CATEGORICAL ASSESSMENT OF DEPRESSION BASED ON HIGH LEVEL FEATURES

by

VAZAKOPOYLOY K. CALLIOPE-MARINA

BSc. Technological Educational Institute of Crete, 2013

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

DEPARTMENT OF INFORMATICS ENGINEERING

SCHOOL OF ENGINEERING

TECHNOLOGICAL EDUCATIONAL INSTITUTE OF CRETE

Approved by:
Major Professor
Professor Manolis Tsiknakis

Heraklion, Crete
2018

Copyright

Calliope-Marina Vazakopoulou

2018

Abstract

Depression is one of the leading causes of ill health and disability worldwide. In the US nearly 50 million people were living with depression in 2015. According to the latest estimates from the World Health Organization (WHO), more than 300 million people worldwide are living with depression, an increase of more than 18% between 2005 and 2015[1]. This fact seems to be confirmed by an EU Green Paper [1], which manifests that one to four civilians suffer from a mental illness at some point during their lifetime, sometimes even causing suicidal tendencies.

Objective measures of depressive symptomatology could be advantageous for clinicians, in the context of a clinical decision support system.

Automatic recognition of emotions through intelligent systems, established to assess facial expressions is a vital issue that needs to be addressed for understanding human behavior, interpersonal relationships, and most importantly for mental health assessment. According to the literature in the field of clinical research, facial expressions were used to assess possible deficiencies or weaknesses in emotional expression as well as in social diagnosis in psychiatric-psychological disorders [2] [3]. Major Depressive Disorder (MDD) is the most common mood disorder, varying in terms of severity, impairing an individual's functionality and ability to deal with their daily routine. The figures provided by the EU Green Paper [4], [1] are alarming, as it is argued that by 2020 MDD is expected to become the main cause of disability.

The aim of this dissertation is to develop a framework capable of detecting objective signs of MDD, to support the clinical care of patients. More specifically, the proposed system aims to identify signs related to MDD as portrayed in the facial expression of an individual. This has been designed within the terms of the AViD-Corpus (AVEC'14) [5], which provides audio-visual recordings annotated for the level of depression based on self-reports. The proposed framework was evaluated through experimental tests, designed to detect patterns related to depression in the facial expression.

In the proposed implementation the emotions of happiness and sadness, two of the six basic emotions according to Ekman and Friesen [6], [7], are examined more carefully, as they seem to be respectively negatively and positively correlated with MDD. For example, the emotion of

[1] http://www.paho.org/hq/index.php?option=com_content&view=article&id=13102%3Adepression-lets-talk-says-who-as-depression-tops-list-of-causes-of-ill-health&catid=740%3Apress-releases&Itemid=1926&lang=en

sadness appears to be more prominent in individuals suffering by MDD compared to that of happiness.

In summary, the primary purpose of this thesis is to design and develop an application in MATLAB for facial image analysis, with the ultimate goal of detecting visual signs of depression through video recordings. As it is mentioned above, Facial expression is significant in human interaction and communication since it contains critical information regarding to emotion analysis. High level information about the facial features was extracted with the use of OpenFace toolkit [8], while special focus has been given to the AViD-Corpus (AVEC'14) for testing the developed algorithms. Subsequently, since high-level information was extracted by the facial features, the proper machine learning algorithms were selected, in order to examine the sensitivity and specificity of this proposed framework. Several classification algorithms were tested, namely: Discriminant Analysis, Random Forest Tree, Naïve Bayes, Linear and k-nearest Neighbor Classifiers. The best performing method involved the Discriminant Analysis classifier, with a runtime of approximately 4 minutes, to be precise 4,29 minutes. Additionally, by selecting an approximately 1 second-window, with Leave-One-Subject-Out cross-validation an accuracy of 72.57% was achieved for our depression assessment framework.

Περίληψη

Η κατάθλιψη είναι μια από τις κύριες αιτίες κακής υγείας και δυσλειτουργίας παγκοσμίως. Στις ΗΠΑ, περίπου 50 εκατομμύρια άνθρωποι είχαν διαγνωστεί με κατάθλιψη έως το 2015. Σύμφωνα με τις τελευταίες εκτιμήσεις του Οργανισμού Παγκόσμιας Υγείας (WHO), περισσότεροι από 300 εκατομμύρια άνθρωποι παγκοσμίως ζουν με κατάθλιψη, δηλαδή αύξηση άνω του 18% μεταξύ 2005 και 2015. Το γεγονός αυτό φαίνεται να επιβεβαιώνεται από την Πράσινη Βίβλο της ΕΕ [1] , η οποία καταδεικνύει ότι ένας στους τέσσερις πολίτες υποφέρουν από κάποια ψυχική ασθένεια κάποια στιγμή κατά τη διάρκεια της ζωής τους, προκαλώντας μερικές φορές ακόμη και αυτοκτονικές τάσεις.

Οι αντικειμενικές εκτιμήσεις της καταθλιπτικής συμπτωματολογίας θα μπορούσαν να είναι επωφελείς για τους κλινικούς ιατρούς, στο πλαίσιο ενός συστήματος υποστήριξης κλινικών αποφάσεων.

Η αυτόματη αναγνώριση των συναισθημάτων μέσω ευφυών συστημάτων, η οποία αξιολογεί τις εκφράσεις του προσώπου είναι ένα ζωτικής σημασίας ζήτημα που πρέπει να αντιμετωπιστεί για την κατανόηση της ανθρώπινης συμπεριφοράς, των διαπροσωπικών σχέσεων, και το πιο σημαντικό για την αξιολόγηση της ψυχικής υγείας. Σύμφωνα με τη βιβλιογραφία στον τομέα της κλινικής έρευνας, οι εκφράσεις του προσώπου χρησιμοποιήθηκαν για την αξιολόγηση πιθανών ελλείψεων ή αδυναμιών στη συναισθηματική έκφραση, καθώς και στην κοινωνική διάγνωση σε ψυχιατρικές και ψυχολογικές διαταραχές [2] [3]. Η Μείζονα Καταθλιπτική Διαταραχή (MMD) είναι η πιο συχνή διαταραχή της διάθεσης, η όποια ποικίλλει από την άποψη της σοβαρότητας, αλλοιώνοντας την ικανότητα του ατόμου να ασχοληθεί με την καθημερινή ρουτίνα του. Τα στοιχεία που παρέχονται από την Πράσινη Βίβλο της ΕΕ [1], [4], καταδεικνύουν ότι η κατάθλιψη είναι μια από τις κύριες αιτίες της κακής υγείας και της δυσλειτουργίας παγκοσμίως. Στις ΗΠΑ, περίπου 50 εκατομμύρια άνθρωποι είχαν διαγνωστεί με κατάθλιψη έως το 2015. Σύμφωνα με τις τελευταίες εκτιμήσεις του Οργανισμού Παγκόσμιας Υγείας (WHO), περισσότεροι από 300 εκατομμύρια άνθρωποι παγκοσμίως υποφέρουν από κατάθλιψη, καταγράφεται δηλαδή, μια αύξηση άνω του 18% μεταξύ 2005 και 2015. Το γεγονός αυτό φαίνεται να επιβεβαιώνεται από την Πράσινη Βίβλο της ΕΕ [1], στη οποία καταδεικνύεται ότι ένας στους τέσσερις πολίτες υποφέρουν από κάποια ψυχική ασθένεια κάποια στιγμή κατά τη διάρκεια της ζωής τους, προκαλώντας μερικές φορές ακόμη και αυτοκτονικές τάσεις.

Οι αντικειμενικές εκτιμήσεις της καταθλιπτικής συμπτωματολογίας θα μπορούσαν να είναι επωφελείς για τους κλινικούς ιατρούς, στο πλαίσιο ενός συστήματος υποστήριξης κλινικών αποφάσεων.

Η αυτόματη αναγνώριση των συναισθημάτων μέσω ευφυών συστημάτων, η οποία αξιολογεί τις εκφράσεις του προσώπου είναι ένα ζήτημα ζωτικής σημασίας, που πρέπει να αντιμετωπιστεί συμβάλλοντας στην κατανόηση της ανθρώπινης συμπεριφοράς, των διαπροσωπικών σχέσεων, και κυρίως για την αξιολόγηση της ψυχικής υγείας. Σύμφωνα με τη βιβλιογραφία στον τομέα της κλινικής έρευνας, οι εκφράσεις του προσώπου χρησιμοποιήθηκαν για την αξιολόγηση πιθανών ελλείψεων ή αδυναμιών στη συναισθηματική έκφραση, καθώς επίσης και στην κοινωνική διάγνωση σε ψυχιατρικές και ψυχολογικές διαταραχές [2], [3]. Η Μείζονα Καταθλιπτική Διαταραχή (MMD) είναι η πιο συχνή διαταραχή της διάθεσης, η όποια ποικίλλει ως προς τη σοβαρότητα, αλλοιώνοντας την ικανότητα του ατόμου να καταπιαστεί με την καθημερινή ρουτίνα του. Τα στοιχεία που παρέχονται από την Πράσινη Βίβλο της ΕΕ, είναι ανησυχητικά καθώς διατυπώνεται ότι ως το 2020 αναμένεται η Μείζονα Καταθλιπτική Διαταραχή (MMD) να αποτελεί τη βασική αιτία δυσλειτουργίας.

Στόχος αυτής της διπλωματικής είναι να αναπτυχθεί μια μεθοδολογία, που θα είναι ικανή να ανιχνεύει αντικειμενικά σημάδια Μείζονας Καταθλιπτικής Διαταραχής, συμβάλλοντας έτσι στην κλινική φροντίδα των ασθενών. Ειδικότερα, το προτεινόμενο σύστημα στοχεύει στην αναγνώριση ενδείξεων στις εκφράσεις του προσώπου ενός ατόμου, που σχετίζονται με την Μείζονα Καταθλιπτική Διαταραχή. Στα πλαίσια της παρούσας εργασίας που πραγματοποιήθηκε σχεδιαστήκαν αλγοριθμικές διαδικασίες για τον εντοπισμό προτύπων, ικανών να ανιχνεύουν συγκεκριμένες εκφράσεις του προσώπου που συνδέονται με την κατάθλιψη και μελετήθηκε η συμπεριφορά τους μέσω πειραματικών ελέγχων. Συγκεκριμένα, η λειτουργία της μεθοδολογίας και των αλγορίθμων που αναπτύχθηκαν επικυρώθηκε με το AviD-Corpus (AVEC’14), ένα σύνολο οπτικοακουστικής καταθλιπτικής γλώσσας σώματος [5],το οποίο παρέχει οπτικοακουστικό υλικό συνοδευόμενο από ετικέτες με τον αντίστοιχο βαθμό κατάθλιψης του κάθε συμμετέχοντα.

Στην προτεινόμενη υλοποίηση εξετάστηκαν προσεκτικά τα συναισθήματα της χαράς και της λύπης, που είναι δυο από τα έξι βασικά συναισθήματα, σύμφωνα με τους Ekman και Friesen [6] [7], καθώς φαίνεται να συνδέονται θετικά και αρνητικά αντίστοιχα με την MMD. Παραδείγματος χάριν, το συναίσθημα της λύπης εμφανίζεται να είναι κυρίαρχο στα άτομα που υποφέρουν από MMD από εκείνο της χαράς.

Συνοψίζοντας ο κύριος σκοπός της παρούσας εργασίας είναι να σχεδιαστεί και να αναπτυχθεί μια εφαρμογή με το MATLAB με στόχο την ανάλυση της εικόνας του προσώπου, με απώτερο στόχο την ανίχνευση οπτικών ενδείξεων της κατάθλιψης μέσω καταγεγραμμένων βίντεο. Όπως αναφέρθηκε και παραπάνω, οι εκφράσεις του προσώπου είναι καθοριστικές στην ανθρώπινη αλληλεπίδραση και επικοινωνία καθώς παρέχουν σημαντικές πληροφορίες όσων αφορά την ανάλυση των συναισθημάτων. Υψηλού επιπέδου πληροφορίες σχετικά με τα χαρακτηριστικά του προσώπου αντληθήκαν μέσω του OpenFace toolkit [8], ενώ δόθηκε ιδιαίτερη έμφαση στο AviD-Corpus (AVEC’14), ώστε να ελεγχθούν οι υλοποιημένοι αλγόριθμοι. Αφού εντοπίστηκαν οι απαιτούμενες υψηλού επιπέδου πληροφορίες που είναι σχετιζόμενες με τα χαρακτηριστικά του προσώπου, επιλέχτηκαν οι κατάλληλοι αλγόριθμοι μηχανικής μάθησης. Απώτερος στόχος ήταν να εξεταστεί η «Ευαισθησία» (‘sensitivity’) και η «Ειδικότητα» (‘specificity’) του παρόντος εργαλείου υποστήριξης για τη διάγνωση της κατάθλιψης. Ορισμένοι από τους αλγόριθμους κατηγοριοποίησης που εξεταστήκαν είναι οι παρακάτω: Διαχωριστική Ανάλυση (Discriminant Analysis), Δέντρα αποφάσεων (Random Forest Tree), Απλός Κατηγοριοποιητής Bayes (Naïve Bayes), Γραμμικός (Linear) και τον Κατηγοριοποιητή Εγγύτατου Γείτονα (k-nearest Neighbor Classifiers). Η αποτελεσματικότερη μέθοδος περιλαμβάνει τον Κατηγοριοποιητή Διαχωριστική Ανάλυση (Discriminant Analysis) με χρόνο εκτέλεσης τα 4 λεπτά, κατά προσέγγιση, για την ακρίβεια 4,29 λεπτά. Καταληκτικά, επιλέγοντας κατά προσέγγιση το 1 λεπτό- παράθυρο και αξιοποιώντας το πρωτόκολλο Leave-One-Subject-Out της μεθόδου cross-validation, ακρίβεια 72,57% για το εργαλείο υποστήριξης της διάγνωσης της κατάθλιψης.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Initially, I would like to express my appreciation to Dr. Manolis Tsiknakis, my major Professor, for accepting me as his master student. He has definitely been a source of inspiration from the beginning. The encouragement and trust he has shown to me were crucial factors that enabled the completion of this thesis.

My thanks also goes to my co-advisor Ms. Pampouchidou Anastasia, since she has been mentoring me throughout my dissertation. Many thanks to all my colleagues and those from the Biomedical Informatics & eHealth Laboratory, but especially Dr. Matthew Pediaditis and Mrs. Evaggelia Maniadi for being supportive. Also, I would like to thank Eirini Christinaki for her willingness, Flora for the construction of the image pattern through Photoshop, and Effie for being around whenever I needed her.

I firmly believe that we should never forget where we first started from. So, special thanks go to Prof. George Kornaros, my undergraduate supervisor, as he provided me with the courage and the supplies for the next steps.

Last but not least, I would like to thank my sister Eleni and Nasos, for helping me on the final steps of preparation of this dissertation.

Calliope-Marina Vazakopoulou
2018

Dedication

My bottomless appreciativeness for their love and many years of supporting goes to my precious parents, Konstantinos and Eirini. So countless thanks for being patient, always next to me and encouraging me during my entire journey in life. Also, I would like to express my gratitude to my sister Eleni for her persistence that I should participate in the postgraduate process (I owe it to you!). I could not omit thanking my four leg doggo-therapist Rastos for being around. So, a monumental thank to my family for believing in me when I did not do so myself.

Dedicated to those close to my heart!
My beloved family …

Chapter 1 - INTRODUCTION

This era is mainly characterized by the global instability, both in economic and political terms, and as expected the evidence presented is neither positive nor encouraging. The latter consists an essential factor causing strain to a person's state of living, both personally and socially, but also economically. The aforementioned factors can be associated with the mood disorder of depression, which affects 350 million people worldwide. Speaking of numbers, around 58,000 deaths is the result of suicide annually, which is more than the respective deaths resulting from fatal illnesses, road accidents or murders today in the EU. According to EU Green Paper [1], depression is considered to be the fourth most significant cause of suffering and disability. Statistics show that depression will be the main factor of mental illness by 2020 in the developed world. More than 27% of adult Europeans are estimated to experience at least one form of mental ill health over a year, classifying anxiety disorders and depression as the most common forms of mental illness in the European Union [9].

According to clinical literature, depression can be mild, moderate or severe, and the most typical symptoms of a depressed individual are described below. Emotions such as deep sadness may last for weeks or months depending on the severity, which ultimately acts as a brake in daily life. Motivation can be affected too, and people may experience thoughts about life, for example if they are worth living or not, which in some cases may lead to a suicidal behavior. The cause of stress, anxiety or depression may not necessarily be work-related. The person concerned may be experiencing difficulties outside the workplace such as bereavement, financial problems, relationship breakdown or other family problems. Indeed, non-work related stress, anxiety, and depression cause more sickness absence than work-related difficulties. At such times, work may be the ‘safe place’: the place where they feel supported and appreciated, where their role will be an important one.

After reviewing the state-of-the-art methods related to mood disorder detection and mainly the Major Depression, we concluded that those related to this disorder are still in primary levels, despite the fact that automatic affective computing recognition has become an active research area in the past decade. Although treatment of depression disorders has proven to be effective in most cases, misdiagnosing depressed patients is a common obstacle [10], [11]. Not only because depression manifests itself in different ways, but also because clinical interviews and self-reported history are currently the only ways for diagnosis. This raises a risk of subjective prejudices, either from the patients self-report or the clinical estimation. Thus, by utilizing the advancements of affective computing techniques, the goal is to develop an objective facial image analysis framework, so as to support mental health professionals during the diagnosis of clinical depression. Several statistical analyses have shown that situations such as unemployment, job losses, and low living standards, may show signs of a decrease in a person's self-esteem, resulting in depression. All experts describe depression as a common mood disorder with a main symptom of a persistent sadness. At this point, it is worth mentioning another phenomenon which is none other than the refugee flow happening at the European level. Thus, the refugees can be further categorized as marginalized groups, which seem to be at an increased risk of mental illnesses, not only because of the conditions they have already experienced before arriving in the host countries, but also due to the situations they find with their arrival. The main objective of this work was to implement a facial image analysis application with the ultimate goal of detecting visual signs of depression that would work supportively for mental health professionals, in order to help vulnerable groups.

Facial Expressions can be considered as a kind of nonverbal communication. It is worth mentioning that when people are communicating, 55% of the message is conveyed through facial expression, 38% provided by vocal cues, and only the remaining 7% is carried by verbal cues [12]. Therefore, observing facial expressions, can be beneficial in understanding each other. Facial expressions play a further decisive role on everyone's daily life, as their absence would not allow harmonic interaction with other people. It is found that a distressed face seems to persist even after the recovery from depression, increasing the risk of depressive episodes in the future [13], [14], [15]. That is why we have been extensively examining the characteristic features of the person where the sadness is depicted.

It is generally accepted that depressed people are characterized by 'drowsiness', so the body gestures that involve the whole body or individual parts may also contribute on carrying out cues. Also, other symptoms that have been recorded as signs of depression are behavioral nervousness, accumulated anxiety and nervous movement of the foot, depending on the type of depression. In addition, other contributing factors, which have been extensively used to assess depression are reported to be the pose of the head (head pose), its motion and direction. The inactivity seems to be a typical characteristic of depressed people, motivating the examination of both the existence and the frequency of movement on specific facial features. In terms of non-verbal manifestations of depression, the facial features such as the eyes and the mouth have been reported to transmit important information about the psycho-emotional status of a person. As far as it concerns, some relative features of the eyes: eye openings/blinking the motion of the iris, direction of sight, the absence of eye contact and the activity of the eyelids, can also be considered. In the case of the mouth, we should focus more closely on the following features. It is a fact that phenomena such as smile intensity, smile duration, smiling while not talking and lack of smiles are additional useful characteristics to draw conclusions, whether the patient shows signs of depression [16]. Taking into account all the above factors we utilized two basic emotions. The emotion of happiness, which is reduced or even absent from a depressed person, in contrast to the emotion of sadness, which is frequently evident.

Until today, the audio-visual systems for assessing automated depression, which either utilize a continuous or a categorical model, have exclusively been evaluated in research level. However, a main drawback is that these recognition systems have not been applied to general populations in order to assess their feasibility. For the first time in October 2013 the Audio/Visual Emotion Challenge, referred to as “AVEC'13” [17] took place and about a year later, i.e., in November 2014 another challenge followed, referred to as “AVEC'14” [5]. The challenge of depression required the prediction of an individual score in the Beck Depression Inventory (BDI) [10], therefore the assessment of the participants' depression rate was based on their videotaped recordings (video and speech available). Given the dataset provided by the aforementioned challenges (“AVEC'13” & “AVEC'14”), is, to the best of our knowledge, the only publically available dataset, also consisted by video recordings, and annotated with a depression labels, it was the proper choice for the accomplishment of our project. Therefore, the AVEC dataset was employed during the development of the proposed system, for testing the developed algorithms in order to extract high-level features related to the non-verbal manifestations as described herein. The OpenFace toolkit (software) will be used for locating facial landmarks, utilized for extracting high-level information about the facial features.

Summarizing, we present a computational framework for categorical assessment of depression based on high-level features that emerged from the AVEC dataset and therefore the analysis of facial expression changes, which occur to each participant. In the case of feature extraction, we examined the two specific categories, namely geometric (Linear and Eccentricity features) and motion-based techniques. The second phase of our implementation addresses the classification problem. We experimented with a number of classifiers, with the ultimate goal of answering which one provides the best confidence rates. Moreover, we cross-examined the completion time of each classifiers, because it is not suffice to be slow with satisfying results or the opposite, that is, fast but with mild results. Namely we execute Discriminant Analysis, Random Forest Tree, Naïve Bayes, Linear and k-nearest Neighbor Classifiers. Our top performing method has emerged with Discriminant Analysis classifier, with an accuracy of 72.57%with an execution time of approximately 4 minutes for our depression assessment framework. By selecting an approximately 1 second-window, with Leave-One Subject-Out cross-validation, our experiments were accomplished. As it is mentioned earlier, the target group of this implementation is addressed to all population groups, although special focus is given to those who currently show signs of depression. The major benefit of such a tool is that it could be part of a decision support framework for depression assessment, aiding clinicians with the diagnosis by detecting relevant non-verbal cues.

1.1 Thesis Outline

In Chapter 1a brief description of depression and the target group of people that this mental illness affects were presented, along with a summary of the employed methodology. The rest of the thesis is organized as follows. Chapter 2 provides the clinical perspective of depression. More specifically, the symptomatology according to DSM-5 [18], the classification of mood disorders, and treatment of depression are demonstrated. In Chapter 3 a review of the literature about Clinical Depression definitions and related work about 2D Facial landmark Detection implementations is listed. As well as the databases related to this mood disorder, related to our research. Chapter 4 presents the proposed methodology: preprocessing, features extraction, and classification of high-level information about the facial features. Furthermore, the experiments and the results obtained are discussed in Chapter 5. The different configurations of the tested methods are included, although special emphasis is given to best results of our investigation. Finally, Chapter 6 presents the conclusions, and discusses directions for future work for further improvements in performances and utility of the proposed framework.

1.2 Motivation and Research Questions

The aim of this thesis is to perform R&D in the domain of affective computing. More specifically, the objective of the current thesis is to design and develop an application in MATLAB for facial image analysis with the purpose to detect visual signs of depression from video recordings. As a result, the main research question posed in the context of the present master thesis can be expressed as follows: "Which is the optimal method to provide a tool that could be part of a decision support framework for depression assessment?" The main research question can be decomposed into a number or related sub-questions:

1. "Which is the optimal feature set to maximize the performance of the implementation?"
2. "Which machine learning algorithm generates the best classification results, in terms of sensitivity and specificity for this application?"

1.3 Research Methodology-Libraries Utilization

The review conducted with the use of online databases (Springer[2], Google Scholar[3], IEEE[4], ResearchGate[5], National Center for Biotechnology Information [(NCBI) - PMC][6] and Scopus[7], ACM[8]) is dated from 2010 to 2017. With the exception of the topics of and BDI and Affective computing which are respectively dated since 1996 and 1997. Primarily, the search was complemented by a manual review by using keywords in order to detect papers related to the topic. The articles which were eventually considered were selected based on their relevance with the specific scientific field, i.e. 2D- Geometrical Facial Feature Landmarks Detection in terms of Depression and Emotions.

During this study all features, data, and signals were processed in MATLAB which is a software used for signal processing mainly for prototyping purposes. All tests for the different methods of classification were also performed in MATLAB, which includes a large collection of inbuilt classification functions.

In the process of our implementation we focused on estimating the optimal feature set used to correctly classify the presence of facial signs relevant to depression. In other words we aim to compute the maximum success of f1-score, namely to achieve the optimal recognition on whether or not the candidate indicates signs of depression, based on the provided labels (BDI score). After the many different experiments, an optimal feature set for classification purposes was reached.

[2]http://www.springer.com/gp/
[3]https://scholar.google.gr
[4]https://www.ieee.org/index.html
[5]https://www.researchgate.net
[6]https://www.ncbi.nlm.nih.gov/pmc/
[7]https://www.scopus.com
[8]https://dl.acm.org

With the main objective as stated above, we proceed with the completion of this work by analyzing a number of factors and their interaction which are: the quality of AVEC dataset, the size of the training dataset, and the window size. Finally, conclusions have been drawn and findings have been presented.

1.4 List of Publications

The work reported in the present dissertation contributed to the following scientific publications:

i. A. Pampouchidou, O. Simantiraki, C. - M. Vazakopoulou, C. Chatzaki, M. Pediaditis, A. Maridaki, K. Marias, P. Simos, F. Yang, F. Meriaudeau, M. Tsiknakis,"Facial Geometry and Speech Analysis for Depression Detection", 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 11-15 July 2017 (1433:1436)

ii. A. Pampouchidou, M. Pediaditis, A. Maridaki,M. Awais, C.-M.Vazakopoulou, S.Sfakianakis, M,Tsiknakis,P. Simos, K. Marias, F. Yang, F. Meriaudeau, “Quantitative comparison of motion history image variants for video-based depression assessment” submitted to EURASIP Journal on Image and Video Processing, 2017. (1), 64.

iii. A. Pampouchidou, O. Simantiraki, C. - M. Vazakopoulou, K. Marias, P. Simos, F. Yang, F. Meriaudeau, M. Tsiknakis, " Détection de la dépression par l’analyse de la géométrie faciale et de la parole", GRETSI - Traitement du Signal et des Images Juan-Les-Pins - 5/8, Septembre 2017.

iv. C. - M. Vazakopoulou, A. Pampouchidou, F. Yang, F. Meriaudeau, K. Marias, M. Tsiknakis, “Détection de la dépression par analyse de la géométrie faciale et apprentissage automatique”, Le Congrès National de Recherche des IUT (CNRIUT), Aix-en- Provence, CNRIUT’2018.

Chapter 2 - CLINICAL PERSPECTIVE OF DEPRESSION

In order to achieve the goal of this work, which is the implementation of a supportive framework for clinical health professionals (see Chap.4), anticipated to recognize whether the patient shows signs of depression, it is primarily necessary to have knowledge of the attributes that are identified in a depressed person. This chapter corresponds to a brief historical review of clinical depression and the degree of disease emergence in modern society, as well as the findings that researchers have come up with over time. Afterwards a brief description of clinical depression, classifications of clinical depression based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) and a general idea of depression symptomatology is given.

2.1 Historical Review

In this subchapter, we quote a brief historical background on the spread of depression over the years. More particularly, scientific opinions are summarized and the severity of the mental disorder is confirmed by the percentage rates mentioned below.

2.1.1 Review of Depression over Time

Remarkable is the fact that depression had already been assessed twenty centuries ago, with Aaron Beck classifying it as a disease state since 1967, i.e. as a clinical case, not just as emotional deficiency. Regier et.al also found that about 9.5% of the U.S. population from 18 years old and older (18.8 million American adults) are affected by mental disorders in any given year [19]. In the mid-1990s, further research was added to the previous ones, giving also less encouraging statistics. Indeed, the American Psychiatric Society in 1994 claimed that the risk of MDD was fluctuated between 10%-25% for women and 5%-12% for men. According to Oakes et.al [20], there was another study from the previous decade demonstrated in 2003 by Kessler et al. [21], concluding that the 16.2% of the global population had suffered at least one clinical depression episode in the previous year. The same year, Sheline et al. [22], found the recurrence of depression ranges between 20-80% within five years despite the previous successful treatment completion [20]. Below, a diagram (Figure 2) is given concerning the suicide and self-inflicted injury for males and females, from selected European countries, in 2006 or latest available year [9], which confirms the severity of this mental illness.

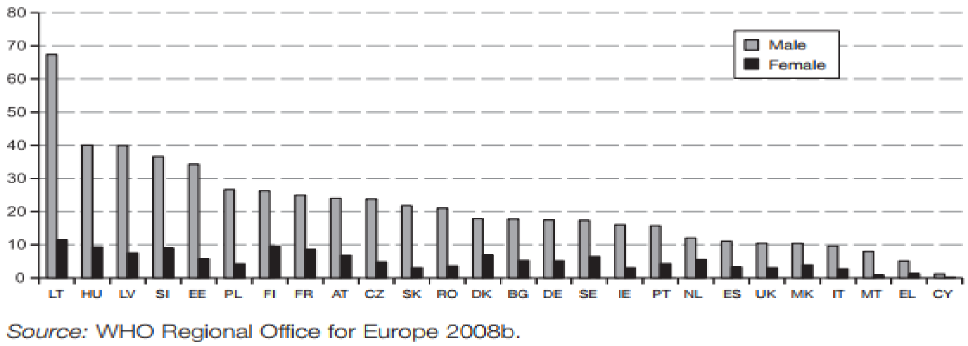


Figure 1: Standardized death rates per 100 000, all ages, from suicide and self-inflicted injury for males and females, selected European countries, 2006 or latest available year.

In the years that followed, the cause of depression remained largely unknown. In the recent past, and specifically a decade ago, Hasler pointed out that the rates of depression vary between 30-40% and are associated with genetic factors. However the genes involved are not clearly identified. Until now, it is stated that scientists have difficulty identifying a specific brain region or neurotransmitter path as the only cause of depression [20]. In agreement with The World Mental Health Survey that took place, just five years ago, in 17 countries it is stated that the average number of people who reported the occurrence of at least one episode of depression, during the previous year, reached 1 in 20. Thus, depression has been considered the leading cause of disability worldwide, while attention has been focused on seeking to reduce clinical depression and other mental health conditions [23].

The findings of the last two years, reporting the outbreak of depression worldwide could be considered far from encouraging. It is still manifested to be as one of the 'hot' and more extensive medical problems of the era. This view is confirmed by Cleary et al, who claims that depression is the leading cause of disability, estimated to affect up to 10–15% of the general population [20]. According to the World Health Organization (WHO) it is denoted that more than 300 million people are affected in a global level from this disorder. Another remark is that depression could lead to suicide. More specifically, close to 800.000 people commit suicide every year, making suicide the second leading cause of people’s death, aged 15 to 29 years old [24]. As announced by the National Institute of Mental Health (NIMH) in the US, the documented cases of individuals with clinical depression holds primacy in appearance compared to other mental illnesses, by affecting 10 million of the population each year [19].

2.2 The Clinical Definition of Depression based on DSM-5

Description of DSM and ICD-10

The Diagnostic and Statistical Manual of Mental Disorders (DSM) is a tool that provides diagnostic pattern for disorders, as well as standard criteria for the classification of mental disorders. The fundamental professional organization of psychiatrists and trainee psychiatrists known as the American Psychiatric Association, which is located in the U.S., publishes through years the DSM [18].

ICD-10, is another manual of Classification of Mental and Behavioral Disorders, released by the World Health Organization Health Organization (WHO) [25]. The ICD-10 seems to be more often proposed as a guide for clinical diagnosis rather than the DSM that appears to be applied chiefly at research level when it comes solely on mental disorders[9]. So when it comes to the following collected information for clinical depression, we have chosen the most recent manual, known as DSM-5, i.e., the latest version, released on May 18, 2013 in order to be more accurate and up-to-date.

Definition of Clinical Depression

Occasionally several definitions have been formulated by experts and non-experts on the term of clinical depression. Below brief definitions are given, based on DSM-5 [18], which lists the commonly accepted characteristics that define a mental disorder, but also those basic features that should be treated as clinical depression and not just as sadness.

At first, the meaning of the term "mental well-being" has to be made clear, to help us find out what those key features missing from a depressed person are. In the definition according to WHO, mental health is given as: “a state of well-being in which the individual realizes their abilities, can cope with the normal stresses of life, can work productively and fruitfully, and are able to make a contribution to their community” [25]. By definition, it can easily be established that depression is in conflict with mental well-being. Therefore, people suffering from clinical depression are categorized as mentally ill and in need of therapeutic support and treatment.

As reported in DSM-5: "A mental disorder is a syndrome characterized by clinically significant disturbance in an individual's cognition, emotion regulation, or behavior that reflects a dysfunction in the psychological, biological, or developmental processes underlying mental functioning.

[9] http://journal.ahima.org/2016/08/10/dsm-5-vs-icd-10-cm/

Mental disorders are usually associated with significant distress in social, occupational, or other important activities. An expectable or culturally approved response to a common stressor or loss, such as the death of a loved one, is not considered a mental disorder." [18].

A common assumption of mental health professionals is that the severity of the illness, due to the long period of its existence, deteriorates the ability of the person to respond to everyday life. The UMHS Depression Guideline [26], introduces depression as a typical mental mood disorder that needs to be seriously addressed because of the high percentage of morbidity and lethality. Indicatively, it is reported that the financial burden due to the major depressive disorder had reached 59 billion dollars in 2006. Overall, the above guideline mentioned that the rates of hospitalized patients for severe depression, causing mortality due to suicide are estimated to have reached 15%. Additionally, in case of the female population, the rate of depression appears to be higher than that of the males. That is corroborated, since the disorder appears to be between 20% and 25% for women, while men are exposed at danger of 7% - 12%. According to Oakes et.al [20], depression seems to be the number one cause of disability, affecting up to 15% of the population, classifying this one as the most frequent psychiatric mood disorder.

2.3 Classifying Mood Disorders

This particular sub-chapter outlines the categories of mood disorders (Figure 1), as well as some of the major changes of DSM-5 version, including the introduction of two new disorders: the disruptive mood dysregulation disorder and the premenstrual dysphoric disorder. Though the Bereavement Exclusion[10] is not further manifested as a condition of clinical mood disorder.

[10] https://pro.psychcentralcom/dsm-5-changes-depression-depressive-disorders/004259.html

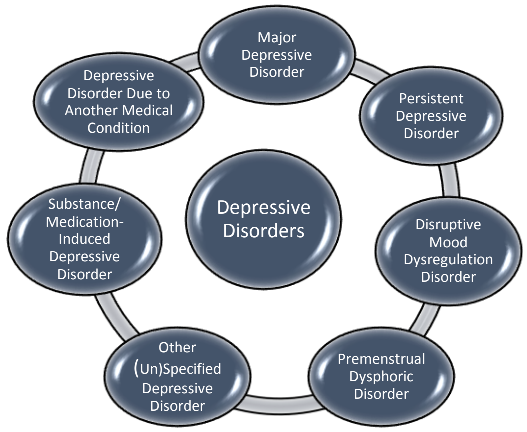


Figure 2: Classification of Depressive Disorders via DSM-5.

Disruptive Mood Dysregulation Disorder

A new type introduced in the DSM-5 [18], the Disruptive Mood Dysregulation disorder, can be diagnosed in children between 6 to 18 years old, when the age of initiation appears before the 10th year of life.

In order to diagnose such a disorder, the DSM-5 mentions that the following conditions must be encountered. Initially, fundamental circumstances for the clinical diagnosis seem to be the duration of symptoms that must be at minimum 3 times a week. Also the severity and tenacious irritability or anger of attitude, as well as the recurrent temper eruption must be present almost daily. It is also suggested that great attention be given when symptoms are constantly present for at least 12 months. A further observation is that during this period the child should not have been three or more consecutive months without clinical symptoms.

Major depressive Disorder

In order to achieve a comprehensive diagnosis of major depressive disorder (MDD), it is necessary to examine the number of episodes that occur, as well as the severity of these events. So, with regard to the number of episodes, the division of the following two main categories for Major depressive disorder (MDD) has been decided. The first category refers to episodes occurring once (single depressive episode), with absence of a history of manic or hypomanic episode. The second category responds to recurring episodes (recurrent depressive episode), where the criteria are the same as those of the MDD in individual episodes, but with the minimum occurrence of 2 major depressive events.

An additional categorization that works supportively for the diagnosis, related to the severity of major depressive episodes, is the following: mild, moderate, severe, with psychotic features, and partial or full remission. Furthermore, the DMS-5 refers the following 9 specifiers, giving the final and complete clinical profile of the patient:

Specify with:

- anxious distress
- mixed features
- melancholic features
- atypical features
- mood-congruent psychotic features
- mood-incongruent psychotic features
- catatonia
- peripartum onset
- seasonal pattern.

In order to diagnose an individual with MDD, five, or more of the following symptoms must be present for 14 days without remission: loss of pleasure in all or almost all activities, weight loss or gain (more than 5% of bodyweight), insomnia or hypersomnia, psychomotor agitation or retardation, fatigue, depressed mood, diminished interest. Also, feeling of worthlessness or unreasonable guilt, anhedonia[11], disability to concentrate, recurrent thoughts of death or suicide could be present. At least, one of the symptoms must be either depressed mood or diminished interest or pleasure. Another statement declaims that Major depression is a combination of symptoms that interfere with the ability to work, sleep, eat and enjoy pleasurable activities. These disabling episodes of depression can occur once, twice or several times in an individual existence [27].

Persistent Depressive Disorder (Dysthymia)

Dysthymia is replaced with the term "persistent depressive disorder." A recent report on the characteristics of dysthymia was given by Iyer et.al [27], who proposes that Dysthymia, a less severe type of depression, implicates chronic symptoms that do not disable the patient, though deprive them from being functional or feeling well. Sometimes people with Dysthymia may also experience major depressive episodes. Devanand demonstrated that "the prevalence of dysthymic

[11]http://www.medicinenet.com/script/main/art.asp?articlekey=17900
"Anhdonia: Loss of the capacity to experience pleasure. The inability to gain pleasure from normally pleasurable experiences. Anhedonia is a core clinical feature of depression, schizophrenia, and some other mental illnesses."

disorder is approximately 2% in the elderly population where subsyndromal depressions of lesser severity are more common" [28]. The new condition includes both chronic major depressive disorder and the previous dysthymic disorder. In agreement with "Depressive Disorders: DSM-5(r) Selections 1st Edition" [29], the change is due to: "An inability to find scientifically meaningful differences between these two conditions led to their combination with specifiers included to identify different pathways to the diagnosis and to provide continuity with DSM-IV".

Premenstrual Dysphoric Disorder

An additional formal and independent category for diagnosis in DSM-5 is now considered the Premenstrual Dysphoric Disorder. In order for a female patient to be diagnosed with this disorder she must be distinguished by at least 5 mood symptoms that occurred 7 days before the beginning of the majority of menstrual cycles during the previous year. Pursuant to DSM-5 [18] the following symptoms are nominally referred: marked affective emotional instability (mood swing), marked irritability or anger or interpersonal conflict, marked depression, and clear anxiety or tension, anhedonia, poor concentration, fatigue, variance in appetite, insomnia or hypersomnia, sense of feeling out of control/overwhelmed, and physical symptoms, i.e., aches and pains. Moreover, it is still mentioned that the risks and prognostic factors associated with environmental, genetic, and physiological and course modifiers, with the onset of this disorder are likely at any point after the menarche.

Substance/medication-induced Depressive Disorder

Another finding made by clinical health professionals, lies in the observation of the occurrence of depressive symptoms during, or soon after exposure to a substance or medication. In agreement with DSM-5 [18], the symptoms of this disorder are not due to an independent depressive disorder caused by non-substance or medication and do not appear only during a delirium. The core symptoms of this disorder are: marked, persistent, and function-impairing depressed mood or anhedonia caused by a substance or medication.

Depressive Disorder due to another Medical Condition

It is proven that depressive disorder may be caused by another medical condition. In order to diagnose an individual with the type of depressive disorder due to another medical condition the following criteria must be met. The main characteristic of this disorder is the decrease in patient function due to depression or anhedonia with the main feature of persistence. When it comes to the symptoms, it is mentioned that they are not triggered by another mental disorder and do not appear solely throughout a delirium.

Other Specified and Unspecified Depressive Disorders

This mood disorder was previously mentioned (DSM-IV) as depressive disorder not otherwise specified (NOS).An example of the other specified depressive disorder is the recurrent brief depression, in which four symptoms of depression must be present, with a duration of 2-13 days for a year. Additionally, another specifier is the short-duration (4-13 days) [18] depressive episode, which is associated with major depressive disorder, as four of the symptoms of MDD are repetitive for more than four days. Lastly, the depressive episode with insufficient symptoms is associated with clinically significant distress or impairment that persists for at least 14 days and one of the eight symptoms of MDD appears.

The unspecified depressive disorder does not determine the reason for not meeting criteria for other depressive disorder categories. More specifically, in an unspecified depressive disorder a few or all of the symptoms of individual types of mood disorder are indicated. The different combination of symptoms makes it more difficult to diagnose an unspecified depressive disorder, even though the signs of a problem are evident.

Bereavement Exclusion

It is controversial that bereavement exclusion for the diagnosis of major depression has been removed from the revised DSM-5. Supporters of DSM-5 defend their position by considering that there is no clinical basis for blocking patients from the diagnosis of MDD, when the condition occurs shortly after the death of a loved one (mourning), which in most cases typically ceases after a period of two months [30], [31]. The opposite view refers to the removal of the bereavement exclusion, which will result in the encouragement of over-prescribing antidepressants and the "medication" of an ordinary sadness [31]. As reported by A.P.A. [29], since MDD (Major Depressive Disorder) roused under bereavement, it subjoined a risk for suffering, suicidal ideation[12], poorer somatic health, feeling of worthlessness, both worse interpersonal, functioning and an increased risk for persistent complex bereavement disorder, which is described in DSM-5.

[12] https://psychcentral.com/encyclopedia/suicidal-ideation/

"Suicidal ideation, or suicidal thinking, is the contemplation of ending one's own life. These types of thoughts may arise in people who feel completely hopeless or believe they can no longer cope with their life situation. Suicidal ideation can vary greatly from fleeting thoughts to preoccupation to detailed planning."

2.3.1 Depression Symptomatology

For many people with depression, symptoms usually are severe enough to cause noticeable problems in day-to-day activities, such as social activities, work, school, or relationships of any kind, or even co-existence with others. Despite the fact that depression seems to happen once in a person's life, it is not uncommon for patients to suffer from numerous episodes. General symptoms include: tenacious sadness, anxiety or moodiness, and weight loss/gain or reduced energy. Sense of pessimism or culpability, worthlessness, weakness, or uninterested in activities once seem to be enjoyable (Anhedonia). Thoughts of death or suicide, suicide attempts, difficulty in concentrating or decision-making. In the following subsections, we will ascertain that there is enough overlapping of symptoms, depending on the age of the patients.

Chapter 3 - LITERATURE REVIEW

In the following sections, , we are going to refer to the work carried out in the recent past on this research area, but also on the three most valid clinical datasets created for this purpose, which are none other than DAIC-WOZ, BlackDog and AViD-Corpus dataset (AVEC'14). The chapter is concluded by presenting our novelties in the area of 2D facial feature detectors, aiming at innovating in clinical care by providing a system for diagnosis of depression.

3.1 Related Work

The bibliographic review of our research has been divided into two subsections. We initially approached the wider context of the various applications, associated to automatic facial expression analysis (AFEA) [32], using methods that have been implemented recently in the field of emotional recognition. In a second phase, both by taking into account the previous fields of research, as well as the research conducted for depression [33], we examined the behavior of the object of our analysis, i.e. those emotions (imprinted frequent feeling of sadness and absence of that of joy [34]) usually found in a person suffering from depression. The reason we worked on recognizing emotions is because we have less resources of already developed tools depicting people in depression. That is confirmed by the distinguished in that field, Cohn [35], studies intensively the automatic analysis of emotions for clinical purposes such as the recognition of depression.

Implemented Emotion Recognition Methods

At this point, related work carried out in the recent past, corresponding to the proposed method is being reviewed. The automatic recognition of the person's expressions and their categorization into six emotions (Figure3), defined by Paul Ekman [6], [7], seems to attract the interest of the researchers for clinical purposes.

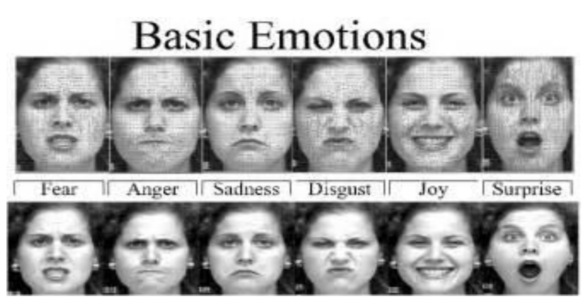


Figure 3: The Internationally Recognized Facial Expressions of the 6 Basic Emotion according to Ekman.

An implementation with geometric-based method in order to make the optimal classification of emotion was proposed by Vonikakis et al. [36]. Zhang et al. [37], explored the impact of jointly modeling both an individual's self-report and the perceived labels of others. This was done by using deep belief networks, so as to learn a representative feature space and model the potential complementary relationship between intention and the perception using multi-task learning. The same year, Repenga et al. [38], found out the relation between eye actions and smiles by using spontaneous datasets, which are the Cohn-Kanade(CK+) and the DISFA. Another method for geometrical facial features was proposed by Loconsole et al. [39], who manifested that their method provides great accuracy in terms of the optimal classification of the basic emotions. Gacan et al. [40], presents a geometric-based method, with spatial features using two combination of landmarks. The extracted features carried out from Cohn-Kanade (CK+) dataset. This is based on facial and vocal expression in naturalistic video recordings.

In order to detect the mental state of people, a number of behavioral signals have been investigated. More precisely, this was achieved by utilizing signals like the rhythm/tone and the intensity of the speech, blood pressure, heart rate, facial expression, and brain signals [41], [42], [43], as transmitters for psychophysiological information. Also, there are cases where results derived from the combination of more than one of the previously noted methods [13]. Thus, clinicians manage to get knowledge, so as to come up with findings of mental dysregulations or/and depression [44]. However, in this dissertation we have been exclusively involved with facial expression (non-verbal cues), (see Chapter 4.3, for more details), related to those a depressed person displays. The current chapter provides a brief description of the already deployed applications in this field, on which we have based both the development and the completion of this work.

Methods of Depression Assessment or Mental Disorders

Both the communication and the understanding of other people's emotions lead to the need for a proper "study" of their facial expressions. Automatically, anyone can notice the importance of facial expressions but also the reason why they play an integral role in order to detect the symptoms of Depression. Lately, the automatic detection of depression [45], [46], [47], [48], [49] seems to be the new state-of-the-art method in line with the bibliographic review preceded [50], [51]. Such methods are based on machine learning techniques, utilizing audio or video signals, or the combination of them, i.e. multimodal approaches.

Wang et al. [52], presented an automated video-based facial feature detector, which is capable of being applied to patients with any mental dysregulation. Alghowinem et al. [53], detected that the average distance between the eyelids was significantly smaller and the average duration of blinks was significantly longer in depressed individuals, by making use of facial landmarks detected on individuals' faces from their video recordings. Scherer et al. [54], recommended 4 nonverbal behavior descriptors, namely: Vertical Head Gaze, Vertical Eye Gaze, Smile Intensity, and Smile Duration that could assess mental disorders such as depression, anxiety and/or PTSD. Their evaluation arose by visual signals, extracted from the Distress Assessment Interview Corpus (DAIC) dataset, by giving increased emphasis on features like pose of the head, eye-gaze and smiles.

Later on, Scherer et al. [55] proposed a number of nonverbal behavior descriptors that could automatically utilize audiovisual signals, to support and improve clinical assessments. These are the Visual Behavior Descriptors including: Vertical Head Gaze, Vertical Eye Gaze, Smile Intensity and Smile Duration. Also they added Acoustic Behavior Descriptors, namely: Normalized Amplitude Quotient (NAQ), Quasi-Open Quotient (QOQ), Spectral Stationarity, and Intensity variation. Gupta et al [56], developed a multimodal system, with audio, visual and linguistic cues, that was able to predict the depression level in continuous-time.

More recently Nasir et al. [57] , proposed a multimodal classification system for depression as part of the AVEC2016. According to their results that system performs the best audio modality, while polynomial parameterization of facial landmarks along with geometrical features turned out to be the best video feature set. Pampouchidou et al. [58], through the implementation of the article "Depression Assessment by Fusing High and Low Level Features from Audio, Video, and Text." utilizes a combination of high level and low level features. An artificial intelligent system was proposed for automatic depression scale prediction from Jan et al. [59]. Joshi et al. [60] , claim that their experiments, through the analysis via histograms of relative body parts movement prove the effectiveness of the proposed method. The same year it was Joshi et al., who developed an automated framework depression analysis. This was achieved by taking into account features as facial dynamics, head movement, relative body part movement etc. of patients suffering from major depressive disorders. Bhatia et.al [33], presented in her PhD a research, which provides an improved approach capable of measuring the severity of depression from multimodal channels.

3.2 Depression Datasets

When it comes to reliability test of our framework, inevitably we concluded in the necessity of being able to apply it, so as to ensure the flexibility of the framework. It is essential that frameworks are able to be performed with the same or better accuracy in more than one video samples. The absence of clinically valid datasets, in the research community, causes problems such as, the lack of accurate recognition systems.

However, in this dissertation this role is taken by the AVEC'14 dataset, which is the only openly available dataset, with the only exclusion not to be traded on the internet, for purposes of personal data protection. In this subsection, we will briefly mention some of the databases that exist, as well as their characteristics. It is worth pointing out that these datasets are intended for clinical research purposes, while some mainly concern the detection of depression. Subsequently, a general description of AVEC is given, (see Chapter 4 for extensive list of attributes) since the proposed method has been tested on that specific dataset, for evaluating the decision support framework for depression assessment and derive to conclusions for our work.

AViD-Corpus dataset (AVEC'14)

The AVEC[13] is a subset of the audio-video depressive language corpus (AViD-Corpus), composed by 340 video recordings of 292 subjects, where individuals are interacting with a human-computer interface, while spoken language during all tasks took is the German language [17]. The participants were recorded one to four times, during fourteen days, where the total number of people is 84 and the age ranging from 18 to 63 years old. The AVEC subset includes 150 sessions, which are divided equally into training, development and testing sets. The recordings consists of activities such as, reading, singing, sustained vowels, task solving, counting, etc. (see Section 4.1.1 for more details). Furthermore the self-reported depression scores from the individuals were specified according to the Beck depression rating scale (BDI) [61], and are provided only for the training and the development parts [5], that is 100 sessions.

DAIC-WOZ

The Distress Analysis Interview Corpus (DAIC)[14] database is part of a larger corpus and sustains the diagnosis of distress conditions such as anxiety, depression, and post-traumatic stress disorder, via clinical interviews [62]. This dataset consists of 189 tasks with an interaction ranging from 7

[13] https://avec2013-db.sspnet.eu
[14] http://dcapswoz.ict.usc.edu

to 33min. The participants had to fill in the PHQ-9 questionnaire, which is a clinical questionnaire concerning the assessment of depression. It has been showed that the 29% of the participants suffer from depression. Each task contains participants' audio files, transcription of the interaction, and facial features.

BlackDog

The Black Dog Institute[15] is recognized internationally as a leader in the detection, prevention and the cure of mental illness, and it was established in 2002. The dataset of audio-video recordings include men and women aged between 21 to 75 years old [63]. The dataset is composed by130 subjects, where the 60 were diagnosed as patients who had been diagnosed with severe depression, and 70 turned out to be healthy participants, carefully assessed by clinicians [33]. These audio-video sets contain several parts, including a read sentences task and an interview with the subjects i.e., 20 sentences with negative and positive meaning. The interview was conducted by asking specific open questions (in 8 question groups), where the subjects were asked to describe events that had aroused significant emotions [64].

3.3 Contributions & Novelties

As established in the extensive review we conducted, mental disorders are highly prevalent, and are considered as major and serious illnesses. In the case of proposed work, the contribution to the field of affective computing, and more specifically in the automatic recognition of depression, lies in the fact that despite the exclusive use of non-verbal features, the achieved performance is fairly compared to the state-of-the-art. More precisely, the achieved accuracy reached 72.56%, with a runtime of approximately 4 minutes of depression recognition on the person being examined. The main contribution of this thesis is that besides the most often geometric features that exploited, few new facial features characteristics that have been exported in case of depression assessment. At this point it is worth mentioning that these facial features have not been used in the past, namely these features were extracted via the eccentricity method. As far as it concerns the proposed framework, the best performance was achieved with the use of Discriminate Analysis classifier, which could be considered as another novelty, also by using the method of leave-one-out cross-validation. Our best performing method involved the Discriminant Analysis classifier, by selecting

[15] https://www.blackdoginstitute.org.au/clinical-resources

an approximately 1 second-window, and the Leave-One Subject-Out cross-validation, for our depression assessment framework.

Chapter 4 - METHODOLOGY

In Chapter 4, we present in detail the most frequently used, but also the most common facial expressions in people with depression problems, which we have utilized with the ultimate goal of achieving the most optimal implementation of our framework. In addition, the current chapter provides a more detailed description of the software platforms (OpenFace toolkik & Matlab) we have used, in order to implement the framework, as well as the characteristics of the AVEC dataset, the Beck Depression Inventory and other questionnaires assessing depression given.

4.1 Preprocessing

In order to be able to derive the necessary facial features from participants' video recordings, based on which we implemented the first part of our work (facial feature extraction). Thus, we had to go through the preprocessing phase, where we needed to find out the precise location of the landmark points. This was achieved by utilizing the OpenFace toolkit [8], which is a freely available and open-source software, for facial landmark detection, facial action unit recognition, eye-gaze and head pose.

4.1.1 AVEC'14 Dataset &BDI Remarks

Among the previous existing datasets, only the AVEC dataset could be shared under a privacy agreement. Thus, this was the main reason that we utilized this dataset. As it is mentioned in the previous chapter this database consists of 340 video clips of 292 subjects, where only one person appears in each clip, i.e. some subjects appear in more than one clip. The average age of participants is 31.5 years old, with a standard deviation of 12.3 years old and ranges from 18 to 63 years old. The speakers were recorded between one to four times, within a period of fourteen days [65]. The AVEC depression database consists of audio and video of the participants, who were involved in a human-computer interaction experiment, led by a presentation Power Point that included the following tasks:

- to talk loudly when solving a task
- to number from 1 to 10
- to read speech: excerpts of a novel and a fable
- to sing a German nursery rhyme

- to tell a story from their past: the best event until the present and the most sad fact in childhood
- sustained vowel phonation, sustained loud vowel phonation, and sustained smiling vowel phonation
- to say an imagined story applying the Thematic Apperception Test (TAT)

At this point, it is worth mentioning that we specifically used the AVEC 2014 dataset, where two specific tasks have been selected, namely FreeForm and NorthWind. We have not utilized the full recordings like AVEC2013 dataset. This is due to the fact that we only have the corresponding labels, which indicate the depression score, only for the 200 subjects.

Beck Depression Inventory

The Beck Depression Inventory, which is a self-reported 21 multiple choice inventory, was utilized for assessment of the depression severity [66]. The BDI scores range from 0 to 63. More precisely, the score 0-13 refers to minimal depression, the 14-19 score shows mild depression, moderate depression indicates the 20-28 score, while the score 29-63 points out severe depression. The average BDI-level in the AVEC dataset is 15 points (standard deviation = 12.3). As far as it concerns the hardware specifications, the setup was consisted by a webcam and a microphone, and took place in a number of quiet settings. The audio was recorded via a headset linked to the sound card of a laptop, with a sampling rate of 44, 100Hz, 16 bit. The last action which took place was the re-sampling of the initial videos, in order to obtain same specifications for all recordings 30 fps [56], at 640 × 480 pixels [67].

Other Questionnaires assessing depression

The K10[16,17] and PHQ-9 are the most generally utilized and approved validated assessment tools in health care. The questions in the K10 cover depression, anxiety and general psychological well-being[18] and this apparatus is typically favored for use as an underlying appraisal, in contrast with the PHQ-9, which is particularly use aids in diagnosis of depression . The PHQ-9[19] is initially filled by the patient, and then scored by the clinician. When the score is more than 10, it translates

[16] https://www.beyondblue.org.au/the-facts/anxiety-and-depression-checklist-k10
[17] http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4817.0.55.001Chapter92007-08
[18] https://bpac.org.nz/BPJ/2009/adultdep/assessment.aspx
[19] http://www.cqaimh.org/pdf/tool_phq9.pdf

to major depression with 88% of sensitivity and 88% specificity. More specifically, when the score is 5, 10, 15, and 20, respectively corresponds to mild, moderate, severe and severe depression[20] .

4.1.2 OpenFace toolkit

OpenFace is an OpenSource an application which provides a series of options for facial video processing by Baltrusaitis et al. [8]. The features extracted from a facial video recordings with the use of OpenFace include AUs prediction, 2D and 3D facial landmarks, Histogram of Oriented Gradients (HOG), yaw pitch and roll [68].Features provided by OpenFace where computed on the AVEC dataset, and stored for further processing in terms of depression assessment. More specifically, the total number of landmarks is 68: 51 are within the face, outlining facial features (mouth, eyebrows and nose), and the rest 17 landmarks are spotted on the face outline.

The proposed work exploits mainly the 2D facial landmarks, and thus the process for facial landmarks detection implemented in [8] is described briefly. OpenFace in order to achieve facial landmark detection and tracking, utilizes the Conditional Local Neural Fields (CLNF). CLNF, which is an instance of CLM, cope with the issues of feature detection in complex scenes [69]. More specifically, CLNF uses more advanced patch experts and an optimization algorithm [70]. The core components of CLNF are the Point Distribution Model (PDM) which captures landmark shape variations, and the patch experts that capture local appearance variations of each landmark. Fig. 4 illustrates the OpenFace pipeline.

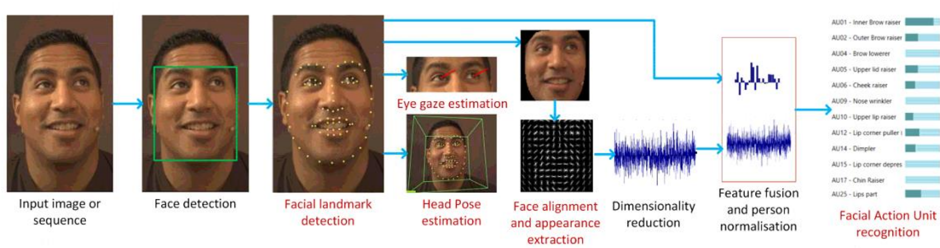


Figure 4: OpenFaze Pipeline (Taken from: [8]*).*

4.2 Software Description

Both the implementation and evaluation of the proposed methodology have been implemented on the same software platform. Concerning the implementation, MATLAB was used for building

[20]Kroenke K, Spitzer R, Williams W, The PHQ-9: Validity of a brief depression severity measure.JGIM,2001,16:606-616

functions and scripts needed for producing the necessary feature files. The produced files, "feature vectors", ('.mat') were the inputs to the same software platform, for the second part which refers to the classification part (see Chap.4.4 for more details.), which provides the final output of the methodology.

4.2.1 MATLAB®

MATLAB is a scientific programming language and numerical analysis environment. The name MATLAB[21] stands for matrix laboratory that was originally written by maintaining simple access to matrix software.

Utilization of MATLAB provides matrix calculations, creates user interfaces and data visualization, develops and executes algorithms, by giving a set of application-specific solutions within the tool-boxes, which is the core feature of MATLAB. Signal processing, neural networks, wavelets, simulation, fuzzy logic, control systems and more, are the fields where tool-boxes are available. There are various sets of MATLAB functions (m-files) in Toolboxes, which contribute in solving problems of a particular class, like those mentioned above. Since 2017, MATLAB appears to be used by over 2 million users across academia and industry[22].

4.3 Feature Extraction Methods & Algorithms

After reviewing the literature we concluded to four types of feature extraction methods applied to facial images: geometric, appearance, motion, and color based [71], [72], [73], [74]. Loconsole et al., [39], Palestra et al., [75], and Murino et al., [76] suggest two fundamental types of features: appearance and geometrical. In the proposed work, two specific categories of feature extraction have been employed: geometric, and the motion. The flow diagram of the framework implemented by the proposed work is illustrated by Fig.5.

[21] https://www.mathworks.com
[22] https://en.wikipedia.org/wiki/MATLAB

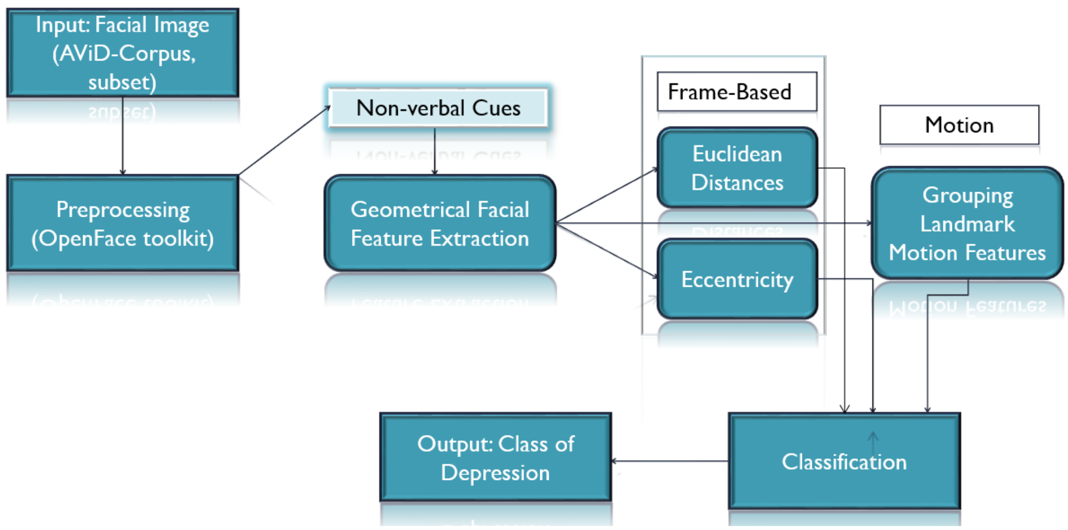


Figure 5: Flow Diagram of the proposed depression assessment framework.

4.3.1 Geometrical Feature Extraction Methodology

In the following sub-chapters we will explain in detail the methodology that relates to the first part of our work, i.e., the methods we used to carry out the facial feature extraction. The types of features that have been extracted include geometric (also referred to as shape), and motion (grouping specific features frame-based, and motion-based). The geometrical facial features proposed in this dissertation are: linear and eccentric.

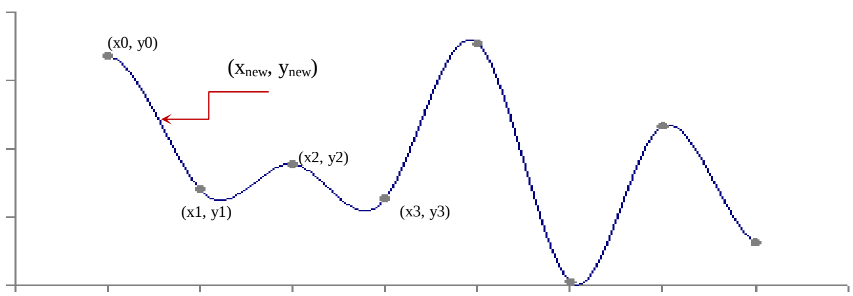
As far as it concerns the normalization of the linear features the following method has been evaluated. In order to take into account the diversity of individual anatomy, a scale of variability should be introduced. This is accomplished by normalizing each feature, with the coordinates of the landmark points horizontally at the height of the ears, according to the scale of each person respectively. However, the Eccentric features are by default normalized in the range [0, 1]. The normalization results are necessary in order to avoid the affect by people anthropometric traits dependencies to the extracted features. After normalizing the groups of landmarks we now need to extract the features we want to use as a feature vector for the next task, which is the classification part.

The geometric features were mainly implemented following the papers [36] and [39] respectively. In the case that the coordinates from particular landmarks were needed, which were not provided by OpenFace, Cubic Interpolation was performed, which is explained next. The basic goal of utilizing the specific linear features proposed below is to give a measure of the "relative movements between facial landmarks while expressing depression".

Cubic interpolation method

Before we proceed with the extraction of features, we noticed the need to create new landmarks, in order to get the extra facial features needed for the framework. This was achieved by utilizing the Cubic interpolation[23] method (Figure 6), which is a method that offers continuity between the segments. As such it requires more than just the two endpoints of the segment but also the two points on either side of them. So the function requires 4 points in all labelled y0, y1, y2, and y3. The. mu still behaves the same way for interpolating between the segment y1 to y2.

Figure 6: Explanation of Cubic Interpolation Method. (Taken from: [77])



This does raise issues for how to interpolate between the first and last segments. According to the researcher Paul Bourke: “A common solution is the dream up two extra points at the start and end of the sequence, the new points are created so that they have a slope equal to the slope of the start or end segment” [77]. See the Appendix chapter for the relevant code.

4.3.1.1 Frame-Based Features

After completing the bibliographic review, we resulted to the conclusion that certain facial expressions like distress and sadness are highly prevalent in depression. The loss of the feeling of joy was another finding. Consequently, appearances of less smiles as well as more sadness are apparent. Thus, in order to create the final feature vector we extracted features in respect to this type of manifestation of depression.

For the needs of the implementation of the proposed depression assessment framework, we needed to use a pattern, to evaluate our geometrical methodology on facial image analysis, in order to detect visual signs of depression through video recordings. Therefore, all considered distances were very carefully chosen. The frame-based features that were utilized can be partitioned into following groups shown in Figure 8, and Figure 9 in the case of Euclidean distances, and Figure

[23]http://paulbourke.net/miscellaneous/interpolation/

10, in terms of eccentricity method, respectively. The pattern that we used in order to extract our facial features is shown in Figure 7 to the right.

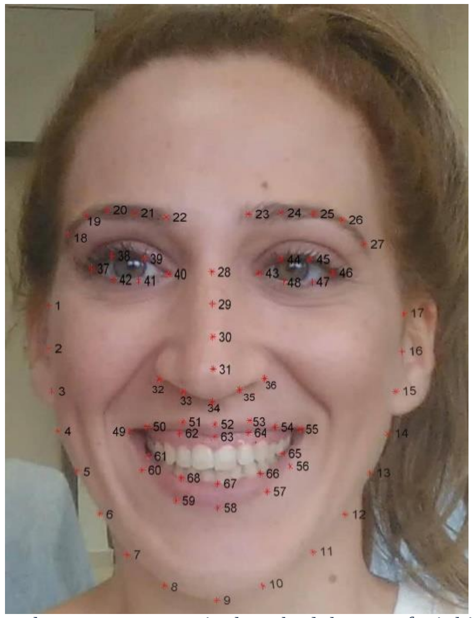


Figure7: Pattern used, to evaluate our geometrical methodology on facial image analysis, in order to detect visual signs of depression through video recordings.

Euclidean Distance Formula

In a Euclidean space, x,y, $\in \mathbb{R}$ ıe distance between points is given by:

Formula: $d = \sqrt{(x2 - x1)^2 + (y2 - y1)^2}$

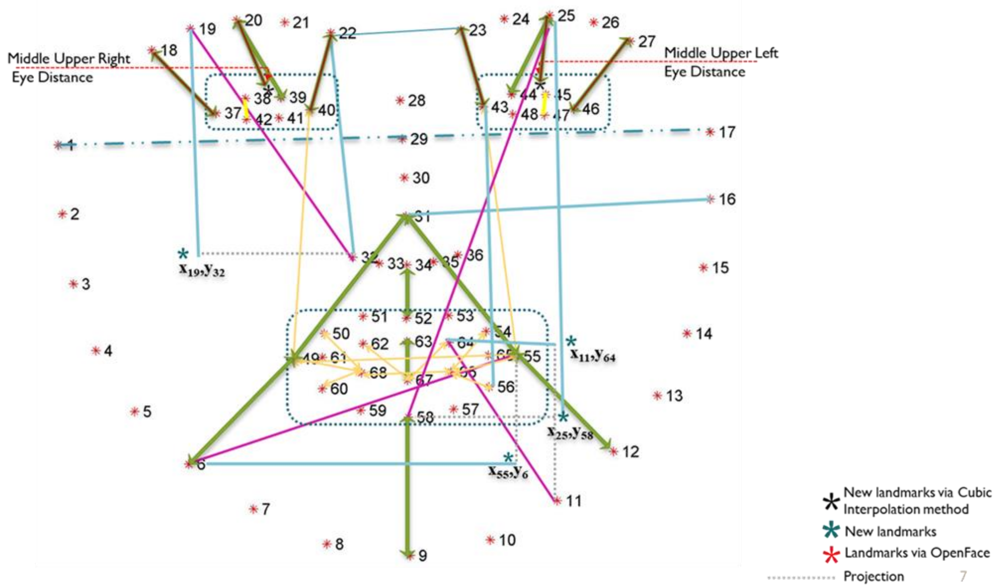


Figure 8: Frame-based features extracted based on Euclidean distance and Cubic Interpolation method.

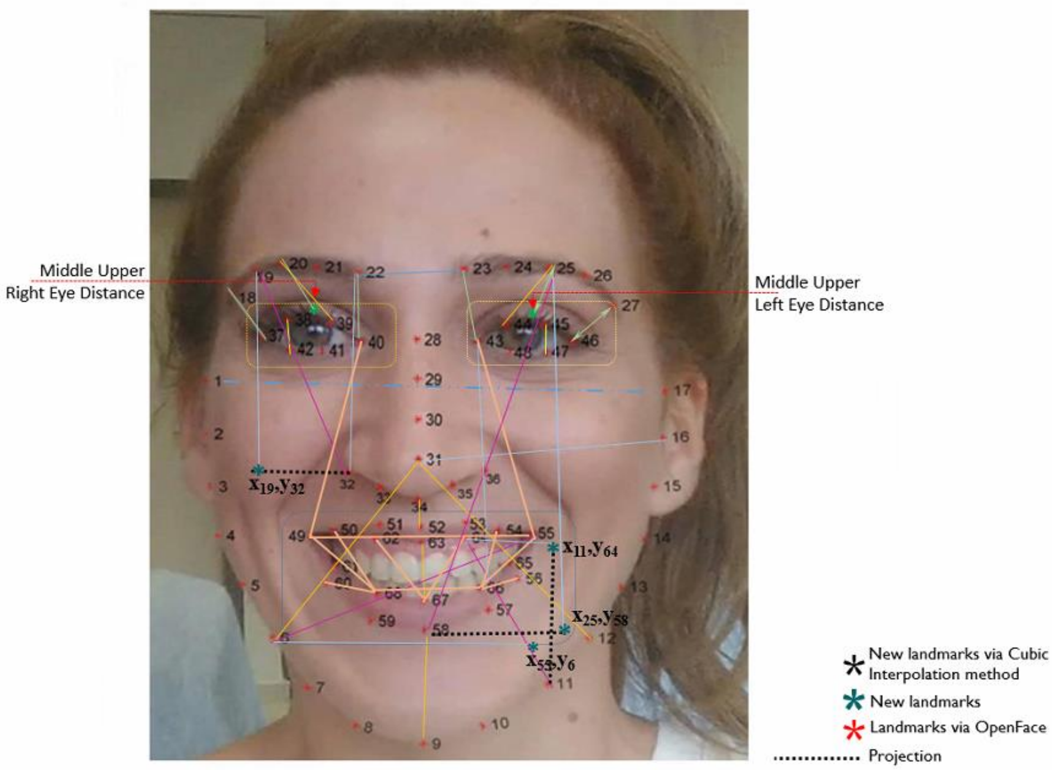


Figure 9: Mapping the Euclidean distances on real face to extract the required facial features for the depression assessment framework.

Moreover, we present the distances we examined in a form of grouping in correspondence to the landmarks we got from the OpenFace, based on the pattern of the Figure 7(c.f. Table 1, and Table 2).

Table 1: Total combinations of landmarks, given by OpenFace, related to facial features.

Number of features	**Facial features**	**Corresponding Landmarks' (OpenFace toolkit)**
1.	eyeliddistanceRight	*L38, L42*
2.	eyeliddistanceLeft	*L45, L47*
3.	eyebrowdistance	*L22, L23*
4.	Mouth_Eye_Right_dist	*L40, L49*
5.	Mouth_Eye_Left_dist	*L43, L55*
6.	Mouth_Nose_Right_dist	*L31, L49*
7.	Mouth_Nose_Left_dist	*L43, L55*
8.	Mouth_Chin_Right_dist	*L6, L 49*
9.	Mouth_Chin_Left_dist	*L12, L55*
10.	Eyebrow_Eye_Right_dist	*L20, L39*
11.	Eyebrow_Eye_Left_dist	*L25, L44*
12.	Mouth_height_dist	*L48-L68*
13.	LeftEye_height_dist	*L43-48*
14.	RightEye_height_dist	*L37-42)*
15.	Eyebrow_exter_Nose_dist	*L19, L32*
16.	Eyebrow_inter_Nose_dist	*L22, L32*
17.	Eye_inter_Mouth_dist	*L43, L56*
18.	Eyebrow_Mouth_dist	*L25, L58*
19.	Ear_Nose_dist	*L16, L31*
20.	Chin_Right_Mouth_dist	*L6, L55*
21.	Chin_Left_Mouth_dist	*L11, L64*
22.	Inner_hor_Mouth_dist	*L63, L67*
23.	Inner_ver_Mouth_dist	*L66, L68*
24.	Nose_upper_Mouth_dist	*L34, L52*
25.	Chin_Lower_Mouth_dist	*L9, L58*
26.	Inner_RighteyebrowEyedist	*L22, L40*
27.	Exter_RighteyebrowEyedist	*L18, L37*
28.	Inner_LefteyebrowEyedist	*L23, L43*
29.	Exter_LefteyebrowEyedist	*L27, L46*
30.	Midl_RighteyebrowEyedist	$L20$, New $L_{(MidlUpper\_Eye\_Right\_dist)}$
31.	Midl_LefteyebrowEyedist	$L25$, new$L_{(MidlUpper_Eye_Left_dist)}$

Table 2: Overall grouping of mouth distances corresponding to specific landmarks shown below.

Number of features extracted via mouth	Corresponding Landmarks' (OpenFace toolkit)
1.	*L55, L49*
2.	*L68, L50*
3.	*L66, L54*
4.	*L67, L62*
5.	*L67, L64*
6.	*L68, L49*
7.	*L68, L60*
8.	*L66, L56*

Eccentricity Methodology

In order to achieve the Extraction of the eccentricity Features, the following pattern has been utilized (Figure 10). In the case of an ellipse, the eccentricity is defined as the ratio of the in-between distance of the two foci, to the length of the major axis or equivalently.

Formula: $e = \frac{\sqrt{a^2+b^2}}{a}$

With the $a = \frac{B_{Mx}-A_{Mx}}{2}$, and $b = A_{My} - U_{m1y}$ respectively.

More precisely, Fig.10 illustrates the role of eccentricity and the way this method is implemented in the proposed work.

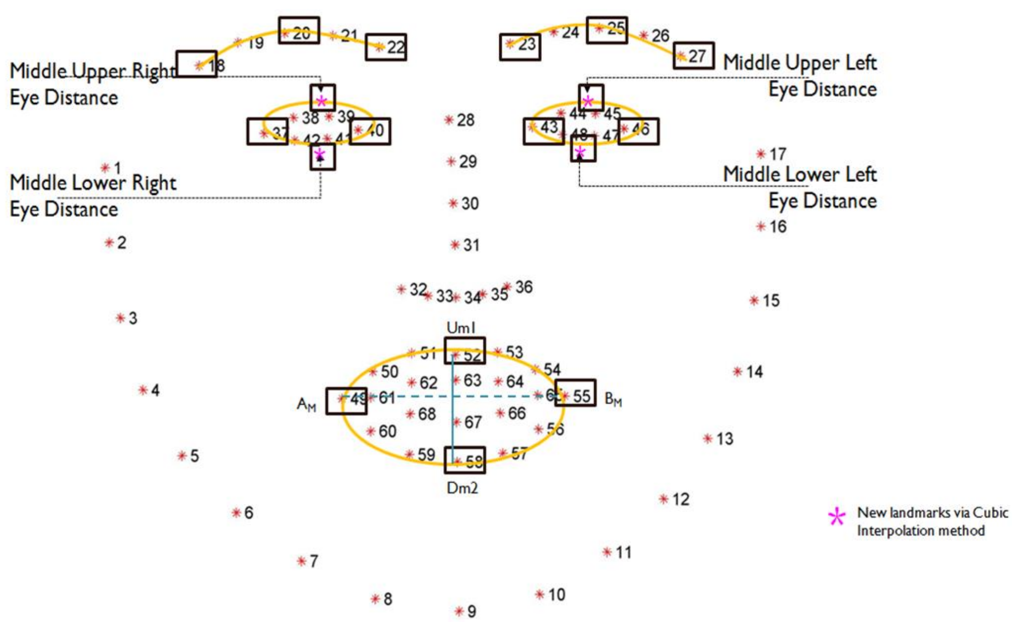


Figure 10: Eccentricity features extracted based on eccentricity formulas and Cubic Interpolation method

In addition, the facial features examined for the eccentricity method are presented in Table 3, corresponding to landmarks derived from OpenFace, based on the pattern of the Figure 9.

Table 3: Number of Eccentricity features extracted, along with their correspondence to the OpenFace landmarks.

Number of features	**Facial features**	**Corresponding Landmarks' (OpenFace toolkit)**
1.	Upper Eye Right	*L37, New $L_{(MidlUpper\_Eye\_Right\_dist)}$, L40*
2.	Lower Eye Right	*L37, New $L_{(MidlLower\_Eye\_Right\_dist)}$, L40*
3.	Upper Eye Left	*L43, New $L_{(MidlUpper\_Eye\_Left\_dist)}$, L46*
4.	Lower Eye Left	*L43, New$L_{(MidlLower_Eye_Left_dist)}$, L46*
5.	Upper Mouth	*L49, L52, L55*
6.	Lower Mouth	*L49, L58, L55*
7.	Eyebrow Right	*L18, L20, L22*
8.	Eyebrow Left	*L23, L25, L27*

After completing the extraction of frame-based features, time domain metrics were calculated for each one of these features. Below, in the subchapter 4.3.3, a full explanation of these time domain metrics is provided.

Note: All four new landmarks, both for the Euclidean distances and for the implementation of eccentricity method have emerged through the utilization of a function we constructed from scratch. This function implements the Cubic Interpolation method (see. Chap Appendix for the code), which is explained in a previous chapter.

4.3.2 Grouping Landmarks Features

It is desirable to group the landmarks by type or decision, i.e., according to the conclusions that we want to extract, in order to have a more efficient framework. The more custom landmarks that are set in our system, the more effective the usage will be. Thus, we implemented a more automated system, which aids to search more quickly, or to rearrange the data.

4.3.2.1 Motion Features

Our goal in this work was to divide the data into important parts for analysis to a greater extent. In Automatic recognition of emotions, and more specifically in the feature extraction, and in the proposed work those related to depression, the clustering of extracted features could refer to the segmentation of the desired objects outside those that do not provide extra information, i.e., the background. The regions of interest are the eyes, the eyebrows, the mouth, as well as the face as a whole. Besides, this segmentation emphasizes the extraction of features that will help in the optimization of the proposed application, so that it is more sensitive to the recognition of signs of depression.

Segmentation Technique – "Windowing"

The windowing corresponds to a method, which is applied to specific frames, so that it can easily be implemented in order to execute the available technique we designed for the construction of a feature set. To be more accurate an extended signal is multiplied with a window function of certain period (length), providing bound length weighted version of the primer signal[24].This technique multiplies each frame with a window function, which we define initially to improve the representation of frequency bands.

A reason for choosing the windowing technique is that is reduces the loss of energy, produced by the overlapping frame method [80]. Another reason for choosing the windowing method, as opposed to the overlapping approach is that it makes it possible to display and enhance the contrast in selected segments of the total range of pixel values. One more benefit of the window method, unlike the overlay approach, is that it provides the appearance and enhancement of contrast to specific portions of the total pixel range. This can be compared to the limitations of images shown on film where the full range of exposure is displayed in one image and cannot be changed. By selecting the windowing method it is easy to create many displayed images, each one "focusing on" a particular range of pixel values[25].

For our framework testing we extracted 12 feature vectors with 12 different windows. In other words we constructed a function, which is able through various windowing (i.e., in our case Window = {10, 15, 20, 25, 30, 35, 45, 50, 55, 60, 75, 80}) to conclude to the best combination in order which one is most optimal. At this point it should be given an explanation on how the number of frames translates into time. More specifically, since we have 30 frames per second, then for window 60 we mean that we have information with a 2 second window and so on.

A function namely: "landmarksmotion" was created from scratch, generating displacement, velocity and acceleration. The values produced where summed creating three time series concerning each of these regions. Time domain metrics (explained below) were also calculated in

[24]http://www.cs.tut.fi/kurssit/SGN-4010/ikkunointi_en.pdf

[25] http://www.sprawls.org/resources/DIGPROCESS/module.htm

the same way as the frame-based features. Below the total amount resulting from the addition of corresponding combinations of specific landmarks is given, and is also illustrated in Figure 11.

1. *Total facial movement*

   Sum of distance, velocity, and acceleration corresponding between landmarks 1 and 68

2. *Right Eye movement*

   Sum of distance, velocity and acceleration corresponding between landmarks 37 and 40

3. *Left Eye movement*

   Sum distance, velocity and acceleration corresponding between landmarks 43 and 48

4. *Overall moth movement*

   Sum distance, velocity and acceleration corresponding between landmarks 49 and 68

5. *Movement of new projected landmarks*

   Sum distance, velocity and acceleration corresponding between landmarks:

   i. x19,y32 and 19
   ii. X55,y6 and 6
   iii. X25,y58 and 58
   iv. X11,y64 and 64

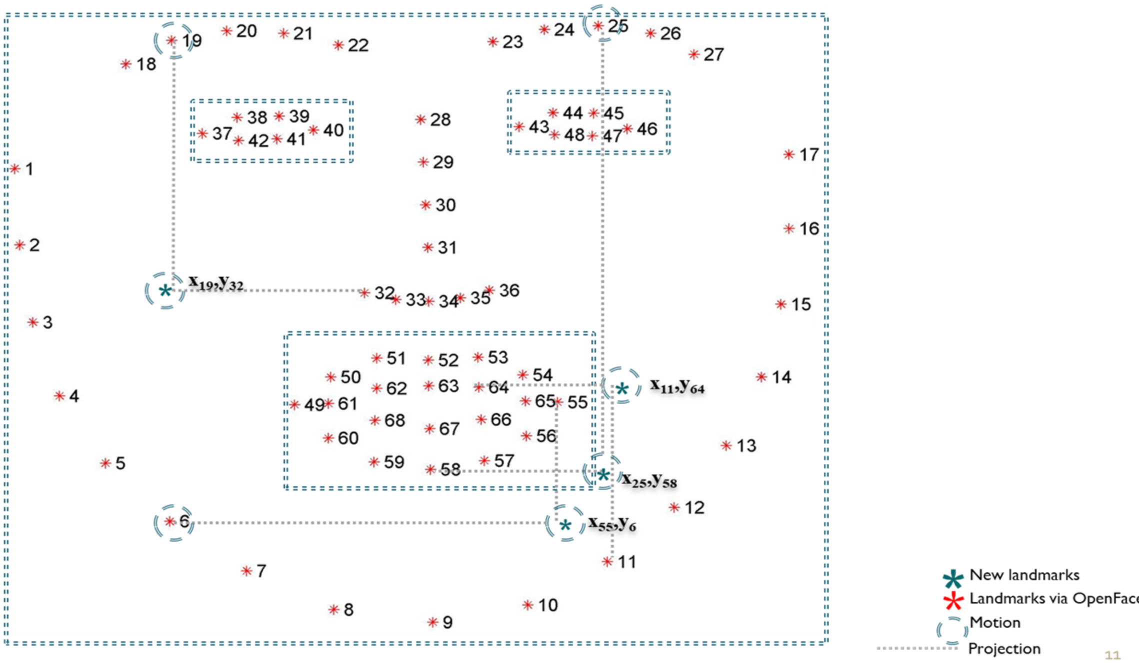


Figure 11: Total grouped features in terms of motion, i.e., distance, velocity, and acceleration

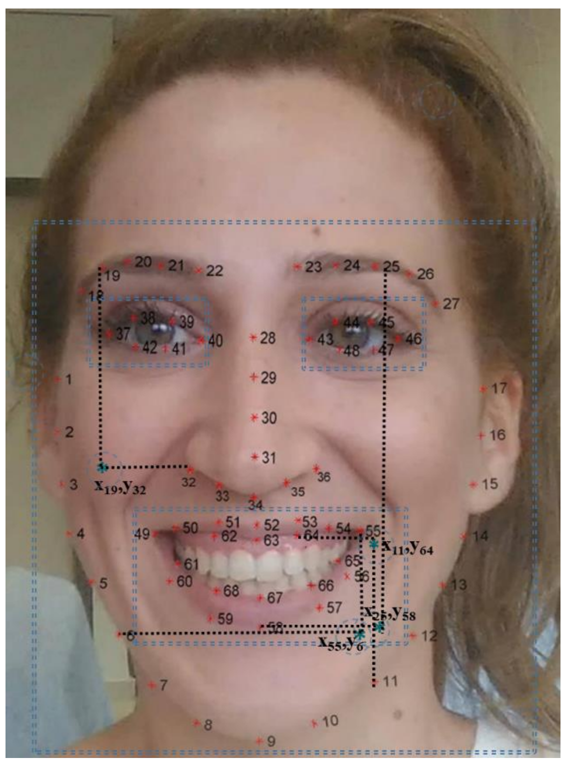


Figure 12: Representation of grouped landmarks on a facial image, in terms of motion-based features

4.3.3 Statistics in Terms of Time Series

Essentially, Signal Processing is a field that analyzes analog and digital signals, corresponding to variables that vary depending the "epoch", through birth of time series. Consequently, the Time Series are a collection of data vectors ordered over time, which are distinguished from randomly sampled data. The ultimate purpose of the time series is to represent the time-evolution of a dynamic population or process. Linear sorting of time series gives them a distinct place in data analysis, with a specialized set of methods[26].Furthermore, time series analysis concerns areas such as: Identifying patterns, Modeling patterns, and/or Forecasting values. More precisely, in Biomedicine time series that represent physiological signals (EEG), heart-rate, and patient temperature could play a crucial role to draw optimal outcomes.

In line with Karagiannis et al. [78], "Biomedical signal measurement, parameter identification and characterization initiate by the acquisition of diagnostic data in the form of image or time series that carry valuable information related to underlying physical processes. *The aforementioned researcher supplements: "*The last stage of this a typical measurement system refers to digital signal analysis with a higher level of sophistication techniques that extract features out of the digital signal or make a pattern recognition and classification in order to deliver useful diagnostic information".

According to Vigo et al. [79], when we execute the plot command on time domain signals, "we obtain a time-amplitude representation of the signal". Another manifestation is the following: "In time domain analysis, the interval between adjacent normal R-waves is measured over the period of recording", given by Yadav et al. [80]. It is a fact, therefore, that a variety of statistical variables can be calculated from these intervals, depending on the conclusions that are needed to be drawn. A signal could be measured like a function of time, as time-domain signals are represented in their raw format. Hence, when we decide to plot a specific signal, what we get is that on one of the axes is time and on the other is usually the amplitude [81]. In case of our implementation, the plot of a signal correspond to the following, one of the axes is the time and the other is the variables amplitude of a specific facial feature. In the next subchapter we demonstrate in detail the statistic variables that we examined in case of our implementation.

[26]https://www.mathworks.com/help/matlab/data_analysis/introduction.html

Syntax

Time Series is a collection of observations x_t, each one being recorded at time t. The time of the corresponding values could be discrete, t = 1,2,3,…,n or continuous t > 0. To be more accurate this method evaluates compression of the data so as to provide compact description of the data[27].

Formula: Time Series =$[\{\{t_1,v_1\},\{t_2,v_2\}\ldots\}]$,

The above syntax represents a time series specified by time-value pairs $\{t_i,v_i\}$[28].

Utilization of Inbuilt Matlab Mathematical Functions

In case of already inbuilt functions of Matlab we have also performed a list of statistical[29] measurements, in order to create the final feature vector for the classification part. All those formulas correspond to each facial feature. The name of the function that contains all those mathematical functions/formulas, which corresponds to the time domain features in terms of signal processing is: "`timeSeriesStatistics()`" and the statistics are listed in *Table 4* below .

Table 4: Utilization of inbuilt statistics of Matlab for the Facial Features of the framework.

Inbuilt Statistics in case of features	
1. mean(): mean value of the signal.	2. var(): square of the standard deviation
3. median():median signal value.	4. std(): mean deviation of the data sample compared to the average.
5. mode():value that appears most frequently in the signal.	6. skewness(): difference of delay or time between two rising edges of two signals.
7. max():largest value of value of the signal	8. kurtosis():peakedness's degree distribution toward the mean of the signal.
9. min():minimum value of value of the signal	10. entropy(): a scalar value representing the entropy of the signal.
11. mad():mean absolute deviation of the values in the signal	12. iqr():returns the interquartile range of the signal
13. range(): difference between the largest and smallest values of the signal	14. corrcoef(): a matrix of correlation coefficients for the signal, where the columns of the signal represent random variables and the rows represent observations.
15. meanfreq():mean normalized frequency, of the power spectrum of the signal	

Implemented Functions

In the case of implemented functions from scratch, the need to implement a number of other functions that would contribute to integration of the Geometric Facial Feature Extraction emerged.

[27]http://www.stat.columbia.edu/~rdavis/lectures/Session6.pdf
[28]http://reference.wolfram.com/language/ref/TimeSeries.html
[29]https://www.mathworks.com/help/matlab/descriptive-statistics.html

In line with Acharjee et al. [82], we are going to explain the *"energy"*, and *"bandpower"* functions (see *Table 5*) that were implemented in our framework, respectively.

Table 5: Implemented Time domain statistics of each one of the facial features

Implemented Statistics in case of facial features	
1. energy(): energy measure, sum of the squares of the absolute magnitude of the signal.	2. bandpower(): average power of the input signal.

The formulas for *energy* (1), and *bandpower* (2) developed according to [82] are given below.

Mathematical Formula for the *energy ()*:$\varepsilon_x = \sum_1^n |x[n]|^2$ (1)

and for the *bandpower()*:$\varepsilon_x = \sum_{i=t1}^{i=t2} |[x_i]|^2$ (2), respectively.

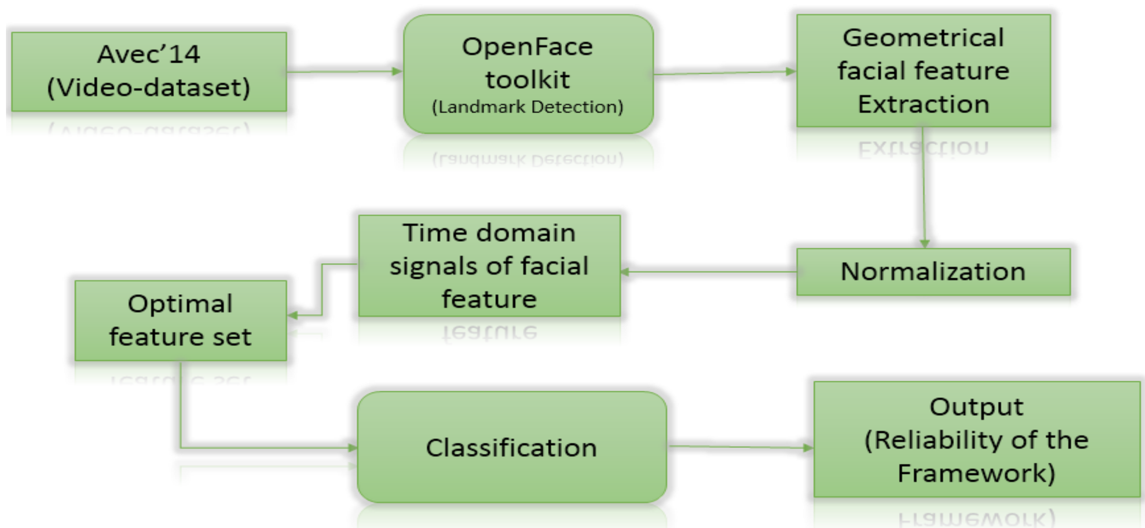


Figure 13: Diagram of the implemented system, in order to evaluate the depression assessment framework.

Below are some examples of time series, based on specific facial features, extracted in terms of time-domain signals, which are related to particular facial states of a depressed person. Namely, less eye blinking, not smiling, and veraguth fold state, which is a facial feature attributed to melancholy and depression, shown in Figures 14, 15, and 16 respectively.

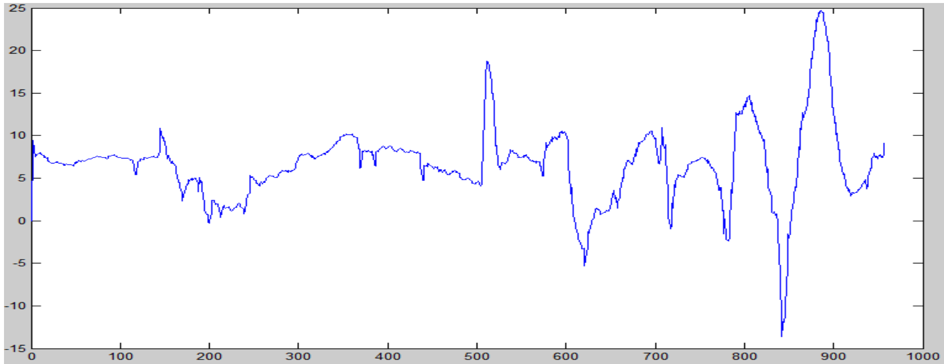


Figure 14: Time domain signal of eye-blinking

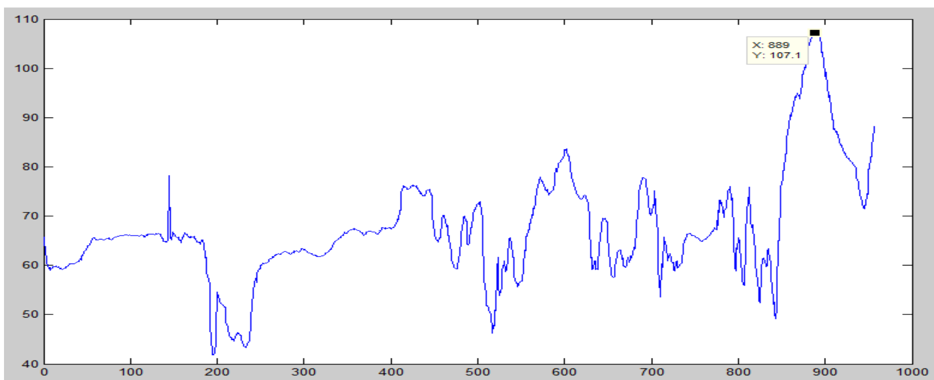


Figure 15: Time domain signal of mouth corners

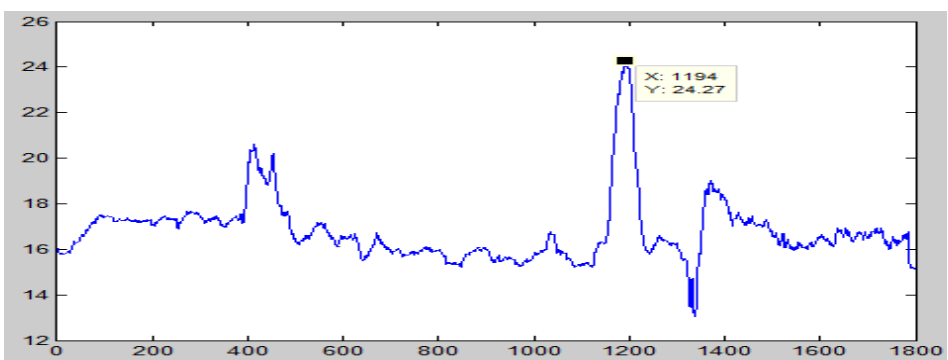


Figure 16: Time domain signal if frowns. Distance between inner eyebrows

Thus, the Final Feature Vector arose from a linear combination of the Frame-Based Features, Eccentricity Features, Landmarks Motion (Figure 16), from which we have calculated the aforementioned Time domain statistics for each one of the above facial features. Thus, we ended up with a feature vector of 200 samples with 1819 features for each individual (size: 200x1819).

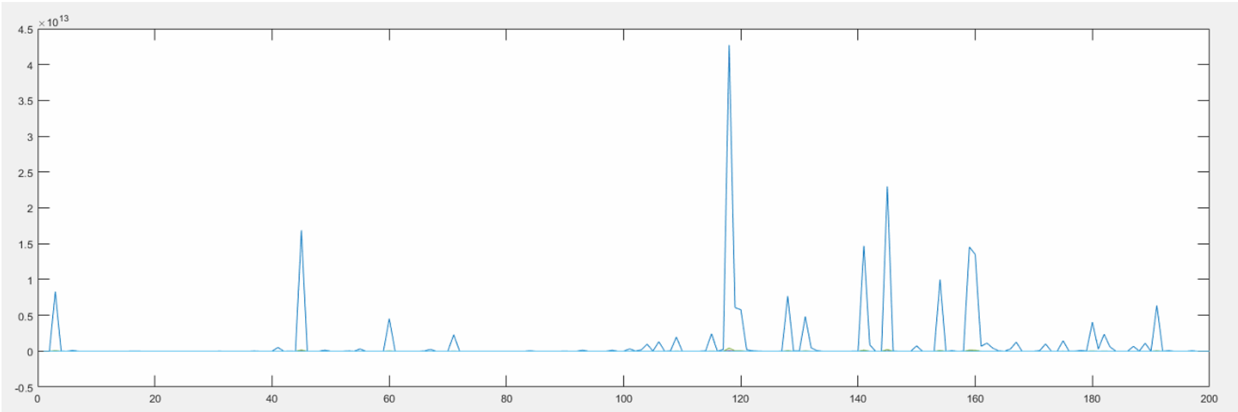


Figure 17: Indicative final feature vector imprinted as emerged with a window of 10.
The final feature vector consists of 200 samples, and 1819 features for each individual (i.e., size: 200 x 1819).

4.3.4 Dimensionality Reduction Technique

After exploring the state-of-the-art on dimensionality reduction techniques currently available, we have concluded to the following types i.e., Principal Component Analysis, Missing Values Ratio, Low Variance Filter, High Correlation Filter, Backward Feature Elimination, and Forward Feature Construction [83]. Another common method is the discrete cosine transformation (DCT). However this method is said to be more complex and time consuming [84], which are main drawbacks for the performance of applications. Thus, this work has centered on the Principal Component Analysis (PCA) method, which is commonly employed in image processing, for dimensionality reduction in an efficient manner.

PCA (Principal Component Analysis) on extracted Facial Features

The key issue of choosing a method or a technique to execute is: availability, accuracy, time-limitations and, process-speed. So, the PCA is selected because it is really the simplest and easiest method to implement, with extremely fast computation time. PCA (Principal Component Analysis) is an algorithm that tries to construct a computational model by taking into account the absolute most relevant information that best describes the best feature set, which is based on a linear transformation of the initial features. Since we have calculated the time domain statistics (see previous sub-chapter) of each extracted facial feature, denoted below, we then apply the PCA. The goal is to reduce the dimension of the original feature vector, and to introduce the reduced one to the classifier, for concluding to the final results.

Definition of PCA

According to [85] the definition of PCA is: "Principal Component Analysis, or PCA, is a statistical method commonly used to reduce the number of variables in a dataset. It does so by lumping highly correlated variables together". The main advantage of utilizing the Principal component analysis (PCA) is that this technique emphasizes variation and brings out strong patterns in a dataset. Moreover, it is often preferred in order to make data easy to explore and visualize. In other words, PCA (Principal Component Analysis) is useful for reducing the dimensions of a feature set [86].

Process Flow of PCA

We proceed to reduction of the dimensions of a set of n-dimensional data (original feature vector), displaying it as an m-dimensional subspace (m <n), because we need to increase computational efficiency while ensuring that there will be no loss of information. A brief description is given below of the steps that Principal component analysis (PCA) technique executes in order to identify the final patterns of reduced dataset.

- Determining data (via extracting the facial features).
    - Decide the number of data, e.g., $x_1, x_2, x_3, ..., x_f$, Vx1 vectors
        - In our case, we worked on a 200 x 1819 array, where V=200 is the number of samples based on the video recordings from AVEC'14 and f=1819, the number of features for each sample.
- Exporting the Eigenvalues and the Eigenvectors from correlation or covariance matrix, or execution of Singular Vector Decomposition.
- Classify the eigenvalues in descending order and select the equals "m" equivalent to the larger eigenvalues "m". The "m" is the number of dimensions of the new subset of attributes ($m \leq n$).
- Develop the projection matrix 'J' through the chosen "m" eigenvectors.
- Transforming the primal matrix 'P' of the dataset through 'J', in order to dominate the m-dimensional subset of attributes i.e., the matrix 'F'.

Summarizing, the PCA technique, provided by a statistical algorithm, alters a set of correlated variables to a set of non-correlated linear recombination of those variables. These are named principal components. So, the principal components,'F', is a linear recombination of the variables in our data set, 'P'.

4.4 Classification

Related areas[30]of implemented classifiers observed, appear to be the medical assessment, and the treatment effectiveness analysis. It is based on examining the attributes of a new object (unclassified), which is based on these attributes, and is assigned to a predetermined set of classes. The categorization process consists of assigning each record to some of the predefined categories. The objects to be categorized are generally represented by the database entries. As we mentioned earlier, since high-level information was extracted by the facial features (see subchapter "Evaluation of Facial Feature Extraction"), the proper machine learning algorithms were selected, in order to examine the sensitivity and specificity of this proposed framework. This was achieved by testing a series of algorithms like: Discriminant Analysis, Random Forest Tree, Naïve Bayes, Linear and k-nearest Neighbor Classifiers. More details about the classifiers are given bellow.

4.4.1 Machine Learning Methods

Machine learning [87] is a data science technique that makes computers capable of using existing data in order to forecast future behaviors, and outcomes. Using machine learning, computers learn without being explicitly programmed. Machine learning is considered a subcategory of Artificial Intelligence (AI). Forecasts or predictions from machine learning can make apps and devices more efficient.

Unsupervised Learning

Unsupervised Learning provides non-labeled input data and, does not have prior knowledge of the result. A model is prepared by deducing structures present in the input data. This may be to extract general rules. It may be through a mathematical process to systematically reduce redundancy, or it may be to organize data by similarity. These problems are solved by clustering, dimensionality reduction and association rule learning techniques. Example algorithms include the Apriori algorithm and k-Means.

Semi-Supervised Learning

While Unsupervised Learning provides non-labeled input data, in the case of Semi-Supervised Learning, the data given is a combination of labeled and non-labeled data. Another difference concerning the previous method is that though there is a desired prediction query that must be

[30]http://mmlab.ceid.upatras.gr/courses/data_mining/, Course: 'Data Mining and Learning Algorithms', Department of Computer Engineering & Informatics University of Patras, Greece.

answered, the corresponding pattern must learn the structures in order to sort the data, and predict results. These problems are solved by classification and regression techniques. Essentially, these algorithms make assumptions on how to model the non-labeled data and then draw conclusions.

Supervised Learning

As stated above in Semi-Supervised Learning, the data given is a combination of labeled / non-labeled data, however the Supervised Learning has only labeled input data, which are called training data, and know result each time. This method has the training process, which is required to make predictions, and correct the pattern when those predictions are incorrect. The operation of training process is paused when the model achieves a desired level of accuracy on the training data. These problems are solved by classification and regression algorithms. These algorithms include Logistic Regression and the Back Propagation Neural Network.

Note: It is important to mention that a check had to be made for the correct correlation of the labeled values given by the AVEV’14 dataset concerning the participants' depression assessment, which resulted from the extensively used BDI questionnaire, in relation to the implementation of the algorithm we implemented, which is associated with the facial features extraction. To be more accurate, each participant has its own unique label that corresponds to the value of depression. So, when the algorithm is executed, in order to extract the desired facial features from each of the 200 videos, the classification has to be executed in the same order, because in all other cases we will get the wrong results. The crosscheck conducted through the inbuilt functions provided by Matlab, known as ‘*isequal’, and ‘ismember’*. So, after confirming that there was a proper match then we proceeded to the next stage, which is the classification part.

4.4.1.1 Selected Classifiers

In the current sub-chapter we list the key features and methods for testing the five selected classifiers respectively, in order to complete the experimental part, with the final goal to detect the most suitable one for the depression assessment framework. Descriptions of commands that apply the below classifiers have been obtained from MathWorks.

Discriminant Analysis

In the case where the problem requires to determine which variables are distinguished between two or more occurring groups, then this solution is implemented through the use of the method, known as the Discriminant function analysis[31].

For instance, such an issue could emerge for a cytologist who could initially record diverse attributes of comparable kinds of cells (gatherings/groups), and after that perform a discriminant function analysis, to decide the arrangement of qualities that takes into account the best segregation between the sorts/groups. An additional illustration could be α medicinal specialist who may record diverse factors identifying with patients' backgrounds to realize which factors best anticipate whether a patient is probably going to recoup totally (first gathering), mostly (second gathering), not in the slightest degree (third gathering).

Description

The discriminant analysis[32,33] model is described as follows:

General method: *mdl = fitcdiscr(X,Y).*

Expanded method: *mdl = fitcdiscr(___,Name,Value).*

Each class (Y) generates data (X) utilizing a multivariate normal distribution. Essentially this means that the model assumes X has a Gaussian mixture distribution.

- The linear discriminant analysis model has the same covariance matrix for each class; only the means vary.
- In contrast, the quadratic discriminant analysis model presents means and covariances variations of each class.

In line to this modeling assumption, *"fitcdiscr"* infers the mean and covariance parameters of each class.

- Linear discriminant analysis, computes the sample mean of each class. Then it computes the sample covariance by first subtracting the sample mean of each class from the observations of that class, and taking the empirical covariance matrix of the result.

[31] http://www.statsoft.com/Textbook/Discriminant-Function-Analysis
[32] https://www.mathworks.com/help/stats/discriminant-analysis.html
[33] https://www.mathworks.com/help/stats/creating-discriminant-analysis-model.html

- Quadratic discriminant analysis computes the sample mean of each class. Then it computes the sample covariances by first subtracting the sample mean of each class from the observations of that class, and taking the empirical covariance matrix of each class.

In order to minimize the expected classification cost the command '*predict*' is utilized via the following method.

General Formula[34]: $\hat{y} = arg_{y=1,\dots K} \min \sum_{k=1}^{K} \hat{P}(\mathrm{k}|\mathrm{x})\mathrm{C}(\mathrm{y}|\mathrm{k}),$

With:

- $\hat{y}$ represents the predicted classification
- K defines the number of classes.
- $\hat{P}$(k|x) relates to the posterior probability of class k for observation x.
- C(y|k) is the cost of classifying an observation as y when its true class is k.

The fit method does not use prior probabilities or costs for fitting. In the case of the proposed work we have dealt with the second method, i.e., the quadratic discriminant analysis model. Besides, this classifier was the one that proved to provide the best performance, in terms of F1-score, for our application. In other words, it was the Discriminant Analysis classifier that resulted to an F1-score equal to 72.57% for our depression assessment framework.

Below the basic idea of the description of the other four classifiers is given, as well as their syntax that we have used until we have the optimal desired result.

K-nearest neighbors (KNN)

A nearest-neighbor[35] classification object, where both distance metric ("nearest") and number of neighbors can be modified. The object classifies new observations utilizing the predict method. The object contains the data used for training, so can compute resubstitution predictions. The k-nearest-neighbor model is described below.

Description

Returns a classification model based on the input variables X and output (response) Y.

General method: *mdl = fitcknn(X,Y).*

Expanded method: *mdl = fitcknn(___,Name,Value).*

[34] https://www.mathworks.com/help/stats/classificationdiscriminant-class.html
[35] https://www.mathworks.com/help/stats/classificationknn-class.html?requestedDomain=true

Random Forest tree

Description

Fit binary classification decision tree for multiclass classification. The random forest tree[36]model is described below:

I. Returns a fitted binary classification decision tree based on the input variables contained in matrix X and output Y. The returned binary tree splits branching nodes based on the values of a column of X.

General method*: tree = fitctree(X,Y)*

II. Fits a tree with additional options specified by one or more name-value pair arguments, using any of the previous syntaxes. For example, you can specify the algorithm used to find the best split on a categorical predictor, grow a cross-validated tree, or hold out a fraction of the input data for validation.

Expanded method: *tree = fitctree(___,Name,Value)*

Naïve Bayes

Description

The present classifier trains multiclass naive Bayes model. The Naïve Bayes[37] model is described as follows.

I. Returns a multiclass naive Bayes model (Mdl), trained by predictors X and class labels Y.

General method*: mdl = fitcnb(X,Y).*

II. Returns a naive Bayes classifier with additional options specified by one or more Name,Value pair arguments, using any of the previous syntaxes. For example, you can specify a distribution to model the data, prior probabilities for the classes, or the kernel smoothing window bandwidth.

Expanded method: *mdl = fitcnb(___,Name,Value).*

Linear - Support Vector Machine

Description

Fit linear classification model *"fitclinear"* trains linear classification models for two-class (binary) learning with high-dimensional, full or sparse predictor data. Available linear classification models include regularized support vector machines (SVM) and logistic regression models. The fitclinear

[36]https://www.mathworks.com/help/stats/fitctree.html?s_tid=doc_ta
[37]https://www.mathworks.com/help/stats/fitcnb.html?s_tid=doc_ta

minimizes the objective function using techniques that reduce computing time (e.g., stochastic gradient descent). A high-dimensional data set includes many predictor variables. Although such a data set can consume a significant fraction of memory, it must fit in the MATLAB® Workspace. For low- through medium-dimensional predictor data sets, see Alternatives for Lower-Dimensional Data.

The Linear - Support Vector Machine[38] model is described below.

I. Returns a trained linear classification model object that contains the results of fitting a binary support vector machine to the predictors X and class labels Y.

General method*: mdl= fitclinear(X,Y)*

II. Returns a trained linear classification model with additional options specified by one or more Name,Value pair arguments. For example, you can specify that the columns of the predictor matrix correspond to observations, implement logistic regression, or specify to cross-validate. It is good practice to cross-validate using the KfoldName,Value pair argument. The cross-validation results determine how well the model generalizes.

Expanded method: *mdl = fitclinear(X,Y,Name,Value).*

This particular classifier proved to be much faster than the other studied algorithms, whilst also delivering better results, i.e. F1-score =75, 86%. To be more accurate, by training the *fitclinear* with the following Values: i.e., Learner = svm, Prior = uniform, and Solver = lbfgs it emerged that we get same F1-score, despite the different values of PCA, and windowing (see: Segmentation technique, explained above). This could be an indication of either the existence of overfitting or the fact that it is a truly robust algorithm. In determining which of the two circumstances we are against at, it will require additional research and experimental work.

[38]https://www.mathworks.com/help/stats/fitclinear.html

Chapter 5 - EXPERIMENTAL RESULTS/METRICS

5.1 Evaluation Method& Parameters

This section presents the evaluation methods, and the measurements that emerged during the development of the second part of the framework evaluation, i.e., the classification. Thus, in the present chapter we will analyze the methods and parameters used in each of the five classifiers, which we mentioned in Chapter 4, so that we can perform the evaluation for the proposed depression assessment system, based on the facial features from the AVEC'14 dataset participants.

5.2 Cross-Validation Method

The fundamental thought of the cross-validation method is the accompanying. A portion of the information (data) is expelled before the training starts. At that point when the training is accomplished, the data that was excluded could be utilized to test the performance of the learned model on the "new" information. Along these lines, the point in the cross-validation[39] method is to guarantee that every case from the first dataset has a similar shot of showing up in the preparation and testing set.

Basic protocols:

i. N-fold cross-validation: partition the data into pieces and train times, treating an alternative piece as the holdout set each time.
ii. Leave-one-out validation: as the n-fold cross-validation method, except that pieces comprise a single datapoint.

In the case of this work we dealt with a variation of the second method. Thus, below is a brief definition of this technique.

Definition of Leave -One-Subject-Out Method

Leave-one-out cross-validation is a special case of cross-validation where the number of folds equals the number of instances in the data set. Thus, the selected learning algorithm is applied once for each one of the 200 subjects. In other words we always exclude all instances of the same Subject, i.e. this one that will be tested, using all other subjects as a training set and using the selected subject as a single-item test set [88].

[39] http://users.sussex.ac.uk/~christ/crs/ml/lec03a.html

5.3 Analysis Measures of Performance

Since the part of the implementation of the feature extraction has been completed, we have moved on to the second part of the classification, which refers to the examination of the accuracy and the sensitivity of the proposed system. In order to draw the necessary conclusions, we proceeded to the evaluation by calculating the following measures.

Terminology of statistical measurements by Confusion Matrix

In the field of machine learning and specifically the problem of statistical classification, a Confusion Matrix[40], (or Error Matrix), is a specific table layout that allows visualization of the performance of an algorithm [89]. Typically in a supervised learning method we make use of the previous mentioned matrix, while a Matching Matrix is utilized in unsupervised learning method [90]. That is the reason we have chosen the Confusion Matrix. More precisely the confusion matrix is a tool in predictive analytics and it is composed of a 2D table. Each row represents the instances in an actual class, while each column represents the instances in a predicted class. The confusion matrix provides a variations of measures by depicting the number of true positives, true negatives, false positives and false negatives. Once we decided the final model of each implemented classifier, and proceeded to the desired parameterizations, for each one of them, we examined the following statistical measurements:

Accuracy: the number of instances that have been correctly classified.

$$\textbf{Accuracy} = \frac{(\mathrm{TP} + \mathrm{TN})}{(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})}$$

Sensitivity (additionally True Positive Rate or Recall): the proportion of positives that are correctly identified as positives.

$$\textbf{Sensitivity} = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FN})}$$

Specificity (additionally True Negative Rate): the proportion of negatives that are correctly identified as negatives.

$$\textbf{Specificity} = \frac{\mathrm{TN}}{(\mathrm{TN} + \mathrm{FP})}$$

Precision (additionally Positive Predictive Value): the degree of the predicted positive values.

$$\textbf{Precision} = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FP})}$$

[40]https://www.mathworks.com/help/stats/confusionmat.html

F-score (additionally F1score or F-measure): qualifies the harmonic mean of precision and sensitivity.

$$\textbf{F1-score} = 2 * \frac{\mathrm{TP}}{\frac{1}{\mathrm{Recall}} + \frac{1}{\mathrm{Precision}}} = 2 * \frac{\mathrm{Precision} * \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}$$

Explanation of variables in relation to the statistical measurements.

<u>True Positives (TP)</u>

Acceptable outcome about confirming the actual classification, when the predicted result is acceptable to be in a particular category, which corresponds to the actual label.

<u>False Positives (FP)</u> or miss or false reject

Incorrectly classified outcome (misclassified), when the predicted result is discarded from a particular class, where the actual label is in this class.

<u>True Negatives (TN)</u>

Another acceptable outcome, when the predicted result is discarded from a particular class, where the actual label does not corresponds not in that class.

<u>False Negatives (FN)</u> or false alarm or false acceptance

Another incorrectly classified outcome (misclassified), when the predicted result is acceptable to be in a particular class, where the actual label does not corresponds in that class.

The different performance metrics obtained allow more detailed analysis than by taking into account only the correct guesses (accuracy). Exclusive accuracy measurement is not reliable for the actual performance of a classifier, as it would produce misleading results in the case of an unbalanced dataset (i.e., when the number of samples in different categories varies). That is the main reason we proceeded to the above measures.

5.4 Results of Proposed Evaluation

The decision about which was ultimately the best classifier for our framework arrived after the completion of a large set of experiments with different classifiers. We initially trained each of the five classifiers with the corresponding parameters found to give their best results. So we have reached the following conclusions, which are given schematically in the following tables. We then identify the impact of the results by altering the features extraction set, also the changes in the size and impact of the segmentation technique known as "windowing".

Comparison of the Classifiers

As it was mentioned earlier, in order to be able to find out the best result for our framework in terms of accuracy, we performed several tests with the five different classifiers. Namely they are the following: Linear, Discriminant Analysis, Random Forest tree, k-nearest neighbor and Naive Bayes. At the first stage, our goal was to find the maximum f1-score for each of the five classifiers, which emerged depending on various windows independently of the PCA. The findings that derived via specific parameterization (Table 6) are shown in the following table (Table 7.) and diagram (Figure 15.), respectively.

Table 6: Parameters that were the more optimal results of each classifier.

Classifiers	**Parameters**
Linear	Learner = [{'svm'}]; Prior=[{'uniform'}];Solver = [{'lbfgs'}]
Discriminant Analysis	discrimantypes=[{'pseudoQuadratic'}],prior= [{'empirical'}], scoreTransf=[{'doublelogit'}]
Random Forest-tree	predictSelect= [{'allsplits'}], scoreTransf=[{'doublelogit'}], AlgorithmForCategorical=[{'Exact'}], MergeLeaves=[{'off'}], pior= [{'uniform'}]
k-nearest Neighbor	numberofneighbours= [14], distancetypes=[{'seuclidean'}]
Naïve Bayes:	distribNames=[{'kernel'}],pior=[{'empirical'}],scoreTransf=[{'invlogit'};],Kernel=[{'nornal'}]

Table 7: Total Maximum F1-score variance, of the five classifiers tested, with non-fixed PCA, and different windowing.

Window/ Classifiers	10		15		20		25		30		35		50		60		75		80	
	PCA	f1-score	PCA	f1-score	PCA	f1-score	PCA	f1-score	PCA	f1-score	PCA	f1-score	PCA	f1-score	PCA	f1-score	PCA	f1-score	PCA	f1-score
Linear	10	63,45	10	63,45	10	63,45	10	**75,86**	10	64,86	10	62,02	130	10,71	10	63,45	10	63,45	10	62,98
Discriminant Analysis	150,0	68,83	110,0	72,20	105,00	**72,57**	105,0	68,53	130,0	67,91	110,0	70,50	110,0	69,77	120,0	68,62	110,0	71,02	105,0	70,78
Random Forest tree	60	61,81	40	54,81	90	56,7	50	**64,39**	90	57,73	10	52,53	140	55,34	10	59,96	50	61,17	80	56,72
k-nearest neighbor	0	0	60	51,34	60	48,91	50	51,31	10	50	30	52,41	50	49,74	0	0	20	50,27	30	**54,84**
Naïve Bayes	105	45,45	120	56,16	105	53,73	150	58,49	140	54,73	130	**58,72**	130	57,14	105	52,46	105	53,63	105	44,05

Note: In case of the k-nearest Neighbor classifier the following observations emerged. The reason that the comparison was chosen to be done with a number of neighbors equal to the value of 14 is due to the fact that the knn provides the maximum f1-score with the particular combination, as shown in the above diagram. Moreover, in case of window=10, PCA selected feature =105, and numberofneighbours=6 the f1-score was f1=35, 22. Also, for the combination of window=60, PCA selected feature =105, and numberofneighbours=6 the f1-score was f1=41, 21 (see Table 8).

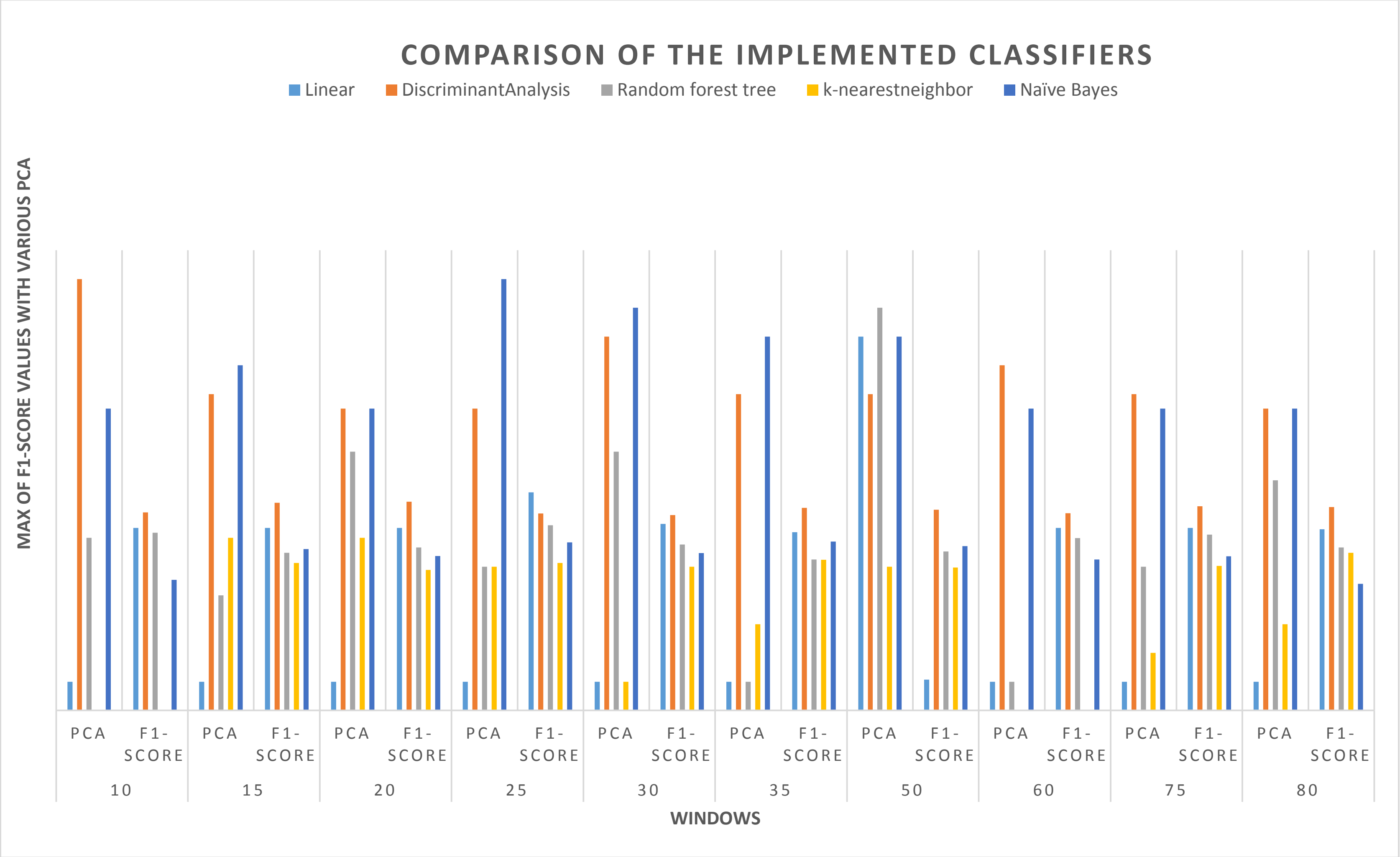


Figure 18: Comparison in between the five classifiers, tested on the framework. Namely: Linear, Discriminant Analysis, Random Forest tree, k-nearest Neighbor and Naive Bayes with non-fixed PCA.

The graph above (Figure 17) represents the maximum value achieved by all the classifiers in respect to the f1-score. The first stage control was carried out independently to the PCA coeffs, also with different windows, i.e. the various vectors we derived from the first part of the implementations, which relates to the export of facial features. By observing the diagram one can see the maximum F1-score of each classifier with the corresponding PCA, grouped with color combinations per window.

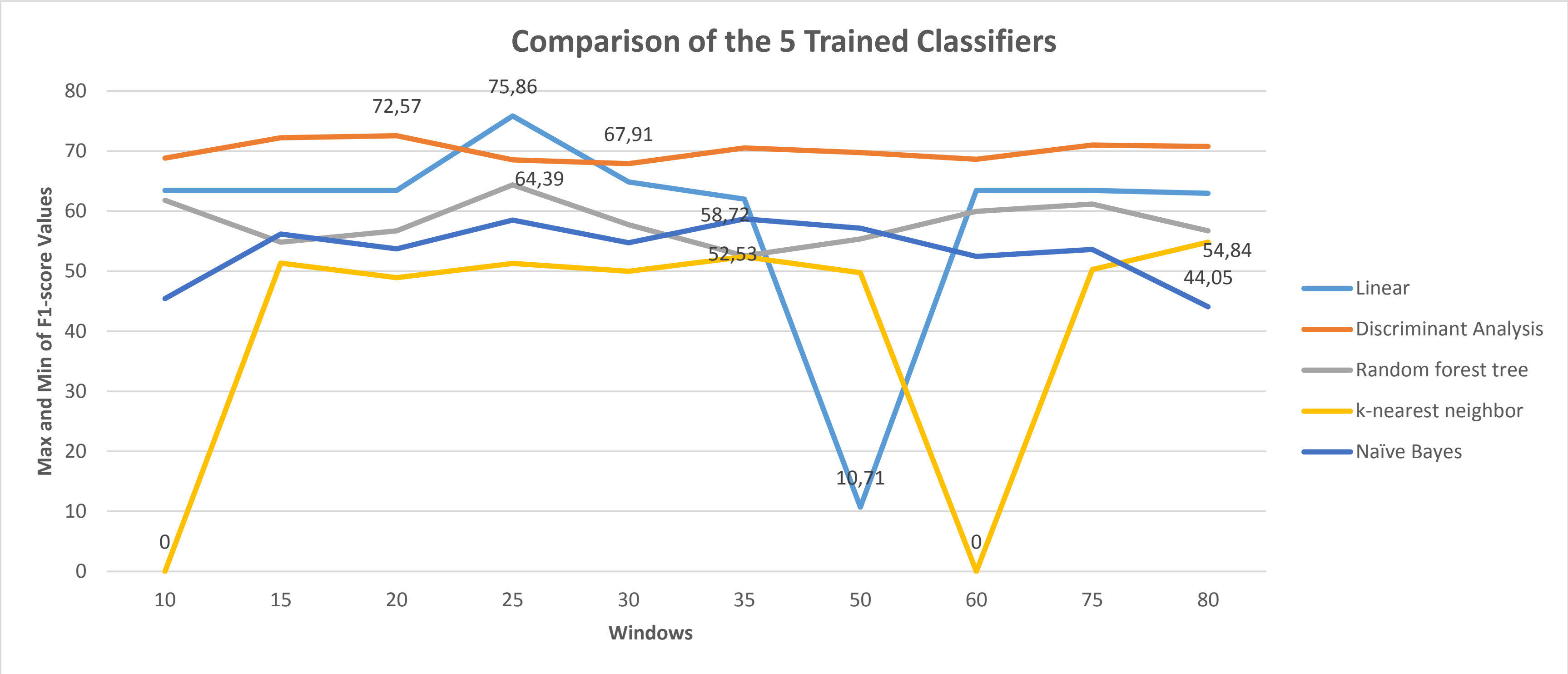


Figure 19: Comparison between the five classifiers, tested on the framework, regardless of the PCA. Results of the f1-score fluctuation of the:: Linear, Discriminant Analysis, Random Forest tree, k-nearest neighbor and Naive Bayes

Due to the fact that we observed a large drop in the value of the f1-score (Figure 18), we decided to test two new vectors with the aim of seeing even more intuitively the values corresponding to the f1-scores. For this reason, we focused on checking what happened before and after window 50, by training the classifiers with windows 45 and 55 respectively. In addition, it should be noted that this test was made with a constant PCA since the highest score was derived with (PCA*) Selected Feature = 105* based on the Discriminant Analysis classifier. This is confirmed by both the previous table and diagram and those that follow.

The table (see Table 8) below depicts the overall fluctuation of the F1 score, as demonstrated by each one of the five classifiers that were trained, with the PCA (Principal Component Analysis) fixed at 105 (the first 105 more efficient features of the *Final Sumfeature Vector*, which occurred from the facial feature extraction part) and for different windows (10-80). This has been done because we derived on our best results in case of the 105 coefficients, which emerged via the PCA technique.

Table 8: Total F1-score variance, of the five classifiers tested, with fixed PCA-105 for different windowing (10-80).

	10	15	20	25	30	35	45	50	55	60	75	80
Linear	*63,45*	*63,45*	*63,45*	***75,86***	*64,86*	*62,06*	*72,25*	*10,71*	*0*	*63,45*	*63,45*	*62,98*
Discriminant Analysis	*63,68*	*69,37*	***72,57***	*68,53*	*67,80*	*69,02*	*65,60*	*68,77*	*60,16*	*62,13*	*67,19*	*70,78*
Random Forest tree	***59,11***	*49,48*	*54,82*	*57,45*	*54,84*	*42,33*	*50,26*	*50,72*	*49,2*	*44,57*	*47,83*	*54,36*
k-nearest neighbor	*35,22*	*36,94*	*31,33*	***40,68***	*31,37*	*45,78*	*24,29*	*34,54*	*39,74*	*33,96*	*38,36*	*34,97*
Naïve Bayes	*45,45*	*54,82*	*53,73*	*53,52*	*44,09*	***57,14***	*44,2*	*50,49*	*57,14*	*52,46*	*53,63*	*44,05*

The performance of Linear classifier might seem quite better, though for the reasons explained in a previous subchapter, we propose the second classifier i.e., the Discriminant analysis. The measurements that are shown below (Figure 19) refer to the overall fluctuation of the F1-score of the five classifiers trained, in terms of different windows as shown in the above panel (10-80) with fixed PCA, Selected feature = 105.

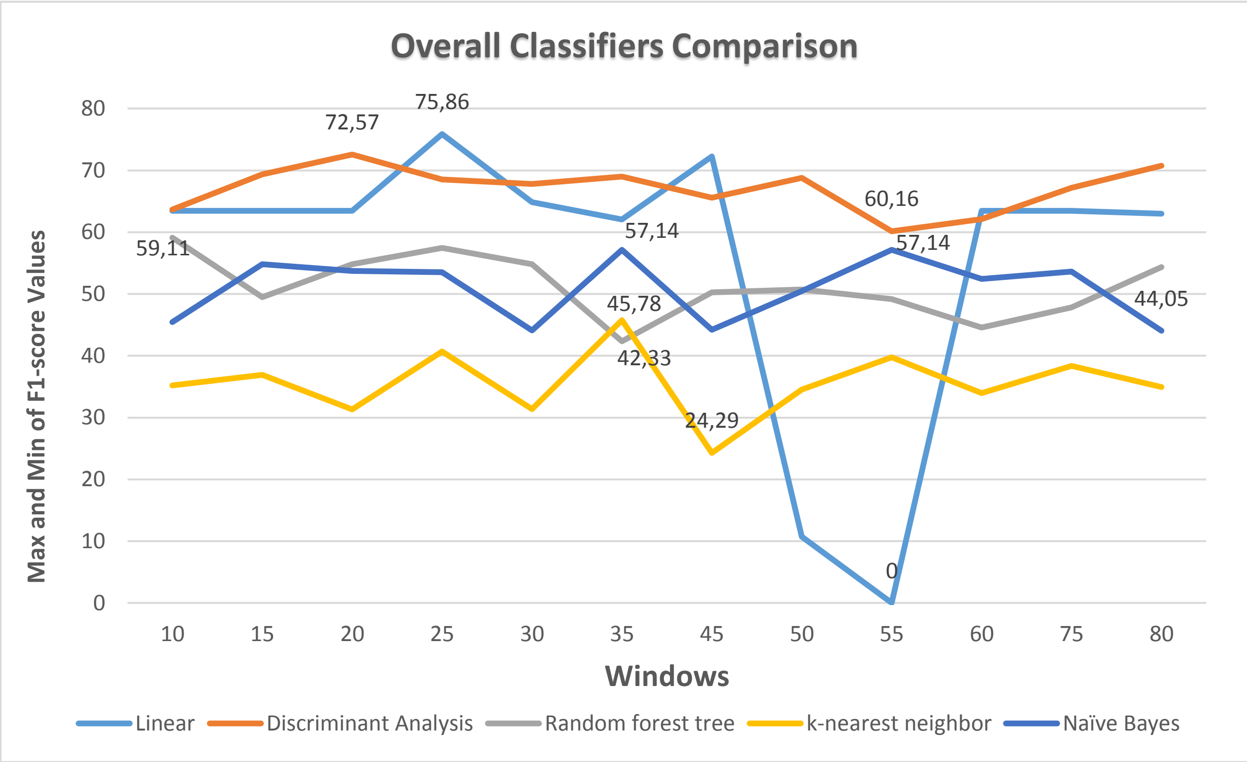


Figure 20: Comparison in between the five classifiers, tested on the framework. Namely: Linear, Discriminant Analysis, Random Forest tree, k-nearest neighbor and Naive Bayes

Discriminant Analysis Classifier

As it emerged by a series of experiments that preceded, i.e., by both various combinations of parameters and different windowing, the Discriminant Analysis classifier seemed to prevail over the rest of the classifiers, and to generate the best results in terms of sensitivity and accuracy. As referred previously, the input arguments that utilized these best results namely are: *discriminant types* = *pseudoQuadratic*, *prior* = *empirical*,*scoreTransf* = *doublelogit* respectively. Below a more detailed description of these arguments is given[41].

- ✓ Discriminant type, specified as the comma-separated pair composing of 'DiscrimType' and a character vector in this table. The 'pseudoquadratic' value is portrayed as QDA (Quadratic discriminant analysis).This parameter provides a Predictor Covariance Treatment with the subsequent characteristics. The covariance matrices may vary among categories. The software modifies the covariance matrix utilizing the pseudo inverse.

[41] https://www.mathworks.com/help/stats/fitcdiscr.html

- ✓ Prior probabilities for each class, determined as the comma-separated pair consisting of 'Prior' and a value in this table. The 'empirical' value implies that the class prior probabilities are the class relative frequencies in Y.
- ✓ Score transformation, indicated as the comma-separated pair comprising of 'ScoreTransform', and either a character vector or a function handle. The 'doublelogit' value is defined through the equation: 1/(1 + e–2x)

The corresponding tables and diagrams confirming this fact are listed in *Table 9* and *Figure 20* below.

Table 9: Best results of framework with discriminant analysis classifier.

	10	15	20	25	30	35	45	50	55	60	75	80
PCA-100	38,60	60,09	47,76	56,39	53,47	49,06	47,93	60,34	60,16	37,89	59,03	56,19
PCA-105	63,68	69,37	72,57	68,53	67,80	69,02	65,60	68,77	67,42	62,13	67,19	70,78
PCA-110	5,79	72,20	65,69	64,00	63,57	68,75	62,60	69,77	64,93	64,79	71,02	63,28

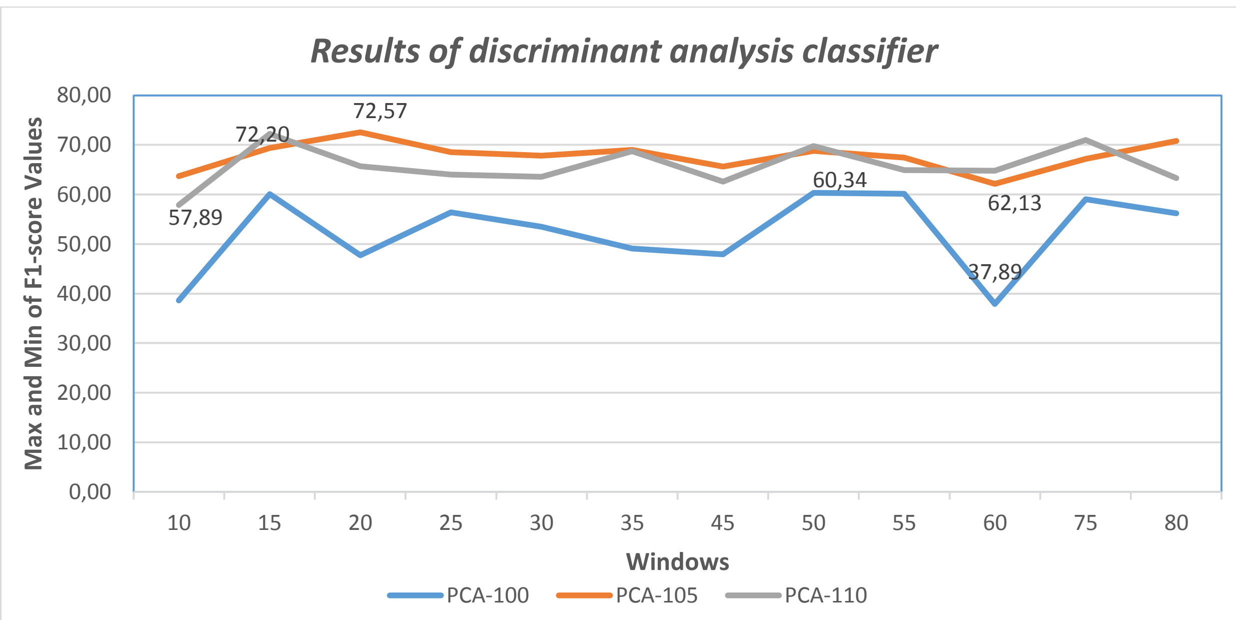
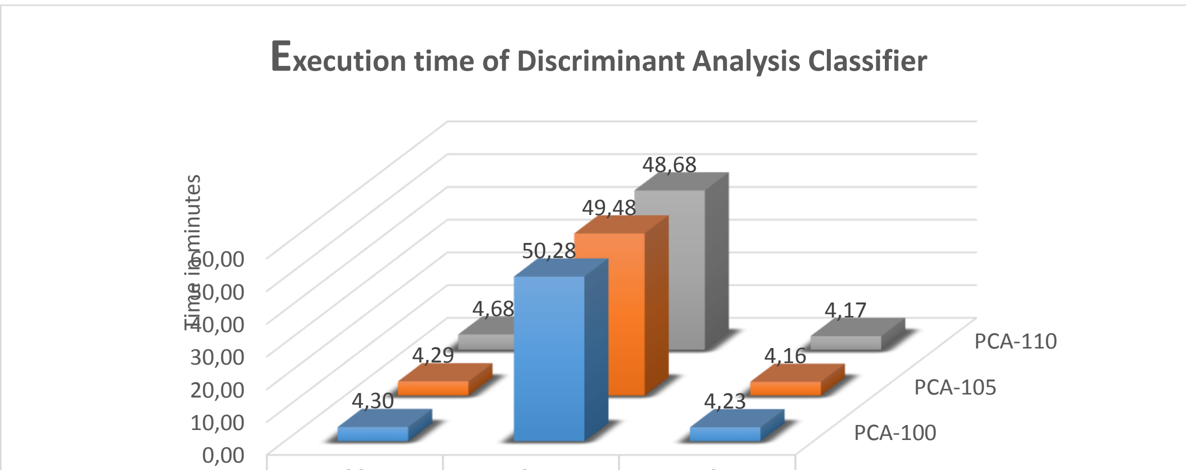


Figure 21: Diagram of the fluctuation of the F1-score depending specific principal components, i.e., the PCA-100, PCA-105, and PCA110. Max and Min values of F1-score about Discriminant Analysis classifier.

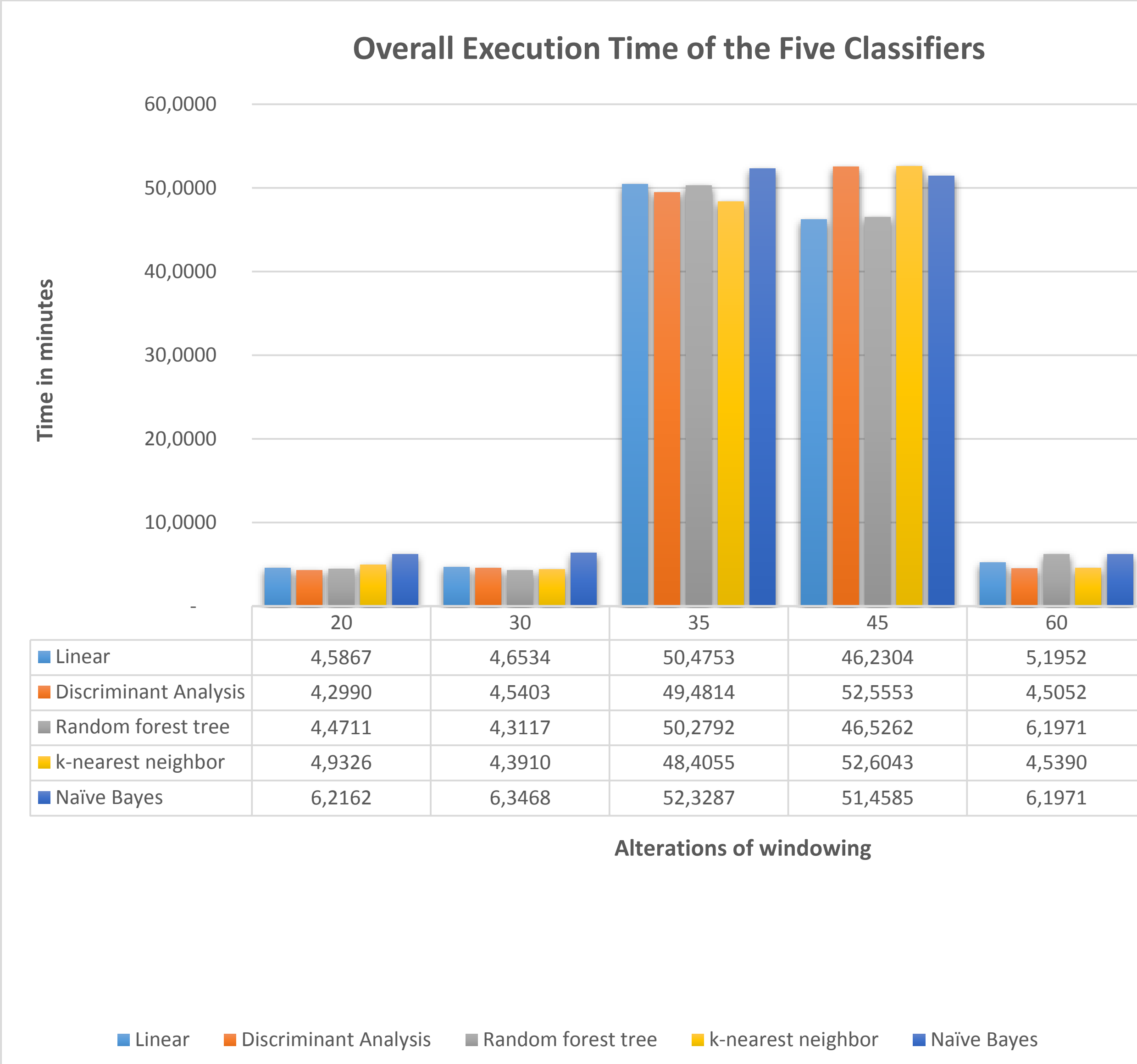
Yet another issue that we dealt with in this dissertation, was the execution time of each one of the five classifiers. The reason we faced this task is because it would be inappropriate to train a classifier either with a slow execution time, but with satisfactory results, either fast but with mild results. Our final goal is to detect the optimal classifier, both in terms of execution time, as well as the reliability of the results. This can easily be understood for the following reasons. Firstly it would be unsuitable for any tool, i.e., in our case for the assessment framework of depression, to expect long time to carry out the final results even if these results are satisfactorily. While it is prohibitive to build a tool that indeed will be fast but with very low rates of accuracy. This is because, although there will be low energy consumption, such a tool could not be considered useful as it will lead to the wrong conclusions. Therefore the reason in dealing with this task was because it is desirable to make such optimizations that will provide a quick speed and reliable tool when extracting results. It is worth mentioning that all the experiments were tested on a 64-bit Operating System consisting of a processor Intel core i5-2500 CPU on 3.30GHz, and an installed 10,0GB RAM. Below is the relevant chart (Figure 21) that respond to the concerns we just mentioned.



	20	35	50
PCA-100	4,30	50,28	4,23
PCA-105	4,29	49,48	4,16
PCA-110	4,68	48,68	4,17

Figure 22: Execution time of the proposed implemented framework. Discriminant Analysis classification with different values of PCA, i.e., principal component. The execution time is reflected in minutes.

Finally, we made a total estimation of the execution time regarding all the classifiers that were trained with a fixed value for the PCA, with a Selected feature =105, and those parameters that gave the maximum results for the f1-score, for each one of them respectively.



	20	30	35	45	60
Linear	4,5867	4,6534	50,4753	46,2304	5,1952
Discriminant Analysis	4,2990	4,5403	49,4814	52,5553	4,5052
Random forest tree	4,4711	4,3117	50,2792	46,5262	6,1971
k-nearest neighbor	4,9326	4,3910	48,4055	52,6043	4,5390
Naïve Bayes	6,2162	6,3468	52,3287	51,4585	6,1971

Figure 23: Execution time of all machine learning algorithms trained with fixed PCA value, i.e., the coeffs=105.

The increase of the time concerning the windows 35 and 35 respectively, depicted in the diagram above (Figure 23) it is likely due to the phenomenon of periodicity [91]. Periodicity scanning is an obstacle that has addressed a lot of attention for analyzing purely periodic signals especially in the case that it is desirable to consider time series. Our goal was to examine whether the Discriminant Analysis classifier, who derived the best results both in terms of correct assessment of depression, but is also the most optimal in the case of execution time. The findings we have reached indeed agree that the Discriminant Analysis classifier was the top choice in contrast to the rest machine learning algorithms, *Fig.22* given above, both with the graphs and with the time table of each classifier, confirms the previous statement. With an exception of the Naïve Bayes being slower, the rest seems to be very close to execution times, however, the percentages of f1-score are quite weaker relative to the Discriminant Analysis classifier, so this lead us to the conclusion that the latter is the most suitable for our methodology.

Chapter 6 - CONCLUSION&DISCUSION

6.1 Summary

Facial expression is significant in human interaction and communication since it contains critical information regarding to emotion analysis. In summary, the primary purpose of this dissertation was to design and develop an application in MATLAB for facial image analysis, with the ultimate goal to detect visual signs of depression through video recordings. High level information about the facial features was extracted with the use of OpenFace toolkit, while special focus has been given to the AViD-Corpus (AVEC'14) for testing the developed algorithms. Subsequently, since high-level information was extracted by the facial features, the proper machine learning algorithms were selected, in order to examine the sensitivity and specificity of this proposed framework. Several classification algorithms were tested, namely: Discriminant Analysis, Random Forest Tree, Naïve Bayes, Linear and k-nearest Neighbor Classifiers. The best performing method involved the Discriminant Analysis classifier, with an execution time of approximately 4 minutes. Additionally, by selecting an approximately 1 second-window, with Leave-One Subject-Out cross-validation an accuracy of 72.57% was achieved for our depression assessment framework.

6.2 Future Work

It would be desirable if the proposed system would be complemented by speech-based features, in order to have multimodal depression assessment, as well as if it is tested on the data of AVEC'16 in order to obtain comparative performance. Another expansion could be to consider the detection of even more efficient combinations of Feature Extraction. The method of Interpolation of x,y coordinates in-between lost frames, in order to eliminate the possibility of lost information is another thought. Also the possibility of different classification parameters, could by chance give even better results. Another expansion might be testing the framework on other datasets. Detection of a more sufficient geometrical facial feature vector, related to emotions like fear [92], [93], in order to obtain the optimal utilization of the framework is another question to be answered.

Bibliography

[1] E. COMMISSION, "Green Paper:Improving the mental health of the population:Towards a strategy on mental health for the," Brussels, 14.10.2005.

[2] H. Berenbaum and F. T. Oltmann, "Emotional experience and expression in schizophrenia and depression.," *Journal of Abnormal Psychiatry Research,* vol. 101, no. 1, pp. 37-44, 1992.

[3] M. Mandal, R. Pandey and A. Prasad, "Facial expressions of emotions and schizophrenia:a review.," *Schizophrenia Bulletin,* vol. 24, no. 3, pp. 399-412, 1998.

[4] European Commission, "The State of Mental Health," European Commission, Health & Consumer Protection Directorate-General, 2004.

[5] M. Valstar, B. Schuller, K. Smith, T. Almaev, F. Eyben, J. Krajewski, R. Cowie and M. Pantic, "AVEC 2014 –3D Dimensional Affect and Depression Recognition Challenge," in *AVEC '14 Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge*, Orlando, Florida, USA, November 07 - 07, 2014.

[6] P. Ekman, Emotion in the Human Face: Guide-Lines for Research and an Integration of Findings, Pergamon, 1972, p. 191.

[7] P. Ekman, J. C. Hager and H. Oster, Emotion in the Human Face, United States: MALOR BOOKS, 2013, p. 439.

[8] T. Baltrusaitis, P. Robinson and L. P. Morency, "OpenFace: an open source facial behavior analysis toolkit," 2016.

[9] P. Mladovsky, S. Allin, C. Masseria, . C. Hernández-Quevedo, D. McDaid and E. Mossialos, "Health in the European Union Trends and analysis," European Observatory on Health Systems and Policies, Observatory Studies Series No 19, United Kingdom..

[10] R. Dimitrovska and B. Dejanova, "Diagnostic challenges for major depression in early breast cancer survivors," *Europian Neuropsychopharmacology,* vol. 26, no. 2, p. 401, 2016.

[11] C. E. Merten, J. C. Cwik, J. Margraf and S. Schneider, "Overdiagnosis of mental disorders in children and adolescents (in developed countries)," *Child Adolesc Psychiatry Ment Health,* vol. 5, no. 11, 2017.

[12] "The Role of Nonverbal Communication in Preschool Classroom Interactions, [Thesis]," Oslo and Akershus University College of Applied Sciences, Dublin, Malta, 15th August, 2012.

[13] A. Stuhrmann, . T. Suslow and . U. Dannlowski, "Facial emotion processing in major depression: a systematic review of neuroimaging findings," *Biol Mood Anxiety Disord,* vol. 1, no. 10, 2011.

[14] . J. Joormann and H. I. Gotlib , "Selective attention to emotional faces following recovery from depression," *Abnorm Psychol,* vol. 116, no. 1, pp. 80-5, 2007.

[15] L. A. Bouhuys , E. Geerts and C. M. Gordijn , "Depressed patients' perceptions of facial emotions in depressed and remitted states are associated with relapse: a longitudinal study," *J Nerv Ment Dis.* , vol. 187, no. 10, pp. 595-602, 1999 .

[16] H. Ellgring, Non-verbal Communication in Depression, Cambridge: Cambridge University Press, 2007.

[17] M. Valstar, B. Schuller, K. Smith, F. Eyeben, B. Jiang, S. Bilakhia, S. Schnieder, R. Cowie and M. Pantic, "AVEC 2013: The Continuous Audio/Visual Emotion and Depression Recognition Challenge, 3rd ACM," in *ACM-Multimedia,* Barcelona, Spain, 2013.

[18] A. P. Association, Diagnostic and Statistical Manual of Mental Disorders , Fifth Edition, Arlington, VA: Association, American Psychiatric, 2013.

[19] B. Kaplan, S. Lee and E. Leibenluft, "National Institutes of Health (NIH)," Department of Health and Human Services, [Online]. Available: https://www.nimh.nih.gov/health/statistics/index.shtml. [Accessed 3 October 2017].

[20] P. Oakes, M. Loukas, R. J. Oskouian and S. Tubbs, "The Neutroanatomy of Depression: A Review," *Clinical Anatomy,* vol. 30, no. 1, pp. 44-49, 2016.

[21] R. C. Kessler, P. Berglund, O. Demler , R. Jin , D. Koretz , . K. R. Merikangas, A. J. Rush , E. E. Walters and P. S. Wang , "The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R)," *National Comorbidity Survey Replication,* vol. 289, no. 23, pp. 3095-3105, 2003 Jun 18.

[22] Y. I. Sheline, M. . H. Gado and H. C. Kraemer, "Untreated depression and hippocampal volume loss," *The American Journal of Psychiatry,* vol. 160, no. 8, pp. 1516-1518, August 2003.

[23] M. Marcus, M. T. Yasamy, M. v. Ommeren, D. Chisholm and S. Saxena, "Depression: a global public health concern," *WHO Department of Mental Health and Substance Abuse,* p. 247, 2012.

[24] W. H. Organization, " Depression," WHO, February 2017. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs369/en/. [Accessed 9 October 2017].

[25] W. H. Organization, The ICD-10 Classification of Mental and Behavioural Disorders: "Blue Book“ (Clinical descriptions and diagnostic guidelines) und "Green Book“ (Diagnostic criteria for research), WHO.

[26] T. L. Schwenk, L. B. Terrell, R. V. Harrison, A. L. Temper, M. A. Valenstein and J. R. Bostwick, "UMHS Depression Guidline Update," University of Michigan, Faculty Group Practice, Quality management Program, Michigan, 2011.

[27] K. Iyer and A. Z. Khan, "Depression-A Review," *Research Journal of Recent Sciences,* vol. 1(4), pp. 78-87, 2012.

[28] D. P. Devanand, "Dysthymic disorder in the elderly population," *International Psychogeriatric Association*, vol. 26, no. 1, pp. 39-48, 2014.

[29] A. P. Association, Depressive Disorders: DSM-5(r) Selections 1st Edition, Arlington: American Psychiatric Publishing, 2015.

[30] A. Iglewicz, K. Seay, S. D. Zetumer and S. Zisook, "The removal of the bereavement exclusion in the DSM-5: exploring the evidence," *Current psychiatry reports,* vol. 15, no. 11, p. 413, 18 October 2013.

[31] R. Pies, "The Bereavement Exclusion and DSM-5: An Update and Commentary," *Innov Clin Neurosci,* Vols. v.11(7-8), p. 19–22, Jul-Aug 2014.

[32] A. A. Salah, M. Pantic and A. Vinciarelli, "Recent developments in social signal processing," in *2011 IEEE International Conference on Systems, Man, and Cybernetics*, Anchorage, AK, USA, 9-12 Oct. 2011.

[33] S. Bhatia, Affect Analysis using Multimodal Cues, Thesis, Canberra, Australia: University of Canberra, Human-Centred Technology Research Centre, 2017.

[34] L. K. Yoon, J. Joormann and I. H. Gotlib, "Judging the Intensity of Facial Expressions of Emotion: Depression-Related Biases in the Processing of Positive Affect," *HHS Author Manuscripts,* vol. 118, no. 1, p. 223–228, 2009 Feb;.

[35] J. F. Cohn, "Automated Facial Expression Analysis for Research and Clinical Use," National Institutes of Health(NIH), National Institute of Mental Health (NIMH), University of Pittsburgh, Pittsburgh, PA, United States, 2017.

[36] V. Vonikakis, Y. Yazici, V. D. Nguyen and S. Winkler, "Group happiness assessment using geometric features and dataset balancing," in *18th ACM International Conference on Multimodal Interaction*, Tokyo, Japan, 2016.

[37] B. Zhang, G. Essl and E. M. Provost , "Automatic recognition of self-reported and perceived emotion: does joint modeling help?," in *18th ACM International Conference on Multimodal Interaction*, Tokyo, Japan, 2016.

[38] M. Rupenga and H. . B. Vadapalli, "Investigating the temporal association between eye actions and smiles," in *Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech), 2016*, 2016.

[39] C. Loconsole, C. R. Miranda, G. Augusto, A. Frisoli and V. Orvalho, "Real-Time Emotion Recognition: a Novel Method for Geometrical Facial Features Extraction," in *Computer Vision Theory and Applications - VISAPP*, Lisbon, Portugal, 2014.

[40] C. Gacav, . B. Benligiray and C. Topal, "Greedy Search for Descriptive Spatial Face Features," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, USA, 2017.

[41] R. U. Acharya, V. . K. Sudarshan, H. Adeli, . J. Santhosh, J. E. W. Koh and . A. Adeli, "Computer-Aided Diagnosis of Depression Using EEG Signals," in *European neurology*, Basel, 2015.

[42] O. FAUST, P. C. . A. ANG, S. D. PUTHANKATTIL and P. K. JOSEPH, "DEPRESSION DIAGNOSIS SUPPORT SYSTEM BASED ON EEG SIGNAL ENTROPIES," *Journal of Mechanics in Medicine and Biology*, vol. 14 , no. 03, p. 20 , June 2014 .

[43] G. M. Slavich and M. R. Irwin, "From Stress to Inflammation and Major Depressive Disorder: A Social Signal Transduction Theory of Depression," *HHS Author Manuscripts*, vol. 140, no. 3, p. 774–815, 2014 May.

[44] A. Doryab, K. J. Min, J. Wiese, J. Zimmerman and J. Hong, "Detection of behavior change in people with depression," in *Workshop on Modern Artificial Intelligence for Health Analytics (MAIHA)*, Québec, Canada, 2014.

[45] J. M. Girard and . J. F. Cohn, "Automated Audiovisual Depression Analysis," *Current Opinion in Psychology*, pp. 75-79, December 2014.

[46] J. F. Cohn, "Automatic Multimodal Affect Detection for Research and Clinical Use," National Institute of Mental Health (NIMH), University of Pittsburgh, 2017.

[47] M. Valstar, "Automatic Behaviour Understanding in Medicine," in *ICMI RFMIR*, Istanbul, Turkey, 2014 .

[48] J. Joshi, "Depression Analysis: A Multimodal Approach," in *ICMI'12 Santa Monica*, California, USA, October 22–26, 2012.

[49] M. Mahmoud and P. Robinson, "Towards automatic analysis of gestures and body," in *PervasiveHealth* , Cancun, Mexico, 2016, May 16-19.

[50] A. Pampouchidou, S. Panagiotis, K. Marias, F. Meriaudeau, F. Yang, M. Pediaditis and M. Tsiknakis, "Automatic assessment of depression based on visual cues: A systematic review.," *IEEE Transactions on Affective Computing* , 2017.

[51] A. Pampouchidou, M. Pediaditis, A. Maridaki, M. Awais, C.-M. Vazakopoulou, S. Sfakianakis and M. Tsiknakis, "Quantitative comparison of motion history image variants for video-based depression assessment.," *EURASIP Journal on Image and Video Processing,* vol. 1, no. 64, 2017.

[52] P. Wang, F. Barrett, E. Martin, M. Milonova, . R. E. Gur, C. R. Gur, C. Kohler and R. Verma, "Automated video-based facial expression analysis of neuropsychiatric disorders," *Journal of Neuroscience Methods,* vol. 168, no. 1, pp. 224-238, 2008.

[53] S. Alghowinem, R. Goecke, M. Wagner, G. Parker and M. Breakspear, "Eye movement analysis for depression detection," in *Image Processing (ICIP), 2013 20th IEEE International Conference*, Melbourne, VIC, Australia, 15-18 Sept. 2013.

[54] S. Scherer, G. Stratou, M. Mahmoud, J. Boberg, J. Gratch, A. Rizzo and L. P. Morency, "Automatic behavior descriptors for psychological disorder analysis," in *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops* , Shanghai, China, 22-26 April 2013.

[55] S. Scherer, G. Stratou, G. Lucas, M. Mahmoud, J. Boberg, J. Gratch, A. Rizzo and L.-P. Morency, "Automatic audiovisual behavior descriptors for psychological disorder analysis," *Image and Vision Computing,* vol. 32, no. 10, pp. 648-658, 2014.

[56] R. Gupta, N. Malandrakis, B. Xiao, T. Guha, M. V. Segbroeck, M. Black, A. Potamianos and S. Narayanan, "Multimodal Prediction of Affective Dimensions and Depression in Human-Computer Interactions," in *AVEC '14 Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge,* Orlando, Florida, USA, 2014.

[57] M. Nasir , A. Jati, P. G. Shivakumar, N. S. Chakravarthula and P. Georgiou, "Multimodal and multiresolution depression detection from speech and facial landmark features," in *6th International Workshop on Audio/Visual Emotion Challenge,* Amsterdam, The Netherlands, 2016.

[58] A. Pampouchidou, O. Simantiraki, A. Fazlollahi, M. Pediaditis, D. Manousos, A. Roniotis and G. Giannakakis, "Depression Assessment by Fusing High and Low Level Features from Audio, Video, and Text.," *In Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge,* pp. 27-34, 2016.

[59] J. Asim , M. Hongying , G. A. Yona Falinie and Z. Fan, "Artificial Intelligent System for Automatic Depression Level Analysis through Visual and Vocal Expressions," in *IEEE Transactions on Cognitive and Developmental Systems*, 31 July 2017 .

[60] J. Joshi, A. Dhall, R. Goecke and J. F. Cohn, "Relative body parts movement for automatic depression analysis.," *In Affective Computing and Intelligent Interaction (ACII), Humaine Association Conference,* pp. 492-497, 2013.

[61] R. A. Steer, R. Ball and W. F. Ranieri, "Comparison of Beck Depression Inventories -IA and -II in Psychiatric Outpatients," *Journal of personality Assesment,* vol. 67, no. 3, pp. 588-597, 1996.

[62] J. Gratch, . R. Artstein, G. Lucas, G. Stratou, S. Scherer, A. Nazarian, R. Wood, . J. Boberg, D. DeVault, S. Marsella, D. Traum, S. Rizzo and L.-P. Morency, "The Distress Analysis Interview Corpus of human and computer interviews," in *In Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 2014), LREC*, Reykjavik, Iceland, 2014.

[63] S. Alghowinem, R. Goecke, M. Wagner, J. Epps, M. Breakspear and G. Parker, "Detecting depression: A comparison between spontaneous and read speech," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference,* Vancouver, BC, Canada, 26-31 May 2013.

[64] G. McIntyre, R. Göcke, M. Hyett, M. Green and M. Breakspear, "An Approach for Automatically Measuring Facial Activity in Depressed Subjects," in *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference*, Amsterdam, Netherlands, Netherlands, 10-12 Sept. 2009.

[65] V. Mitra, E. Shriberg , M. McLaren , A. Kathol, C. Richey , D. Vergyri and M. Graciarena , "The SRI AVEC-2014 Evaluation System," in *Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge*, Orlando, FL, USA, November 7, 2014, .

[66] A. T. Beck, R. A. Steer, R. Ball and W. F. Ranieri, "Comparison of Beck Depression Inventories -IA and -II in Psychiatric Outpatients," *JOURNAL OF PERSONALITY ASSESMENT*, pp. 588-597, 19 January 1996.

[67] J. R. Williamson, T. F. Quatieri, B. . S. Helfer, G. Ciccarelli and D. D. Mehta, "Vocal and Facial Biomarkers of Depression based on Motor Incoordination and Timing," in *AVEC '14 Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge*, Orlando, Florida, USA , November 07 - 07, 2014 .

[68] B. Amos, B. L. and M. Satyanar, "OpenFace: A general-purpose face recognition library with mobile applications," School of Computer Science Carnegie Mellon University Pittsburgh, June 2016.

[69] T. Baltrusaitis, . P. Robinson and L. P. Morency, "Constrained Local Neural Fields for robust facial landmark detection in the wild," in *ICCVW*, 2013.

[70] D. Cristinacce and T. Cootes, "Feature detection and tracking with constrained local model," in *BMVC*, 2006.

[71] S. Phimoltares, C. Lursinsap and K. Chamnongthai, "Face detection and facial feature llocalization without considerning the appearance of image context," *Image and Vision Computing,* vol. 25, no. 5, pp. 741-753, 2007.

[72] Z. Yu-Jin, Advances in Face Image Analysis: Techniques and Technologies: Techniques and Technologies, New York, United States of America: Medical Information Science Reference, 2010.

[73] A. Martinez and D. Shichuan, "A Model of the Perception of Facial Expressions of Emotion by Humans: Research Overview and Perspectives," *Machine Learning Research,* vol. 13, no. 1, pp. 1589-1608 , 2012.

[74] K. K. Jageshvar and D. P. Mahendra, "Facial Feature Expression Based Approach for Human Face Recognition: A Review," *IJISET - International Journal of Innovative Science, Engineering & Technology*, vol. 1, no. 3, 2014.

[75] G. Palestra, A. Pettinicchio, D. M. Coco, P. Carcagnı, M. Leo and C. Distante, "Improved Performance in Facial Expression Recognition Using 32 Geometric Features," in *Image Analysis and Processing*, Genoa, Italy, 2015.

[76] V. Murino and E. Puppo, Image Analysis and Processing — ICIAP 2015: 18th International Conference, Springer, 2015.

[77] P. Bourke, "Interpolation methods," December 1999.

[78] A. Karagiannis, P. Constantinou and D. Vouyioukas, "Biomedical Time Series Processing and Analysis Methods: The Case of Empirical Mode Decomposition," in *Advanced Biomedical Engineering*, G. D. Gargiulo and A. McEwan, Eds., INTECT, August 23, 2011.

[79] D. E. Vigo, S. . M. Guinjoan, M. Scaramal , L. . N. Siri and D. P. Cardinali, *Autonomic Neuroscience: Basic and Clinical,* vol. 123, p. 94 – 100, 2005.

[80] G. S. Yadav, S. Yadav and Prachi , "Time and Frequency Exploration of ECG Signal," *International Journal of Computer Applications,* vol. 67, no. 4, p. 0975 – 8887, 2013.

[81] N. Y. Belova, S. V. Mihaylov and B. G. Piryova, "Wavelet transform: A better approach for the evaluation of instantaneous changes in heart rate variability," *Autonomic Neuroscience,* vol. 131, no. 1-2, pp. 107-122, January 2007.

[82] D. Acharjee, A. Mukherjee, J. K. Mandal and N. Mukherjee, "Activity recognition system using inbuilt sensors of smart mobile phone and minimizing feature vectors," *Microsyst Technol*, 12 May 2015.

[83] R. Silipo, "Seven Techniques for Data Dimensionality Reduction," [Online]. Available: https://www.kdnuggets.com/2015/05/7-methods-data-dimensionality-reduction.html. [Accessed 21 November 2017].

[84] S. Mahajan and A. Singh, "Integrated PCA & DCT Based Fusion Using Consistency Verification & Non-Linear Enhancement," *International Journal Of Engineering And Computer Science* , vol. 3, no. 3, pp. 4030-4039, March, 2014.

[85] A. Keshvani, "PRINCIPAL COMPONENT ANALYSIS IN 6 STEPS, under Statistics," COOLSTATSBLOG, Economics and statistics in plain English, 21 March 2015. [Online]. Available: https://coolstatsblog.com/2015/03/21/principal-component-analysis-explained/. [Accessed 10 November 2017].

[86] V. Powell, "Principal Component Analysis," [Online]. Available: http://setosa.io/ev/principal-component-analysis/. [Accessed 03 Octomber 2017].

[87] J. Brownlee, "A Tour of Machine Learning Algorithms," 25 November 2013. [Online]. Available: https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/. [Accessed 15 Octomber 2017].

[88] C. Sammut and G. I. Webb, Encyclopedia of Machine Learning, Springer Science+Business Media, 2011.

[89] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 2006, p. 861–874, 19 December 2005.

[90] J.-M. Torres-Moreno, Automatic Text Summarization, John Wiley & Sons, 10 November 2014 , p. 320 .

[91] D. Tominaga, "Periodicity Detection Method for Small-Sample Time Series Datasets," *Bioinform Biol Insights*, vol. 4, p. 127–136, 2010 Nov 22..

[92] S. Yoon, H. S. Kim, J.-I. Kim, S. Lee and S.-H. Lee, "Reading simple and complex facial expressions in patients," *Psychiatry and Clinical Neurosciences*, vol. 3, no. 70, p. 151–158 , 2016.

[93] P. Gilbert, K. McEwan, F. Catarino and R. Baião, "Fears of Negative Emotions in Relation to Fears of Happiness, Compassion,Alexithymia and Psychopathology in a Depressed Population: A Preliminary Study," *Depression & Anxiety*, 13 May 2014.

[94] A. Martinez and S. Du, "A Model of the Perception of Facial Expressions of Emotion by Humans: Research Overview and Perspectives," *Journal of Machine Learning Research,* vol. 13, no. 2012, pp. 1589-1608, 2012.

APPENDIX

Cubic Interpolation Method

```
doubleCubicInterpolate(
double y0,double y1,
double y2,double y3,
double mu)
{
double a0,a1,a2,a3,mu2;

   mu2 = mu*mu;
   a0 = y3 - y2 - y0 + y1;
   a1 = y0 - y1 - a0;
   a2 = y2 - y0;
   a3 = y1;

return(a0*mu*mu2+a1*mu2+a2*mu+a3);
```