increase in eye blinks during the target offset period. This pattern resulted very different for truth tellers showing that there is a significant difference in eye blink behavior between truthful and deceptive behavior.
Singh et al. in [59] show that while lying there is an increase in cognitive load and a significant decrease in eye blinks, directly followed by an increase in blink rate as soon as the cognitive demand ceases, after telling the lie. A threshold is set by the authors for this study, either at 26 blinks/minute or it is calculated personally using the average blink rate from a blink detection algorithm. Blink detection is done with MATLAB using the HAAR Cascade algorithm.
Lim et al. study eye gaze [41] to investigate the relation with lie detection. The result supports the theory that cognitive load decreases the number of eye movements.
Bhaskaran et al. measure deception by the deviation from normal behavior [11] at critical points during an investigative interrogation. For starters a dynamic Bayesian model of the eye movement is trained during a normal conversation with each of the 40 subjects of the experiment, then the remainder of the conversation is broken into pieces and each piece is tested against the normal behavior. The deviation from

normality are observed during critical points in the interrogation and used to deduce the presence of deceit, obtaining an accuracy of 82.5%.
In [55] Proudfoot et al. using latent growth curve modeling, the authors research how the pupil diameter changes over the course of an interaction with a deception detection system. The assumption is that anxiety changes the pupil diameter. The subjects are presented with crime-relevant target items (possibly incriminating) and non relevant items. The results indicate that the trends in the changes are indicative of deception during the interaction, regardless if incriminating items are shown.

Neuroscience

Electroencephalogram (EEG) and functional Magnetic Resonance Imaging (fMRI) have been employed for lie detection with good results for a long time, but at the cost of invasiveness, since both methods require big machinery, a suitable environment, and a subject willing to participate.
EEG is a monitoring method that records brain activities based on its potential.
In [58] Simbolon et al, use Event Related Potentials (ERP) to measure brain response directly from thought or perception. Among the many types of signals that constitute the ERP signal, P300 is the most critical for lie detection, as it appears as a response to meaningful rare stimuli (called odd ball stimuli). Eleven males of age between 20 and 27 took part in the study. The gathered data were then divided into training and test sets to produce different models. The highest accuracy of 70.83% was reached by a SVM classifier in detecting lying subjects.
In [37] twenty people of ages between 22 and 24 years old were subject to a card test using an EEG machine. EEG signals were collected using electrodes attached to the subject's head. The authors used the EEG signals to identify useful frequency bands and to measure lying state based on spectral analysis, with the use of fuzzy reasoning, obtaining 89.5% detection accuracy.
Arasteh et al. [8] utilize an alternative approach to the polygraph, the P300 Guilty Knowledge Test (GTK). GTK is based on the amplitude of P300 ERP wave as an index for the subject's recognition of concealed information. The Guilty Knowledge Test works on the assumption that among many similar unfamiliar topics, the recognition of familiar ones will be followed by a different response. 62 subjects were part of this experiment and participated in a mock crime followed by the P300 GTK. The authors used empirical mode decomposition (EMD) to extract features from the EEG signal and modeled them trough matlab. A genetic algorithm was then utilized for the feature selection and to handle the dimensionality increase of this approach. The classification accuracy of guilty and innocent subjects was 92.73%.
Another neuro-scientific approach to lie detection is functional Magnetic Resonance Imaging (fMRI). fMRI measures brain activity by detecting changes associated with blood flow. This technique relies on the fact that cerebral blood flow and neuronal activation are coupled. When an area of the brain is in use, blood flow to that region also increases [68].
In most of the experiments with fMRI and lie detection, candidates are instructed to lie and tell the truth in specific situations, and the brain activity from these instances is compared to a baseline condition. The regions showing greater activation for lies

than truth are supposed to be the most significant for deception detection. In a recent study it is shown that there is considerable agreement on the significant areas of th brain that regard lying [26].

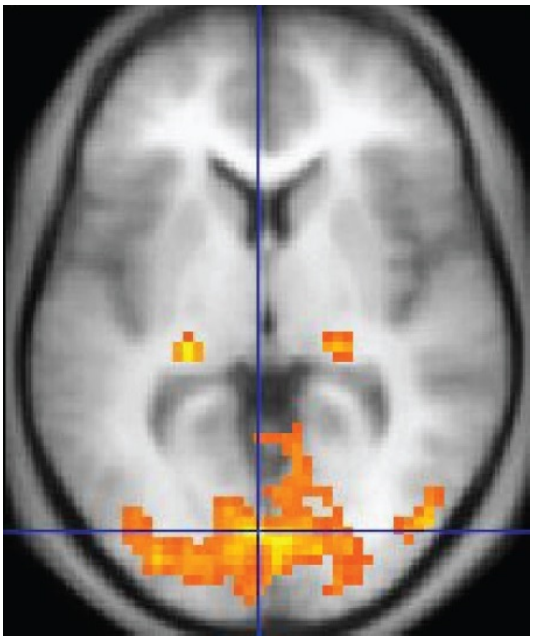

Figure 1.3. fMRI image with yellow areas showing increased activity compared with a control condition [68]

As much as fMRI is useful for lie detection, it presents some shortcomings [20][1]: many fMRI studies are small, not replicated and done with just a few subjects; there are contradicting results between some studies; most of the studies are not done in a contest of high stake deception, but in a controlled environment where subjects are asked to lie about some topic or event, without necessarily a real interest in being deceitful. Another important point is that the fMRI approach requires collaboration and expensive equipment to be carried out, so this is a very limiting factor.

Thermal

In thermal imaging, thermal features are extracted from the face using a high definition thermal camera. The objective is to analyze what kind of differences occur when a subject responds truthfully or deceptively to particular stimuli.

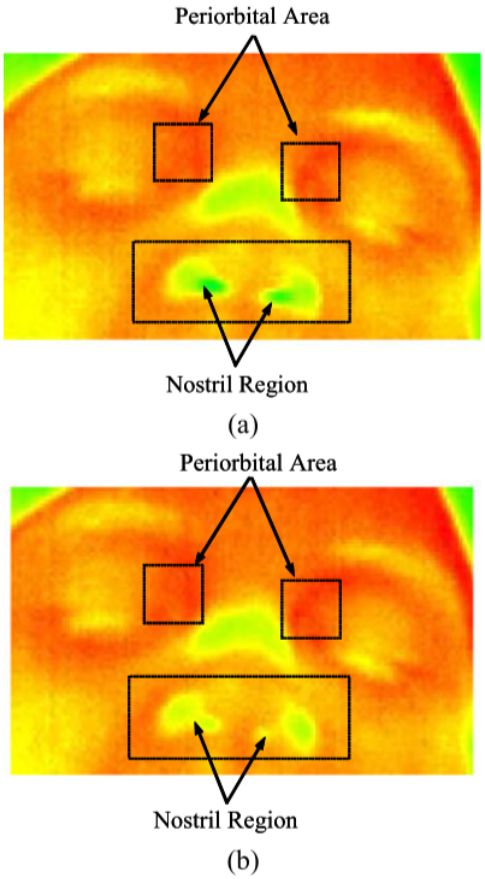

Figure 1.4. Examples of thermal images during (a) questioning and (b) answering [64]

According to a recent study [4] examining 30 subjects, the most relevant zones for deception detection in the facial area are the forehead and periorbital regions. In this study the subjects were registered with a thermal camera for one minute to extract the baseline features, and after that the interviewers asked a series of questions. A thermal map was created from the registrations using the Hue Saturation Value to differentiate between lies and truths.
In [56] the authors set up an experiment to collect 492 responses from 25 participants, using a deception scenario requiring the subjects to learn a story provided by the authors and practice their lies before hand, so that cognitive load would be increased in the interview. At the beginning of the interview four baseline questions were asked to register the initial thermal state of the subjects, and then a series of questions were asked, with answers both present and missing from the provided story. After extracting the periorbital region's thermal variation, a k-nearest neighbor classifier was used, with an 87% accuracy in predicting lies or truths.
In [64] data are gathered non-intrusively from the nostril and periorbital regions using two dimensional far infrared cameras. The study lasted for two years and covered 18 tests on subjects involved in real crime cases. The temperature is extracted and converted in change in blood flow velocity, and a signature of the respiration pattern is determined in terms of the ratio of the measured maximum and minimum

temperatures in the nostril area. The classification rate for this study is 88.5%.

Multimodal

After seeing these different techniques to detect lies, it's only natural for researchers to try and fuse or aggregate some of them to try and get better results [5].
In [3] Abouelenien et al. collect data from a dataset of 30 subjects to examine thermal and visual clues of deception. Their aim is to identify the regions that offer higher capability of detecting deceit. The method employed uses the CERT (Computer Expression Recognition Toolbox) to detect facial expression and encodes them with AU (Action Units). To extract thermal features they create a thermal map using gray-scale and Hue Saturation Value. They also calculated normalized blinking rates and the mean head orientation angle along the entire length of the response. In addiction over 60 physiological features are extracted and stored with the use of sensors and other RBG cameras. The experimental results show that the non-contact feature fusion model outperforms traditional physiological measurements, and that the forehead region is one of the most promising areas to gather information for deception detection using thermal imaging.
In a following paper [2] Abouelenien et al. explore a multimodal deception detection approach comprised of physiological, linguistic, and thermal features on a new dataset of 149 recordings. They set out to determine the most discriminative region of the face based on thermal imaging, and perform feature analysis using a decision tree model. The result show that the forehead could be a better indicator of deceit than the periorbital area. The physiological features did not contribute very much, while the linguistic feature played a critical role, where self-referencing and exaggeration words where big indicators of deceit. The overall accuracy of the system is 70%
Another example of multimodal detection is found in [73] where Wu et al. develop a framework to automatically detect deception in videos of trials. They utilize three modalities: vision, audio and text. For vision, they employ various classifiers trained on low level video features to predict on human micro-expressions, to successively predict deception. Interestingly, IDT (Improved Dense Trajectory) features, often used to recognize actions in videos, are good predictors of deceptive behavior. The authors decided to fuse the score of the classifiers on IDT and micro-expressions to boost the performance. Regarding text, the transcript of the considered videos are analyzed, but the performance increase is very marginal. For speech, they integrate the vision side with MFCC features analysis from the audio, boosting the performances significantly, reaching an AUC of 0.877.
Noje et al. [47] set up a study with ten subjects to observe the potential of head movements in lie detection. They built an application to detect head movement and position by performing a frame to frame analysis on a video stream. A correlation was made between head movement/position and the identification of lies. The results of the study are not concluding as this data can't be utilized without being incorporated with other modalities such as voice, gaze, words, expressions et cetera.

Facial Expression and Micro-Expressions

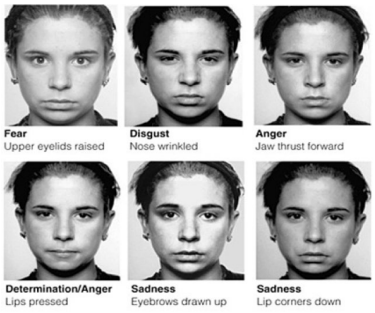

Figure 1.5. Facial Expressions

Facial Expressions are one of the main methods that we use to express our emotions. But what happens when we want to hide our emotions instead?
Facial micro-expressions are very fast (1/2 to 1/25 of a second) and involuntary expressions that come up on the human face when they are trying to suppress or hide an emotion, and are very difficult to control just using one's willpower [24]. Studying and classifying micro expression is very valuable and has many applications, especially in psychology and forensic sciences, but it is a hard feat as the duration is very short and the intensity is low. Micro-expressions have been studied since 1966 to recognize and distinguish real or fake emotions, initially by Haggard and Isaacs [30], and three years later by Ekman and Frisen [23].
Substantial work on Micro-Expressions has been done by Pfister, Li et al. In [51] they collaborated with psychologists to design an induced emotion suppression experiment. The data was collected with a high speed camera to be able to register the micro expressions of the subjects. A Temporal interpolation model was used to counter the shortness of the video length, while multiple kernel learning, random forest and SVM were used to classify micro expressions reaching an accuracy of 71.4%.
The lack of a database was one of the biggest hindrance to research, to solve this problem in [39] they unveil a new dataset, the Spontaneous Micro-expression Database (SMIC), which includes 164 micro-expression video clips taken from a group of 20 participants. They used two high speed cameras to record the face of the subjects while they were watching a selection of videos that induced strong emotional response, and they had to try and suppress those emotions. After the video the subjects had to answer about the emotions they felt while watching it. The data were then segmented and labeled by two annotators.
A study of spontaneous micro expression spotting and recognition methods is done in [40]. A new training-free method, based on feature difference contrast for recognizing micro-expressions is presented. The features are extracted from the video using Local Binary Pattern (LBP) and Histogram of Optical Flow. To recognize Micro expressions the authors performed face alignment and temporal interpolation, then they trained an SVM classifier. This micro-expression framework was tested on the SMIC and CASMEII database with very good results. After combining micro-expressions spotting and recognition they released a new micro-expression analysis

system (MESR) that is able to recognize micro-expressions from spontaneous video data.
Owayjan et al. [49] designed a lie detection system using micro-expressions. At first an embedded video system is used to record the subject interview. The video stream is converted into frames, and each frame is processed in four stages: converting the images, filtering out useless features, applying geometric templates and finally extracting the measurements to detect the micro-expressions. Results show that up to eight facial expressions can be recognized, and that lies can be discerned with high precision.
In [36] Kawulok et al. explore how to exploit fast smile intensity detectors to extract temporal features using a SVM classifier. Using exclusively a face detector, without localizing or tracking facial landmarks, they analyze the smile intensity time series. They then employ an SVM classifier to improve training from weakly labeled datasets. Then, to train the smile detectors, they use uniform local binary pattern features. This allows to detect, in real time, between spontaneous or posed expressions.
Su et al. [63] aim to test the validity of facial clues to deception detection in high-stakes situations using computer vision approaches. By using invariant 2D features from nine separate regions of the face they perform facial analysis on eye blink, eyebrow motion, wrinkle occurrence and mouth motion, integrated with a facial behavior pattern vector. Training a Random Forest to classify the patterns into deceptive or truthful, they achieved a 76.92% accuracy.

1.3.1 Action Units

Action Units (AUs) are defined as a contraction or relaxation of one or more muscles. They have been used in the Facial Action Coding System (FACS), a system developed initially by Hjorztsjö [34] and then improved by Freisen and Ekman [22], that categorizes all the facial movements by their appearance on the face. Using FACS it's possible to code nearly any facial expression by deconstructing them into Action Units. For example the Duchenne smile (felt smile) is a combination of AU6 (orbicularis oculi, pars lateralis) and AU12 (zygomatic mayor). There are 44 AUs categorized with five level of intensity (A to E). Most of the AUs have an onset, peak and offset phase. By using AUs is possible to recognize emotions based on the combination of AUs displayed (Fig. 1.6).

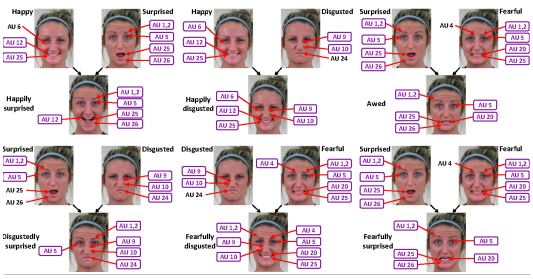

Figure 1.6. AUs of six compound facial expressions of emotion. The AUs of the basic emotions are combined to produce the compound category [21].

Substantial work on AU classification and intensity estimation has been done in [9] by Baltrusaitis et al. (Fig. 1.7) while developing the OpenFace [10] system. Baltrusaitis et al. developed a real-time Facial Action Unit occurrence and intensity detection system based on appearance and geometric features, using Histogram of Oriented Gradients, Shape Parameters and Landmark Locations. They achieved good results by using a median based feature normalization technique so that they could account for person specific neutral expressions.

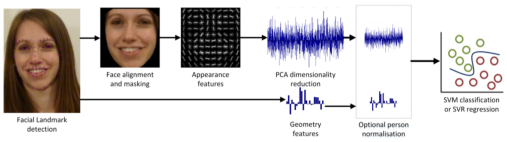

Figure 1.7. OpenFace AU detection and intensity estimation pipeline [9].

In [32] Hao et al. explore the dependencies between AUs, and propose a new AU recognition method consisting of a three layer Bayesian network where the top two layers are latent regression Bayesian networks (LRBN). LRBN are graphical models consisting of a visible layer representing the ground truth for AUs and a latent one. LRBN is able to capture the dependencies between AUs. They test this system on the CK+, SEMAINE and BP4D database, demonstrating that their approach can accurately capture AU relationships.
In [18] the authors try to model three fundamental aspects of automated AU detection: spatial representation, temporal modeling, and AU correlation. They proposed an approach using a hybrid network architecture. Spatial representation is extrapolated by using a Convolutional Neural Network (CNN). For temporal dependencies Long Short-Term Memory Neural Networks (LSTM) are stacked on top of the CNN. The output of LSTMs and CNNs are then fused together to predict 12 AUs on a frame to frame basis. This network was then tested on the GFT and BP4D dataset, gaining improvements over standard CNN.
De la Torre et al. [16] [17] tackle this problem by personalizing a generic classifier without requiring additional labels for the test subject (unsupervised). They use a transductive learning method, referred as Selective Transfer Machine (STM), that personalizes a generic classifier by attenuating people specific biases. This is done by learning a classifier and re-weighting the training samples most relevant to the subject. The performance of STM were compared to generic classifiers and cross-domain learning methods and evaluated on CK+, GEMEP-FERA, RUFACS and GFT dataset, and STM is shown to outperform the generic classifiers on all datasets.
Many approaches to automatic facial AU detection fail to account for individual difference in morphology and behavior for the target person. This is a hard problem because it's difficult for classifiers to generalize to very different subjects, and training a person-specific classifier is neither feasible (very low training data) nor has particular theoretical interest.

1.4 My Contributions

1) This thesis has set a precedent for future research on high-stakes deception detection using facial action units. As far as we know, there is no computer vision research that has testified the validity of facial action unit as indicators of high-stakes deception. As will be seen in Chapter X, our results are promising, implying the research potential of this topic in the future. 2) The proposed method is at the forefront of analyzing facial expressions in unconstrained environments. Since the videos in our database were mostly collected from YouTube, certain uncontrollable factors add to the difficulty of their analysis.
These totally unconstrained and spontaneous videos are subject to temporal variations in illumination, head pose and facial occlusion. However, in the current literature, little research has been conducted to address these issues with regard to facial expression analysis. Instead of seeking a solution to solve them, almost all of the studies to date have excluded certain data that were not ideal for the purpose of analysis. In comparison, the proposed method seeks to address the deception detection problem, but in the presence of exactly these factors.

FUTUREU USE - app per telefono video -> result

Chapter 2

Architecture

In this chapter we will give a brief introduction of Machine Learning (2.1), explaining Classification (2.1.3), Regression (2.1.4), Supervised (2.1.1) and Unsupervised (2.1.2) Learning, proceeding then to explain SVM (2.3) which is the most important technique used for this thesis.

2.1 Machine Learning

Machine Learning is a subfield of Artificial Intelligence that uses statistical techniques to provide computers with the ability to progressively improve performance on different tasks using data, while not being explicitly programmed [19].
The application for machine learning are huge and diverse, and range from character recognition to email filtering, with lots of application in computer vision, such as image recognition and classification.
Machine learning also focuses on making prediction on data, by utilizing techniques taken from mathematical optimization. This has many applications, ranging from health care by predicting risk factors for diseases or gaining insights for prevention, to sports or politics where actual results can be predicted.
The field of machine learning is subdivided in two broad categories (Fig. 2.1), supervised learning (2.1.1) and unsupervised learning (Fig. 2.1.2) based on whether the data is labeled or not.

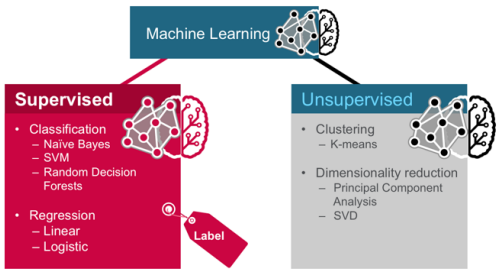

Figure 2.1. Machine Learning Subfields [44]

Another classification is based on the kind of output that the users wish to

obtain, mainly Classification (2.1.3), Regression (2.1.4) or Clustering.

2.1.1 Supervised Learning

When using supervised learning we want to learn a function that maps an input to an output, based on example input-output pairs [62]. This means that we need labeled training data, consisting of a vector of input objects and an output value. The objective is to correctly classify the new data based on the analyzed training pairs.
The general steps to solve a supervised learning problem are (Fig. 2.2):

1. Understand what kind of training example to use.
2. Gathering a training set.
3. Model the training set to be fed as input to the algorithm by choosing which features to use and how to represent the data.
4. Choose what kind of algorithm can best train the model.
5. Run the algorithm and evaluate the resulting accuracy on the test set

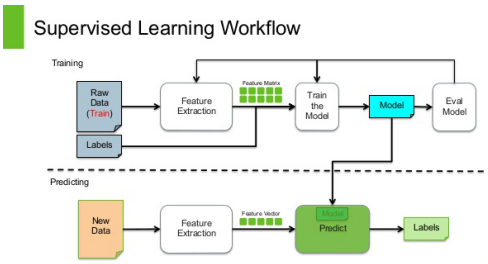

Figure 2.2. Supervised Learning Workflow [43]

There are important considerations to make when using a supervised learning approach:

- **Dimensionality of Input**: When the input feature vectors are very big there could be problems in learning the function, even if not all features contribute significantly to the function. This happens because the data depends on too many variables and this could cause high variance.
  To avoid this, it is important to reduce the number of features through manual removal or using feature selection algorithms. This usually improves the accuracy of the classifier.

- **Overfitting and Underfitting** [14]: Overfitting (or overtraining) happens when the algorithm adapts too much on the training data and is no longer able to make accurate predictions on the test data (Fig. 2.3). This usually happens when there is an excessive number of parameters than can be justified by the data [25].
  Underfitting is the opposite: a model is not able to correctly capture the structure of the data, for example when fitting non linear data with a linear model.
  A common way to avoid overfitting is to resample the data using different

techniques, commonly k-fold cross validation or leave-one-out. Other methods include feature removal, early stopping, regularization.

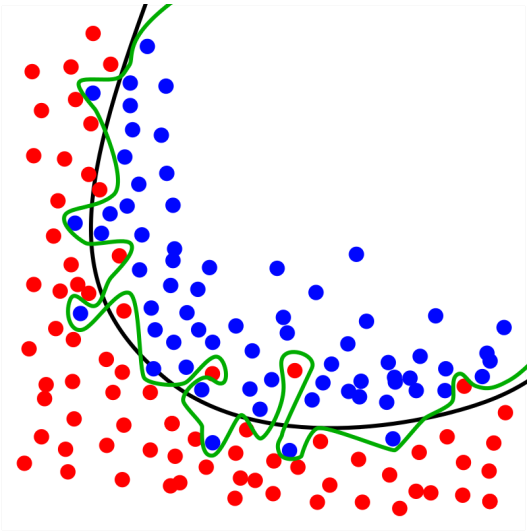

Figure 2.3. Example of Overfitting (green line) [71]

- **Bias-Variance Tradeoff** [57]: suppose we have different (but equally good) training set. An algorithm is biased for input x if, when trained on each of these data sets, it is consistently incorrect at predicting the correct output for x.
  A learning algorithm has high variance for a particular input x if it predicts different output values when trained on different training sets.
  The prediction error of a classifier is related to the sum of bias and variance, so generally there is a tradeoff between them. Low bias means that it fits the new data well, but if the bias is too low it will fit each training set differently and so result in high variance.

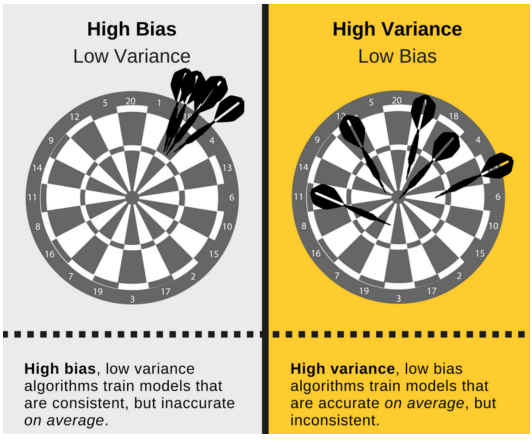

Figure 2.4. Example of Bias-Variance Tradeoff [12]

There are many algorithms used to perform supervised learning tasks, the most

commonly used are:

- Linear Regression (2.1.4)
- Logistic Regression
- Naive Bayes
- Linear Discriminant Analysis
- Decision Trees
- k-Nearest Neighbor
- Neural Networks
- Support Vector Machines (2.3)

2.1.2 Unsupervised Learning

Unsupervised learning is the subfield of Machine Learning tasked with inferring a function from the analysis of unlabeled data (not classified). Being unclassified there is also a difficulty in to evaluate the accuracy of the model.
Usually items are grouped by some measure of similarity, like for example in k-means clustering (Fig. 2.5).

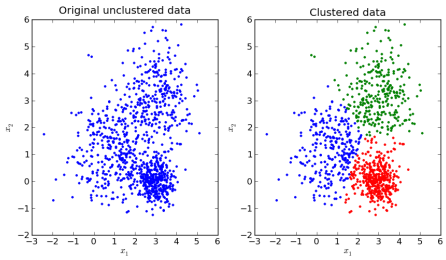

Figure 2.5. Example of Unclustered and Clustered data [35]

These are the most widely used unsupervised learning algorithms:

- Clustering
  - k-means
  - mixture models
  - hierarchical clustering
- Anomaly detection
- Neural Networks
  - Autoencoders
  - Deep Belief Nets
  - Hebbian Learning
  - Generative Adversarial Networks
  - Self-organizing map
- Expectation–maximization algorithm (EM)
- Method of moments
- Blind signal separation techniques
  - Principal component analysis,

– Independent component analysis,
– Non-negative matrix factorization
– Singular value decomposition.

2.1.3 Classification

In machine learning, classification is the problem of identifying in which of a set of categories a new observation belongs, based on a training set of data containing observations whose category is known in advance. A common example is classifying an email as spam or not.
Classification is considered an instance of supervised learning, based on instances where a training set is available. As for unsupervised learning, classification would be clustering since it groups data into categories based on similarity, but without knowing the label of the data.
The observations, called features or explanatory variables, take different types based on the value. They can be categorical, numerical or ordinal, or be compared by similarity between previous observations using some kind of distance function.
The observations representing the categories to be predicted are called explanatory variables (or regressors, or independent variables).
The classifier is the algorithm that implements the classification, or a function that maps input data to a category.

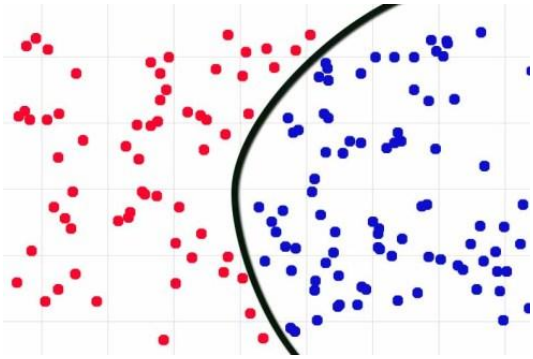

Figure 2.6. Example of data division by a Classification algorithm

Many classification algorithm, such as SVM (2.3), logistic regression, LDA (Linear Discriminant Analysis) or perceptron, can be described using a linear function assigning a score to each category c, by doing the dot product of the feature vector of an instance with a vector of weights, thus combining them. The predicted category will be the one with the highest score. This function is called linear predictor function and has this general formula:

$$score(X_i, c) = \beta_c \cdot X_i \tag{2.1}$$

where X_i is the feature vector of instance i and β_c is the vector of weights of category c. This kind of algorithms are known as linear classifiers.

2.1.4 Regression Analysis

Regression analysis is used in statistics and machine learning to estimate the relationships between variables. It focus on the relationship between one dependent variable and more independent variables (also called predictors) and it is subsequently used to make predictions on how changing some of the independent variables will affect the dependent one. [72]
Regression is generally used to estimate the average value of the dependent variable when the independent variables are fixed. In doing that a regression function is calculated. Another use of regression analysis is to understand the relationship among the independent and dependent variables. In regression analysis, it is also of interest to characterize the variation of the dependent variable around the prediction of the regression function using a probability distribution.

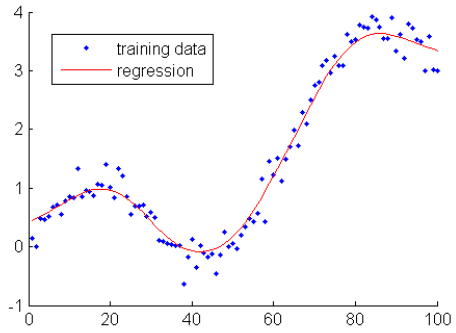

Figure 2.7. Example of Regression Analysis

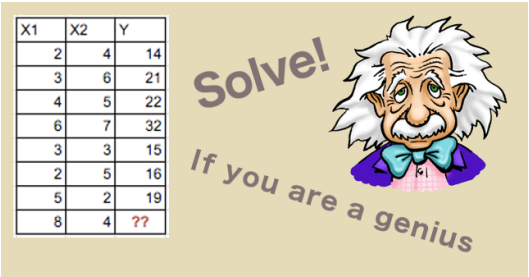

Figure 2.8. Regression Analysis can solve this! [7]

Regression Model
A general regression model uses the following variables and parameters: The unknown parameters, β, that represents either a scalar or a vector.
The independent variables, X.
The dependent variable, Y.

A regression model relates Y to a function of X and β.

$$Y \approx f(X, \beta) \tag{2.2}$$

This is usually formalized as

$$E(Y|X) = f(X, \beta) \tag{2.3}$$

If β is of length k and the number of observed data points is enough ($N > k$), then it's possible to estimate a unique value for β that best fits the data. If this is the case, regression analysys provides the means to find a solution for unknown parameters of β to, for example, apply the method of least squares.

To apply regression the data must abide by some assumption, generally:

- The sample is representative of the population for the inference prediction.
- The error is a random variable with a mean of zero.
- The independent variables are measured with no error.
- The independent variables (predictors) are linearly independent.
- The errors are uncorrelated.
- The variance of the error is constant across observations.

2.1.5 Linear Regression

Linear regression is the most basic case of Regression Analysis, and it is a useful tool for predicting a quantitative response. Also most of the newer approaches to regression are often a generalization or extension of linear regression.
Linear regression is a linear approach to modeling the relationship between a dependent variable and one or more independent variables. The simplest case of linear regression (LR) is when there is only one independent variable, and is called simple linear regression [69] and has this form:

$$y_i \approx \beta_0 + \beta_1 x_i + \varepsilon_i \tag{2.4}$$

where ϵ_i is the error for the i-th observation. The error is a catch-all for what we miss with this simple model, since it's very probable that the true relationship is not linear, and that there may be other variables that influence y.
The point of LR is to estimate the β coefficients to make predictions. To do so, we utilize of the training dataset to produce estimates $\widehat{\beta_0}$ and $\widehat{\beta_1}$ for the model coefficients. We can then make predictions by computing:

$$\widehat{y_i} \approx \widehat{\beta_0} + \widehat{\beta_1 x_i} \tag{2.5}$$

$e_i = y_i - \widehat{y_i}$ is the difference between the true value and the prediction for an observation, and it is called residual (Fig. 2.9).

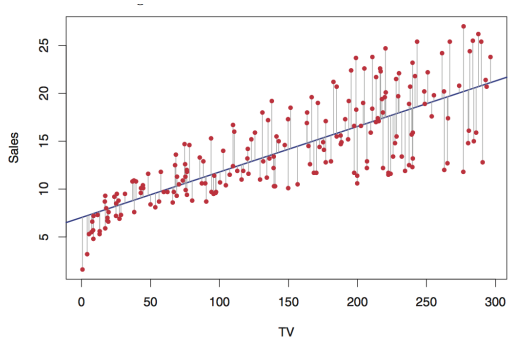

Figure 2.9. The fit is found by minimizing the sum of squared errors. Each gray line segment represents an error, and the prediction makes a compromise by averaging their square [28]

The most common method for estimation is called least squares. This method obtains parameter estimates that minimize the sum of squared residuals (SSR):

$$SSR = \sum_{i=1}^{n} e_i^2 \tag{2.6}$$

The minimizers are

$$\widehat{\beta_1} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} \tag{2.7}$$

$$\widehat{\beta_0} = \bar{y} - \widehat{\beta_1}\bar{x} \tag{2.8}$$

where $\bar{x}$ and $\bar{y}$ are the mean of x and y.

If we assume that the error term has constant variance, the estimate of the variance of the error is given by:

$$\widehat{\sigma}_{\varepsilon}^2 = \frac{SSR}{n-2} \tag{2.9}$$

And is called mean square error (MSE) of the regression. The denominator is the sample size reduced by the number of model parameters estimated from the same data, $(n-p)$ for p regressors or $(n-p-1)$ if an intercept is used [61]. In the case of simple linear regression $p = 1$, so the denominator is $n-2$.

We can then estimate the standard errors for the parameters, that tell us the average amount that the estimate differs from the actual value.

$$\widehat{\sigma}_{\beta_1} = \widehat{\sigma}_{\varepsilon}\sqrt{\frac{1}{\sum(x_i - \bar{x})^2}} \tag{2.10}$$

$$\widehat{\sigma}_{\beta_0} = \widehat{\sigma}_{\varepsilon}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2}} \tag{2.11}$$

Standard errors can be used to compute confidence intervals. A 95% confidence interval is defined as a range of values such that with 95% probability, the range will contain the true unknown value of the parameter.

The general multiple regression model, where there are p independent variables, follows this formula:

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i \tag{2.12}$$

where x_{ij} is the i-th observation on the j-th independent variable.

2.2 Random Forest

2.3 SVM

Support Vector Machines (SVM) are a supervised machine learning algorithm used for both classification and regression.
The main idea is to find the optimal hyperplane for linearly separable data, and then extend this idea to data that are not linearly separable by mapping this data in a new space using a kernel function.
The definition of an hyperplane for p-dimensions is:

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p = 0 \tag{2.13}$$

Support vectors are the data points that lie closest to the hyperplane (Fig. 2.10), they also are the data points most difficult to classify and have direct bearing on the optimum location of the hyperplane.

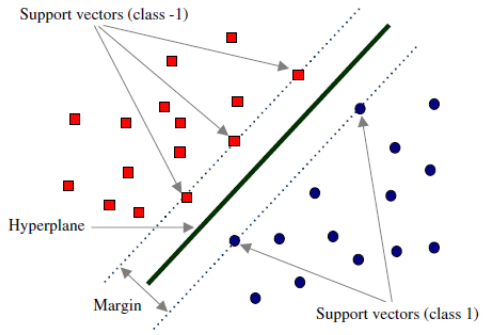

Figure 2.10. Example of Support Vectors

The distance between the hyperplane and the nearest data point from either set is known as the margin. The best hyperplane is the one that maximizes the margins for the data we are classifying.

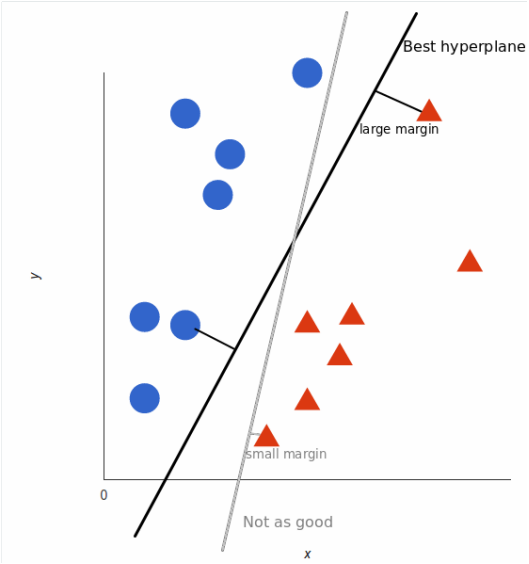

Figure 2.11. Different hyperplanes with different margins [60]

So the objective is to choose a hyperplane with the largest possible margin between it and the support vectors, since the larger is the margin, the lower the generalization error of the classifier.
In essence, support vectors are the elements of the training set that would change the position of the dividing hyperplane if removed. This makes the support vectors the critical elements of the training set.

Finding the maximal margin hyperplane based on a set of training observations $x_1, \dots x_n \in R^p$ and with class lables $y_1 \dots y_n$, translates to an optimization problem:
Maximize M

$$\beta_0, \beta_1, \beta_2, \dots, \beta_p, M \tag{2.14}$$

subject to

$$\sum_{j=1}^{p} \beta_j^2 = 1 \tag{2.15}$$

$$y_i(\beta_0 + \beta_1 X_i 1 + \beta_2 X_i 2 + \dots + \beta_p X_i p) \geq M \forall i = 1, \dots, n \tag{2.16}$$

2.3 and 2.4 ensure that each observation is on the correct side of the hyperplane and at least a distance M from the hyperplane. Hence, M represents the margin of our hyperplane, and the optimization problem chooses $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ to maximize M.
Unfortunately this hyperplane does not necessarily exists, but we can extend this concept to find a hyperplane that almost separates the classes using a soft margin. This is what an SVM does.

It is very probable that the data is not linearly separable, as we can see in figure 2.12.

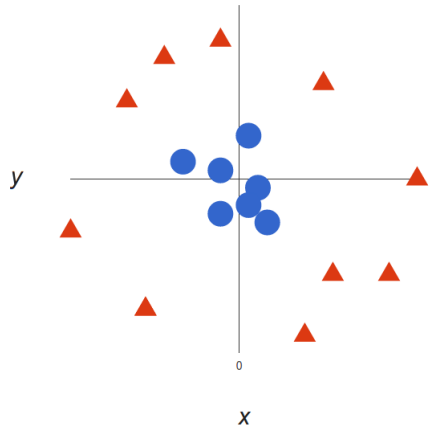

Figure 2.12. Linearly not separable data [60]

In this case, for example, we can add a new dimension and separate the data.

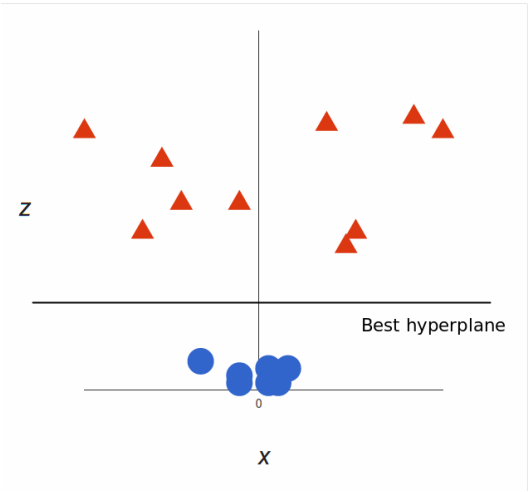

Figure 2.13. Three dimensional separable space [60]

And then map back to two dimension.

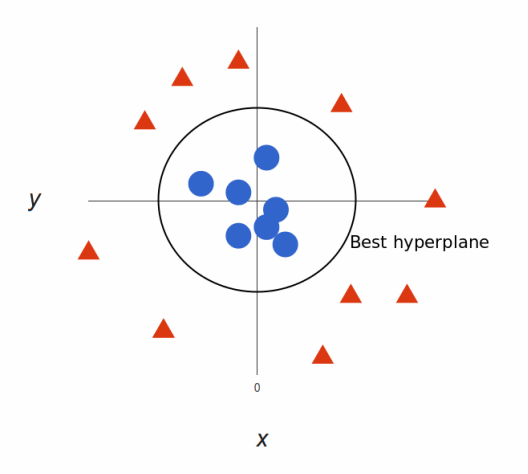

Figure 2.14. Separated data [60]

Calculating the transformation can be very computationally intensive, but SVM just needs the dot product between the vectors. This is called the kernel function. Kernels can be of different types, like linear or

Chapter 3

Experiments

In this section I will explain what I've done and the stack used for this thesis: I will start by reviewing the OpenFace tool (3.1), the database (3.2) utilized and all the experiments and techniques used, such as.....

3.1 OpenFace

The OpenFace [65] toolkit is a tool for machine learning and computer vision researchers created ba Baltrusaitis et al. to perform facial landmark detection, head pose estimation, action unit recognition and eye gaze estimation.

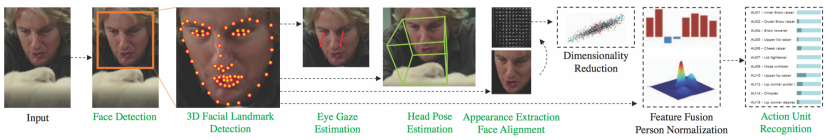

Figure 3.1. OpenFace Pipeline [65]

3.2 Real Life Trial DataBase

3.3 GLM

3.4 LDA

3.5 QDA

3.6 SVM

3.7 Correlations

?

Chapter 4

Results

Results obtained

Chapter 5

Conclusions

Conclusion

Bibliography

[1] Sean A. Spence. “Playing Devil’s advocate: The case against fMRI lie detection”. In: 13 (Feb. 2008), pp. 11–25.

[2] M. Abouelenien et al. “Detecting Deceptive Behavior via Integration of Discriminative Features From Multiple Modalities”. In: *IEEE Transactions on Information Forensics and Security* 12.5 (May 2017).

[3] Mohamed Abouelenien, Rada Mihalcea, and Mihai Burzo. “Analyzing Thermal and Visual Clues of Deception for a Non-Contact Deception Detection Approach”. In: *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. 2016. URL: http://doi.acm.org/10.1145/2910674.2910682.

[4] Mohamed Abouelenien, Rada Mihalcea, and Mihai Burzo. “Trimodal Analysis of Deceptive Behavior”. In: *Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection*. 2015. URL: http://doi.acm.org/10.1145/2823465.2823470.

[5] Mohamed Abouelenien et al. “Deception Detection Using a Multimodal Approach”. In: *Proceedings of the 16th International Conference on Multimodal Interaction*. 2014. URL: http://doi.acm.org/10.1145/2663204.2663229.

[6] Baker Alysha, ten Brinke Leanne, and Porter Stephen. “Will get fooled again: Emotionally intelligent people are easily duped by high-stakes deceivers”. In: *Legal and Criminological Psychology* 18.2 (), pp. 300–313. DOI: 10.1111/j.2044-8333.2012.02054.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2044-8333.2012.02054.x.

[7] Tarek Amr. *Predict the Future with Regression Analysis*. URL: https://medium.com/@gr33ndata/learn-regressions-analysis-23b789bf2c36.

[8] A. Arasteh, M. H. Moradi, and A. Janghorbani. “A Novel Method Based on Empirical Mode Decomposition for P300-Based Detection of Deception”. In: *IEEE Transactions on Information Forensics and Security* 11.11 (Nov. 2016).

[9] T. Baltrušaitis, M. Mahmoud, and P. Robinson. “Cross-dataset learning and person-specific normalisation for automatic Action Unit detection”. In: *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*. Vol. 06. May 2015, pp. 1–6.

[10] T. Baltrušaitis, P. Robinson, and L. P. Morency. “OpenFace: An open source facial behavior analysis toolkit”. In: *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. Mar. 2016, pp. 1–10.

[11] N. Bhaskaran et al. “Lie to Me: Deceit detection via online behavioral learning”. In: *Face and Gesture 2011*. Mar. 2011.

[12] *Bias-Variance Tradeoff*. URL: https://elitedatascience.com/bias-variance-tradeoff.

[13] Marilyn G. Boltz, Rebecca L. Dyer, and Anna R. Miller. “Jo Are You Lying to Me? Temporal Cues for Deception”. In: *Journal of Language and Social Psychology* 29.4 (2010), pp. 458–466. DOI: 10.1177/0261927X10385976. URL: https://doi.org/10.1177/0261927X10385976.

[14] D. R. Burnham K. P.; Anderson. *Model Selection and Multimodel Inference (2nd ed.), Springer-Verlag*. 2002.

[15] Jr. Charles F. Bond and Bella M. DePaulo. “Accuracy of Deception Judgments”. In: *Personality and Social Psychology Review* 10.3 (2006), pp. 214–234. DOI: 10.1207/s15327957pspr1003_2. URL: https://doi.org/10.1207/s15327957pspr1003_2.

[16] W. S. Chu, F. D. L. Torre, and J. F. Cohn. “Selective Transfer Machine for Personalized Facial Action Unit Detection”. In: *2013 IEEE Conference on Computer Vision and Pattern Recognition*. June 2013.

[17] W. S. Chu, F. D. l. Torre, and J. F. Cohn. “Selective Transfer Machine for Personalized Facial Expression Analysis”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.3 (Mar. 2017). DOI: 10.1109/TPAMI.2016.2547397.

[18] W. S. Chu, F. De la Torre, and J. F. Cohn. “Learning Spatial and Temporal Cues for Multi-Label Facial Action Unit Detection”. In: *2017 12th IEEE International Conference on Automatic Face Gesture Recognition (FG 2017)*. May 2017.

[19] Wikipedia contributors. *Machine learning — Wikipedia, The Free Encyclopedia*. [Online; accessed 16-July-2018]. 2018. URL: https://en.wikipedia.org/w/index.php?title=Machine_learning&oldid=849817385.

[20] Langleben Daniel D. “Detection of deception with fMRI: Are we there yet?” In: *Legal and Criminological Psychology* 13.1 (), pp. 1–9. DOI: 10.1348/135532507X251641. URL: https://onlinelibrary.wiley.com/doi/abs/10.1348/135532507X251641.

[21] Shichuan Du, Yong Tao, and Aleix M. Martinez. “Compound facial expressions of emotion”. In: *Proceedings of the National Academy of Sciences of the United States of America* (2014).

[22] P. Ekman and W. Friesen. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. 1978.

[23] P. Ekman and W. V. Friesen. “Nonverbal leakage and clues to deception”. In: *Psychiatry* 32.1 (1969), pp. 88–106.

[24] Paul Ekman. *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. 2007.

[25] Skrondal A. Everitt B.S. *Cambridge Dictionary of Statistics*. 2010.

[26] Martha J. Farah et al. “Functional MRI-based lie detection: scientific and societal challenges”. In: *Nature Reviews Neuroscience* 15 (Jan. 2014). URL: http://dx.doi.org/10.1038/nrn3665.

[27] Kyosuke Fukuda. “Eye blinks: new indices for the detection of deception”. In: *International Journal of Psychophysiology* 40.3 (2001), pp. 239–245.

[28] Trevor Hastie Gareth James Daniela Witten and Robert Tibshiran. *An Introduction to Statistical Learning*. URL: http://www-bcf.usc.edu/~gareth/ISL/index.html.

[29] S. George et al. “Eye blink count and eye blink duration analysis for deception detection”. In: *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. Sept. 2017.

[30] E. Haggard and K. Isaacs. “Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy”. In: *Methods of research in psychotherapy. New York: Appleton-Century-Crofts* (1966), pp. 154–165.

[31] Jeffrey Hancock. “Digital deception: When, where and how people lie online”. In: (Jan. 2012), pp. 287–301.

[32] L. Hao et al. “Facial Action Unit Recognition Augmented by Their Dependencies”. In: *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*. May 2018.

[33] Maria Hartwig et al. “Detecting deception in suspects: verbal cues as a function of interview strategy”. In: *Psychology, Crime & Law* 17.7 (2011), pp. 643–656. DOI: 10.1080/10683160903446982. URL: https://doi.org/10.1080/10683160903446982.

[34] Carl-Herman Hjortsjö. *Man’s face and mimic language*. 1969.

[35] *k-means data clustering*. URL: https://towardsdatascience.com/k-means-data-clustering-bce3335d2203.

[36] Michal Kawulok et al. “In Search of Truth: Analysis of Smile Intensity Dynamics to Detect Deception”. In: *Advances in Artificial Intelligence - IBERAMIA 2016*. 2016.

[37] Ying-Fang Lai, Mu-Yen Chen, and Hsiu-Sen Chiang. “Constructing the lie detection system with fuzzy reasoning approach”. In: *Granular Computing* (Nov. 2017). URL: https://doi.org/10.1007/s41066-017-0064-3.

[38] Sharon Leal and Aldert Vrij. “Blinking During and After Lying”. In: *Journal of Nonverbal Behavior* 32.4 (Dec. 2008), pp. 187–194. URL: https://doi.org/10.1007/s10919-008-0051-0.

[39] Xiaobai Li et al. “A Spontaneous Micro Facial Expression Database: Inducement, Collection and Baseline”. In: *Face and Gesture (FG)*. 2013.

[40] Xiaobai Li et al. “Reading Hidden Emotions: Spontaneous Micro-expression Spotting and Recognition”. In: *IEEE Trans. Affective Computing (TAFFC)*. 2015.

[41] Kai Keat Lim et al. “Lying Through the Eyes: Detecting Lies Through Eye Movements”. In: *Proceedings of the 6th Workshop on Eye Gaze in Intelligent Human Machine Interaction: Gaze in Multimodal Interaction*. 2013. URL: http://doi.acm.org/10.1145/2535948.2535954.

[42] Bella M. DePaulo et al. “Lying in Everyday Life”. In: 70 (June 1996), pp. 979–95.

[43] *Machine Learning 101 what is machine learning*. URL: https://hydrasky.com/network-security/machine-learning-101-what-is-machine-learning/.

[44] Carol McDonald. *Demystifying AI, Machine Learning and Deep Learning*. URL: https://mapr.com/blog/demystifying-ai-ml-dl/.

[45] Rada Mihalcea, Verónica Pérez-Rosas, and Mihai Burzo. “Automatic Detection of Deceit in Verbal Communication”. In: *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*. 2013. URL: http://doi.acm.org/10.1145/2522848.2522888.

[46] H. Nasri, W. Ouarda, and A. M. Alimi. “ReLiDSS: Novel lie detection system from speech signal”. In: *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*. Nov. 2016.

[47] D. I. Noje and R. Malutan. “Head movement analysis in lie detection”. In: *2015 Conference Grid, Cloud High Performance Computing in Science (ROLCG)*. Oct. 2015.

[48] Paola Noreña Cardona. “A COMPENDIUM OF PATTERN RECOGNITION TECHNIQUES IN FACE, SPEECH AND LIE DETECTION”. In: 24 (Nov. 2015).

[49] M. Owayjan et al. “The design and development of a Lie Detection System using facial micro-expressions”. In: *2012 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA)*. Dec. 2012.

[50] Verónica Pérez-Rosas et al. “Deception Detection Using Real-life Trial Data”. In: *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. 2015.

[51] Tomas Pfister et al. “Recognising Spontaneous Facial Micro-expressions”. In: *International Conference on Computer Vision (ICCV)*. 2011.

[52] Stephen M Porter, Leanne ten Brinke, and Brendan B Wallace. “Secrets and Lies: Involuntary Leakage in Deceptive Facial Expressions as a Function of Emotional Intensity”. In: 2012.

[53] Stephen Porter and Leanne ten Brinke. “The Truth About Lies: What Works in Detecting High-Stakes Deception?” In: 15 (Feb. 2010), pp. 57–75.

[54] Stephen Porter and Mary Campbell. “A. Vrij, Detecting Lies and Deceit: The Psychology of Lying and Implications forProfessional Practice”. In: 7 (Sept. 1999), pp. 227–232.

[55] J. G. Proudfoot et al. “Deception is in the eye of the communicator: Investigating pupil diameter variations in automated deception detection interviews”. In: *2015 IEEE International Conference on Intelligence and Security Informatics (ISI)*. May 2015.

[56] B. Rajoub and R. Zwiggelaar. “Thermal facial analysis for deception detection”. In: *IEEE Transactions on Information Forensics and Security*. 2014.

[57] E. Bienenstock S. Geman and R. Doursat. *Neural networks and the bias/variance dilemma*. 1992.

[58] A. I. Simbolon et al. “An experiment of lie detection based EEG-P300 classified by SVM algorithm”. In: *2015 International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology (ICACOMIT)*. Oct. 2015.

[59] B. Singh, P. Rajiv, and M. Chandra. “Lie detection using image processing”. In: *2015 International Conference on Advanced Computing and Communication Systems*. Jan. 2015.

[60] Bruno Stecanella. *An introduction to Support Vector Machines (SVM)*. URL: https://monkeylearn.com/blog/introduction-to-support-vector-machines-svm/.

[61] R.G.D Steel and J. H Torrie. *Principles and Procedures of Statistics with Special Reference to the Biological Science, McGraw Hill*. 1960.

[62] Peter Norvig Stuart J. Russell. *Artificial Intelligence: A Modern Approach, Third Edition - Prentice Hall*. 2010.

[63] Lin Su and Martin Levine. “Does “lie to me” lie to you? An evaluation of facial clues to high-stakes deception”. In: *Computer Vision and Image Understanding* 147 (2016). URL: http://www.sciencedirect.com/science/article/pii/S1077314216000345.

[64] S. Sumriddetchkajorn et al. “Simultaneous Analysis of Far Infrared Signals From Periorbital and Nostril Areas for Nonintrusive Lie Detection: Performance From Real Case Study”. In: *Journal of Lightwave Technology* 33.16 (Aug. 2015).

[65] T. usaitis et al. “OpenFace 2.0: Facial Behavior Analysis Toolkit”. In: *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*. May 2018.

[66] Bruno Verschuere et al. “The ease of lying”. In: 20 (Nov. 2010), pp. 908–11.

[67] Aldert Vrij et al. “Police officers ability to detect deception in high stakes situations and in repeated lie detection test”. In: 20 (Sept. 2006), pp. 741–755.

[68] Wikipedia contributors. *Functional magnetic resonance imaging — Wikipedia, The Free Encyclopedia*. [Online; accessed 20-June-2018]. 2018. URL: https://en.wikipedia.org/w/index.php?title=Functional_magnetic_resonance_imaging&oldid=842791829.

[69] Wikipedia contributors. *Linear regression — Wikipedia, The Free Encyclopedia*. [Online; accessed 5-August-2018]. 2018. URL: https://en.wikipedia.org/w/index.php?title=Linear_regression&oldid=851432903.

[70] Wikipedia contributors. *Mel-frequency cepstrum — Wikipedia, The Free Encyclopedia*. 2018. URL: https://en.wikipedia.org/w/index.php?title=Mel-frequency_cepstrum&oldid=835415759.

[71] Wikipedia contributors. *Overfitting — Wikipedia, The Free Encyclopedia*. [Online; accessed 17-July-2018]. 2018. URL: https://en.wikipedia.org/w/index.php?title=Overfitting&oldid=848185507.

[72] Wikipedia contributors. *Regression analysis — Wikipedia, The Free Encyclopedia*. [Online; accessed 3-August-2018]. 2018. URL: https://en.wikipedia.org/w/index.php?title=Regression_analysis&oldid=850022021.

[73] Zhe Wu et al. "Deception Detection in Videos". In: *CoRR* abs/1712.04415 (2017). URL: http://arxiv.org/abs/1712.04415.

List of Figures