People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
Mohamed El Bachir El Ibrahimi University of Bordj Bou Arreridj
Faculty of Mathematics and Informatics
Informatics Department

DISSERTATION

Presented in fulfillment of the requirements in order to obtain the degree

Master's in Informatics

Specialty: Information and Communication Technology

THEME
Micro Features Descriptor for Face Expression Analysis

Presented by:

BOUDROUAZ Roumaissa

MEKHALFI Aya Hibet Errahmane

Publicly defended on: 22/06/2023

In front of the jury composed of:

President: Pr. BOUZIANE Abderraouf

Examiner: Dr. ZITOUNI Sihem

Supervisor: Dr. REGOUID Meryem

2022/2023

إهداء

إلى قدوتي، حبيبتي أمــي...

إلى سندي، عزيزي أبــي...

إلي أختَي و نجمتي، صفا و مروى...

إلي أخوَي و قمري، محمد و أحمد...

إلى عائلتــي الداعمتين، بــودرواز و فضيــل ...

إلى صديقتــي و بسمتي ، بسمة ...

إلى شريكة العمل، ايــة هبة الرحمان ...

إلى زملاء و زميلات دفعتـــــي ...

إلى الذين مروا بنا و مررنا بهم ...

إلى هؤلاء جميعاًــا، أهدي عملي هذا.

روميساء

Dédicace

À mes chers parents,

qui m'ont inculqué des valeurs précieuses,

Vous avez été mes premiers professeurs, mes guides vertueux,

Votre confiance en moi a été ma plus grande richesse,

Vos sacrifices ont été une source d'inspiration inépuisable.

je vous suis éternellement reconnaissante pour cette tendresse.

À mes frères et sœurs bien-aimés,

Votre confiance en moi a été une source d'inspiration et de force tout au long de cette aventure.

À mes nièce et neveu WASSIM et ASMA,

qui illuminent ma vie de leur innocence et de leur joie contagieuse,

À Naima, ma chère amie,

pour les moments partagés sans pareille,

À mon binôme Roumaissa,

À mes cousins et cousines,

À tous ceux qui nous aiment, qui ont cru en moi et qui ont été les témoins de ma progression.

Je vous suis reconnaissante pour chaque mot d'encouragement et chaque geste de tendresse.

À vous tous, je dédie ma mémoire de fin d'études.

Avec toute ma gratitude et mon amour,

Aya hibet errahmane

Acknowledgment

First and foremost, we would like to express our gratitude to almighty **ALLAH** for granting us the strength, perseverance, and wisdom throughout our journey in completing this thesis. Without His blessings and guidance, this accomplishment would not have been possible.

Our heartfelt thanks to our supervisor **Dr.REGOUID Meryem**, for her support, invaluable expertise, and guidance throughout this research. Her dedication to excellence, patience, and insightful feedback have been invaluable in shaping the direction of this thesis and enhancing our understanding of the subject matter.

We would like to extend our immense appreciation to the professors who have served as role models and idols throughout our academic journey. Their profound knowledge, passion for teaching, and commitment to research have inspired us to strive for intellectual growth and pursue this thesis with utmost dedication.

Lastly, We would like to express our deepest appreciation to all the researchers, scholars, and authors whose work we have referenced and relied upon throughout this thesis. Their contributions have significantly enriched the body of knowledge in the field of memory research and have served as a foundation for our own work.

We are indebted to the numerous individuals who have contributed to our growth as a researcher and have played a vital role in the completion of this memory thesis. Their support and encouragement have been invaluable, and we are sincerely grateful for their presence in our academic journey.

Many thanks go to our family and friends for their unwavering support, encouragement, and understanding throughout this endeavor. Their belief in our abilities and their willingness to provide a listening ear during challenging times have been a constant source of strength and motivation.

Abstract

Facial expressions are a fundamental aspect of human communication .They have gained attention as a biometric system. This research focuses on micro facial expressions, which reveal hidden emotions and intentions through subtle facial muscle contractions. The objective behind this work is to develop a highly accurate micro facial expression recognition system using CNN deep learning method and ALBP method for feature extraction. A CNN is trained on a curated dataset to detect and interpret these subtle emotional cues. Our study aims to advance micro facial expression recognition, benefiting emotion analysis, lie detection, and understanding human behavior.

The proposed CNN-based model, shows promise in accurately classifying and recognizing subtle emotional cues. In which will contribute to the potential improvement of psychology, human-computer interaction, and developing empathic systems.

However, limitations related to lighting conditions, image quality, and facial structure variations are acknowledged, requiring further investigation for enhanced generalizability. Future work involves integrating multimodal inputs, exploring advanced CNN models, and augmenting the dataset with diverse examples to improve reliability and effectiveness in real-world scenarios.

Keywords: Biometrics, Facial Micro-Expression Recognition, Convolutional Neural Network, Feature Extraction.

Résumé

Les expressions faciales sont un aspect fondamental de la communication humaine. Elles ont attiré l'attention en tant que système biométrique. Cette recherche se concentre sur les micro-expressions faciales, qui révèlent des émotions et des intentions cachées à travers de subtiles contractions musculaires du visage. L'objectif de ce travail est de développer un système de reconnaissance des micro-expressions faciales extrêmement précis en utilisant la méthode d'apprentissage en profondeur CNN et la méthode ALBP pour l'extraction des caractéristiques. Un CNN est entraîné sur un ensemble de données sélectionné pour détecter et interpréter ces indices émotionnels subtils. Notre étude vise à faire progresser la reconnaissance des micro-expressions faciales, ce qui bénéficie à l'analyse des émotions, à la détection du mensonge et à la compréhension du comportement humain.

Le modèle proposé basé sur CNN montre des résultats prometteurs pour classifier et reconnaître avec précision ces indices émotionnels subtils. Cela contribuera potentiellement à l'amélioration de la psychologie, de l'interaction homme-machine et du développement de systèmes empathiques.

Cependant, des limitations liées aux conditions d'éclairage, à la qualité de l'image et aux variations de la structure faciale sont reconnues, nécessitant des recherches supplémentaires pour une meilleure généralisation. Les travaux futurs impliquent l'intégration d'entrées multimodales, l'exploration de modèles CNN avancés et l'augmentation de l'ensemble de données avec des exemples variés pour améliorer la fiabilité et l'efficacité dans des scénarios réels.
Mots clés : biométrie, reconnaissance de la micro-expression faciale, réseau de neurones convolutifs, extraction de caractéristiques.

ملخص

تعابير الوجه هي أحد الجوانب الأساسية للتواصل البشري ، وقد اكتسبت الاهتمام كنظام قياس حيوي. يركز هذا البحث على تعابير الوجه الدقيقة ، والتي تكشف عن المشاعر والنوايا الخفية من خلال التقلصات الدقيقة لعضلات الوجه. الهدف من هذا العمل هو تطوير نظام دقيق للغاية للتعرف على تعبيرات الوجه باستخدام طريقة التعلم العميق للشبكة العصبية التلافيفية وطريقة النمط الثنائي المحلي التكيفي لاستخراج الميزات. يتم تدريب الشبكة العصبية التلافيفية على مجموعة بيانات منظمة لاكتشاف وتفسير هذه الإشارات العاطفية الدقيقة. تهدف دراستنا إلى تعزيز التعرف على تعبيرات الوجه الدقيقة بالاستفادة من تحليل المشاعر ، وكشف الكذب ، وفهم السلوك البشري

يُظهر النموذج المقترح المستند إلى الشبكة العصبية التلافيفية والذي يستخدم النمط الثنائي المحلي التكيفي لاستخراج الميزات ، نتائج واعدة في تصنيف الإشارات العاطفية الدقيقة والتعرف عليها مؤخراً يساهم كذالك في علم النفس البشري ،تفاعل الكمبيوتر ، والحوسبة العاطفية ، و يحتمل أن تحسن التفاعل البشري وتطوير أنظمة التعاطف

.ومع ذلك ، فإن القيود المتعلقة بظروف الإضاءة وجودة الصورة وتباين بنية الوجه معترف بها وتتطلب مزيداً من التطوير من أجل تعزيز التعميم. مما يستدعي إجراء مزيد من البحوث لتعزيز القابلية العامة للنموذج. يشمل العمل المستقبلي دمج المدخلات المتعددة الوسائط ، واستكشاف نماذج متقدمة لشبكات التعلم العميق، وتوسيع مجموعة البيانات بأمثلة متنوعة لتحسين الموثوقية والفعالية في سيناريوهات العالم الحقيقي،
الكلمات المفتاحية: القياسات الحيوية ، التعرف على تعبيرات الوجه الدقيقة ، الشبكة العصبية التلافيفية ، استخراج الميزات

Table of contents

List of Figures **x**

List of Tables **xi**

1 General Introduction **1**

1.1 Context 1

1.2 Problematic 1

1.3 Objectives 2

1.4 Methodology and results 2

1.5 Report outline 2

2 State of The Art **4**

2.1 Introduction 4

2.2 Definition of Biometrics 5

2.3 Biometric Criteria Requirements 6

2.4 Biometric Applications 7

2.5 Design of Biometric Systems 7

2.6 Face Anatomy 9

2.7 Face Expression 10

2.8 Micro Expression 10

2.9 Micro Vs. Macro Expressions 11

2.10 Related Work 13

2.10.1 Region-based approach 13

2.10.2 Feature extraction methods 14

2.10.3 Deep learning architectures 14

2.11 Existing Dataset 15
2.11.1 Spontaneous micro-expression datasets 16
2.11.2 Non-spontaneous micro expression datasets 17
2.12 Conclusion 18

3 Conception and Realization **19**
3.1 Introduction 19
3.2 The architecture of the proposed micro-expression recognition system 19
3.2.1 Preprocessing 19
3.2.2 Feature Extraction 20
3.2.3 Classification 21
3.3 Conclusion 29

4 Experiments and Results **30**
4.1 Introduction 30
4.2 Technical Requirements 30
4.2.1 Hardware Requirements 30
4.2.2 Software Requirements 31
4.3 Micro Expression System GUI 31
4.4 Numerical Results 34
4.4.1 Validation Metrics 34
4.4.2 Results 35
4.5 Discussion 37

5 General Conclusion **43**
5.1 Contributions 43
5.2 Limitations 43
5.3 Future work and perspectives 44

References **44**

List of Figures

2.1 Informations Sets[1]. 5
2.2 Biometrics Applications. 7
2.3 Biometric System Modules.[2] 9
2.4 Muscles of facial expression.[3] 10
2.5 Competing Neural Pathways in High-Risk Scenarios in Micro-Expressions.[4] . 11

3.1 CNN Architecture SMIC 24
3.2 CNN Architecture CASIA 27

4.1 Welcome Interface. 32
4.2 SMIC Interface. 32
4.3 Dialogue window. 32
4.4 Displaying Original Image 33
4.5 Displaying Preprocessed Image 33
4.6 Displaying ALBP Features 33
4.7 Displaying Classification class for SMIC Dataset. 34
4.8 Displaying Classification class for CASIA Dataset 34
4.9 Confusions Matrix 37
4.10 Training and Validation Accuracy graphs 37

List of Tables

2.1 Main differences between macro- and microexpressions.[4] 12

2.2 Summary of some posed and spontaneous micro-expression dataset[5] 18

4.1 SMIC Result. 36

4.2 CASIA Result. 36

Chapter 1

General Introduction

1.1 Context

In recent years, the field of biometrics has become increasingly vital for identification and authentication, relying on unique physical or behavioral characteristics. Biometric systems have traditionally focused on modalities such as fingerprints, iris patterns, and voice recognition. depending on the number of these modalities the biometric system is divided into multimodal systems if the modalities are two or more and unimodal systems if the modality used is unique. facial expressions have gained significant attention as a unimodal system due to their ability to convey emotions and intentions. Facial expressions offer a rich source of information in biometric analysis, enabling a deeper understanding of individuals beyond their physical traits. Within the realm of facial expressions, there are two distinct categories: macro facial expressions and micro facial expressions. Macro facial expressions are the easily recognizable, overt movements of facial muscles that accompany pronounced emotions. In contrast, micro-facial expressions refer to the subtle, fleeting contractions of facial muscles that reveal hidden emotions and intentions.

1.2 Problematic

Humans' faces make very brief, involuntary facial expressions known as "facial micro-expressions" when they unintentionally hide an emotion. Since micro-expressions last for a shorter period of time than macro-expressions, efforts to train people to recognize them have had very poor results. This makes it more difficult for humans to recognize micro-expressions even when they were taught to do as well.

1.3 Objectives

The goal of this research project is to develop a high-accuracy micro facial expression recognition system that surpasses the limitations of existing methods[31][39][40][41]. Current approaches often struggle with low accuracy rates and difficulties in capturing subtle facial cues. Therefore, there is a pressing need for an improved system that can accurately detect and interpret mi- croexpressions. By achieving unprecedented levels of accuracy, we aim to advance the field of micro facial expression recognition and unlock new possibilities in emotion analysis, lie detection, and understanding human behavior.

1.4 Methodology and results

To accomplish our goal, we will utilize a deep learning approach, specifically employing a convolutional neural network (CNN). We will train the CNN on a carefully curated dataset consisting exclusively of micro facial expressions. Additionally, we will employ the Adaptive Local Binary Patterns (ALBP) method to extract discriminative features from the facial images, as it has shown promising results in capturing facial texture. This combination of deep learning and ALBP feature extraction aims to enhance the accuracy and robustness of micro facial expression recognition. By focusing on the challenges and nuances associated with these subtle emotional cues, we aim to contribute to the advancement of biometric recognition systems.

1.5 Report outline

In this study, we present a comprehensive investigation of micro facial expression recognition in biometrics. The research is organized as follows:

- In Chapter 2, we explore the field of biometrics focusing on the criteria requirements of a biometric system and examine the diverse applications of biometrics across various sectors and shifting focus also to facial expression analysis, starting with describing face anatomy, The discussion extends to different types of facial expressions and draws comparisons between them.

- Chapter 3 focuses on the conception and realization of our proposed approach. We detail the methodology we developed, which combines deep learning techniques, starting with describing the datasets we used then we detailed the architecture of the proposed system.
- Chapter 4 presents the experiments conducted and the results obtained from our proposed approach. We discuss the accuracy performance, achieved by our model in recognizing micro facial expressions. Additionally, we present our graphical interface with it different modules.
- Finally, in Chapter 5, we provide a general conclusion that summarizes the key findings of our study. We discuss the implications of our research, and its potential applications in real-world scenarios, and suggest avenues for future research and improvements in micro facial expression recognition.

Chapter 2

State of The Art

2.1 Introduction

With the extensive automation we live in this decade, security problems are increasing, and providing a good security level is a necessity. The development of biometric technology has advanced rapidly in recent years, giving us more accurate and reliable systems for identifying individuals that's why the world opts for The Biometric system Furthermore, the research of facial expression recognition, which is an essential component of biometric systems, is a fast-expanding topic that still requires significant contributions. Despite this, several state-of-the-art solutions have been presented. Because micro and macro expression recognition approaches are constructed and conceptualized similarly, this chapter will present an overview of research in both disciplines. The chapter also includes a primer on face anatomy. It then describes the distinctions between micro and macro expressions and examines possible ways from the beginning of the research to the present. in this chapter will have an overview of the state of the art including biometric definition its different types, How can we develop a biometric system? and what is its design? we will see also face expressions and differences between micro and macro, Additionally, the chapter discusses the different research of this field and the different datasets used in it. Finally, the chapter concludes with a summary of the findings.

2.2 Definition of Biometrics

Nowadays, with our automated lives, security problems are getting crucial, and allowing access issues, and authoring sensitive actions are serious questions. There are two ways of responding: the first one is based on what a person has, which is called "Ownership factors", ID card for example; the Second way is based on what a person knows; like passwords, called "Knowledge factors"[1].

Moreover, the two previous methods could be stolen or copied, So users should have many IDs or at least memorize countless Passwords.

Most recently, Jingdong, which is an e-commerce platform, divulged nearly 12G users' privacy data; Gmail, which is an email product provided by Google, had its 5 million passwords accounts breached.[2] Here comes biometrics to solve the issue of personal identification and to distinguish between an allowed and unlawful individual. Biometric characteristics provide the benefits listed below:

- They are unique to each person.
- They are difficult to forget, steal, borrow, share, or observe.
- They change constantly and are always available.
- They cannot be simply conveyed, or transformed to another person[1].

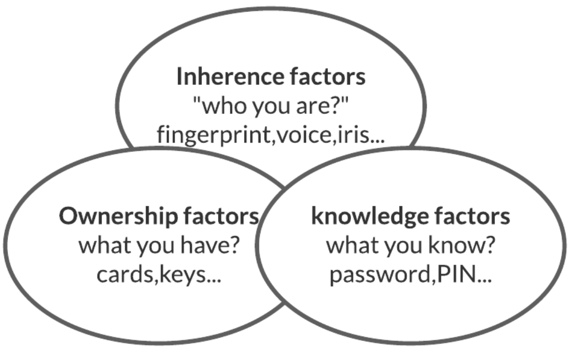


Figure 2.1: Informations Sets[1].

2.3 Biometric Criteria Requirements

A biometric criteria is a measurable physical or behavioral characteristic of an identifiable individual as discussed in the definition (section 2.2). It determines how a person is recognized. To design a practical biometric system it is important to know what characteristics the system should use to make decisions about someone's identity and each one of these characteristics has its own strengths and weaknesses. The choice of which characteristic is usually depending on the application domain and possibly the population identified. In some cases, multiple traits are selected. Jain, Flynn and Ross[6] identified several requirements that a typical biometric function must meet.

1. Universality: The individuals accessing the application must own those characteristics, we can't identify blind people by their iris for example

2. Uniqueness: The characteristics must be different enough for people to be able to distinguish between them.

3. Permanence: The biometric characteristics shouldn't be changeable cross time or minimum for a duration of the operating recognition system. we can't consider a trait that changes as a biometric trait.

4. Measurability: The biometric characteristics need to be a measurable quantity to be further processed by a machine, and transferred later to the recognition system, Suitable devices connected to the machine are in charge to acquire and digitize those biometric traits.

5. Performance: An acceptable degree of performance is required for the application that uses the biometric characteristics. including the matching accuracy/time in addition to resources to create the entire recognition system.

6. Acceptability: This represents persons ready to comply with the system by giving their biometric characteristics that they are defined by.

7. Circumvention: For evaluating the system's strength; i.e. if it is easy to fool the system by making it take wrong decisions

2.4 Biometric Applications

The need for secure, reliable identity validation and confirmation has driven the adoption of biometric technologies in a diverse range of applications. Biometric applications can be found in many domains. here's a list of where utilization and deployment are possible: Law Enforcement (Forensics), Background Checks, Surveillance, Border Control, Fraud Reduction, Trusted Traveler, Physical Access ControlTime and Attendance, Consumer Recognition, Remote Authentication: Biometrics provide a safe way of authentication for remote access to critical information by precisely identifying mobile device users. Previously, fingerprint and voice recognition were used in deployments. Asset Protection, Logical Access Control[7].

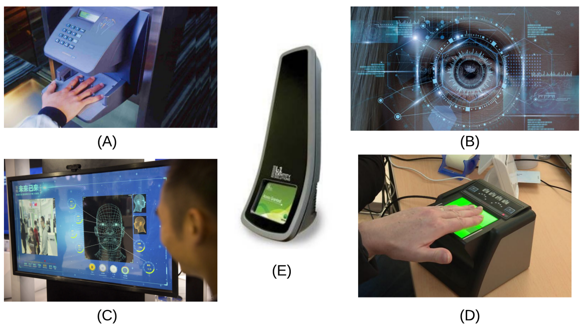


Figure 2.2: Biometrics Applications.
(A): Hand geometry, (B): Iris Recognition, (C): Face Recognition, (D): Multiple Fingerprints Recognition, (E): One Fingerprint Recognition

2.5 Design of Biometric Systems

A generalized biometric system is a functional combination of five main following subsystems[8]:

1. **Data Acquisition:** This subsystem is in charge of acquiring a sample of biometric traits from individuals, such as images. This biometric sample is uncompressed data, known as raw biometric data, and it is captured by a sensor. This component is the single point at which the user and biometric system communicate, and the process is also known as biometric presentation.

2. **Processing Module:** This subsystem is responsible for preprocessing data and extracting features from a biometric sample in order to generate a digital representation known as a biometric template or reference, which represents the uniqueness of the sample while

also being somewhat invariant in relation to multiple samples created from the same individual over time. Sample enhancement, quality evaluation (segmentation), and feature extraction are all part of the processing module. The result of quality control tests (segmentation and feature extraction) is a quality score, which reflects the sample's quality based on the success of the feature extraction method.

3. **Data Storage:** This module is in charge of storing the biometric template in a shared database, which is preserved for future procedures and is also known as a reference in the biometrics domain. that they are produced and stored during the enrollment mode that we will mention afterward.

4. **Matching Module:** This one depends on the application purpose, each newly created sample template is then compared with one in verification mode or more reference templates in identification mode by comparison algorithm. The result of the comparison algorithm is a comparison score or similarity (respectively dissimilarity) score, indicating how similar the templates are. The comparison score is then transferred to a decision-making module.

5. **Decision Module:** This subsystem compares a verification or an identification threshold with the result score from the matching module. The threshold is a specified ratio that is typically defined by the administrators of a biometric system. If the score from the comparator (template comparison) meets the threshold, the compared templates are matched (or accepted), otherwise, the compared templates are not-matched (or rejected) if the resulting score is under the threshold. the systems can be either highly secure or not secure at all, based on the threshold settings so it has the most important role in the security of systems.

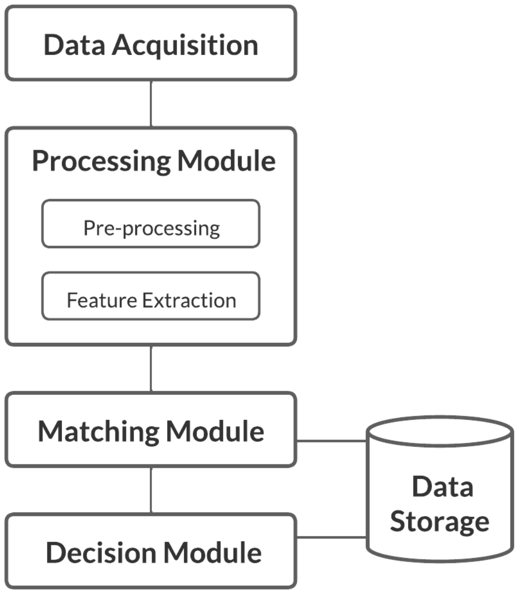


Figure 2.3: Biometric System Modules.[2]

2.6 Face Anatomy

The facial muscles, also known as cranial muscles, are a collection of around 20 thin muscle tissue located under the skin of our face and head. they surround facial apertures (mouth, nose, eye, and ear) or wrap around the head and neck. As a result, these muscles are classified into groups[3];

- Buccolabial group (Muscles of the mouth)
- Nasal group (Muscles of the nose)
- Orbital group (Muscles of the eyelid)
- Epicranial group(Muscles of the skull and neck)
- Auricular group (Muscles of the external ear)

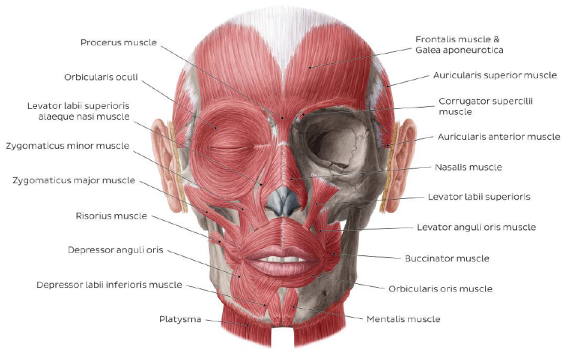


Figure 2.4: Muscles of facial expression.[3]

2.7 Face Expression

Facial expressions are often divided into macro and micro-expressions. The six most common facial expressions, which typically last between 0.5 and 4 seconds, are referred to as macro-expressions. Humans intentionally express macro-expressions, which can be easily detected with the naked eye without prior training. Several psychological studies have found that macro-expressions may not accurately reflect the emotional state of the individual expressing them. Micro-expressions, on the other hand, are quick (short duration) and subtle (low intensity), which are facial expressions that last between 0.04 to 0.2 seconds and are frequently localized in certain areas of the face. Furthermore, micro-expressions are typically subconsciously expressed and hence show true feelings. Because microexpressions are subtle and brief in duration, it is difficult for humans to recognize them; even educated specialists achieve a detection rate of only 47%[9].

2.8 Micro Expression

The movement of facial skin and connective tissue causes facial expressions. Facial nerve nuclei, which regulate these motions, stimulate the facial muscles. cortex and subcortical upper motor neuron circuits in turn regulate these motor neurons. Two separate neural pathways (located in various brain regions) are involved in influencing facial behavior, according to a neuropsychological investigation of the facial expression of W. E. Rinn [10]. The cortical circuit is primarily in charge of posed facial expressions and is found in the cortical motor strip (i.e., voluntary facial actions). In addition, the brain's subcortical regions, where the subcortical circuit is located, are principally in charge of generating spontaneous facial expressions (i.e., involuntary emotion), When people try to hide or suppress their reactions in a stressful situation[11].

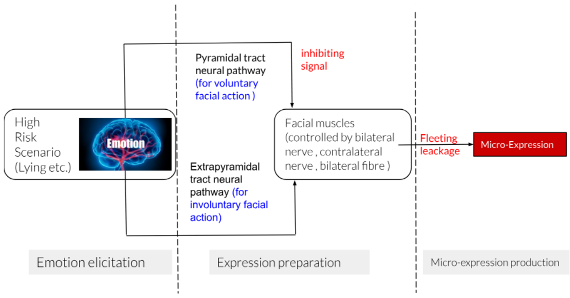


Figure 2.5: Competing Neural Pathways in High-Risk Scenarios in Micro-Expressions.[4]

The pyramidal and extrapyramidal tract neural pathways are two separate brain pathways for transmitting facial behavior. The pyramidal pathway is in charge of Macro expressions with voluntarily performed facial movements, whilst the extrapyramidal pathway is in charge of spontaneous facial expressions. Both circuits are activated and engaged in a back-and-forth conflict in high-risk situations like lying, which causes the short release of real emotions in the form of microexpressions[8].

Both systems are likely to be active in an emotional circumstance, which causes the momentary release of real emotions in the form of microexpressions 2.5. Localized facial deformations known as micro-expressions are brought on by the involuntary contraction of facial muscles. In contrast, macro-expressions use more muscle across a wider area of the face and have significantly stronger muscular motion [12].

micro-expressions characteristics. 7 basic emotions: fear, sadness, anger, happiness, disgust, surprise, and contempt can be expressed via micro-expressions[12]. According to Ekman[13], there are some muscles in the face that are accurate indications of the presence of associated emotions because they cannot be controlled in a conscious way. Micro-expressions may include all or just a portion of the muscular movements that form conventional expressions.

2.9 Micro Vs. Macro Expressions

Micro expressions differ from macro expressions in that they are neuroanatomically based, have a very short duration, little fluctuation, and have fewer action sites on the external facial features. This distinction can also be drawn from the hiding mechanism: When people strive to hide their emotions, their real feelings might easily "leak out" and emerge as microexpressions. Additional investigations have demonstrated that the precentral gyrus, which is dominated by the bilateral nerve, controls the motion of the muscles in the upper face (forehead and upper eyelid) during false expressions, whereas the lower face is regulated by the contralateral nerve[14][15]. The subcortical structure, whose bilateral fiber regulates the movement of the muscles, controls spontaneous expressions. This also supports the restraint hypothesis of C. Darwin and P. Prodger [16], which claims that spontaneous expressions controlled by the extrapyramidal tract engage in a conflict with random expressions controlled by the pyramidal tract neural pathway. Bilateral fibers predominate in the extrapyramidal tract throughout this procedure. Moreover, if extrapyramidal tract activity is mirrored on the face, it affects the muscles that line the upper face, resulting in uninhibited facial expressions. Moreover, this supports the idea that diverse facial features can exhibit micro-expressions[17]. According to the theory mentioned in “Frontiers in Psychology” [18] the feedback from the top and lower facial areas affects the ability to recognize micro-expressions differently. Following that, three additional studies were provided to emphasize the 3 functions of face feedback in determining the minute motions of micro expressions. First, microexpressions with a duration of 450 MS are simpler to recognize once feedback in the upper face is reinforced by a limiting gel. Second, the accuracy of detecting micro-expressions (timing conditions of 50, 150, 333, and 450 ms) is decreased when lower face feedback is enhanced. Third, preventing lower facial feedback increases the precision with which microexpressions may be recognized[4].

Micro-expressions, which are shorter-lived than macro-expressions and marked by more restriction of facial muscle action, can represent a person’s real emotions and are more challenging to regulate. The main distinctions between macro- and micro-expressions are described in Table 2.1.

Table 2.1: Main differences between macro- and microexpressions.[4]

Difference	**Micro-expression**	**Macro-expression**
Noticeability	Easy to ignore	Easy noticed
Time interval	Short duration (0.065-0.5 seconds)	Long duration (0.5-4 Seconds)
Motion intensity	Slight variation	Large variation
Subjectivity	Involuntary (uncontrollable)	Voluntary (under control)
Action areas	Fewer	Almost all areas

Short duration, the most crucial aspect of micro-expressions is thought to be their brief duration. The majority of psychologists today concur that microexpressions don't actually endure longer than 30 to 60 milliseconds. In order to analyze microexpressions, Yan et al.[19] carried out a nice analysis using distribution curves of time amount and start duration. According to these authors, a microexpression's start-up time (the interval between the start and apex frames) is typically 0.26 seconds or less. Also, due to the psychological suppression of human automatic response, the muscular movement in the micro-expression is very faint and the expression itself would be uncontrolled. Dynamic features. The face's left side reflects emotions that occur more intensely, according to the neuropsychology of B. Bhushan [20], whereas the right side reflects emotions that are less intense. According to research on asymmetry, the right side of the face prominently expresses social context indicators whereas the left side shows more introspective emotions. These studies add to the body of evidence that distinguishes artificial from genuine facial expressions and so implicitly explains the dynamic properties of micro-expressions[4].

2.10 Related Work

2.10.1 Region-based approach

ME, defined as modest and transient motion, has been demonstrated to be part region-dependent P. Hus'ak et al[21], since its appearances are brief, unintentionally X. Li et al [22], but most don't occur throughout the face. As a result, they split the face into a particular number of part classes. Nevertheless, certain face regions, like the cheeks, do not play a significant role in the ME analysis Y.-H. Oh et al. [23]. To solve this problem, A. Davison et al. [24] discovered that the face should be divided into a few region-of-interests (ROIs), and the ROIs that include facial movements activated whenever one or many action units of facial expression coding system (FACS) are active should be considered Y.-H. et al. [23]. ROIs have been demonstrated to be more effective than the complete face in ME detection [24] [25]. Studying specific ROIs is commonly used in spotting activities since it aids in the correct detection of geometrical characteristics [26]. Nevertheless, only a small amount of research has used local patching in the process of ME identification, such as Zhao and Yu [27] who used essential morphological patches. Wang et al. [28] exploited ROIs in conjunction with a micro-attention mechanism to achieve a 66% accuracy rate for a ME recognition objective.

2.10.2 Feature extraction methods

Finding competent extracted features is always regarded as a vital aspect of automatic ME recognition activities. Numerous research has used hand-crafted extraction of features algorithms. The two most often used approaches are LBP-based and optical flow-based. One appearance-based technique, LBP: local binary pattern by T. Ojala et al. [29], helps in extracting local texture modifications from circular areas using binary codes encoded in a bar chart. Many LBP variations have been developed, however, the most often used approach, which has also been provided as the basic assessment system method in a large number of studies, is LBP-TOP by G. Zhao and M. Pietikainen [30]. It is a more advanced variant of LBP that is based on the spatial and temporal field and three orthogonal dimensions. This enables LBP-TOP to utilize temporal limitations as well as spatial neighborhoods dynamically. Optical flow-based approaches, on the opposite hand, assess brightness variation in video clip sequences, depending largely on temporal changes [23]. Different approaches for feature extraction are also often used, such as the histogram of picture gradient orientation [31]. These feature descriptors use video clips of spatiotemporal features to derive temporal dynamics and texture. Earlier studies used visual features in their methodologies to enhance ME recognition. Pfister and his colleagues extracted features using LBP on three orthogonal planes (LBP-TOP), proceeded by several kernel learning algorithms for binary classification.Li X. et al. [22] used three commonly used feature descriptors for their experimental studies: HOG, HIGO-TOP, and LBP-TOP. Apart from typical feature extraction techniques, a few studies used deep learning to learn hidden features. D. Patel et al. [32] transferred ImageNet models for model training.

2.10.3 Deep learning architectures

Deep learning and CNN's quick progress over time has been unquestionably and dramatically beyond expectation. Deep learning has been developed and improved in a variety of applications, like micro expression (ME) analysis, and is now one of the most used approaches, particularly for facial recognition. Systems for deep learning have been presented in the machine vision domain, including VGG-Net [33], ResNet [34], and Senet50 [35]. Although those models were addressed relatively separately in terms of computational concept and architectural design, they all have the capability and ability to acquire spatial characteristics in hidden layers that resulted in high-end image implementations. ME recognition studies have frequently turned to CNN, either to obtain a better approach or to achieve a higher recognition classification accuracy. CNN was first developed in the research of D. Patel et al. [32] to extract useful features from video sequences. That's why Peng and his colleagues [36] enhanced the usage of CNN for ME recognition using an end-to-end network called DTSCNN: Dual Temporal Scale Convolutional Neural Network. Deep learning has indeed advanced to a new level, outperforming existing approaches in the computer vision field. It is envisaged that the ME recognition process task, in particular, would overcome the restriction of poor datasets and exceed the limitations imposed by the demanding challenges provided.

2.11 Existing Dataset

Micro-expression analysis approaches have relied heavily on well-established datasets that appropriately labelled actual truth. Micro-expressions are challenging to evoke in a supervised condition since of their inherent properties such as involuntariness, brief duration, and minor fluctuation. Thus yet, just a few microexpression datasets were created. Yet, as previously said, the majority of these are still containing different shortcomings in terms of their elicitation paradigm, labelling methodologies, or limited data size. Several early research generated microexpressions by creating high-stakes circumstances, such as encouraging participants to deceive by suppressing negative effects induced by an irritating video while imitating pleasurable ones USF-HD , Polikovsky's dataset , York DDT , SMIC , CASME , CASME II , SAMM , and CAS(ME)2, are typical micro-expression datasets created since 2011. Published micro-expressions may be found in both USF-HD and Polikovsky's datasets: Participants in Polikovsky's dataset were asked to replicate the motion of a micro-expression while those in USF-HD were instructed to make both macro- and micro-expressions. It should be highlighted that these posed microexpressions are distinct from those that occur naturally. Moreover, York DDT contains unprompted microexpressions with great ecological validity. However, the data in York DDT was combined with non-emotional facial movement brought on by conversation. However, none of the three datasets—USF-HD, Polikovsky's dataset, and York DDT—are accessible to the general public[37].

The paradigm of elicitation by constantly lying has two significant flaws:

1. micro-expressions are frequently contaminated by unrelated (i.e., talking) facial movements

2. the types of elicited micro-expressions are severely constrained (for example, the happiness type can never be evoked).

Contrarily, it is generally known that seeing an emotional film while keeping your face neutral (i.e., suppressing your emotions) is a successful technique for eliciting microexpressions. This elicitation paradigm was employed in 5 datasets: SMIC, CASME, CASME II, SAMM, and CAS(ME)2. MEVIEW employed a new elicitation paradigm entirely, creating a high-stakes scenario through the use of poker games or challenging TV interviews[4]. We will contrast and analyze the pertinent publically accessible datasets for facial microexpression research. The datasets are categorized based on the type of micro-expressions they contain, as well as whether they are spontaneous or nonspontaneous, Table 2.2 summarizes some of these datasets.

2.11.1 Spontaneous micro-expression datasets

2.11.1.1 The Spontaneous Micro-expression Corpus (SMIC) dataset:

This dataset contains videos of subjects exhibiting natural, unconscious micro-expressions in response to emotional stimuli. It includes 164 video clips of 16 subjects, with each clip lasting between 3 to 5 seconds. The dataset includes seven facial expressions, including happiness, surprise, disgust, fear, anger, sadness, and contempt.

2.11.1.2 The Chinese Academy of Sciences Micro-expression (CASME) dataset:

This dataset consists of 200 video clips of 26 subjects, with each clip lasting between 1 to 5 seconds. The dataset includes 14 facial expressions, including happiness, surprise, disgust, fear, anger, sadness, and contempt.

2.11.1.3 The SMIC-HS dataset:

This dataset is a high-speed version of the SMIC dataset and includes 100 video clips of 16 subjects. The dataset includes seven facial expressions.

2.11.1.4 The SMIC-NIR dataset:

This dataset includes 29 video clips of 16 subjects captured under near-infrared lighting conditions. The dataset includes six facial expressions, including happiness, surprise, disgust, fear, anger, and sadness

2.11.2 Non-spontaneous micro expression datasets

2.11.2.1 USF-HD:

contains 100 posed microexpressions captured at 29.7 frames per second. The participants were instructed to imitate several microexpressions in the order of their choice after being given examples of them. By leaving out disgust, fear, and contempt from the seven universal expressions, the categories for micro-expressions are smile, surprise, rage, and sad. Additionally, this dataset has not been made accessible for use in public studies.

2.11.2.2 YorkDDT (York Deception Detection Test):

This dataset aimed to analyze micro-expressions in participants who truthfully or deceptively described emotional or non-emotional film clips. The researchers recorded 20 video clips at a resolution of 320x240 and a frame rate of 25 fps. Out of the 9 participants (3 male and 6 female), micro-expressions were observed in their responses. A total of 18 micro-expressions were extracted for analysis, 7 from the emotional scenario and 11 from the non-emotional version. However, due to the unavailability of the data for public access, further examination and study of these micro-expressions are currently not possible

2.11.2.3 Polikovsky dataset:

The Polikovsky dataset is a non-spontaneous dataset that contains both posed and spontaneous micro-expressions. This dataset includes over 7,000 video clips of 32 subjects, with each clip lasting between 1 to 5 seconds. The dataset includes both macro and micro-expressions, which can be useful for developing more robust recognition algorithms[37].

Table 2.2: Summary of some posed and spontaneous micro-expression dataset[5]

FACS:Facial Action Coding System; **FPS**:Frame Per Second;Sp:Spontaneous

Dataset	Subject	Samples	FACS	FPS	Classes	Resolution	Type
Polkovsky	11	42	NO	200	6	640 x 480	Posed
USF-HD	—	100	NO	29.7	4	1280 x 720	Posed
York DDT	9	18	NO	25	2	320 x 240	Posed
SMIC-sub	6	77	NO	100	3	640 x 480	Sp
SMIC-HS	16	164	NO	100	3	640 x 480	Sp
SMIC-VIS	8	71	NO	25	3	640 x 480	Sp
SMIC-NIR	8	71	NO	25	3	640 x 480	Sp
SMIC-E-HS	16	157	NO	100	3	640 x 480	Sp
SMIC-E-VIS	8	71	NO	25	3	640 x 480	Sp
SMIC-E-NIR	8	71	NO	25	3	640 x 480	Sp
CASME-A	7	96	YES	60	7	1280 x 720	Sp
CASME-B	12	101	YES	60	7	640 x 480	Sp
CASME 2	35	247	YES	200	5	640 x 480	Sp
CAS(ME)	22	357	NO	30	4	640 x 480	Sp
SAMM	32	159	YES	200	7	2040 x 1088	Sp
SAMM Long Video	32	502	YES	200	2	2040 x 1088	Sp

2.12 Conclusion

Biometrics is a practical alternative to using passwords or PINs to identify a person. Every individual's qualities are unique. Indeed, all human beings have unique properties. Even fraternal twins do not possess exactly the same characteristics[38].

In this chapter, we provided an overview of biometrics including a brief definition of biometrics and we cited the different types of biometrics, then we stated criteria requirements we mentioned also some domain applications for biometrics and the design of a biometric system and its different modes.

Furthermore, this chapter provides a comprehensive overview of facial anatomy, It also explains the concepts of expression, macro, and micro-expression, and compares their differences. The evolution of microexpression research, from the pioneering work to current approaches, additionally, a discussion on the available datasets were included, highlighting the various types of data that exist in this field. This chapter lays the foundation for the subsequent chapters, which will delve deeper into the methodologies and techniques used to capture, analyze, and interpret our micro expressions system.

Chapter 3

Conception and Realization

3.1 Introduction

In this chapter, we present our micro-facial expression recognition system, developed through the analysis of two distinct datasets. We explore data preprocessing, feature extraction, and classification phases, providing details of the employed CNN architecture. Our aim is to enhance micro-expression classification accuracy and unlock new possibilities for authentic emotion detection and understanding human emotional reactions.

3.2 The architecture of the proposed micro-expression recognition system

3.2.1 Preprocessing

1. **Image resizing**

```
image = cv2.resize(image, (48,48))
```

The purpose of this function is typically to standardize the size of the images in our dataset. Resizing images to a consistent size of 48 pixels in width and height is useful for preparing images for input into our machine learning model.

2. **Convert RGB image to grayscale**

```
image = cv2.cvtColor(image, cv2.COLOR_RGB2GRAY)
```

By converting an image from RGB to grayscale, we transform the image from having three color channels (Red, Green, and Blue) to a single channel, making it easier to perform feature extraction that does not require color information. This conversion is done using the cv2.cvtColor() function.

3. **Image normalization**

```
image = cv2.normalize(image, None, 0, 255, cv2.NORM_MINMAX)
```

Image normalization is a technique used to adjust the pixel values of images to a specific range or distribution. In this case, the pixel values of the image are scaled to the range of 0 to 255.

3.2.2 Feature Extraction

1. **Adaptive Local Binary Pattern (ALBP)**

Local Binary Patterns in general are descriptors whose purpose is to summarize the local structure of the images. The goal is to be able to discriminate between different images. The ALBP method analyzes local intensity variations of pixels in the image using a circle around each pixel. It generates a local binary code for each pixel by comparing the intensity values of samples on the circle with the intensity value of the central pixel. These binary codes are then converted to decimal values to represent the ALBP features of the image. So it compares the intensity value of each neighbor pixel with the average intensity value of all neighbors as follows:

$$\mathrm{ALBP}_P = \sum_{p=0}^{p-1} \left(s(g_p - g_e) 2^P \right), \tag{3.1}$$

$$where \quad g_e = \sum_{p=0}^{p-1} g_p / P$$

$$s(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \geq 0 \end{cases} \tag{3.2}$$

- g_e is the average intensity value of all P neighbor pixels located on a circle in a local neighborhood.

3.2.3 Classification

1. **Convolutional Neural Network (CNN)**

   A neural network is made up of layers, where the neurons are represented by the nodes or units in each layer. Each layer creates new output data using the input data it receives. When neural networks take input in the form of images, they naturally extract features from each image and train a classifier using those features.

2. **The architecture of CNN:**

   2.1 **SMIC Dataset:** The architecture of the SMIC Convolutional Neural Network (CNN) consists of the following layers:

   - 1. Input Layer:

     The input images are expected to have a shape of (target_width, target_height, 1) representing a single-channel (grayscale) image.

   - 2. Convolutional Layers:

     – The first convolutional layer applies 32 filters of size 3x3 with the ReLU activation function.

     – Batch Normalization is applied after each convolutional layer for activation normalization.

     – The second convolutional layer applies 64 filters of size 3x3 with ReLU activation.

     – MaxPooling2D with a pool size of 2x2 is used for spatial dimension reduction.

     – Dropout with a rate of 0.25 is applied to prevent overfitting.

- 3. Convolutional Layers (with regularization):
  - The third convolutional layer applies 128 filters of size 3x3 with ReLU activation.
  - Regularization (L2 regularization with a coefficient of 0.001) is applied to the third convolutional layer.
  - MaxPooling2D with a pool size of 2x2 is used.
  - The fourth convolutional layer applies 128 filters of size 3x3 with ReLU activation.
  - Regularization (L2 regularization with a coefficient of 0.001) is applied to the fourth convolutional layer.
  - MaxPooling2D with a pool size of 2x2 is used.
  - Dropout with a rate of 0.25 is applied.
- 4. Flatten Layer:
  - The output from the previous layer is flattened into a 1-dimensional vector.
- 5. Dense Layers:
  - The first dense layer has 1024 units with ReLU activation.
  - Regularization (L2 regularization with a coefficient of 0.001) is applied to the first dense layer.
  - Batch Normalization is applied after the first dense layer.
  - Dropout with a rate of 0.5 is applied.
  - The second dense layer has 512 units with ReLU activation.
  - Regularization (L2 regularization with a coefficient of 0.001) is applied to the second dense layer.
  - Batch Normalization is applied after the second dense layer.

- Dropout with a rate of 0.5 is applied.
- The third dense layer has 256 units with ReLU activation.
- Regularization (L2 regularization with a coefficient of 0.001) is applied to the third dense layer.
- Batch Normalization is applied after the third dense layer.
- Dropout with a rate of 0.5 is applied.

• 6.Output Layer:

- The output layer has 3 units representing the 3 classes: Negative, Positive, and Surprise.
- The softmax activation function is applied to produce class probabilities.

Overall, the model utilizes convolutional layers to extract hierarchical features from the input images, followed by fully connected layers for classification. Batch normalization is incorporated after each Conv2D and Dense layer for better training stability. Dropout layers help prevent overfitting by randomly disabling some neurons during training. The model utilizes regularization techniques such as L2 regularization with a coefficient of 0.001 to help prevent overfitting the loss function employed is categorical cross-entropy. The accuracy metric is used to evaluate the model during training The model is compiled with the RMSprop optimizer and categorical cross-entropy loss. The model is compiled using the 'rmsprop' optimizer and the categorical cross-entropy loss function. Early stopping is applied during training based on the validation accuracy to prevent overfitting. The architecture is detailed in 3.1

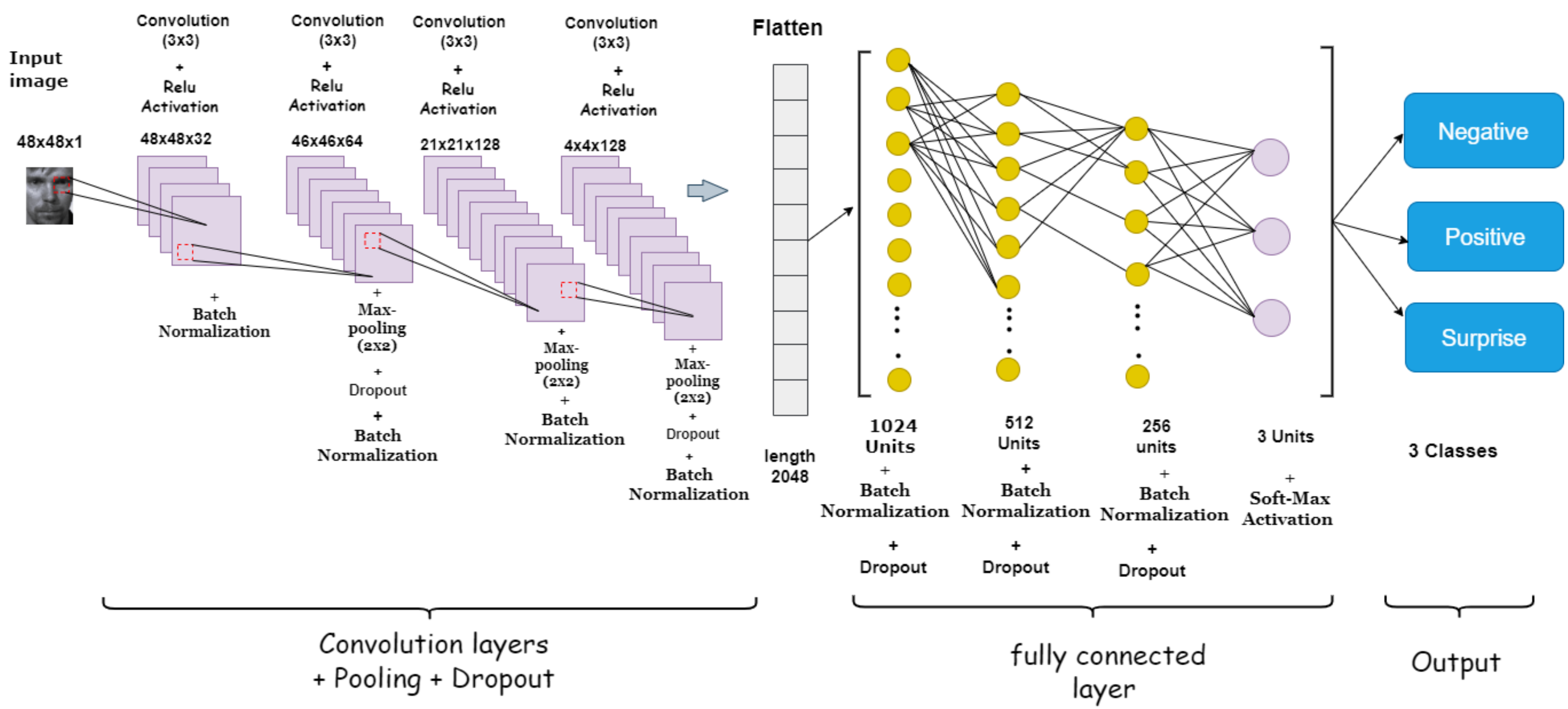


Figure 3.1: CNN Architecture SMIC

2.2 **CASIA Dataset:**

The architecture of the model consists of the following layers and parameters:

- 1. Convolutional Layers:
  - The first convolutional layer (Conv2D) has 32 filters and a kernel size of(3, 3).
  - The second convolutional layer has 64 filters and a kernel size of (3, 3).
  - The third and fourth convolutional layers have 128 filters each and a kernel size of (3, 3).
  - The activation function used in all convolutional layers is'relu', which introduces non-linearity to the model.
- 2. Pooling Layer:
  - A MaxPooling2D layer is used after the first convolutional layer and has a pool size of (2, 2).
  - This layer performs downsampling by taking the maximum value within each 2x2 region.
- 3. Dropout Layers:
  - Two Dropout layers are added after the first and second pooling layers with a dropout rate of 0.25.
  - Another Dropout layer is added after the fourth convolutional layer with a dropout rate of 0.25.
  - Dropout is a regularization technique that randomly sets a fraction of input units to 0 during training, which helps prevent overfitting.
- 4. Flatten Layer:
  - After the last pooling layer, a Flatten layer is used to convert the 2D feature maps into a 1D feature vector.

- This layer prepares the data for the fully connected layers that follow.

- 5. Fully Connected Layers:
  - The first fully connected layer (Dense) has 1024 units/neurons.
  - The second fully connected layer has 6 units, which corresponds to the number of output classes.
  - The activation function used in the fully connected layers is 'relu', except for the output layer.

- 6. Output Layer:
  - The output layer is a Dense layer with 6 units, representing the 6 different micro-facial expressions to be classified.
  - The activation function used in the output layer is 'softmax', which normalizes the outputs into a probability distribution over the classes.
  - The model is compiled with the `'adam'` optimizer, which is an adaptive learning rate optimization algorithm.
  - The loss function used is 'categorical_crossentropy', suitable for multi-class classification problems.
  - The metric used for evaluation is 'accuracy', which measures the percentage of correctly classified samples.

The model architecture consists of multiple convolutional layers for feature extraction, pooling layers for downsampling, dropout layers for regularization, and fully connected layers for classification. The activation function 'relu' is used for most layers except the output layer, which uses 'softmax' for multiclass classification. visual details are showen in figure 3.2

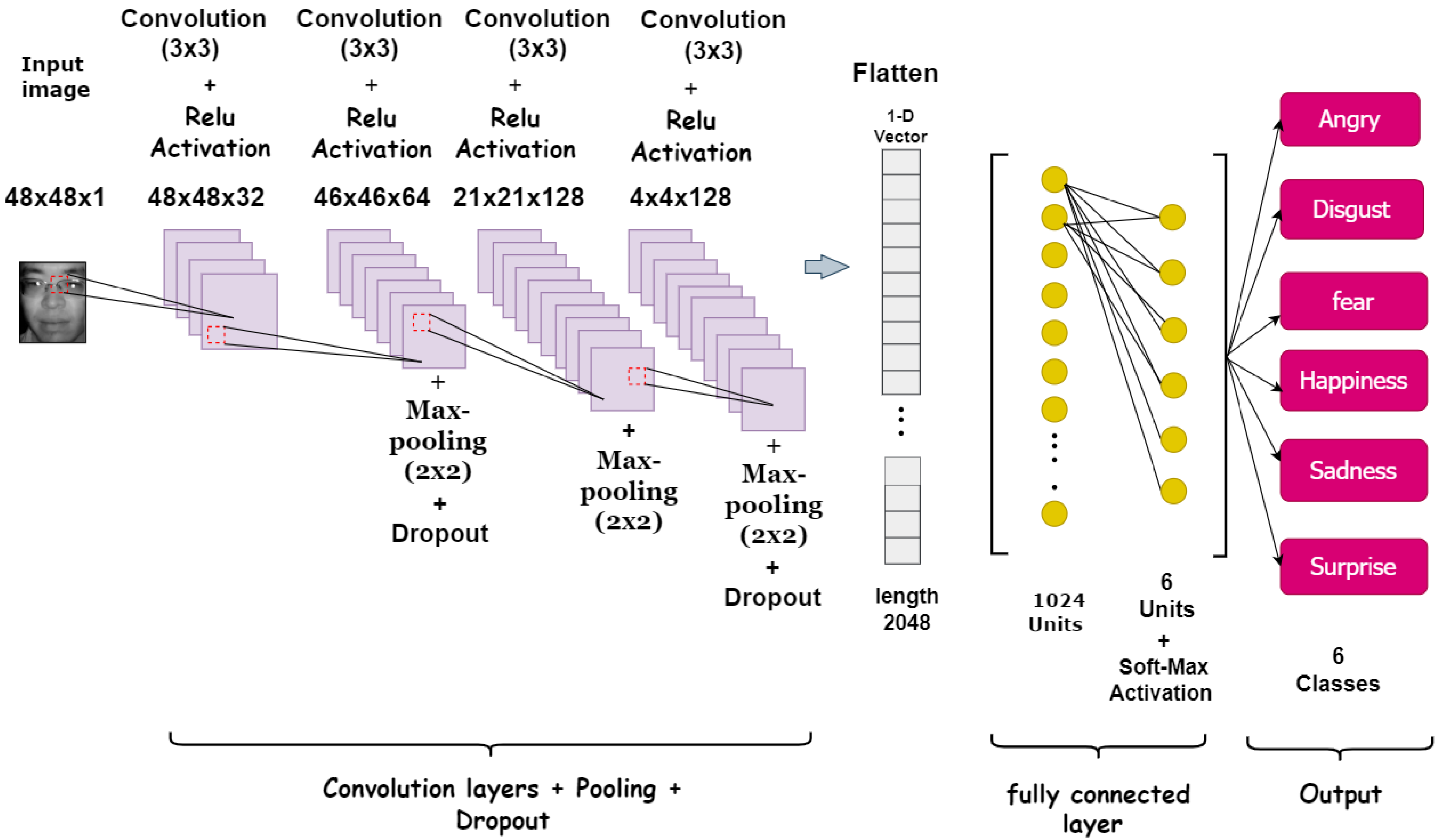


Figure 3.2: CNN Architecture CASIA

3. **Model Training**

   The CASIA model and the SMIC model were compiled with different optimizers. The CASIA model utilized the Adam optimizer, while the SMIC model used the RMSPROP optimizer. Both models employed categorical cross-entropy as the loss function and selected accuracy as the evaluation metric.

   During the training process, the models were fitted to the training images (train_images) and their corresponding labels (y_train) using a batch size of 40. The training was conducted for 40 epochs for CASIA dataset and 30 for SMIC dataset, and the performance was continuously monitored using the validation data (test_images, y_test).

   To prevent overfitting and assess the models' performance, the early stopping callback, represented by 'early', was employed during the training process. This callback monitored the training progress and halted it early if the desired criteria, such as the absence of improvement in the validation loss, were met.

   By following this configuration, both models were able to learn from the training data, adapt their parameters, and minimize the loss function, ultimately improving their accuracy on the given task.

```
# Compile CASIA model
model.compile(optimizer='adam', loss='categorical_crossentropy',
    metrics['accuracy'])
# Compile SMIC model
model.compile(optimizer='rmsprop', loss='categorical_crossentropy',
     metrics=['accuracy'])
# Callback
early=EarlyStopping(monitor='val_accuracy',min_delta=0,patience=25,
    verbose=1, mode='auto')
# Train SMIC model
model.fit(train_images, y_train, batch_size=40, epochs=30,
    validation_data=(test_images, y_test),callbacks=[early])
# Train CASIA model
model.fit(train_images, y_train, batch_size=40, epochs=40,
    validation_data=(test_images, y_test),callbacks=[early])
```

3.3 Conclusion

In this chapter, we have described the implementation of the proposed facial micro-expression recognition using the SMIC and Oulu-CASIA datasets. Our approach involved preprocessing the data to remove noise and artifacts, followed by feature extraction using the Adaptive Local Binary Pattern (ALBP) algorithm. We then used these features to train a classification model, specifically a convolutional neural network (CNN). The architecture of the CNN was designed to adapt to the features of each dataset.

The upcoming chapter will delve into the evaluation and experimental results, providing deeper insights into system performance and validation of the proposed methodologies.

Chapter 4

Experiments and Results

4.1 Introduction

We have done an implementation of the CNN methodology in order to do an experiment. In this chapter, we provide the technical details of the experiment and the results obtained

4.2 Technical Requirements

The technical requirements for this project are:

4.2.1 Hardware Requirements

- **Dell Inspiron 13**
  Processor: Intel(R) Core(TM) i7-7500U @ 2.70GHz 2.90 GHz;
  RAM: 8.00 GB
  System type: 64-bit

- **Dell**
  Processor: Intel(R) Core(TM) i5-8350U CPU @ 1.70GHz 1.90 GHz;
  RAM: 16 GB
  System type: 64-bit

4.2.2 Software Requirements

1. operating system: windows10, windows11
2. Programming Language: Python (version 3.11)
3. Integrated Development Environment (IDE): PyCharm (version 2022.2.2)
4. LaTeX: Typesetting System on Overleaf: An online collaborative platform for writing, editing, and publishing LaTeX documents.
5. Libraries:
   - Keras (version 2.12.0): A high-level neural networks API that simplifies the process of building and training deep learning models.
   - TensorFlow (version 2.12.0): An open-source deep learning framework that provides a wide range of tools and capabilities for developing and deploying machine learning models.
   - Opencv (version 4.7.0.72): A computer vision library that offers a wide range of functions and algorithms for image and video processing.
   - Numpy (version 2.23.5): A fundamental package for scientific computing with Python, providing support for multi-dimensional arrays and mathematical operations.
   - Sklearn (version 0.0.post5): A machine learning library that provides various tools for classification, regression, clustering, and dimensionality reduction.
   - Tkinter: The standard Python interface to the Tk GUI toolkit, used for creating graphical user interfaces (GUIs).

4.3 Micro Expression System GUI

Our graphical user interface (GUI) is designed to classify images from the testing sets that our system has not encountered before. As we have worked with two datasets, the initial window provides us with the capability to select the dataset we wish to process

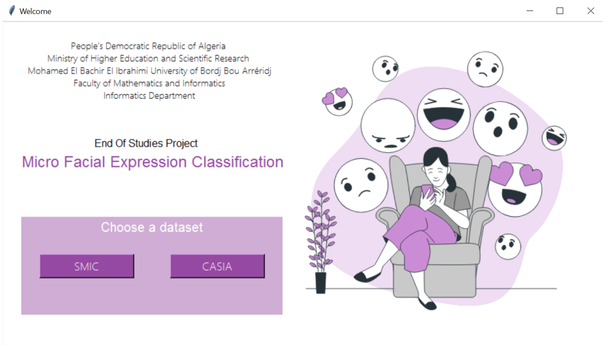


Figure 4.1: Welcome Interface.

- Upon selecting the "SMIC Dataset" button, a secondary window will emerge, presenting the principal modules of our system through a quartet of buttons: browsing, preprocessing, extraction, and classify

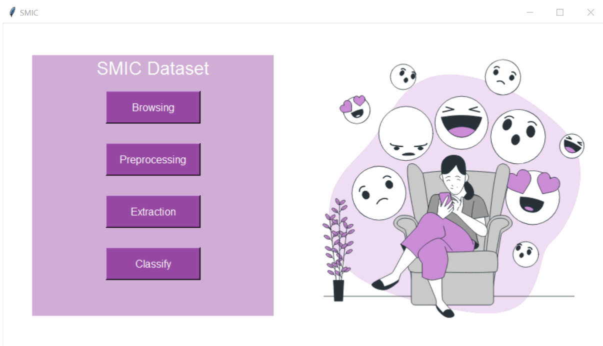


Figure 4.2: SMIC Interface.

- By clicking on the “Browsing ” button dialogue window is open so we can choose an image

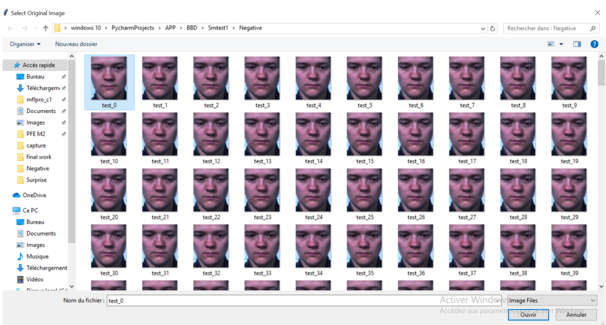


Figure 4.3: Dialogue window.

- After choosing any random image and clicking “open” it’s displayed as follows

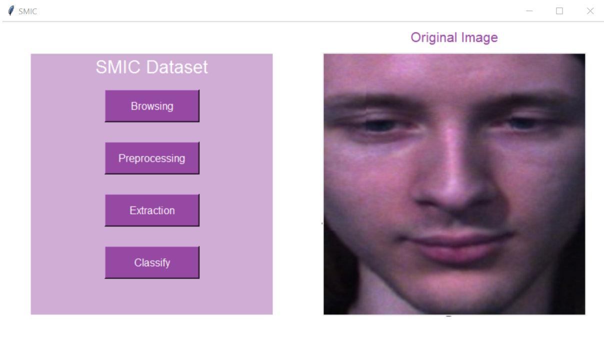


Figure 4.4: Displaying Original Image .

- Then, we click on the "Preprocessing" button to apply our preprocessing method on that specific image like follows

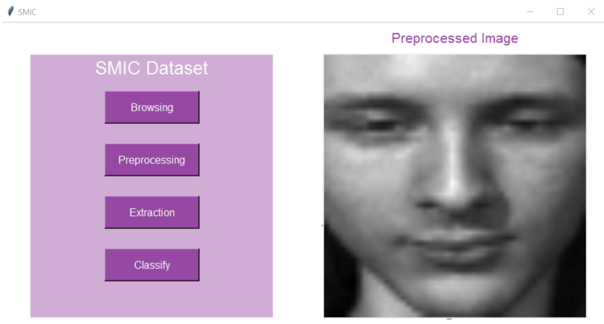


Figure 4.5: Displaying Preprocessed Image .

- After that, we apply our extraction method by clicking on the "Extraction" button result are shown in next figure

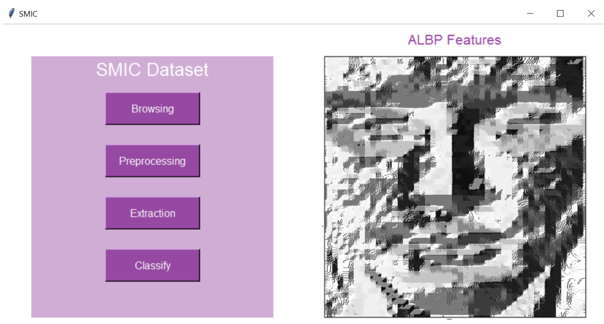


Figure 4.6: Displaying ALBP Features .

- Finally, the classification is made by the button "Classify" which displays the predicted class and also the actual class of the chosen image

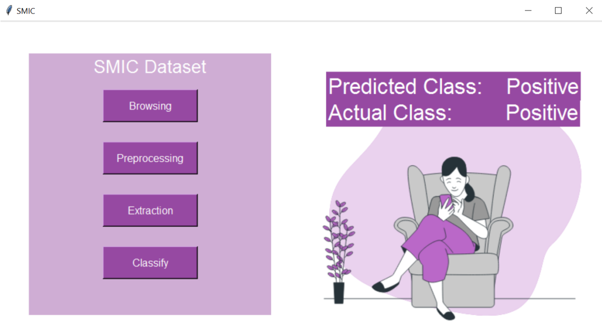


Figure 4.7: Displaying Classification class for SMIC Dataset.

- The same process for the second button a third window is open to apply the same methods on the chosen image from the CASIA dataset

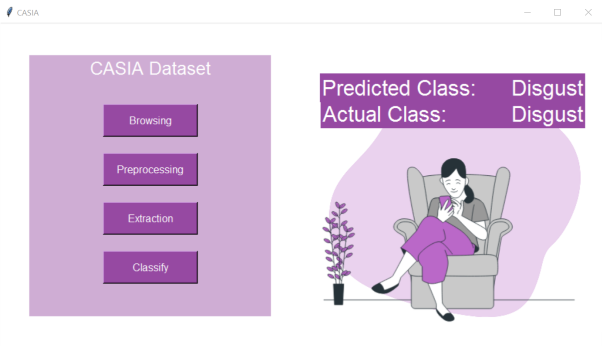


Figure 4.8: Displaying Classification class for CASIA Dataset .

4.4 Numerical Results

4.4.1 Validation Metrics

The performance of our classification model was evaluated using various validation metrics, including accuracy, train accuracy, and validation accuracy. These metrics provide an assessment of the model's overall performance and its ability to accurately classify instances.

Accuracy is calculated as the ratio of correctly classified instances to the total number of instances in the dataset. The equation for accuracy is:

$$\text{Accuracy} = \frac{\text{Number of correctly classified instances}}{\text{Total number of instances}} \times 100\%$$

Furthermore, we monitored the training accuracy and validation accuracy throughout the training process to understand the model's learning and generalization abilities.

4.4.2 Results

We worked with two datasets: OuluCASIA and SMIC

- CASIA Dataset:
  The Oulu-CASIA database includes 80 individuals between the ages of 23 and 58 who expressed surprise, happiness, sadness, anger, fear, and disgust. Males make up 73.8 % of the subjects. The subjects were instructed to sit in the observation room on a chair so that they were facing the camera. About 60 centimeters separate the camera from your face. The photography hardware produces images with a resolution of 320 * 240 pixels at a rate of 25 frames per second.
  In our experiments, nine images per subject for each expression (Anger, Disgust, Fear, Happiness, Sadness, Surprise) are used to carry out and validate our proposed system. A set of 640 images per expression was selected as a training set i.e. 640 * 6 = 3840 and 80 images per expression i.e. 80 * 6=480 was selected as a testing set.

- SMIC Dataset :
  Spontaneous Micro-expression Database SMIC, includes 164 microexpression video clips elicited from 16 participants. The camera records at a rate of 100 frames per second, with an average facial resolution of 160 * 140.Micro Expression detection and recognition performance are provided as baselines. SMIC contains sufficient data to enable comprehensive testing of automatically analyzing micro-expressions.The SMIC dataset includes three emotion labels: positive, negative, and surprise.
  In our study, we used 1428 images PER CLASS i.e. 1428 * 3=4284 for training and 357 per class i.e. 357 * 3=1071 for testing.

Table 4.1 summarizes the results of our proposed micro-expressions recognition system using the SMIC database. We also conducted in this table a comparison between our findings and existing literature in the field that uses the same database. The same comparison is performed in Table 4.2 This table presents the achieved results of the developed micro expressions system using the CASIA database.

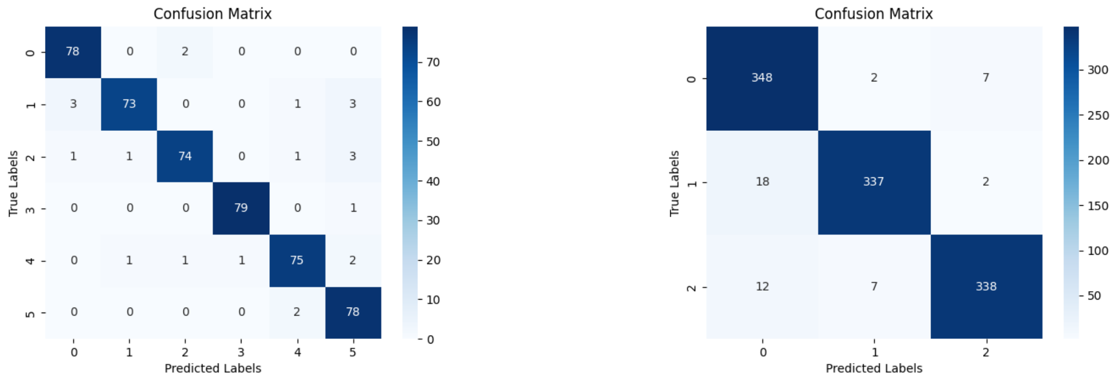
Table 4.1: SMIC Result.

Method	Accuracy
ROI+VGG19[39]	72,62%
HIGO+Mag[31]	81.69%
Riesz pyramid[40]	89,80%
Our approach	95,18%

Table 4.2: CASIA Result.

Method	Accuracy
Static MTL [31]	89,6%
CNN [41]	93,4%
Our approach	95,21%

To gain a deeper understanding of the classification results, we implemented the confusion matrix for both the SMIC and CASIA datasets. The confusion matrix provides a detailed breakdown of the predictions, highlighting the true positives and misclassifications for each emotion category. Figure 4.9b illustrates the confusion matrix for the SMIC dataset, where the labels "0," "1," and "2" represent the "negative," "positive," and "surprise" emotions, respectively. The diagonal elements of the matrix, highlighted in dark blue, represent the true positives, indicating the number of instances correctly classified for each emotion category. Conversely, the cells highlighted in light blue represent the misclassifications, indicating the number of instances that were incorrectly assigned to a different emotion category.
Similarly, Figure 4.9a illustrates the confusion matrix for the CASIA dataset. The dark blue diagonal elements represent the true positives, indicating the instances correctly classified for each emotion category, such as "Anger," "Disgust," "Fear," "Happiness," "Sadness," and "Surprise." However, it is important to note the presence of misclassifications, indicated by the light blue cells. These misclassifications highlight the challenges faced by our model in accurately distinguishing between certain emotion categories. Further analysis is required to understand the underlying factors contributing to these misclassifications.

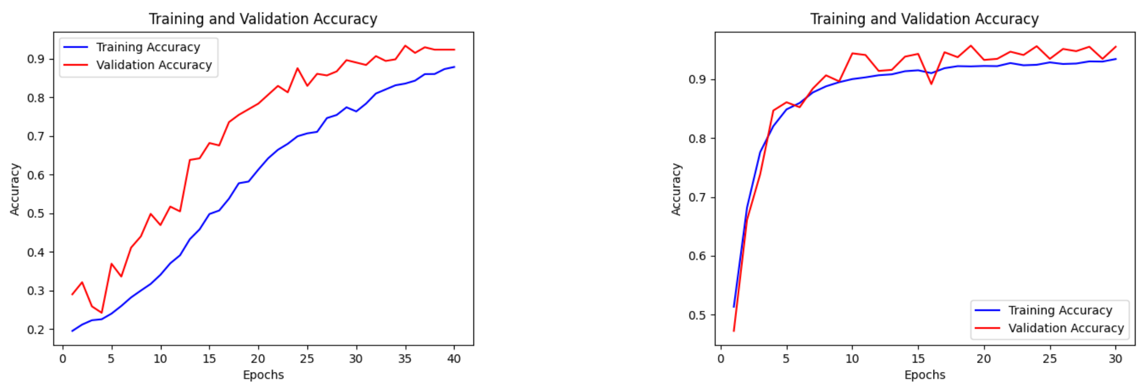


(a) CASIA Confusion Matrix

(b) SMIC Confusion Matrix

Figure 4.9: Confusions Matrix

Furthermore, to assess the learning progress and generalization capability of our model, we plotted the training accuracy and validation accuracy graphs as a function of epochs. These graphs provide valuable insights into the model's performance and its ability to learn from the training data while avoiding overfitting. Figure 4.10b and 4.10a represent the training accuracy graph in blue, which shows the accuracy of our model on the training dataset as training progresses over epochs. Simultaneously, the validation accuracy graph in red, showcases the accuracy of our model on the validation dataset throughout the training process, for SMIC and CASIA dataset, respectively .



(a) CASIA Training and Validation Accuracy graph

(b) SMIC Training and Validation Accuracy graph

Figure 4.10: Training and Validation Accuracy graphs

4.5 Discussion

1. SMIC Result Discussion:

The table of results 4.1 presents the accuracy values obtained from different approaches for comparison. The approaches in the table achieved accuracy results of 72.62%, 81.69%, and 89.80%, respectively.

In contrast, our approach yielded a significantly higher accuracy of 95%. This indicates that our method outperformed the other approaches in accurately classifying instances.

The higher accuracy achieved by our approach can be attributed to several factors. Firstly, it may be the ALBP algorithm for feature extraction that effectively captured the underlying patterns and characteristics of the data. This allowed our model to make more precise predictions and achieve a higher accuracy rate.

Furthermore, it is possible that we applied specific preprocessing techniques that enhanced the discriminative power of the features and improved the overall accuracy of the model. Additionally, it is worth considering the impact of the parameters used in the training of the model, as they can significantly affect the model's performance.

Overall, the significantly higher accuracy achieved by our approach compared to the other approaches in the table suggests that our method excels in accurately classifying instances. This highlights the effectiveness of our chosen feature extraction algorithm, the application of suitable preprocessing techniques, and the optimization of model parameters.

2. SMIC Confusion Matrix Discussion:

    - based on the confusion matrix in 4.9b the Negative Class demonstrates a commendable level of precision in its classification results. Out of the 357 instances belonging to the Negative Class, 348 instances are correctly classified as negatives, indicating a significant number of True Positives (TP). This highlights the model's ability to effectively recognize and identify instances that possess the distinct characteristics associated with the Negative Class. However, it is important to acknowledge that there are nine instances within the Negative Class that have been misclassified. Among these misclassifications, two instances are mistakenly labeled as positive and 7 instances are classified as surprise. These misclassifications emphasize the inherent complexity and challenges in accurately distinguishing between certain classes, where some instances may exhibit overlapping characteristics

    - The Positive Class, as observed, exhibits a notable level of precision in its classification results. Within this class, there are 337 instances that are accurately identified and classified as positives, reflecting a considerable number of True Positives (TP). This indicates that the classification model effectively recognizes the distinctive characteristics associated with the Positive Class, leading to the successful identification of the majority of instances belonging to this class. and in this particular case, two instances are misclassified as 'Surprise'. This isolated misclassification demonstrates the inherent complexity and challenges in accurately distinguishing between certain classes. Additionally, it is worth noting that there are 18 instances that are incorrectly classified as Negative, contributing to the overall misclassification count. Despite these misclassifications, the Positive Class maintains a substantial level of accuracy in classification, with the vast majority of instances being correctly identified and classified as positives.

- The Surprise Class comprises 338 instances that are correctly classified as surprise, indicating True Positives (TP). However, there is a set of 12 instances that are misclassified as negatives, presenting a challenge in accurately distinguishing them. Additionally, the model incorrectly classifies 7 instances as positives. These misclassifications emphasize the complexity involved in correctly identifying and differentiating the Surprise Class from other classes. Nevertheless, the classification model demonstrates a significant success rate in accurately labeling the majority of instances within the Surprise Class as surprise, highlighting its effectiveness in capturing the unique characteristics associated with this class.

3. Graph Analysis:

The training and validation accuracy graph presented in Figure 4.10b showcases the performance of our model on both the training and validation sets as training progresses over epochs.

Throughout the training process, the training accuracy curve, represented by the blue line, shows the model's accuracy on the training set. As expected, we observe an upward trend in the training accuracy, indicating that the model improves its performance as it learns from the training data. This increase in training accuracy signifies the model's ability to capture and understand the patterns and features present in the training dataset. On the other hand, the validation accuracy curve, depicted by the red line, represents the model's accuracy on the validation set. The validation accuracy serves as a metric for assessing how well the model generalizes to new, unseen data.

Analyzing the graph, we observe that both the training and validation accuracy curves show a steady increase and closely follow each other, it indicates that the model is learning effectively and has a good ability to generalize to unseen data. This alignment suggests that the model is not overfitting and can reliably classify instances from both the training and validation datasets.

4. CASIA Result Discussion:

   The table of results 4.2 displays the accuracy values obtained from different approaches for comparison. The approaches in the table achieved accuracy results of 89.6% and 93.4%, while our approach yielded a higher accuracy of 95.21%. This demonstrates that our method outperformed the other approaches in accurately classifying instances.

   The superior accuracy achieved by our approach can be attributed to several factors. Firstly, it may be due to the utilization of advanced algorithms or models that effectively captured the underlying patterns and characteristics of the data. These algorithms could have enhanced the model's ability to make precise predictions and achieve a higher accuracy rate.

   Additionally, we may have implemented specific preprocessing techniques tailored to the dataset, which enhanced the discriminative power of the features and improved the overall accuracy of the model. These preprocessing techniques could have included data normalization, feature scaling, or outlier removal, among others. By ensuring the data was appropriately prepared, our model was better equipped to handle the complexities and variations present in the dataset, leading to improved accuracy.

   Furthermore, the optimization of model parameters played a crucial role in achieving the higher accuracy. By fine-tuning the model's hyperparameters, such as learning rate, regularization parameters, or network architecture, we were able to achieve a better balance between model complexity and generalization, resulting in improved accuracy.

   In conclusion, our approach attained a higher accuracy of 95.21% compared to the results of 89.6% and 93.4% obtained by other approaches in the table. This highlights the effectiveness of our chosen algorithms, preprocessing techniques, and parameter optimization in accurately classifying instances. The superior accuracy achieved by our approach demonstrates our ability to effectively leverage the data and develop a robust model for the task at hand.

5. CASIA Confusion Matrix Discussion:

   - The Anger Class in the CASIA dataset exhibits a high level of accuracy in classification, a substantial number specifically 78 instances, are correctly identified and classified as Anger representing True Positives (TP). However, it is worth noting that there are two instances misclassified as Fear. Despite these misclassifications, the overall performance of the model in correctly identifying the Anger Class is commendable, with a majority of instances being accurately labeled and classified.

   - The Disgust Class within the CASIA dataset demonstrates a respectable level of accuracy in classification. Out of the instances belonging to the Disgust Class, 73 instances are correctly classified, indicating a considerable number of True Positives (TP). However, it is important to acknowledge that there are seven instances misclassified within this class. Among these misclassifications, three instances are mistakenly labeled as Anger, three as Surprise, and one as Sadness. Despite these misclassifications, the majority of instances within the Disgust Class are accurately identified and classified, highlighting the model's effectiveness in recognizing and distinguishing the distinctive characteristics associated with this class.

   - The Fear Class, we observe that it also exhibits a notable level of precision in classification. Out of the instances belonging to the Fear Class, 74 instances are correctly classified as Fear, representing True Positives (TP). However, there are six instances that are misclassified. Among these misclassifications, three instances are classified as Surprise, one as Anger, one as Disgust, and one as Sadness. Despite these misclassifications, the Fear Class maintains a substantial level of accuracy, with the majority of instances being correctly identified and classified as Fear.

   - The Happiness Class, we find that it demonstrates a high level of accuracy in classification. Out of the instances belonging to the Happiness Class, 79 instances are correctly classified as Happiness, representing True Positives (TP). However, there is only one instance that is misclassified as surprise. The overall performance of the model in correctly identifying the Happiness Class is commendable, with the vast majority of instances accurately labeled and classified.

- The Sadness Class, we observe a notable level of precision in its classification results. Out of the instances belonging to the Sadness Class, 75 instances are correctly classified as Sadness, reflecting a significant number of True Positives (TP). However, there are five instances that are misclassified. Among these misclassifications, two instances are classified as Surprise, one as Disgust, one as Fear, and one as Happiness. Despite these misclassifications, the Sadness Class maintains a substantial level of accuracy in classification, with the majority of instances being correctly identified and classified as Sadness.

- The Surprise Class showcases a commendable level of accuracy in classification. Out of the instances belonging to the Surprise Class, 78 instances are correctly classified as Surprise, representing True Positives (TP). However, there are two instances that are misclassified as Sadness. The overall performance of the model in correctly identifying the Surprise Class is impressive, with the majority of instances accurately labeled and classified.

6. Graph Analysis:

   The performance of our model on both the training and validation sets as training develops across epochs is displayed in the training and validation accuracy graph shown in Figure 4.10a.

   The training accuracy curve, shown by the blue line, displays the model's accuracy on the training set over an interval of training. As predicted, we see an increase in the training accuracy, which shows that the model becomes better as it gains knowledge from the training data. The model's capacity to recognize and comprehend the patterns and characteristics included in the training dataset is shown by the rise in training accuracy.

   On the other side, the red line as the validation accuracy curve shows how accurate the model was on the validation set. A gauge for determining how effectively the model generalizes to fresh, untested data is the validation accuracy.

   When we analyze the graph, we note that both the training and validation accuracy curves exhibit continuous growth and closely resemble one another; this shows that the model is successfully learning and has a strong capacity for generalization to unknown data. This alignment indicates that the model is capable of accurately classifying instances from both the training and validation datasets, indicating that it is not overfitting.

Chapter 5

General Conclusion

5.1 Contributions

In this study, we used CNN as the baseline recognition model. Features were extracted using Adaptive Local Binary Patterns (ALBP).The utilization of a CNN model has proven to be instrumental in extracting features and learning intricate patterns, enabling accurate classification and recognition of micro expressions. We have made significant strides in the field of micro-facial expression analysis, paving the way for enhanced emotion recognition and understanding. The findings and methodologies presented here can potentially contribute to various domains, including psychology, human-computer interaction, and affective computing. The ability to decipher subtle emotional cues has the potential to enhance human interaction, improve user experiences, and contribute to the development of more empathetic and intuitive systems.

5.2 Limitations

However, several limitations must be acknowledged. Firstly, the performance of our micro features descriptor may vary depending on factors such as lighting conditions, image quality, and individual differences in facial structure. While our approach has demonstrated robustness under controlled environments, further investigation and refinement are necessary to enhance its generalizability across diverse settings. Addressing these challenges could lead to more robust and accurate micro facial expression classification systems.

5.3 Future work and perspectives

Looking ahead, there are several promising avenues for future work and potential extensions of our research. One direction is to explore the integration of multimodal inputs, such as incorporating audio or physiological data, to enhance the accuracy and robustness of our facial expression recognition system. This integration would enable a more comprehensive and holistic understanding of emotions.

Additionally, the deployment of our model in real-time applications is an area of great interest. Optimizing its computational efficiency and developing efficient inference strategies will enable us to integrate our system into various real-time scenarios, such as interactive virtual environments or emotion-aware systems. Therefore, Future work should focus on how to improve the performance of emotion classification with fewer parameters and advanced CNN models . This could involve conducting extensive experiments with different layer configurations, exploring regularization techniques, and optimizing hyperparameters. Additionally, augmenting the dataset with more diverse and representative examples can help improve the model's generalizability and accuracy. Incorporating data from diverse populations and cultural contexts will enhance the model's performance across different demographics, ensuring its reliability and effectiveness in real-world scenarios.

References

[1] J. Yang, D. S. Park, S. Yoon, Y. Chen, and C. Zhang, *Machine Learning and Biometrics*. Rijeka: IntechOpen, Aug 2018. [Online]. Available: https://doi.org/10.5772/intechopen.71297

[2] P. Cerna, M. C. Abas, H. Gebreziagbher, and M. Gebremariam, "Machine learning biometric attendance system using fingerprint fuzzy vault scheme algorithm and multi-task convolution neural network face recognition algorithm," *International Journal of Computer Applications*, vol. 179, pp. 1–6, 06 2018.

[3] J. V. M. Gordana Sendić MD, "Kenhub," https://www.kenhub.com/en/library/anatomy/the-facial-muscles, 2023, accessed: April 12, 2023.

[4] X. Ben, Y. Ren, J. Zhang, S.-J. Wang, K. Kpalma, W. Meng, and Y.-J. Liu, "Video-based facial micro-expression analysis: A survey of datasets, features and algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 9, pp. 5826–5846, 2022.

[5] "Review of micro-expression spotting and recognition in video sequences," *Virtual Reality Intelligent Hardware*, vol. 3, no. 1, pp. 1–17, 2021.

[6] X. Hongxia, L. Lo, H.-H. Shuai, and W.-H. Cheng, "An overview of facial micro-expression analysis: Data, methodology and challenge," 12 2020.

[7] D. Day, *Biometric Applications, Overview*. Boston, MA: Springer US, 2015, pp. 169–174. [Online]. Available: https://doi.org/10.1007/978-1-4899-7488-4_20

[8] K. Vishi and S. Yildirim Yayilgan, "Multimodal biometric authentication using fingerprint and iris recognition in identity management," pp. 334–341, 10 2013.

[9] U. Saeed, "Facial micro-expressions as a soft biometric for person recognition," *Pattern Recognition Letters*, vol. 143, 01 2021.

[10] W. Rinn, "The neuropsychology of facial expression: A review of the neurological and psychological mechanisms for producing facial expressions," *Psychological bulletin*, vol. 95, pp. 52–77, 02 1984.

[11] H. H. Matsumoto, D., "Evidence for training the ability to read microexpressions of emotion." *Motivation and Emotion SP*, p. 181–191, 2011.

[12] W.-J. Yan, X. Li, S.-J. Wang, G. Zhao, Y.-J. Liu, Y.-H. Chen, and X. Fu, "Casme ii: An improved spontaneous micro-expression database and the baseline evaluation," *PloS one*, vol. 9, p. e86041, 01 2014.

[13] P. Ekman and M. O'Sullivan, "From flawed self-assessment to blatant whoppers: The utility of voluntary and involuntary behavior in detecting deception," *Behavioral sciences the law*, vol. 24, pp. 673–86, 09 2006.

[14] S. Porter and L. ten Brinke, "Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions," *Psychological Science*, vol. 19, no. 5, pp. 508–514, 2008. [Online]. Available: http://www.jstor.org/stable/40064786

[15] Z. Ming, J. Xia, M. M. Luqman, J.-C. Burie, and K. Zhao, "Dynamic multi-task learning for face recognition with facial expression," 2019.

[16] C. Darwin, S. Pozzi, and R. Benoit, *L'expression des émotions chez l'homme et les animaux*. Editions Complexe, 1890. [Online]. Available: https://books.google.dz/books?id=9kwp60oGkiYC

[17] J. Wojciechowski, M. Stolarski, and G. Matthews, "Emotional Intelligence and Mismatching Expressive and Verbal Messages: A Contribution to Detection of Deception," *PLoS ONE*, vol. 9, no. 3, p. e92570, Mar. 2014.

[18] X. Zeng, Q. Wu, S. Zhang, Z. Liu, Q. Zhou, and M. Zhang, "A false trail to follow: Differential effects of the facial feedback signals from the upper and lower face on the recognition of micro-expressions," *Frontiers in Psychology*, vol. 9, 2018. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02015

[19] W.-J. Yan, Q. Wu, Y.-H. Chen, J. Liang, and X. Fu, "How fast are the leaked facial expressions: The duration of micro-expressions," *Journal of Nonverbal Behavior*, vol. 37, 12 2013.

[20] B. Bhushan, *Study of Facial Micro-expressions in Psychology*, 10 2015, pp. 265–286.

[21] P. Husak, J. Cech, and J. Matas, "Spotting facial micro-expressions "in the wild"," in *Proc. Computer Vision Winter Workshop*, 2017, https://cmp.felk.cvut.cz/~cechj/ME/.

[22] Y. Li, X. Huang, and G. Zhao, "Can micro-expression be recognized based on single apex frame?" in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 3094–3098.

[23] Y.-H. Oh, J. See, A. C. Le Ngo, R. Phan, and V. Baskaran, "A survey of automatic facial micro-expression analysis: Databases, methods, and challenges," *Frontiers in Psychology*, vol. 9, 06 2018.

[24] A. Davison, W. Merghani, C. Lansley, C.-C. Ng, and M. H. Yap, "Objective micro-facial movement detection using facs-based regions and baseline evaluation," in *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*, 2018, pp. 642–649.

[25] S. Liong, J. See, K. Wong, and R. Phan, "Automatic micro-expression recognition from long video using a single spotted apex," 03 2017, pp. 345–360.

[26] A. K. Davison, C. Lansley, N. Costen, K. Tan, and M. H. Yap, "Samm: A spontaneous micro-facial movement dataset," *IEEE Transactions on Affective Computing*, vol. 9, no. 1, pp. 116–129, 2018.

[27] Y. Zhao and J. Xu, "An improved micro-expression recognition method based on necessary morphological patches," *Symmetry*, 2019.

[28] C. Wang, M. Peng, T. Bi, and T. Chen, "Micro-attention for micro-expression recognition," 2019.

[29] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[30] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 915–928, 2007.

[31] X. Li, X. Hong, A. Moilanen, X. Huang, T. Pfister, G. Zhao, and M. Pietikainen, "Towards reading hidden emotions: A comparative study of spontaneous micro-expression spotting and recognition methods," *IEEE Transactions on Affective Computing*, vol. PP, pp. 1–1, 02 2017.

[32] D. Patel, X. Hong, and G. Zhao, "Selective deep features for micro-expression recognition," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, 2016, pp. 2258–2263.

[33] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv 1409.1556*, 09 2014.

[34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[35] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7132–7141.

[36] C. Wang, "Dual temporal scale convolutional neural network for micro-expression recognition," *Frontiers in Psychology*, vol. 8, p. 1745, 10 2017.

[37] W. Merghani, A. K. Davison, and M. H. Yap, "A review on facial micro-expressions analysis: Datasets, features and metrics," *ArXiv*, vol. abs/1805.02397, 2018.

[38] M. Sharif, M. Raza, J. Shah, Y. Mussarat, and S. Fernandes, *An Overview of Biometrics Methods*, 07 2019, pp. 15–35.

[39] T. T. Quynh Le, T.-K. Tran, and M. Rege, "Dynamic image for micro-expression recognition on region-based framework," in *2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI)*, 2020, pp. 75–81.

[40] C. A. Duque, O. Alata, R. Emonet, A.-C. Legrand, and H. Konik, "Micro-Expression

Spotting using the Riesz Pyramid," in *WACV 2018*, Lake Tahoe, United States, Mar. 2018. [Online]. Available: https://hal.science/hal-01699355

[41] K. Patel, D. Mehta, C. Mistry, R. N. Gupta, and M. Alazab, "Facial sentiment analysis using ai techniques: State-of-the-art, taxonomies, and challenges," *IEEE Access*, vol. 8, pp. 90 495–90 519, 2020.