University of Mohamed Khider Biskra
Faculty of Sciences and Technology
Department of Electrical Engineering

MASTER'S DEGREE

Sciences and Technology
Telecommunications
Network and Telecommunication

Presented and submitted by:
Mohamed El Amine Laiadi
On : 26 Jun 2022

AUTOMATIC FACIAL EXPRESSION CLASSIFICATION WITH CNN

Jury :

Mr.	Abdelkarim Ouafi	Pr University of Biskra	**President**
Ms.	Meriem Fedias	MCB University of Biskra	**Supervisor**
Mr.	Zine-eddine Baarir	Pr University of Biskra	**Examiner**

Academic Year : 2021-2022

University of Mohamed Khider Biskra
Faculty of Sciences and Technology
Department of Electrical Engineering

MASTER’S DEGREE

Sciences and Technology
Telecommunications
Network and Telecommunication

AUTOMATIC FACIAL EXPRESSION CLASSIFICATION WITH CNN

Presented by:

Mohamed El Amine Laiadi

Favorable opinion of the supervisor:

Ms. Meriem Fedias

Favorable opinion of the Jury President:

Mr. Abdelkarim Ouafi

Stamp and signature

Acknowledgments

I wish to thank my supervisor Fedias meriem who was more than generous with his expertise and precious time, and for guiding me and encouraging me to do professional work. A special thanks to the jury member for criticizing my thesis. I want to thank the following people for helping me with their precious I would like to express my sincere gratitude to Mohamed Khider University and all the teachers in the electrical engineering department for all the considerate guidance. I would like to thank other persons my brother Oualid who encouraged me all the time and for passing a good time with them during my whole education. Finally. I wish to acknowledge the support of my loving family, my parents, brother, sister, because they kept me straight all the time.

A particular thanks go to professors, Mr. Baarir zin-eddine, Mr.Ouafi karim , the jury members for having agreed to review and evaluate this work Finally, I thank my family and my friends for their patience, encouragement, and support which were very useful to me during my work.

Abstract

For the past decades, Human-to-Computer Interfaces (HCI) have been relying n simple interactions through classical devices (e.g., keyboard, mouse, touch-screen, etc). Nowadays, there is a large research effort to integrate new interaction modalities in HCI's in order to develop interfaces more similar to human–human communication and make their use more intuitive, natural and efficient . For example, the user voice can be recorded via a microphone and words can be recognized; the user face can be captured via a camera and facial expressions can be identified; likewise the user hand gesture can be traced and its movement can be interpreted, etc. In this work, we investigate how to recognize facial expressions. This work is primarily motivated by the importance of facial expression in face-to-face human communication. A facial expression is a visible manifestation of an emotional state, cognitive activity, intention, personality, and psychopathology of a person. As pointed out by Mehrabian while communicating feelings,55% of the message is conveyed by the facial expression while only 7% by the linguistic language (verbal) and 38%by the paralanguage like intonation (vocal). Human–computer interaction will definitely benefit from automating facial expression recognition. In his seminal work, Eckman showed that there exist mainly six emotions, namely Joy, Surprise, Disgust, Sadness, Anger and Fear. An additional Neutral emotion, or absence of any emotion, is generally considered.

In this work, we used a framework for extracting facial expressions from a set of facial images. Experimental large-scale results were presented using a standard database: CK + (seven categories of expression) (fear, anger, happiness, disgust, surprise, sadness, contempt) The implementation of the proposed system has been divided into four components In the first component, a region of interest as face detection has been performed from the captured input image. For extracting more distinctive and discriminant features, in the second component, a deep learning-based convolutional neural network architecture has been proposed to perform feature learning tasks for classification purposes to recognize the types of expressions. in the third component, The system categorized the results. In the fourth component, Results are displayed Models that will apply are resnet-34, Densnet-121, Resnet-50, Wideresnet-50

Keywords: Deep learning, Machine learning, Convolution neural network (CNN), Deep neural network architectures, Deep learning applications, Image classification, Transfer learning, Emotion image analysis, Supervised learning,

Résumé

Au cours des dernières décennies, les interfaces homme-ordinateur (HCI) ont reposé sur des interactions simples via des dispositifs classiques (par exemple, clavier, souris, écran tactile, etc.). De nos jours, il existe un important effort de recherche pour intégrer de nouvelles modalités d'interaction dans les IHM afin de développer des interfaces plus proches de la communication homme-homme et de rendre leur utilisation plus intuitive, naturelle et efficace. Par exemple, la voix de l'utilisateur peut être enregistrée via un le microphone et les mots peuvent être reconnus ; le visage de l'utilisateur peut être capturé via une caméra et les expressions faciales peuvent être identifiées ; de même, le geste de la main de l'utilisateur peut être tracé et son mouvement peut être interprété, etc. Dans ce travail, nous étudions comment reconnaître les expressions faciales. Ce travail est principalement motivé par l'importance de l'expression faciale dans la communication humaine face à face. Une expression faciale est une manifestation visible d'un état émotionnel, activité cognitive, intention, personnalité et psychopathologie d'une personne. Comme l'a souligné Mehrabian lors de la communication des sentiments, 55% du message est transmis par l'expression faciale alors que seulement 7% par le langage linguistique (verbal) et 38%par le paralangage comme l'intonation (vocale). L'interaction homme-ordinateur en bénéficiera certainement de l'automatisation de la reconnaissance des expressions faciales. Dans son ouvrage fondateur, Eckman a montré qu'il existe principalement six émotions, à savoir la joie, la surprise, le dégoût, la tristesse, la colère et la peur. Une émotion neutre supplémentaire, ou l'absence de toute émotion, est généralement considérée. Dans ce travail, nous avons utilisé un framework pour extraire des expressions faciales à partir d'un ensemble d'images faciales. Des résultats expérimentaux à grande échelle ont été présentés à l'aide d'une base de données standard : CK + (sept catégories d'expression) (peur, colère, bonheur, dégoût, surprise, tristesse, mépris) La mise en œuvre du système proposé a été divisée en quatre composantes. Dans le premier composant, une région d'intérêt comme détection de visage a été effectuée à partir du image d'entrée capturée. Pour extraire des caractéristiques plus distinctives et discriminantes, dans le deuxième composant, une architecture de réseau neuronal convolutif basée sur l'apprentissage profond a été proposée pour effectuer des tâches d'apprentissage de caractéristiques à des fins de classification pour reconnaître les types d'expressions. dans le troisième volet, le système a catégorisé les résultats. Dans le quatrième composant, les résultats sont affichés. Les modèles qui s'appliqueront sont resnet-34. , Densnet-121, Resnet-50, Wideresnet-50

Contents

List of Figures vii

List of Tables ix

1 Introduction 1

1.1 general Introduction . . . 2

1.2 Biometric data . . . 5

1.3 Fundamental Qualities . . . 5

1.4 Secondary Qualities . . . 5

1.5 Biometric System . . . 5

1.5.1 Enrollment . . . 6

1.5.2 Templates . . . 6

1.5.3 Matching . . . 7

1.6 Biometric Systems – Types and Process . . . 9

1.7 Biometric technologies . . . 10

1.7.1 DNA . . . 11

1.7.2 Ear . . . 11

1.7.3 Face . . . 11

1.7.4 Facial thermogram . . . 13

1.7.5 Fingerprint . . . 14

1.7.6 gait . . . 15

1.7.7 Hand geometry . . . 16

1.7.8 Hand Vein . . . 17

1.7.9 Iris . . . 17

1.7.10 Keystroke Dynamics . . . 19

1.7.11 Odor . . . 19

1.8 Comparison of various biometric technologies . . . 19

2 Convolutional Neural Network (CNN) 33

2.1 Introduction . . . 34

2.2 Related Works . . . 35

2.2.1 Data Pre-Processing . . . 36

2.2.2 Data Imbalance . . . 37

2.3 Feature Selection . . . 37
2.3.1 Extra Tree Classifier . . . 38
2.3.2 Pearson Correlation . . . 38
2.3.3 Forward Stepwise Selection . . . 39
2.4 ML Classifiers Techniques . . . 39
2.4.1 Supervised Machine Learning Models . . . 39
2.4.2 Ensemble Machine Learning Model . . . 41
2.4.3 Neural Network Models . . . 42
2.5 Hyperparameter Tuning . . . 43
2.5.1 Cross-Validation . . . 43
2.6 Dataset Description . . . 44
2.6.1 Data Pre-Processing . . . 46
2.6.2 Feature Selection . . . 49
2.6.3 Model Selection . . . 50
2.6.4 Model Evaluation . . . 50
2.7 Analysis . . . 51
2.8 Conclusion . . . 51

3 Experiments and Results 52
3.1 Introduction . . . 53
3.2 Overall architecture of the application . . . 53
3.3 Image acquisition . . . 53
3.4 Pretraitment . . . 53
3.5 Feature extraction . . . 53
3.6 Convolution neural network . . . 54
3.6.1 pooling . . . 55
3.7 parameters . . . 55
3.7.1 Epoch . . . 55
3.7.2 Batch . . . 55
3.8 Database (CK+) . . . 56
3.9 Parameters tuning . . . 56
3.10 Resnet-34 . . . 58
3.11 Resnet-50 . . . 61
3.12 Densnet 121 . . . 64
3.13 Wideresnet50 . . . 67
3.14 Comparisons . . . 70
3.15 Real-time application . . . 70
3.15.1 Python . . . 70
3.15.2 Pytorch . . . 71

3.16 Visual Studio Code 72
3.17 Application 72
3.18 General conclusion 79

Bibliography **82**

List of Figures

1.1 6

1.2 7

1.3 8

1.4 Face recognition structure. 10

1.5 Figure 5 Vulnerabilities of a biometric system. 24

1.6 Table 3: Advantages and drawbacks of the different protection techniques. 27

2.1 Figure 8.1 Workflow model of proposed system. 36

2.2 Face recognition structure. 39

2.3 Description of each feature in the dataset. 45

2.4 Sample dataset. 46

2.5 Architecture of proposed system. 47

2.6 Original dataset distribution. 47

2.7 Figure 8.5 Target class distribution. 48

2.8 Figure 8.6 Resampled distribution applying SMOTE. 48

2.9 Feature ranking using Extra tree classifier. 49

2.10 p-values of the features. 49

3.1 Architecture of a CNN. — Source https://www.mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks–1489512765771.html. 54

3.2 Max pooling takes the largest values. — Source: http://cs231n.github.io/convolutional-networks/ 55

3.3 CK+ samples corresponding to facial expressions 56

3.4 DenseNet architecture. 57

3.5 DenseNet architecture. 57

3.6 Confusion matrices of mode(1) corresponds to Resnet-34 model. 58

3.7 Results of mode(2) for Resnet-34 model. 58

3.8 Confusion matrices of mode(3) corresponds in Resnet-34 model. 59

3.9 Confusion matrices of mode(4) corresponds to Resnet-34 model. 59

3.10 Confusion matrices of mode(5) corresponds to Resnet-34 model. 59

3.11 Better results for Resnet-34 model. 60

3.12 Confusion matrices of mode(1) corresponds to Resnet-50 model. 61

3.13 Confusion matrices of mode(2) corresponds to Resnet-50 model. 61

3.14 Confusion matrices of mode(3) corresponds to Resnet-50 model. 61

3.15 Confusion matrices of mode(4) corresponds to Resnet-50 model. 62
3.16 Confusion matrices of mode(5) corresponds to Resnet-50 model. 62
3.17 Better results for Resnet-50 model. 63
3.18 Confusion matrices of mode(1) corresponds to Densenet121 model. 64
3.19 Confusion matrices of mode(2) corresponds to Densenet121 model. 64
3.20 Confusion matrices of mode(3) corresponds to Densenet121 model. 64
3.21 Confusion matrices of mode(4) corresponds to Densenet121 model. 65
3.22 Confusion matrices of mode(5) corresponds to Densenet121 model. 65
3.23 Better results for Densnet-121 model. 66
3.24 Confusion matrices of mode(1) corresponds to Wideresnet50 model. 67
3.25 Confusion matrices of mode(2) corresponds) to Wideresnet50 model. 67
3.26 Confusion matrices of mode(3) corresponds to Wideresnet50 model. 67
3.27 Confusion matrices of mode(4) corresponds to Wideresnet50 model. 68
3.28 Confusion matrices of mode(5) corresponds to Wideresnet50 model. 68
3.29 Better results for Wideresnet50 model. 69
3.30 Comparisons between the four models and its corresponding modes. 70
3.31 The main interface of the real-time application. 73
3.32 The four used models in our application. 74
3.33 Action of manage tracking button (a) Click on Manage Tracking button, (b) Resulting form of the clicked button. 74
3.34 Action of start online tracking. 75
3.35 Detection feeling of happiness. 75
3.36 Detection feeling of disgust. 76
3.37 Detection feeling of surprise. 77
3.38 Detection feeling of fear. 77
3.39 Detection feeling of anger. 78
3.40 Detection feeling of sadness. 78
3.41 Detection for the feeling of happiness belonging to two individuals at the same scene (same time). 79

List of Tables

1 Introduction

Contents

1.1 general Introduction **2**

1.2 Biometric data **5**

1.3 Fundamental Qualities **5**

1.4 Secondary Qualities **5**

1.5 Biometric System **5**

1.5.1 Enrollment 6

1.5.2 Templates 6

1.5.3 Matching 7

1.6 Biometric Systems – Types and Process **9**

1.7 Biometric technologies **10**

1.7.1 DNA 11

1.7.2 Ear 11

1.7.3 Face 11

1.7.4 Facial thermogram 13

1.7.5 Fingerprint 14

1.7.6 gait 15

1.7.7 Hand geometry 16

1.7.8 Hand Vein 17

1.7.9 Iris 17

1.7.10 Keystroke Dynamics 19

1.7.11 Odor 19

1.8 Comparison of various biometric technologies **19**

1.1 general Introduction

In the ever-changing world of global data communications, inexpensive Internet connections, and fast-paced software development, security is becoming more and more of an issue. Hence it has become more vital to protect the critical infrastructure and provide security for the smooth functioning of the computing solutions from the attackers. One of the ways of protecting is by providing access control to the existing infrastructure. This can be achieved by providing authentication schemes such as secure login with the help of user name and password, pass phrase, smart cards, PIN ,numbers, biometrics, etc.,

A biometric system is essentially a pattern recognition system that makes use of biometric traits to recognize individuals.

The objective is to establish an identity based on ’who you are or what you produce‘, rather than by ’what you possess‘ or ’what you know‘. The significance of using biometrics has been reinforced by the need for large scale identity management systems. The very purpose of identity management is to accurately determine an individual‘s identity in the context of several different applications. This new technique not only provides enhanced security but also avoids, in authentication the need to remember several passwords and maintain multiple authentication tokens.

Emotion classification is a research area in which there has been very intensive literature production concerning natural language processing, multimedia data, semantic knowledge discovery, social network mining, and text and multimedia data mining. This paper addresses the issue of emotion classification and proposes a method for classifying the emotions expressed in multimodal data extracted from videos. The proposed method models multimodal data as a sequence of features extracted from facial expressions, speech, gestures, and text, using a linguistic approach. Each sequence of multimodal data is correctly associated with the emotion by a method that models each emotion using a hidden Markov model. The trained model is evaluated on samples of multimodal sentences associated with seven basic emotions. The experimental results demonstrate a good classification rate for emotions.

In Chapter 1, we gave a general introduction of the biometric system, And an explanation of all its parts.

In Chapter 2, we discussed the components of the biometric system in detail, as well as the types of models used to extract image properties

In Chapter 3, we explained to me the CNN layers theoretically, as well as experimenting with a database ck+ on 4 models(resnet-34, resnet-50, wideresnet-50, densnet121) and it gave good results.

1.1 History of Biometrics

Biometrics is not a new concept; it is the oldest form of identification. Bertillon Systems (1882) took subject's photography, height, the length of one-foot, an arm and index finger. FBI setup a fingerprint identification division in the year 1924. By 1926, law enforcement officials in several U.S. cities had begun submitting fingerprint cards to the FBI in an effort to create a database of fingerprints from known criminals. In the early 1960‘s the FBI invested a large amount of time and effort into the development of automated fingerprint identification systems. This automation of biometric identification for law enforcement purposes coincided with the development of automated systems for non-forensic applications, such as high-security access control. AFIS installed in 1965 with a database of 810,000 fingerprints. During the 1970‘s a biometric product based on measuring the geometry of the hand was introduced in a number of access control applications. First face recognition paper was published in the year 1972 (Goldstein et al). Interest in biometric identification eventually moved from measuring characteristics of the hand to include characteristics of the eye. In the mid-1980‘s the first system that analyzed the unique patterns of the retina was introduced while, concurrently, work was being performed to analyze iris patterns. In the 1990‘s, research continues on developing identification systems based on a wide variety of biometric patterns, such as the traditional biometrics mentioned above (i.e. fingerprint, hand geometry, iris, and retina), along with the development of voice, signature, palm print, and face recognition systems. A few new, innovative approaches are also being examined for biometric analysis, such as ear shape, DNA, keystroke (typing rhythm), and body odor. The computer industry began using biometrics few years ago. However, as with the first computers, biometric systems were massive. Typically created for a specific function, they lacked the adaptability required to integrate into a variety of environments. This resulted in costly solutions that few were able or willing to incorporate. However, over time, as technology advanced and networking and device standards were created, biometric solutions evolved to be widely recognized as viable options to security solutions (Heath, role). Fraud, security breaches, and human administrative error are helping drive the expansion of biometric technology [22]. Traditionally, passwords (knowledge-based security) and ID cards (token-based security) have been used to moderate access to restricted systems. However, security can be easily breached in these systems when a password is revealed to an unauthorized user or an impostor steals a card. Furthermore, simple passwords are easy to guess (by an impostor) and difficult passwords may be hard to recall (by a legitimate user). The emergence of biometrics has addressed the problems that plague traditional verification methods. Biometrics refers to the automatic identification (or verification) of an individual (or a claimed identity) by using certain physiological or behavioral traits associated with the person [16]. By using biometrics it is possible to establish an identity based on 'who you are', rather than by 'what you possess' (e.g., an ID card) or 'what you remember' (e.g., a password). Current biometric systems make use of fingerprints, hand geometry, iris, retina, face, facial thermograms, signature, gait, palm print and voiceprint to establish a person's identity. While biometric systems have their limitations they have an edge over traditional security methods in that they cannot be easily stolen or shared. Besides strengthening security, biometric systems also enhance user convenience by alleviating the need to design and remember passwords. Biometric-based authentication applications include workstation, network, and domain access, single sign-on, application logon, data protection, remote access to resources, transaction security and Web security [29]. Trust in these electronic transactions is essential to the healthy growth of the global economy. Utilized alone or integrated with other technologies such as smart cards, encryption keys and digital signatures, biometrics is set to pervade nearly all aspects of the economy and our daily lives. Utilizing biometrics for personal authentication is becoming convenient and considerably more accurate than current methods (such as the utilization of passwords or PINs). This is because biometrics links the event to a particular individual (a password or token may be used by someone other than the authorized user), is convenient (nothing to carry or remember), accurate (it provides for positive authentication), can provide an audit trail and is becoming socially acceptable and cost effective. There is a variety of means for identifying a person's identity:

- appearance (how the person looks, e.g. height, gender,weight)
- social behavior (how a person interacts with others) dot, a so-called bullet.
- name (what the person is called)
- codes (what a person is called by an organization)
- knowledge (what the person knows)
- possession (what the person owns)
- bio-dynamics (what the person does)
- natural physiology (who the person is, e.g. facial characteristics)
- Imposed physical characteristics (what the person is now, e.g. tags, collars, bracelets)

The goal of authentication is to protect a system against unauthorized use. This feature enables also the protection of subscribers by denying the possibility for intruders to impersonate authorized users. Authentication procedures are based on the following approaches [18]:

- Proof by Knowledge. The verifier known information regarding the claimed identity that can only be known or produced by a principal with that identity (e.g. passport, password, personal identification number (PIN), questionnaire)
- Proof by Possession. The claimant will be authorized by the possession of an object (e.g. magnetic card, smart card, optical card)
- Proof by Property. The claimant directly measures certain claimant properties using human characteristics (e.g. biometrics)

1.2 Biometric data

Biometric data is different than a password that can be guessed or changed because it relies on a physical or behavioral characteristic of a person. In order for a biometric system to function well, the qualities of the data taken need to be such that all users of the system can be uniquely identified. Fundamental and secondary qualities are listed below.

1.3 Fundamental Qualities

- Universality– Must be some trait that can be taken from many people..
- Uniqueness– Unique per person; quality must not occur in two different individuals
- Uniqueness– Unique per person; quality must not occur in two different individuals.
- Permanence– Quality must be constant over time (eliminates need for re-enrollment).
- Collectability– Characteristic must be able to be measured quantitatively.

1.4 Secondary Qualities

- Performance – How well the biometric balances the various requirements of the systems.
- Acceptability – The acceptance of the users to present the biometric data
- Circumvention – How easy it is to fool the system

1.5 Biometric System

All biometric systems consist of three basic elements:

- Enrollment, or the process of collecting biometric samples from an individual, known as the enrollee, and the subsequent generation of template.

- Templates, or the data representing the enrollee‘s biometric.
- Matching, or the process of comparing a live biometric sample against one or many templates in the system‘s database.

1.5.1 Enrollment

Enrollment is the first stage for biometric authentication because enrollment generates a template that will be used for all subsequent matching. Figure 1 shows the enrollment process.

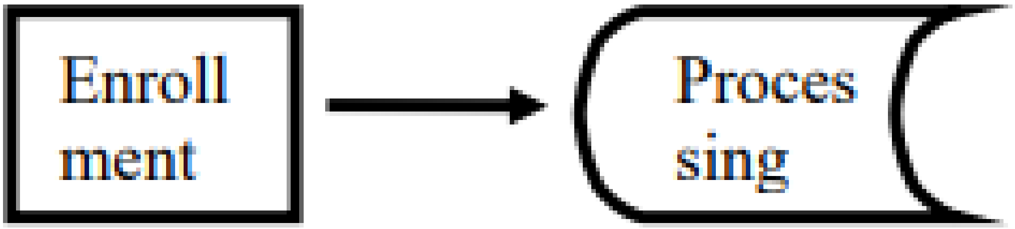


Figure 1 Capturing a biometric sample

Figure 1.1

Typically, the device takes three samples of the same biometric and averages them to produce an enrollment template. Enrollment is complicated by the dependence of the performance of many biometric systems on the users‘ familiarity with the biometric device because enrollment is usually the first time the user is exposed to the device. Environmental conditions also affect enrollment. Enrollment should take place under conditions similar to those expected during the routine matching process. In addition to user and environmental issues, biometrics themselves change over time. Many biometric systems account for these changes by continuously averaging. Templates are averaged and updated each time the user attempts authentication.

1.5.2 Templates

As the data representing the enrollee‘s biometric, the biometric device creates templates. The device uses a proprietary algorithm to extract "features" appropriate to that biometric from the enrollee‘s samples. Figure 2 shows the process of storing a sample in a database. Templates are only a record of distinguishing features, sometimes called minutiae points, of a person‘s biometric characteristic or trait.

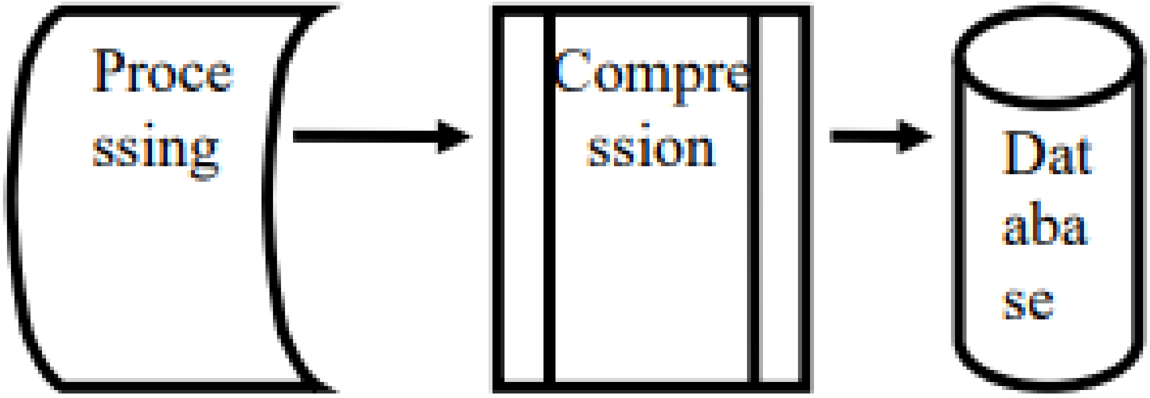


Figure 2 Compression of a biometric sample and storing it in a database

Figure 1.2: .

For example, templates are not an image or record of the actual fingerprint or voice. In basic terms, templates are numerical representations of key points taken from a person‘s body. The template is usually small in terms of computer memory use, and this allows for quick processing. The template must be stored somewhere so that subsequent templates, created when a user tries to access the system using a sensor, can be compared.

1.5.3 Matching

Matching is the comparison of two templates, the template produced at the time of enrollment with the one produced "on the spot" as a user tries to gain access by providing a biometric via a sensor. Figure 3 shows the extraction of live scan and compares it with the biometric template stored in a database to find a match in the stored template

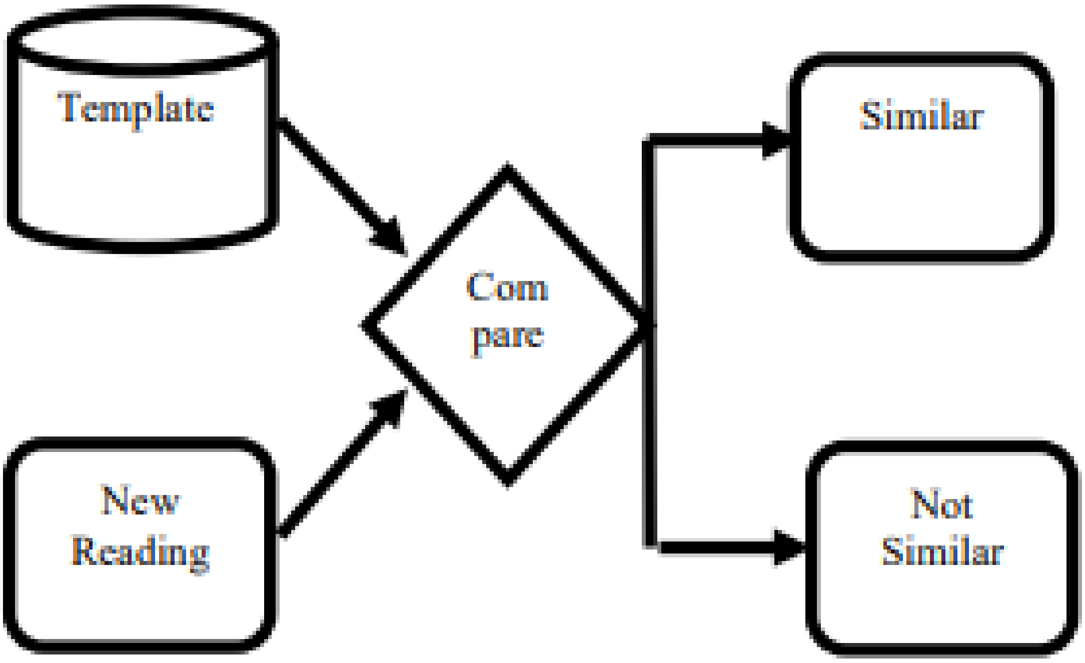


Figure 3 Comparison and evaluation of a biometric sample

Figure 1.3

There are three ways a match can fail:

- Failure to enroll. .
- False match.
- False nonmatch.

Failure to enroll (or acquire) is the failure of the technology to extract distinguishing features appropriate to that technology. In addition, the possibility of a false match (FM) or a false nonmatch (FNM) exists. These two terms are frequently misled "false acceptance" and "false rejection," respectively, but these terms are application-dependent in meaning. FM and FNM are application-neutral terms to describe the matching process between a live sample and a biometric template. A false match occurs when a sample is incorrectly matched to a template in the database (i.e., an imposter is accepted). A false non-match occurs when a sample is incorrectly not matched to a truly matching template in the database (i.e., a legitimate match is denied). Rates for FM and FNM are calculated and used to make tradeoffs between security and convenience. For example, a heavy security emphasis errs on the side of denying legitimate matches and does not tolerate acceptance of imposters. A heavy emphasis on user convenience results in little tolerance for denying legitimate matches but will tolerate some acceptance of imposters.

1.6 Biometric Systems – Types and Process

All biometric systems fall under two categories, which are familiar to those involved in security systems: identification and verification. The process, applications, and challenges are unique for both these categories because of the system-level differences that exist. An identification system is sometimes referred to as "1:N Matching" because a user presents biometric data to the system and the system attempts to identify if the user is enrolled in the system and who the person is. A verification system is referred to as "1:1 Matching" because a person makes a claim to his or her identity, presents biometric data, and the system compares the presented biometric data to the data on file only for the claimed identity. A helpful way to distinguish between these two types of authentication is the two different questions that users are essentially asking: in identification, "Who am I ?"; in verification, "Am I who I claim to be? The process of using a biometric system is designed to be as transparent as possible. To understand what occurs during the verification or identification process, a few sub-processes need to be defined. Presentation is where the user physically presents to the biometric system the data required for capture, such as a fingerprint, iris, or hand. Enrollment is the process when a user is initially registered for access to a system. This requires the user to present his or her biometric data (fingerprint, iris, hand, etc.) so that a template can be formed in the system. This template will serve as a basis for comparison when attempting to gain access at later times. Since the template will be used many times in the future, the quality of the biometric data acquired during this stage is critical. This stage of using a biometric system can be the most tedious. Feature Extraction is an automated process of locating and encoding distinctive characteristics from biometric data to generate a template.

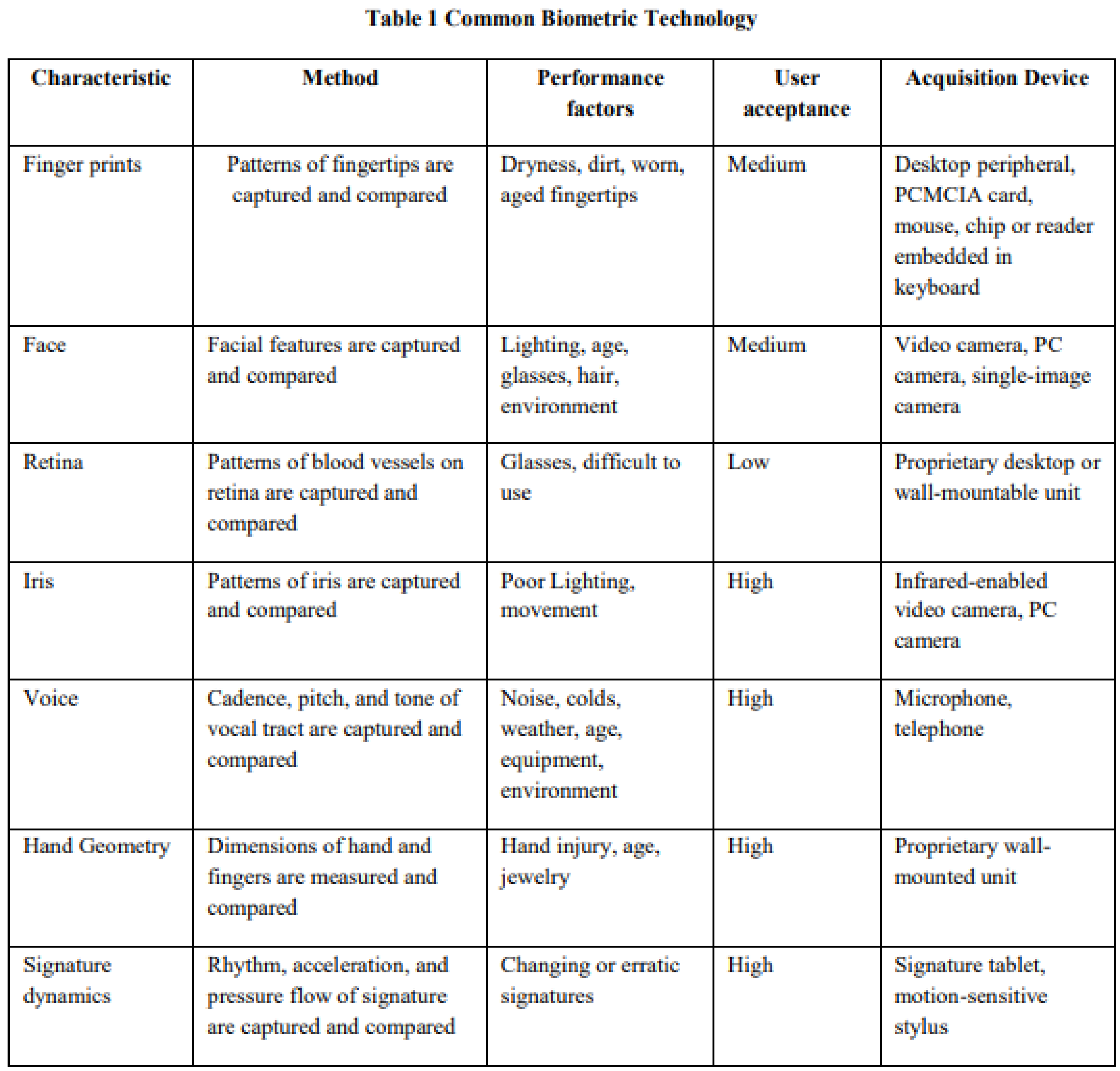


Table 1 Common Biometric Technology

Characteristic	Method	Performance factors	User acceptance	Acquisition Device
Finger prints	Patterns of fingertips are captured and compared	Dryness, dirt, worn, aged fingertips	Medium	Desktop peripheral, PCMCIA card, mouse, chip or reader embedded in keyboard
Face	Facial features are captured and compared	Lighting, age, glasses, hair, environment	Medium	Video camera, PC camera, single-image camera
Retina	Patterns of blood vessels on retina are captured and compared	Glasses, difficult to use	Low	Proprietary desktop or wall-mountable unit
Iris	Patterns of iris are captured and compared	Poor Lighting, movement	High	Infrared-enabled video camera, PC camera
Voice	Cadence, pitch, and tone of vocal tract are captured and compared	Noise, colds, weather, age, equipment, environment	High	Microphone, telephone
Hand Geometry	Dimensions of hand and fingers are measured and compared	Hand injury, age, jewelry	High	Proprietary wall-mounted unit
Signature dynamics	Rhythm, acceleration, and pressure flow of signature are captured and compared	Changing or erratic signatures	High	Signature tablet, motion-sensitive stylus

Figure 1.4: Face recognition structure.

1.7 Biometric technologies

The function of a biometric technologies authentication system is to facilitate controlled access to applications, networks, personal computers (PCs), and physical facilities. A biometric authentication system is essentially a method of establishing a person‘s identity by comparing the binary code of a uniquely specific biological or physical characteristic to the binary code of an electronically stored characteristic called a biometric. The various biometric technologies are DNA, Ear, Face, Facial thermogram, Fingerprint, Gait, Hand geometry,Hand Vein, Iris, Keystroke, Odor, Retina, Signature, Voice, Palmprint. Table 1 gives the summary of some of the common biometric types [7], the method involved, the performance factor, the devices used to acquire the biometric and the user acceptance level.

1.7.1 DNA

Deoxyribonucleic acid (DNA) is the one-dimensional (1–D) ultimate unique code for one‘s individuality —except for the fact that identical twins have identical DNA patterns. It is, however, currently used mostly in the context of forensic applications for person recognition. Anil K. Jain et al [9] gives three issues that limit the utility of this biometrics for other applications:

- contamination and sensitivity: it is easy to steal a piece of DNA from an unsuspecting subject that can be subsequently abused for an ulterior purpose;
- automatic real-time recognition issues: the present technology for DNA matching requires cumbersome chemical methods (wet processes) involving an expert‘s skills and is not geared for on-line noninvasive recognition; and dot, a so-called bullet.
- privacy issues: information about susceptibilities of a person to certain diseases could be gained from the DNA pattern and there is a concern that the unintended abuse of genetic code information may result in discrimination, e.g., in hiring practices.

1.7.2 Ear

It has been suggested that the shape of the ear and the structure of the cartilaginous issue of the pinna are distinctive. The ear recognition approaches are based on Anil K. Jain et al [9] matching the distance of salient points on the pinna from a landmark location on the ear. The features of an ear are not expected to be very distinctive in establishing the identity of an individual.

1.7.3 Face

In the last decade significant advances have been achieved on biometrics, especially on face recognition (Jain et al., 1999). This technology is considered a natural means of biometric identification since the ability to distinguish among individual appearances is possessed by humans. Facial scan systems can range from software-only solutions that process images processed through existing closed-circuit television cameras to full fledged acquisition and processing systems, including cameras, workstations, and backend processors. With facial recognition technology, a digital video camera image is used to analyze facial characteristics such as the distance between eyes, mouth or nose. These measurements are stored in a database and used to compare with a subject standing before a camera.

The face recognition process has two major parts: detection, locating a human face in an image and isolating it from other objects in the frame, and recognition, comparing the face being captured with a database of faces to find a match.

During detection, the hardware/software combination isolates the facial elements of an image and eliminates extraneous information. The software examines the image for typical facial structures (such as eyes and nose), and once it has found them, it calculates the remainder of the face. It then cuts away background details, leaving a close-up of a face inside a rectangular frame called a binary mask. Face recognition uses mainly the following techniques:

- Facial geometry: uses geometrical characteristics of the face. May use several cameras to get better accuracy (2D,3D...)
- Skin pattern recognition (Visual Skin Print)
- Facial thermogram: uses an infrared camera to map the face temperatures.
- Smile: recognition of the wrinkle changes when smiling Facial-scan technology has its advantages and disadvantages.

Advantages

- Easy for humans to verify results
- No contact required
- Commonly available sensors (cameras) dot, a so-called bullet.
- Large amounts of existing data to allow background and/or watch list checks

Disadvantages

- Face can be obstructed by hair, glasses, hats, scarves, etc.
- Sensitive to changes in lighting, expression, and pose.
- Faces change over time.
- Propensity for users to provide poor-quality video images yet to expect accurate results.

One major advantage is that facial-scan technology is one of the biometric capable of identification at a distance without subject complicity or awareness. Another advantage of facialscan technology is the fact that static images can be used to enroll a subject. The main advantage of facial recognition systems is the lack of user interaction needed to perform scans. While users may be required to stand still, facial recognition systems are without a doubt one of the least intrusive on the market. In addition facial recognition is suited to environments where there is significant dirt or potential pathogens, as there is no physical contact required between users and systems.

The disadvantages include acquisition environment and facial characteristic changes that effect matching accuracy and the potential for privacy abuse. Images are most accurate when taken facing the acquisition camera and not sharp angles. The users face must be lit evenly, preferably from the front. Changes in hairstyle, makeup or the wearing of a hat or sunglasses may pose a problem during the verification process. Facial-scanning technology has a poor record in verifying a subject who has had plastic surgery to alter their appearance. The fact that a biometric facial scan can take place without the knowledge or consent of a subject raises privacy concerns among many. Facial-scan technologies have unique advantages over all other biometrics in the areas of surrounding large groups and the ability to use preexisting static images. Its disadvantages include the falsely nonmatching folks when subject appearances change during verification.

Andrea F. Abate, et al [1] discusses the five key factors that significantly affect system face recognition performances.

- Illumination variations due to skin reflectance properties and due to the internal camera control. Several 2D methods do well in recognition tasks only under moderate illumination variation, while performances noticeably drop when both illumination and pose changes occur.

- Pose changes affect the authentication process, because they introduce projective deformations and self-occlusion. Even if methods dealing with up to 32_ head rotation exists, they do not solve the problem considering that security cameras can create viewing angles that are outside of this range when positioned.

- On the contrary, with exception of extreme expressions such as scream, the algorithms are relatively robust to facial expression.

- Another important factor is the time delay, because the face changes over time, in a nonlinear way over long periods. In general this problem is harder to solve with respect to the others and not much has been done especially for age variations.

- At last, occlusions can dramatically affect face recognition performances, in particular if they are located on the upper-side of the face, as documented in literature.

1.7.4 Facial thermogram

Facial thermogram requires an (expensive) infrared camera to detect the facial heat patterns that are unique to every human being. The pattern of heat radiated by human body is a characteristic of an individual and can be captured by an infrared camera in an unobtrusive way much like a regular (visible spectrum) photograph. The technology could be used for covert recognition. A thermogram-based system does not require contact and is noninvasive, but image acquisition is challenging in uncontrolled environments, where heatemanating surfaces (e.g., room heaters and vehicle exhaust pipes) are present in the vicinity of the body. A related technology using near infrared imaging is used to scan the back of a clenched fist to determine hand vein structure [9]. Infrared sensors are prohibitively expensive which is a factor inhibiting wide spread use of the thermograms.

1.7.5 Fingerprint

Fingerprinting technology is the oldest of the biometric sciences and utilizes distinctive features of the fingerprint to identify or verify the identity of individuals [9]. All fingerprints have unique characteristics and patterns. A normal fingerprint pattern is made up of lines and spaces. These lines are called ridges while the spaces between the ridges are called valleys. It is through the pattern of these ridges and valleys that a unique fingerprint is matched for verification and authorization. These unique fingerprint traits are termed "minutiae" and comparisons are made based on these traits. On average, a typical live scan produces 40 "minutiae". The Federal Bureau of Investigation (FBI) has reported that two individuals can share no more than 8 common minutiae.

There are five stages involved in finger-scan verification and identification: fingerprint image acquisition, image processing, location of distinctive characteristics, template creation and template matching. A scanner takes a mathematical snapshot of a user's unique biological traits. This snapshot is saved in a fingerprint database as a minutiae file.

Fingerprint authentication mechanisms rely on the identification of minutiae in the fingerprints - discontinuities in the flow of a user's fingerprints that can come in the form of deltas, pores, islands, and other characteristics. Once systems have isolated the minutiae of a fingerprint, the precision of the matching is based on the numbers of minutiae, which are used to create a positive match.

Fingerprint verification has emerged as one of the most reliable means of biometric authentication due to its universality, distinctiveness, permanence and accuracy [24]. Fingerprint biometrics has a number of benefits. They enjoy good acceptance rates in the general public, and have a positive image. The mechanism is not overly intrusive, and is generally trusted. The mechanism is quite flexible for failure rates, as fingerprints have a high number of potential areas to be mapped for identification. Thus if desired, fingerprint based rollouts can be highly secure.

As a popular and precise mechanism, fingerprinting does have drawbacks. Fingerprinting systems will struggle in areas where users are likely to have either injured or dirty hands. Similarly, the elderly and those with dry skin may struggle to register and make use of these biometrics systems. Finally fingerprinting systems are the most commonly targeted systems for attempts to falsify credentials. A number of methods used to counter this are available, and mechanisms such as liveness testing should be considered when using fingerprinting.

Advantages

- Subjects have multiple fingers.
- Easy to use,with some training dot,a so-called bullet.
- Some systems require little space.
- Large amounts of existing data to allow background
- Has proven effective in many large scale systems over years of use and/or watch list checks
- Fingerprints are unique to each finger of each individual and the ridge arrangement remains permanent during one's lifetime

disavantages

- Public Perceptions.
- Privacy concerns of criminal implications.
- Health or societal concerns with touching a sensor used by countless individuals.
- Collection of high quality nail-to-nail images requires training and skill, but current flat reader technology is very robust
- An individual‘s age and occupation may cause some sensors difficulty in capturing a complete and accurate fingerprint image

1.7.5.1 Issues with use of fingerprints

For a system that uses a fingerprint as its biometric data, it has been found by multiple groups that the use of "gummy fingers" (artificial fingers made from gelatin) can spoof a biometric system. One such study found that it was possible to create a gummy finger from a latent fingerprint, enroll into the system and then verify using the same gummy finger against a live enrolled template.

1.7.6 gait

Gait is the peculiar way one walks and is a complex spatiotemporal biometric. Gait is not supposed to be very distinctive, but is sufficiently discriminatory to allow verification in some low-security applications. Gait is a behavioral biometric and may not remain invariant, especially over a long period of time, due to fluctuations in body weight, major injuries involving joints or brain, or due to inebriety. Acquisition of gait is similar to acquiring a facial picture and, hence, may be an acceptable biometric. Since gait-based systems use the video-sequence footage of a walking person to measure several different movements of each articulate joint, it is input intensive and computationally expensive.

1.7.7 Hand geometry

Among a number of biometric techniques that are used for frequent human identification tasks, hand shape and hand shape-based biometry is an emerging new technique, with certain advantages over the more established competitor techniques. First, human hand data acquisition is less cumbersome, user-friendlier. Furthermore, it is much less susceptible to intrinsic variations and environmental artifacts [9].Hand-based identification/verification systems provide an attractive and growing alternative biometric scheme [30].

Hand geometry employs measurements of various aspects of the hand including width, length, and width and length of fingers. Because of the relatively basic nature of this mechanism, it is suitable for authentication, but not identification. Most people are willing to use their hands and bring a fresh perspective to this technology (as opposed to fingerprint recognition systems, which are associated with criminals being booked).

Hand geometry authentication offers several advantages. Hand geometry verification systems use geometric measurements of hand as the features for the verification of individuals. Hand geometry measurements are easy to collect and non intrusive. The hardware requirements imposed by such a system are not severe. The system just requires properly placed camera that can get the image of the hand. Furthermore, with back illumination the measurements can be made very robust against ambient lighting conditions. Hand geometry acquisition and verification is extremely fast andaccurate enough for verification. It is very well suited forintegration with other biometrics and in particular with fingerprints. It is very easy to envision a system that simultaneously acquires hand geometry and fingerprint images.

Hand geometry, however, does have some negatives. First, hand geometry systems should not be considered for identification purposes, since hands are not unique and accuracy rates cannot be as high as other biometric readers. Rather, they are best used for verification purposes and coupled with a PIN for extra security. Hand geometry systems have been well accepted to date, and can be configured to be quite reliable. They have been deployed in several major areas with no major issues of public acceptance, possibly due to the lack of potential for identification with this mechanism. While hand geometry systems can struggle with equipment issues in dirty environments, authentication should not be an issue as the attributes to be measured are far less detailed than those used for fingerprinting.

Hand geometry systems come at a higher cost. The hardware used for such scanning is dedicated, and systems are proprietary. Different manufacturers employ different mapping standards based on different attributes, so potential buyers would be advised to obtain field results on failure types and enrolment issues. Advantages

- Easy to capture.

- Believed to be a highly stable pattern over the adult lifespan.

Disadvantages

- Use requires some training.
- Not sufficiently distinctive for identification over large databases; usually used for verification of a claimed enrollment identity.
- System requires a large amount of physical space.

1.7.8 Hand Vein

Vein patterns are within you, people don't leave them around (like fingerprints) nor can they be easily observed like Iris patterns or faces. Vein structures are not easily covertly captured or reproduced like other biometric traits.

Vein recognition technology is an emerging biometric technology with great promise. But vein recognition is not without its faults. The technology could be seen as invasive by many people, particularly in the US, where fears surrounding the long-term damage of such internal readings are paramount. In addition, it faces an uphill battle against other biometric technologies that have existed far longer than vein recognition.

1.7.9 Iris

Penny Khaw [9] in his article says that the iris has many features that can be used to distinguish one iris from another. One of the primary visible characteristic is the trabecular meshwork, a tissue which gives the appearance of dividing the iris in a radial fashion that is permanently formed by the eighth month of gestation. During the development of the iris, there is no genetic influence on it, a process known as chaotic morphogenesis that occurs during the seventh month of gestation, which means that even identical twins have differing irises. The iris has in excess of 266 degrees of freedom,i.e. the number of variations in the iris that allow one iris to be distinguished from another.The fact that the iris is protected behind the eyelid, cornea and aqueous humour means that, unlike other biometrics such as fingerprints, the likelihood of damage and/or abrasion is minimal. The iris is also not subject to the effects of aging which means it remains in a stable form from about the age of one until death.The use of glasses or contact lenses (coloured or clear) has little effect on the representation of the iris and hence does not interfere with the recognition technology.

Advantages

- No contact required.
- Protected internal organ; less prone to injury.
- Believed to be highly stable over lifetime.

- Highly protected, internal organ of the eye .
- Externally visible; patterns imaged from a distance.
- Iris patterns possess a high degree of randomness.
- Changing pupil size confirms natural physiology.
- Pre-natal morphogenesis (7th month of gestation).
- Limited genetic penetrance of iris patterns.
- Patterns apparently stable throughout life.
- Encoding and decision-making are tractable.
- Image analysis and encoding time: 1 second .
- Decidability index (d-prime): d' = 7.3 to 11.4.
- Search speed: 100,000 IrisCodes per second on 300MHz CPU.

Disadvantages

- Difficult to capture for some individuals.
- Easily obscured by eyelashes, eyelids, lens and reflections from the cornea.
- Public myths and fears related to —scanning‖ the eye with a light source.
- Acquisition of an iris image requires more training and attentiveness than most biometrics
- Lack of existing data deters ability to use for background or watch list checks
- Cannot be verified by a human
- Small target (1 cm) to acquire from a distance(1 m)
- Moving target ...within another...on yet another.
- Located behind a curved, wet, reflecting surface.
- Obscured by eyelashes, lenses, reflections .
- Partially occluded by eyelids, often drooping.
- Illumination should not be visible or bright.
- Some negative (Orwellian) connotations.

1.7.10 Keystroke Dynamics

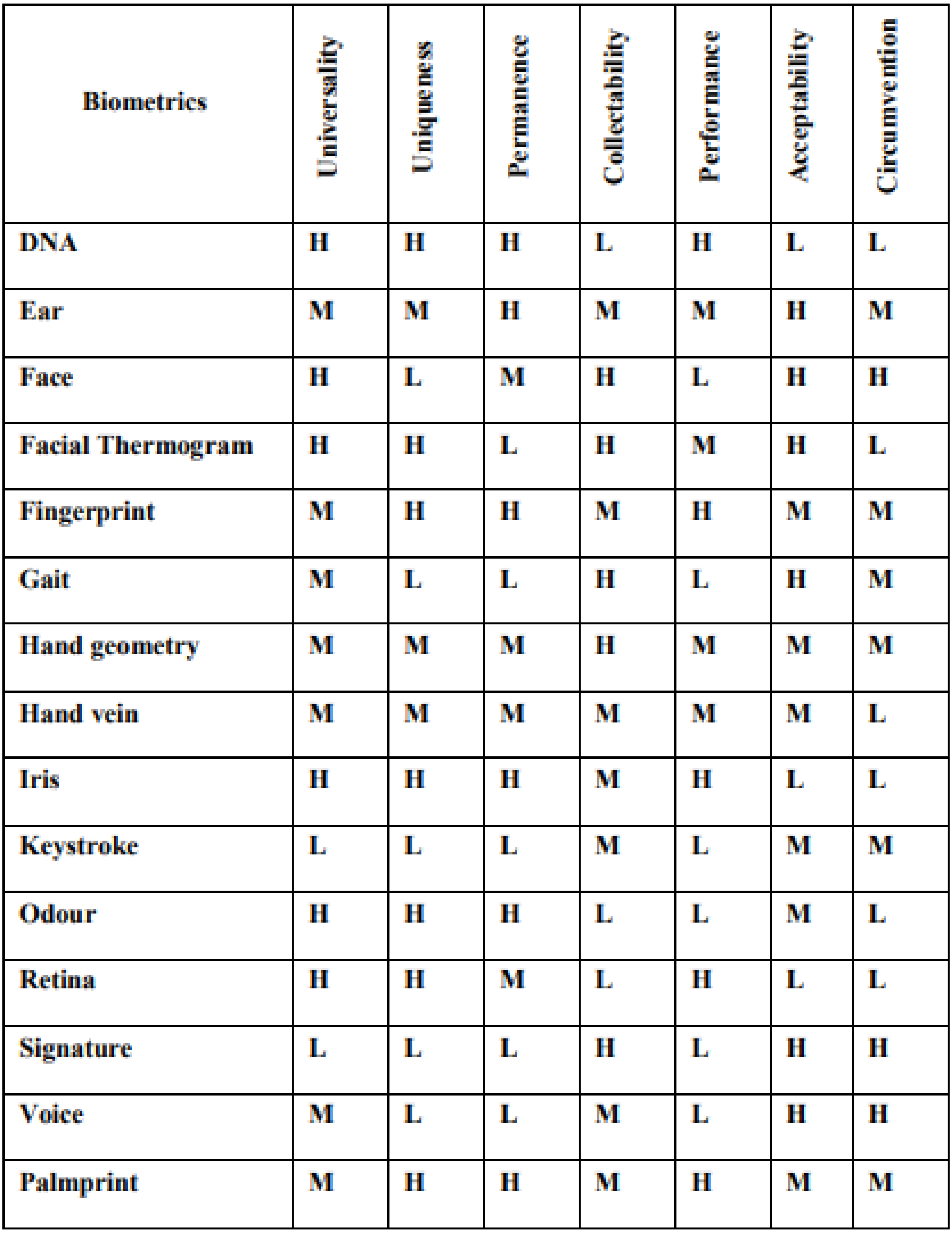
Keystroke dynamics is a biometric method that tries to identify in the typing of different keyboard users. This is an automated method of examining an individual‘s keystrokes on a keyboard. This technology examines such dynamics as speed and pressure, the total time of typing a particular password, and the time a user takes between hitting certain keys. This technology‘s algorithms are still being developed to improve robustness and distinctiveness. Any person knowing the username and password of another user cannot gain access rights for computer system due to the difference of their keystroke patterns. On the other hand, unlike the other biometric methods, Keystroke analysis does not require the aid of extra special tools [9]. Just keyboard and analysis software are sufficient for the biometric analysis. Therefore it is cheaper than biometric authentication methods. However, user authentication via keystroke dynamics remains a difficult task. Because, physiological features such as face, retinal, palm print and fingerprint remains stable over time. But behavioral properties such as keystroke patterns may vary according to the users‘ skill, illness and tiredness.

1.7.11 Odor

It is known that each object exudes an odor that is characteristic of its chemical composition and this could be used for distinguishing various objects. A whiff of air surrounding an object is blown over an array of chemical sensors, each sensitive to a certain group of (aromatic) compounds. A component of the odor emitted by a human (or any animal) body is distinctive to a particular individual. It is not clear if the invariance in the body odor could be detected despite deodorant smells, and varying chemical composition of the surrounding environment [9].

1.8 Comparison of various biometric technologies

Amit Mhatre and et al [11], Anil K. Jain et al [9] discusses the comparison of various biometric technologies. Table 2 shows the results of the comparison.



Biometrics	Universality	Uniqueness	Permanence	Collectability	Performance	Acceptability	Circumvention
DNA	H	H	H	L	H	L	L
Ear	M	M	H	M	M	H	M
Face	H	L	M	H	L	H	H
Facial Thermogram	H	H	L	H	M	H	L
Fingerprint	M	H	H	M	H	M	M
Gait	M	L	L	H	L	H	M
Hand geometry	M	M	M	H	M	M	M
Hand vein	M	M	M	M	M	M	L
Iris	H	H	H	M	H	L	L
Keystroke	L	L	L	M	L	M	M
Odour	H	H	H	L	L	M	L
Retina	H	H	M	L	H	L	L
Signature	L	L	L	H	L	H	H
Voice	M	L	L	M	L	H	H
Palmprint	M	H	H	M	H	M	M

Table 2: Comparison of various biometric techniques H- High, M – Medium, L – Low

2. ISSUES OF BIOMETRICS

While biometric systems can offer greater levels of security, various attacks exist to gain unauthorized access to a system that is protected by biometric authentication. The various issues of the biometric system are dealt here.

2.1System design issues

Biometrics is invariably associated with security; hence the biometric system itself should be reasonably secure and trustworthy. Some of the biometric security issues are

- Rogue sensors and unauthorized acquisition of biometric samples.
- Communications security between sensors, matchers and biometric databases.
- Accuracy.
- Sped.
- Scalability
- Resilience .
- Cost.
- Privacy.

2.2 Authentication or Identification?

The various key issues to be considered during the examination of biometrics for authentication systems are discussed here. First we need to identify whether the system is meant for identifying users, or simply authenticating users. Identification of a user is a much more difficult task.

Authentication is the verification that a user is who they claim to be [11]. For such a situation, the user provides a possible range of users of one. The authentication mechanism then compares the expected credentials for the claimed identity to the credentials it finds on the claimant. If there is a match, the user is authenticated. This is known as a one to one test.

In contrast, identification does not involve a claim of identity at all. Instead, the system is presented with a set of (ideally complete) credentials, and asked to compare this set of credentials against the users it knows of, returning a result, which identifies the user, in question. This is known as a one to many test, and it should be evident that this type of test is both more labour intensive for the system, and more reliant on having a wide range of attributes to compare users with.

2.2.1 Failure Rates

Failure rates are a critical consideration in the configuration and day to day running of a biometrics system. Two types of failure rates must be considered; false acceptance rates, and false rejection rates. These failure rates are a function of how precisely the system attempts to verify each user against the characteristics registered for them. Thus, a system, which is configured, to be very precise and have very low false acceptance rates will almost invariably be performing a higher number of false rejection rates, relative to a balanced system. Similarly a system, which involves lower value access, will likely be granted to be less precise to ensure that the positive customer experience is delivered.

A security professional will tend to move immediately towards a configuration with low false acceptance rates and ignore the false rejection rates. This is not a practical option in all cases, and the configuration depends strongly on the environment in which the biometrics-based system is to be deployed. In a high security military-style setting where security is paramount, users, or at least their administrators will be more accepting of a high false rejection rate which is in support of a low false acceptance rate. In a commercial setting, there will often be less scope for the acceptance of delays and difficulties associated with false rejection rates.

2.2.2 Liveness

A number of attacks on biometrics systems have been proposed over the years, and a number have been quite successful. Fingerprinting systems were originally entirely reliant on fingerprints for authentications, meaning that moulded synthetic fingers with imprints of fingerprints could be used to authenticate users. Similarly, hand geometry systems were entirely reliant on superficial physical attributes.

Since these attacks have been proposed, a new area of biometrics has arisen which focuses entirely on determining whether the authenticating attributes being measured are in fact the attributes of a living being, as opposed to a recording or a synthetic imitation. The mechanisms are varied, relying on things such as prompted user actions to smile for facial recognition. In the area of fingerprinting systems have been developed to measure both perspiration and the pulse of the authenticating user. Hence, liveness testing is becoming a vital part of biometrics systems.

Methods have been proposed to make spoofing biometric systems more difficult. The method that is considered here is the determination of liveness. To determine whether or not a person is live when they present their biometric data to a system can be a difficult task to automate in a fashion that is acceptable to users, and feasible to implement. Many methods exist, such as temperature sensing, detection of pulse in fingertip, pulse oximetry, electrocardiogram, dielectric response, and impedance. Each of these methods have their own challenges in being able to automate and integrate into systems in the most transparent way possible. For example, the extra equipment required to perform some of these tests, such as electrocardiogram, can be expensive and inconvenient for the user.

2.3 Physically Challenged Non-Registrable Users

During the rollout and use of a system based on biometrics, it is inevitable that some users will be found who cannot register for a given system due to their physical characteristics. Thus, a secondary authentication mechanism is always needed for biometric systems. Care must be taken to ensure that the use of biometrics does not mean the complete exclusion of a given group of users.

The mechanics of non-registrable users varies depending on the biometric mechanism being used. Even without actual physical injury, some mechanisms can suffer in particular sections of the population. When examining fingerprinting, for example, it has been found that the elderly often have either very dry skin, or very weak fingerprints simply as a result of aging skin.

2.4 Circumvention

It is key that when rolling out a biometrics system, the view of a system as a chain is maintained, along with the understanding that a system is only as secure as the weakest part of that chain. Biometrics offer a strong means to authenticate users for systems, however attackers will tend to attack systems at their weakest point, not their strongest.It is vital that biometrics serve a supporting role in well designed, properly secured systems. Installing biometrics into a fundamentally weak system is a waste of both resources and time.

2.5 Scalability

There are general concerns related to the scalability of biometrics systems - it is key that any solutions vendor be pressed to prove that the solution offered is going to be appropriately scalable.

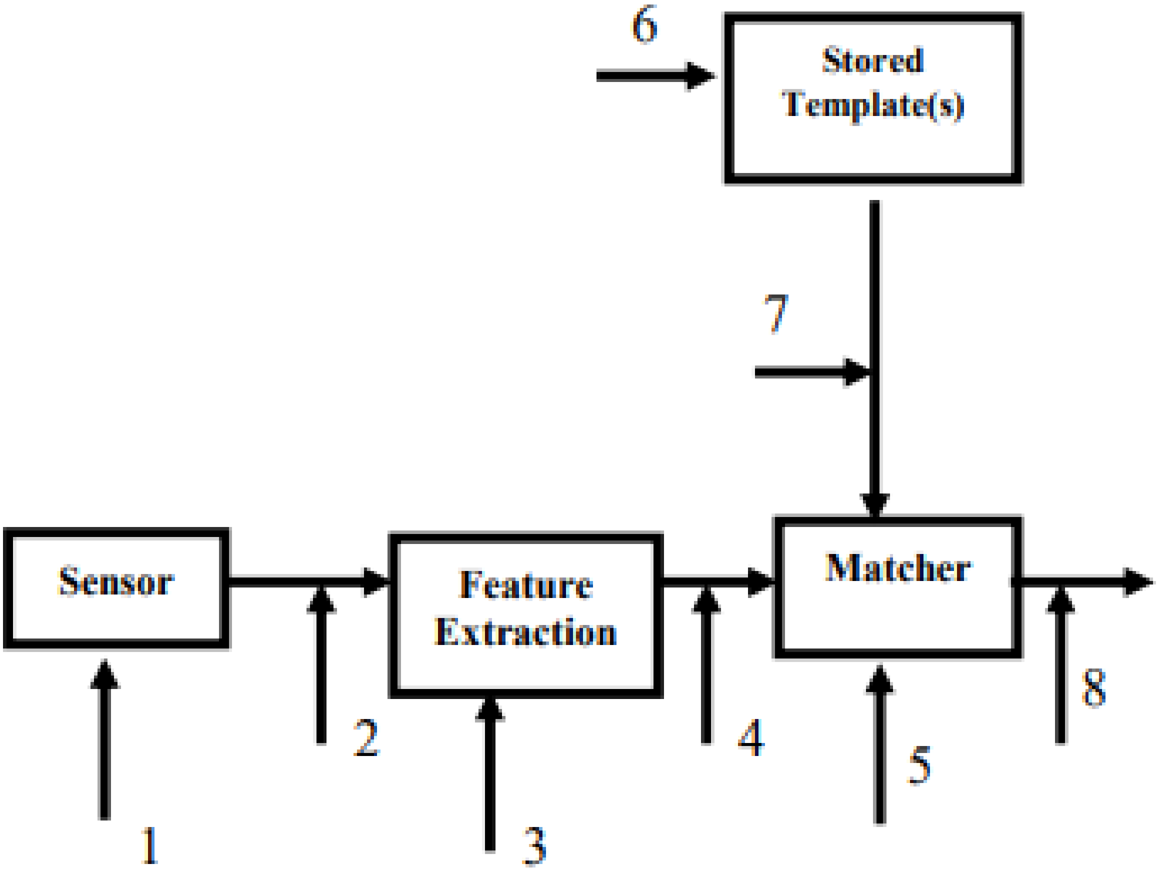


Figure 1.5: Figure 5 Vulnerabilities of a biometric system.

1. Fake biometric.

2. Replay old data.

3. Override feature extractor.

4. Synthesized feature vector.

5. Override matcher.

6. Modify template.

7. Intercept the channel.

8. Override final decision.

While it is highly recommended that test pilots be performed for rollouts of biometrics systems, the issue of scalability is a more difficult one to tackle. Biometrics rollouts across very large user populations may not scale well. Research into this area is lacking at present, and unfortunately the larger rollouts of biometrics to date have involved organizations that are less likely to wish to discuss their authentication mechanisms in depth.

Biometrics come with a complex set of issues which need careful consideration. The issues above illustrate a number of potential pitfalls, and hopefully give an indication of the consequences of ill-considered biometrics deployments.

N.K.Ratha et al [22] identified eight places in the generic biometric system (Figure 4) where attacks may occur.

3. GUIDELINES FOR BIOMETRIC TECHNOLOGY

3.1. Checklist of things to consider when choosing a biometric

Lists are easy to create:

- Physical contact and durability.
- Usage and durability.
- Flexible packaging.
- Infrastructure and interoperability.
- Enrolment.
- Standards.
- Security.
- Proven technology.
- Reliability.
- Accuracy.
- Capture.
- Liveness.
- User friendliness.
- Intrusiveness.
- Context of application.
- Convenience.
- Cost.

3.2 Authentication guidelines

Before enrolling a new user, the system must establish to a proper degree of certainty that a new user is indeed who they claim to be. Without sufficiently strong authentication during the enrolment step, a biometrics authentication system is effectively useless, no matter how strong the actual mechanisms employed.

Authentication for enrolment will always ideally be performed in-person with well-trained and examined staff. The verification process should be explicitly defined in terms of the identification required, and the steps, which must be taken to verify this identification. When possible, authentication should be a two-stage process to diminish the likelihood of insider attacks resulting in deliberately incorrect authentication of new users.

In some models, an in-person verification may not be possible, or may not be seen as necessary for user authentication. For example, a bank rolling out biometrics devices for authentication across the Internet may wish to mail out devices and incorporate their enrolment process into the next session the user establishes through the historic authentication mechanisms. In such cases, additional remote mechanisms must be found to ensure that enrolment is secure. If the chain of authentication to establish biometric authentication credentials contains a step with authentication of a lower strength than the biometrics layer, then an attacker will simply move one layer down and attack systems at the initial authentication/enrolment stage.

3.3. Enrollment Guidelines

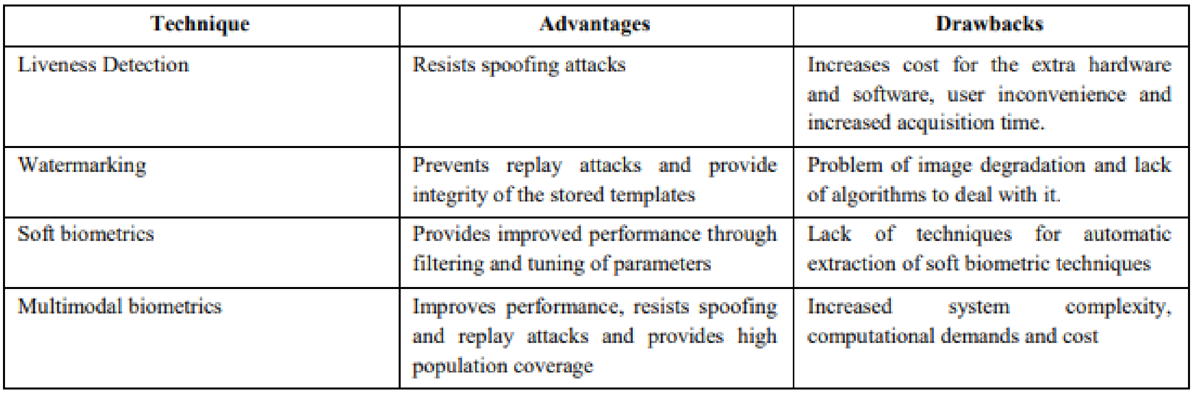
Once authentication has been performed and the user is to authenticate, the system and those administering it must ensure that the user credentials gathered from the mechanism are sufficient to identify the new user accurately. Biometrics systems assemble their views of users through sets of measurements, as described in the techniques section. The precise location of each measurement and number of measurements taken for each user may vary slightly, as each user will have different characteristics to compare. The biometrics system that is used must be configured and employed such that all user credential sets contain a minimum number of points of reference to identify or authenticate that user.

As an example, if a fingerprint based system typically relies on fifteen points of reference, and attempts to enroll a user with worn fingers, the system may simply not be able to enroll this user. If only ten points of reference are found, this must result in a failed enrolment, not simply a weaker profile for this user. Similarly, a user attempting to enroll with damaged fingerprints might enroll with a set of characteristics, which are both temporary and potentially reproducible. Enrolment officers must be qualified to judge both potential cases, and must have clearly defined procedures to deal with these issues.

Similarly, enrolment officers must be accustomed to dealing with equipment and its failures. In the case of mechanisms such as fingerprint authentication, sensors can wear out over time, leading to issues of either complete failure, or repeated generation of user information sets lacking the detail to form a proper authentication set.

3.4. Administration Maintenance

Biometrics systems will ideally be as low-maintenance as possible. If stored locally, the integrity of both authentication data and authentication mechanisms must be maintained. Role separation and tamper proof systems auditing are both controls to be strongly considered.



Technique	Advantages	Drawbacks
Liveness Detection	Resists spoofing attacks	Increases cost for the extra hardware and software, user inconvenience and increased acquisition time.
Watermarking	Prevents replay attacks and provide integrity of the stored templates	Problem of image degradation and lack of algorithms to deal with it.
Soft biometrics	Provides improved performance through filtering and tuning of parameters	Lack of techniques for automatic extraction of soft biometric techniques
Multimodal biometrics	Improves performance, resists spoofing and replay attacks and provides high population coverage	Increased system complexity, computational demands and cost

Figure 1.6: Table 3: Advantages and drawbacks of the different protection techniques.

If systems do require regular maintenance by administrative staff, role-based access controls should be considered to ensure that staff maintaining systems do not have access to either the data, logic, or logs of the systems. Similarly, auditing personnel should not have access to the system whose logs they examine. Maintaining clear separations of both roles and data access will ensure that data and logic functions are kept as securely as possible.

3.5. Increased Security and Perception

Biometrics has a real potential to boost some areas of security in a system, though clearly they are not a magical bullet for all security issues. Biometrics can play a real role in systems where identity theft is an issue, ensuring that each individual user is only present once on a system. Clearly though, this is still limited, as the first user to claim an identity on a given system is then the "owner" of this identity. While biometrics can be used to cut down on account hijacking, issues around fraudulent accounts fall back onto registration procedures, just as is the case with any other authentication system.

On the perception side, it has been suggested that in the casino trade, the use of facial recognition to monitor card counters and the like has been split regarding identifying culprits, and deflecting potential cheats.

3.6. Methods to overcome the Biometric Attacks

3.6.1 Liveness Detection:

Liveness detection refers to the ability of the system to distinguish between a sample feature provided by a live human being and a copy of a feature provided by an artifact. Liveness detection can be implemented using software or hardware means.

- Using extra hardware to acquire life signs like temperature, pulse detection, blood pressure etc for fingerprints and movements of the face for face recognition. Iris recognition devices can measure the involuntary papillary hippos (Constant small constrictions and dilations of the pupil caused by spontaneous movements of the Iris). The drawback is that extra hardware makes the system expensive and bulky.

- Using the information already captured to detect life signs. The only researched method is using information about sweat pores. For this a sensor that can acquire a high-resolution image is required. It is difficult to reproduce the exact size and position of the pores on an artificial mold.

- Using liveness information inherent to the biometric being obtained. For fingerprints, using a side impression near the nail, which has been enrolled earlier, can do this. The advantage is that people do not leave side impressions as latent prints and no major changes in the scanner is needed to acquire this additional information. A system that uses multiple instances of the same biometric can be used for liveness detection by asking the user to provide a random subset of biometric measurements.

3.6.2. Steganographic and Watermarking Techniques

Steganography means secret communication. It involves hiding critical information in unsuspected carrier data.

Steganography based techniques can be suitable for transferring critical biometric information from a client to a server.

Ratha [18] proposes a water marking technique applicable to fingerprinting images compressed with WSQ wavelet-based scheme. The discrete wavelet transform coefficients are changed during WSQ encoding by taking into consideration possible image degradation. This method is used to secure biometric authentication systems for commercial transactions against replay attacks. To achieve this, the service provider issues a different verification string for each transaction. The string is mixed with the fingerprint image before transmission. When the image is received by the service provider it is decompressed and the image is checked for a one-time verification string. Here, the message is not hidden in a fixed location, but is deposited in different places on the structure of the image so that it cannot be easily recovered.

4. Adventages of Biometric tehcnologies

A major motivation for using biometrics is the ability to authenticate the true identity of an individual [9]. Biometric technologies can be applied to areas requiring logical access solutions, and it can be used to access applications, personal computers, networks, financial accounts, human resource records, the telephone system, and invoke customized profiles to enhance the mobility of the disabled. In a business-tobusiness scenario, the biometric authentication system can be linked to the business processes of a company to increase accountability of financial systems, vendors, and supplier transactions; the results can be extremely beneficial.

The global reach of the Internet has made the services and products of a company available 24/7, provided the consumer has a user name and password to login. In many cases the consumer may have forgotten his/her user name, password, or both. The consumer must then take steps to retrieve or reset his/her lost or forgotten login information. By implementing a biometric authentication system consumers can opt to register their biometric trait or smart card with a company‘s businessto-consumer e-commerce environment, which will allow a consumer to access their account and pay for goods and services (e-commerce). The benefit is that a consumer will never lose or forget his/her user name or password, and will be able to conduct business at their convenience. A biometric authentications system can be applied to areas requiring physical access solutions, such as entry into a building, a room, a safe or it may be used to start a motorized vehicle. Additionally, a biometric authentication system can easily be linked to a computer-based application used to monitor time and attendance of employees as they enter and leave company facilities. In short, contactless biometrics can and do lend themselves to people of all ability levels.

5. DISADVANTAGES OF BIOMETRIC TECHNOLOGIES

Some people, especially those with disabilities may have problems with contact biometrics. Not because they do not want to use it, but because they endure a disability that either prevents them from maneuvering into a position that will allow them to make use the biometric or because the biometric authentication system (solution) is not adaptable to the user. For example, if the user is blind a voice biometric may be more appropriate. Some of the disadvantages are listed below:

- Biometric systems are very expensive because, not only the costs for the acquisition of the software and hardware costly but the integration of these in the networks are even more costly. These high costs are coupled with the fact that the returns aren‘t highly encouraging. So, people are not ready to pool in so much money to utilize the latest technology that is available in the market.

- It is an "all or none" technology, i.e. we set up biometric authentication features etc but if we permit the person for a remote login then there is no use incorporating this technology in the network.
- Like every new technology, Biometrics has a low user acceptance rate.
- People consider it to be an invasion of their privacy and thus, it hasn‘t been exploited to its full potential. They don‘t realize the fact that a Biometric system does not copy their fingerprints or any other attributes but goes for a mathematical representation of these attributes that are unique to each person.
- Even though full secrecy is maintained regarding these attributes, even if they get leaked out once, they can be used in exploiting various other areas, like to get credit card and medical information, in banking security systems etc. Even though different biometric systems are highly incompatible with each other, their exploitation may ruin the life of the person who trusted this technology
- Sometimes, a genuine person maybe restricted access to the network and this is very commonly seen in voice recognition patterns where something as small as cold could have the person‘s access rejected.
- Like all systems, even a Biometric system is not foolproof and has its own flaws and can sometimes allow a person who has assumed a fake identity into the network.
- Biometric template data consume more space than the conventional user ID/password combinations.

6. LIMITATIONS OF BIOMETRICS

The main reason for introducing biometric systems is to increase overall security. However, biometric identification is not perfect it is never 100 % certain, it is vulnerable to errors and it can be ’Spoofed’.

The biometric system is only one part of an overall identification or authentication process, and the other parts of that process will play an equal role in determining its effectiveness. That is a biometric sample may be used to authenticate a person before logging into a system, or it may be used in lieu of login. Also in some cases it may be used to verify a person after logging into to the system.

Biometric identification is a statistical process. Variations in conditions between enrolment and acquisition as well as bodily changes (temporary or permanent) mean that there is never a 100 % match. For a password or a PIN, the answer given is either exactly the same as the one that has been stored, or it is not – the smallest deviation is a reason for refusal; for a biometric, there is no clear line between a match and a non-match. Whether a match exists depends therefore not only on the two data sets to be compared, but also on what margin of error is deemed tolerable.

A 90 %probability of a match may or may not be considered acceptable, depending on the implementation of the biometric in question and the application security requirements. Fraudulent reproduction of biometric data is possible; this depends heavily on the modality, application and resources being considered and availability of the data to be reproduced.

Biometric data may be stored on portable media such as smart cards if they will be used in verification mode. This ensures that the data cannot be used without the user‘s own authorization, contrary to what happens with data stored in a central database. Biometric verification/ identification can also be realized through remote access, by transmission of the biometric image or template through a network to the device that will process the decision step. This requires a highly secure connection. Watermarking could be used in this case to ensure that the transmitted data have not been corrupted.

Of course, smart cards can be lost or stolen. For this reason, the data they contain must be encrypted and backed-up. However, if the information is stolen, it is necessary to be able to revoke it and to produce another template which could be used for further identification. Revocation is easy when dealing with pin codes or passwords but not with biometric traits as we cannot change our irises or our fingerprints.

Cancellable biometrics [9] is a new research field and some preliminary propositions have been made. It is possible to generate new facial images for a person by filtering the original image.

7. BIOMETRIC APPLICATIONS

Most biometric applications fall into one of nine general categories:

- Financial services (e.g., ATMs and kiosks).
- Immigration and border control (e.g., points of entry, precleared frequent travelers, passport and visa issuance, asylum cases).
- TSocial services (e.g., fraud prevention inentitlement programs).
- Health care (e.g., security measure for privacy of medical records).
- Physical access control (e.g., institutional, government, and residential).
- Time and attendance (e.g., replacement of time punch card).
- Computer security (e.g., personal computer access, network access, Internet use, e-commerce, e-mail, encryption).
- Telecommunications (e.g., mobile phones, call center technology, phone cards, televised shopping).

- Law enforcement (e.g., criminal investigation, national ID, driver‘s license, correctional institutions/prisons, home confinement, smart gun).

8. CONCLUSION

Biometrics is a promising and exciting area, where different disciplines meet and provide an opportunity for a more secure and responsible world. There are a number of popular biometrics mechanisms currently deployed, some with strong histories, and some relatively new mechanisms. Each mechanism has its own strengths and weaknesses. When properly applied, biometrics can be used to combat fraud, and ensure that timekeeping systems are honest and accurate.

Using one biometric feature can lead to good results, but there is no reliable way to verify the classification. To achieve robust identification and verification two different biometric features can be combined. A multimodal biometrics can provide a more balanced solution to the security and convenience requirements of many applications.

Recent advances in biometric technology have resulted in increased accuracy at reduced costs; biometric technologies are positioning themselves as the foundation for many highly secure identification and personal verification solutions.

Despite the tremendous progress made over the past few years, biometric systems still have to reckon with a number of problems, which illustrate the importance of developing new biometric processing algorithms as well as the consideration of novel data acquisition techniques. Undoubtedly, the simultaneous use of several biometrics would improve the accuracy of an identification system. For example the use of palmprints can boost the performance of hand geometry systems. Therefore, the development of biometric fusion schemes is an important area of study. The possibility of using biometric information to generate cryptographic keys is also an emerging area of study. Thus, there is a definite need for advanced signal processing, computer vision, and pattern recognition techniques to bring the current biometric systems to maturity and allow for their large-scale deployment.

2 Convolutional Neural Network (CNN)

Contents

2.1 Introduction 34

2.2 Related Works 35

2.2.1 Data Pre-Processing 36

2.2.2 Data Imbalance 37

2.3 Feature Selection 37

2.3.1 Extra Tree Classifier 38

2.3.2 Pearson Correlation 38

2.3.3 Forward Stepwise Selection 39

2.4 ML Classifiers Techniques 39

2.4.1 Supervised Machine Learning Models 39

2.4.2 Ensemble Machine Learning Model 41

2.4.3 Neural Network Models 42

2.5 Hyperparameter Tuning 43

2.5.1 Cross-Validation 43

2.6 Dataset Description 44

2.6.1 Data Pre-Processing 46

2.6.2 Feature Selection 49

2.6.3 Model Selection 50

2.6.4 Model Evaluation 50

2.7 Analysis 51

2.8 Conclusion 51

2.1 Introduction

Cardiovascular diseases (CVDs) [26] have become more prevalent in recent times with mortality rates counting as high as 15–17 million per year, which is about 30% 31% of the total deaths worldwide according to the WHO reports. Myocardial infarction, coronary heart diseases, and arrhythmia are some of the most common CVDs. The precarity of CVD makes it one of the most treacherous diseases that prevail presently in society. What accounts for this unpredictability is the diversity of contributing features that range from inevitable factors like inheritance, age, gender to avoidable ones like alcohol abuse and tobacco consumption. Recent medical reports reveal that about 80% of the patients who die premature (below the age of 70) fall under the elderly category and about 25%of the remaining are middle-aged women. Though these factors (age and sex of the patient) are not the best way to discern the worst of the cases, they cannot be entirely disregarded. Other clinical factors like the concentration of serum creatinine and ejection fraction play more prominent roles in patients' survival prediction, especially for those that have been receiving treatment over a period of time.

Considering the diversity of causal factors associated with CVDs, it remains a challenge to sort out the significant features that may aid in timely prediction of a patient's survival. Filtering the prominent factors requires careful perusal through different medical records to note the degree of influence a specific factor exerts on the patient's survival.

Hence, to ensure high efficacy and maximum accuracy, Machine Learning (ML) models could be exploited, as have been in the proposed model. ML models can be used for meticulous prediction of the target variable (which in this case is the survival rate of a patient) by taking into account several affecting features. The affecting features, however, could be a mix of both important as well as less-influential factors. Hence, to overcome this challenge, the proposed model employs feature selection methods.

The dataset used to train and test the ML models was obtained from live medical records [23] collected at the Faisalabad Institute of Cardiology and at the Allied Hospital in Faisalabad (Pakistan) from April to December 2015. The dataset contains the records of 299 heart failure patients, about 105 of which are women and whose age ranges from 40 to 95 years. All the patients listed in the records had suffered left ventricular dysfunction and had a history of cardiac arrests and fall under class III and IV of heart failure stages as deemed by New York Heart Association (NYHA) [6].

The dataset, which entails these live medical records, is normalized to treat outliers and other noise that may interfere with the model's predictive potential and also oversampled as the data showed conspicuous imbalance between classes.

Feature selection methods, like Extra Tree Classifiers, Pearson Correlation, Forward Selection, and Chi-square tests, were performed to sort out the most prominent features. Serum creatinine concentration, ejection fraction, age of the patients, and the duration of treatment were proved to have maximum influence on the patients' death event and were hence used as cardinal predictive variables in the model. Doing this increased the model's prediction accuracy to about 75%.

The proposed model constitutes several ML algorithms like SVM, CNN [27], and other ensemble methods that are trained on the set of medical records. These algorithms were weighed individually based on their accuracy score for different sets of features. The models were trained over three different studies: one with follow-up time, creatinine level, and ejection fraction as major factors, another with age instead of follow-up time and the third over the entire set of features available in the dataset.

2.2 Related Works

Chronic cardiac failure occurs when the fluid accumulates around the heart, causing it to pump insufficiently for the body to function normally. There are two types of chronic cardiac failure based on left ventricular function: systolic heart failure (reduced ejection fraction HFrEF) and diastolic heart failure. For normal people, the percentage of blood that is pumped out by the left ventricle in each contraction is 50%˘70%. The reduced ejection fraction [19] is a case which occurs if the ejection fraction is below 40% without much effective contraction by the heart muscles. The causes of HFrEF are Ischemic cardiomyopathy and dilated cardiomyopathy. The diastolic heart failure [17] occurs when there is a normal contraction in heart muscles but no relaxation in ventricles during ventricular filling. With the increasing severity of CVDs, a need for predicting heart failures is on the rise but the traditional methods employed for CVD-related event prediction, unfortunately have failed to achieve the acme of accuracy.

Given a set of medical records as datasets, ML can be employed to achieve high accuracy in the prediction of patient survival and also in determining the driving factors like ejection fraction and serum creatinine [7] that increase mortality among CVD patients.

According to the NYHA, the heart failure is classified into four classes [4] based on the heart failure symptoms. Class I is where the patient experiences asymptomatic left ventricular dysfunction with normal physical activities, Class II with slight symptoms and limited physical activities, Class III with modest symptoms and very few physical activities, and Class IV with severe symptoms on complete rest. The dataset used in this analysis consists of 299 patients with left ventricular dysfunction belonging to Class III or IV [6].

The survival rate and the heart failure hospitalizations of patients suffering from heart failure due to preserved ejection fraction [5] using various ML models like logistic regression, SVM, random forest, and gradient descent boosting are predicted. Lasso regression and forward selection feature selection techniques [**?**] were used to find that blood urea nitrogen levels, cardiomyopathy questionnaire subscore, and body mass index are the most significant features affecting the survival rate of the patients. The random forest is found to be the best model for predicting the survival rate of the HFpEF patient with an AUC of 0.76.

The decision tree boosting model is trained using 5822 records of the patients in the UCSD Cohort [1, 9] with eight clinical features: diastolic blood pressure, creatinine, blood urea nitrogen, hemoglobin, white blood cell count, platelets, albumin, and red blood cell distribution width. The decision tree is the best classifier with AUC of 0.88. The Cox regression model [2] predicts that age is the most significant feature that affects the survival rate of the patient and with an increase in each year the mortality rate increases by 4%. The next most significant features are ejection fraction and serum creatinine with a p-value of 0.0026, and it is evident that the rate of survival decreases doubly with a one-unit increase in serum creatinine. Serum sodium and anemia were also significant with a p-value of 0.0052 and 0.0096, respectively. Other features like gender, smoking creatinine phosphokinase, and platelets were not significant for predicting the survival rate as these parameters may be significant at the initial stage of heart failure but this analysis contains patients at stage III or IV of heart failure.

The preceding studies presents finding the significant features using statistical methods, which leaves a plenty of space for various ML-based feature selection methods in healthcare services. The aim of this study is to address the issue of efficiency using various feature selection methods to rank the importance of the features in predicting the survival rate. Figure 8.1 shows the workflow model of the proposed methodology.

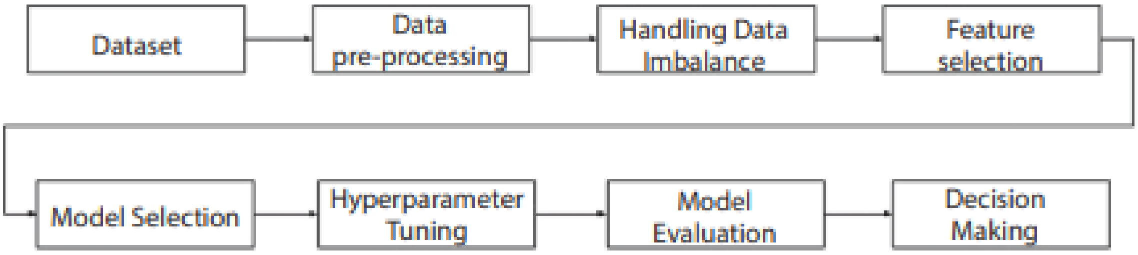


Figure 2.1: Figure 8.1 Workflow model of proposed system.

2.2.1 Data Pre-Processing

Data pre-processing [3] is a cardinal method in data mining technique as it involves cleaning and organizing the data before being used. Processing of data is crucial as it may contain undesirable information such as outliers, noise, duplicate data, and, sometimes, even wrong data which might be misleading. In order to service these complications, data needs to be processed prior to its utilization. Sometimes, data may contain non-numeric and null values or missing attribute values along with imbalanced data. These issues can also be resolved using data pre-processing techniques. The most common procedures involved are data normalization, data wrangling, data resampling, and binning.

2.2.2 Data Imbalance

The distribution of data equally among the classes in the classification problem is one of the most important issues to be addressed to avoid misclassification due to severely skewed classes. There are various techniques to encounter class imbalance problems in the dataset. One of the techniques used for this analysis is Synthetic Minority Oversampling TEchnique (SMOTE).

SMOTE [28] is an oversampling method which is used to scale up the minority class in the dataset. A minority class is a class that is undersampled or in other words has significantly less data samples that fall under it. The presence of such imbalance may cause the training model to often overlook the minority classes; since training models have the propensity to consider the majority classes due to the considerable amount of training data they happen to possess. SMOTE is one of the techniques used to resample the minority classes to achieve equal proportions of data samples as is distributed among other classes. In SMOTE, the oversampling (or resampling) is achieved by considering all the data points that fall under the minority class and connecting each point to its k-nearest homologous neighbors. After repeating this for all the data points, the model artificially synthesizes or creates sample points that lie on the lines joining the original data points. The resampling is done until the data becomes balanced. ADAptive SYNthetic (ADASYN) imbalance learning technique is similar to SMOTE when it comes to the initial oversampling procedures. The data points are connected to each of their k-nearest neighbors and sample points are created on those connecting lines. But in ADASYN, the points are instead placed closer to those lines as opposed to on them, thereby inducing variation in the data samples.

2.3 Feature Selection

Feature selection or variable selection is a cardinal process in the feature engineering technique [13] which is used to reduce the number of dependent variables by picking out only those that have a paramount effect on the target attribute (survival rate in this case). By employing this method, the exhaustive dataset can be reduced in size by pruning away the redundant features [15] that reduces the accuracy of the model. Doing this will help curtail the computational expense of modeling and, in some cases, may also boost the accuracy of the implemented model [12].

The feature selection models applied herein to achieve maximum accuracy in target prediction encompass the following:

1. Extra Tree Classifier.
2. Pearson Correlation .
3. Forward selection.
4. Chi-square.
5. Logit (logistic regression model).

The features that recurred in at least three of the above-mentioned models were assumed as having maximum impact on the target variable and were accordingly classified to fit into the employed ML model.

2.3.1 Extra Tree Classifier

The Extra Tree Classifier or the Extremely Random Tree Classifier [21] is an ensemble algorithm that seeds multiple tree models constructed randomly

from the training dataset and sorts out the features that have been most voted for. It fits each decision tree on the whole dataset rather than a bootstrap replica and picks out a split point at random to split the nodes. The splitting of nodes occurring at every level of the constituent decision trees are based on the measure of randomness or entropy in the sub-nodes. The nodes are split on all variables available in the dataset and the split that results in the most homogenous sub-childs is selected in the constituent tree models. But unlike other tree-based ensemble algorithms like random forest, an Extra Tree Classifier does not choose the split that results in most homogeneity; instead, it picks a split randomly from the underlying decision tree framework. This lowers the variance and makes the model less prone to overfitting.

2.3.2 Pearson Correlation

Pearson Correlation is used to construct a correlation matrix that measures the linear association between two features and gives a value between 1 and 1, indicating how related the two features are to one another. Correlation is a statistical metric for measuring to what degree two features are interdependent and by computing the association between each feature and the target variable, the one exerting high impact on the target can be picked out. In other words, this model helps determine how a change in one variable reflects on the outcome. The measure of linear association between the features is given by the Pearson Correlation Coefficient which can be computed using the equation:

where x_i is the i^{th} value of the variable x, x is the average value of sample attribute x, n is the number of records in the dataset, and x and y are the independent and target variables. A value of 1 indicates positive correlation, 1 indicates negative correlation, and 0 indicates no correlation between the features.

2.3.3 Forward Stepwise Selection

Forward selection is a wrapper model that evaluates the predictive power of the features jointly and returns a set of features that performs the best. It adds predictors to the model one at a time and selects the best model for every combination of features based on the cumulative residual sum of squares. The model starts basically with a null value and with each iteration the best of the attributes are chosen and added to the reduced list. the addition of features continues so far as the incoming variable has no impact on the model's prediction and if such a variable is encountered it is simply ignored.

2.4 ML Classifiers Techniques

Classification models predict the classes or categories of given features using a variety of ML models. The classification algorithms can be classified into different models.

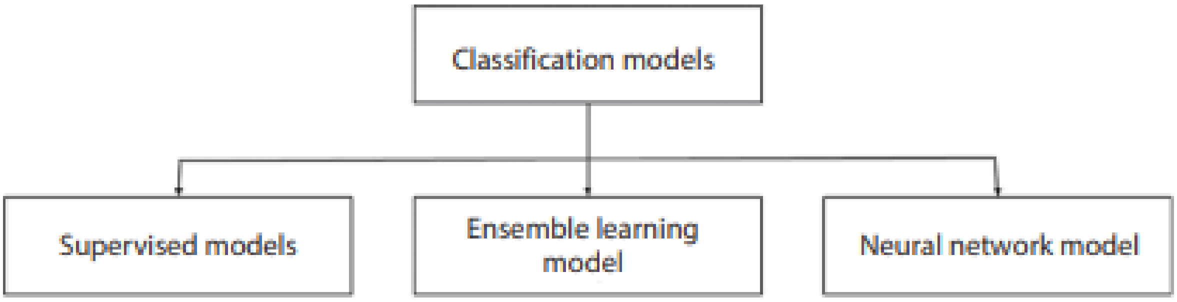


Figure 2.2: Face recognition structure.

2.4.1 Supervised Machine Learning Models

Supervised learning is a branch of ML wherein a model is trained on a labeled dataset to deduce a learning algorithm. This is achieved by allowing the model to make random predictions and to correct and teach itself using the labels available. This happens iteratively until the learning algorithm specific to the data has been deduced and the model has achieved considerable accuracy. Supervised learning is a powerful technique in training ML models and is predominantly used. It can be classified into regression and classification based on whether the data is continuous or categorical.

2.4.1.1 Logistic Regression

As opposed to its name, a logistic regression model is a binary classification algorithm used when the target variable is categorical, or in other words when the target variable could be grouped into two different classes. The prediction is done by fitting the data to an activation function (a sigmoid logistic function) which returns a value between 0 and 1. The returned value determines how strongly a data entry belongs to one of the binary classes. This approach not only helps achieve higher accuracy but also ensures great precision, which makes it an ubiquitous, fundamental, and handy algorithm for binary classification. This statistical model can be mathematically formulated as follows:

where is the logistic function applied to a weighted sum of independent variables xi : i (0, n), where n is the total number of data entries available in the dataset. Conventionally, x0 is assigned the value of 1 which leaves just 0 in the sum, which is considered as a bias. A bias term is included so that even when the model is applied over no independent variable or the sum merely cancels all the terms (positive and negative) the final result does not come to be 0. This makes sense because (0) returns 0.5 which is ambiguous since class prediction becomes vague at that point.

2.4.1.2 SVM

Support Vector Machines (SVM) is a classification algorithm that introduces a best suited decision boundary which splits the dataset accordingly. It is generally used for binary classification but can also be extended to multi-class classification. SVM relies on Support Vectors, which are the data points that are closest to the hyperplane used to classify the dataset. For datasets that cannot be split using a decision boundary, an SVM model uses kernels to extend the data points to a higher dimension and by constructing a hyperplane to separate the classes. The kernels used are of two types: polynomial and radial. SVM models have higher accuracy even when fit to smaller data samples and in some cases may also outperform neural networks.

2.4.1.3 Naive Bayes

Naive Bayes is a fast learning classification algorithm based on Bayes' theorem. The reason it is called naive is because the algorithm supposes the independence of the predictor variables. Or in other words, it assumes that a particular feature is not affected by the presence of other features. For each data point, the algorithm predicts the probability of how related the feature is to that class and the class for which the probability ranks highest is chosen as a probable class [23]. Bayes' theorem can be given as follows:

where P(c|xi) denotes the conditional probability that the data point may belong to class c provided it belongs to the feature xi, i (1, n), n is the number of data entries.

2.4.1.4 Decision Tree

Decision tree is a classification algorithm that splits the dataset into homogenous classes based on the most significant features [16]. With every split, the resulting sub-node is more homogeneously classified than the previous level. As the name suggests, the algorithm uses a tree-like structure to split the datasets. The best split is identified by a series of computations that include several other methods like Gini score, Chi-square, and Entropy, each of which return a value between 0 and 1. Higher the value of Gini and chi-square and lower the value of entropy, more is the resultant class homogeneity among the sub-nodes when split using that particular feature. However, the downside of this algorithm is that it is sensitive to overfitting. But even this can be overcome by appropriate measures like setting height or node constraints and tree pruning.

2.4.1.5 K-Nearest Neighbors (KNN)

KNN is used to classify a data point based on the influence exerted upon it by the k neighboring data points. Here, k is a value often input explicitly which denotes the minimum number of neighbors that influence the class that the data point may belong to. The choice of k should not be too low nor too high as it may result in overfitting and data underfitting, respectively. If the value of k is given as 1, then the model is most likely to overfit since the closest point of influence to a particular data point will be the point itself. Such overfitting increases the variance of the model, thereby deeming it unfit for class prediction. Hence, an optimum k value can be computed by plotting the validation error curve. The k value corresponding to the local (or global) minima of the curve can be chosen since it indicates an optimum number of neighboring class influences that will aid the classification of the data point with maximum accuracy.

2.4.2 Ensemble Machine Learning Model

Ensemble learning is a learning algorithm in which multiple models are constructed and combined into one powerful predictive model [25]. Pooling of models into the ensemble helps achieve maximum accuracy than is often obtained from individually trained models. An ensemble aggregates the result of the models based on which it builds a powerful classifier which has low variance and high accuracy. An ensemble can be used to supplement for weaker models that are prone to overfit. Common ensemble methods involve bagging, boosting, and stacking

2.4.2.1 Random Forest

Random forest classifier is an ensemble tree algorithm which can be applied for both categorical as well as continuous valued data. The “Forest” in the name implies that multiple trees are seeded and grown to the maximum depth, each on a different bootstrap replica of the dataset. Each tree classifier is grown on a randomly selected subset of features from the original dataset and is given the choice to vote for a significant feature. The random forest classifier selects the most-voted class. This algorithm can be extended to both regression as well as classification based models and is prone to outliers and noise in the dataset. While versatile, this algorithm is considered a black-box algorithm due to the fact that most of the random classification, data distribution and “voting” techniques are obscure.

2.4.2.2 AdaBoost

AdaBoost is a boosting algorithm that is used to convert weaker classifiers to strong classifiers. It calculates the weighted sum of all the weak classifiers. The weak classifiers are fit to the dataset and the one that gives the least classification error is chosen. Using this, the weight is calculated and it is updated for every data point. The classifier equation is as follows:

where f i (x) is the weak classifier and i denotes the calculated weight.

2.4.2.3 Bagging

Bagging is again a tree-ensemble algorithm. It grows decision tree CART models on a subset of the data sample. The decision trees are grown parallely to maximum height and are not pruned. Hence, every tree grown as a part of the ensemble model has high variance and low bias. The CART models vote for a class and the most popular class is chosen. The reason the trees are grown deeper is that the concern of overfitting is less in the bagging algorithm. Therefore, presence of noise does not affect the models performance. The data samples used for growing the trees are randomly selected from the dataset so that the data samples are independent of each other. These are called the bootstrap replica. Fitting these bootstrap samples instead of the entire dataset to the trees increases independence of the sample subsets and thereby decreases output variance.

2.4.3 Neural Network Models

Neural network architecture is collection of interconnected nodes/neurons distributed among different layers like Input layer, hidden layer, and output layer. Based on the number of hidden layers in the architecture the neural network is classified as shallow and deep neural networks. The shallow networks consist of a single hidden layer, whereas deep neural networks consist of two or more hidden layers.

2.4.3.1 Artificial Neural Network (ANN)

ANN [10] is a fundamental model in neural networks which works on the basic concepts of Multi-Level Perceptron (MLP). ANN has a series of inputs (continuous or categorical) each with a designated weight. Each of these weighted inputs are summed over multiple latent layers and finally passed through an activation function that generates an output between 0 and 1 or 1 and 1 depending on the function chosen. An ANN is called a Universal Function Approximator due to its potential to fit any non-linear data. Every layer of the network is densely connected, i.e., every neuron is connected to every other neuron and the inputs are processed only in the forward direction; hence, it is also a Feed-Forward Network. The capability to fit any non-linear model is introduced by the activation function, which helps the network uncover complex associations between the input and the output. But ANN is not suitable for back-propagation, in which the weights of the inputs can be altered to minimize the error function at any layer by propagating back.

2.4.3.2 Convolutional Neural Network (CNN)

CNN [14] is a neural network model usually used for visual image classification but can also be extended to fit other categorical data as well. A CNN consists of a series of densely connected latent layers of weights and filters between the input and the output. As in a regular neural network, the process involves the element-wise dot product of the input with a set of two-dimensional array of weights, called filter. The result obtained as a result of this weighted sum is a two-dimensional feature map. The elements of the feature map are applied to an activation function (ReLU or softmax), which returns a dichotomous result of 0 or 1 indicating the class to which the particular data belongs [19].

2.5 Hyperparameter Tuning

Hyperparameters are external model configurations that are used to estimate model parameters that cannot be directly estimated. Hyperparameter tuning is a method used to choose the best set of values for the model parameters in order to achieve better accuracy. In grid search tuning [8] method, the method is built over a range of parameter values that is specified explicitly by the practitioner in a grid format. It goes over every combination of model parameter values (each given as a list or an array) and picks out the one that has better performance than the rest.

Since it builds and evaluates the model over every possible parameter combination, grid search tuning is exhaustive and deemed computationally expensive.

In contrast, random search randomly picks out points from the hyperparameter grid and evaluates the model. Also, the number of iterations or the number of combinations to be tried out can be specified externally. This reduces time and the expense of computational power. The best score returned by the random search is based on the randomly chosen hyperparameter grid values, yet this method outperforms Grid Search with comparable accuracy in less time.

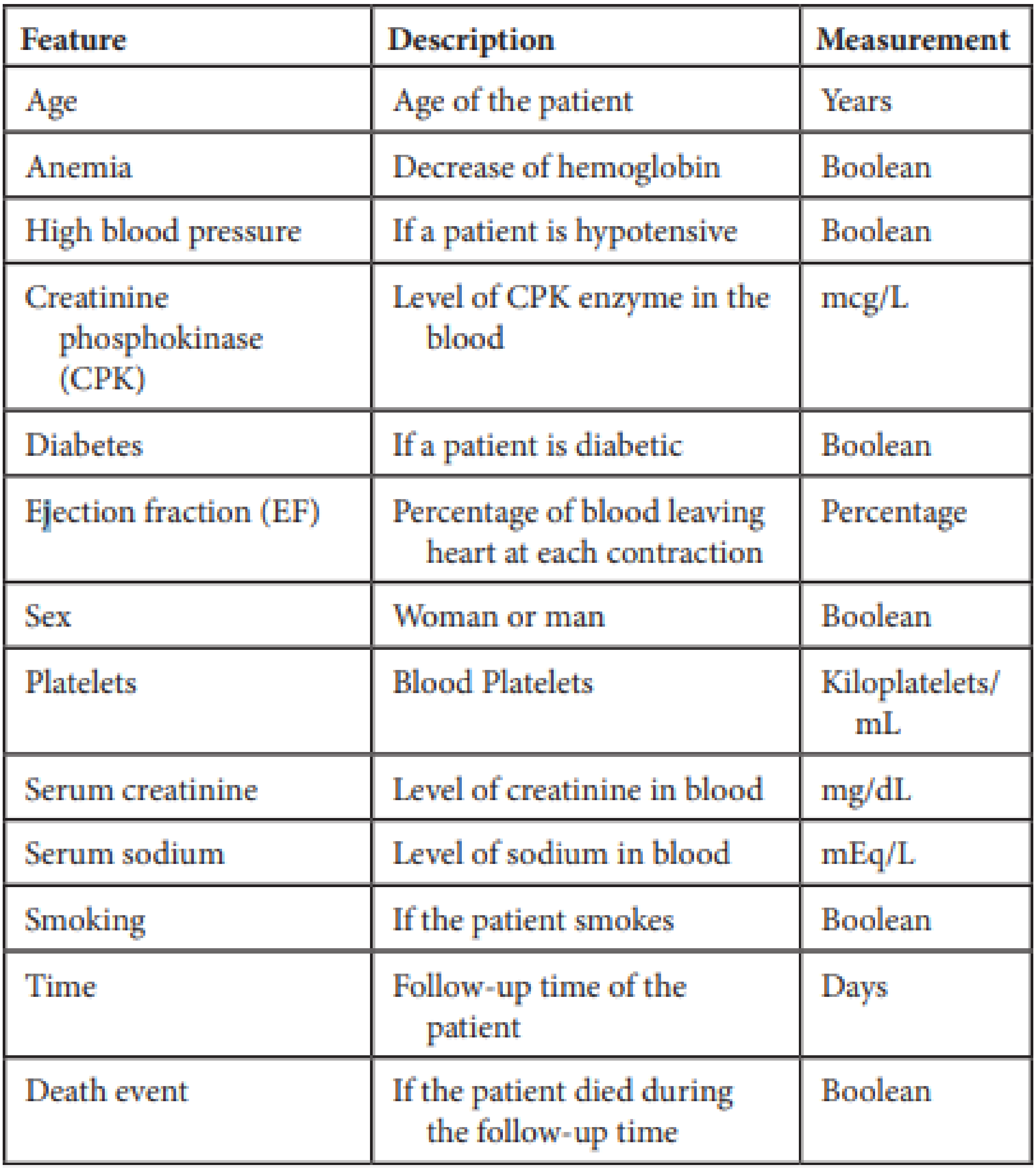
2.5.1 Cross-Validation

Cross-validation is a resampling algorithm used to compare different ML models or to find the best tuning parameter for a model. Basically, the dataset is divided into k blocks where k is provided externally, and every block is tested with other blocks as the training set. This is done iteratively for different models [20]. Cross-validation then returns the model that performed the best and with high accuracy. This method helps figure out the best algorithmic model for a given dataset. It is also used in SVM kernels and other similar models for finding hyperparameters.

2.6 Dataset Description

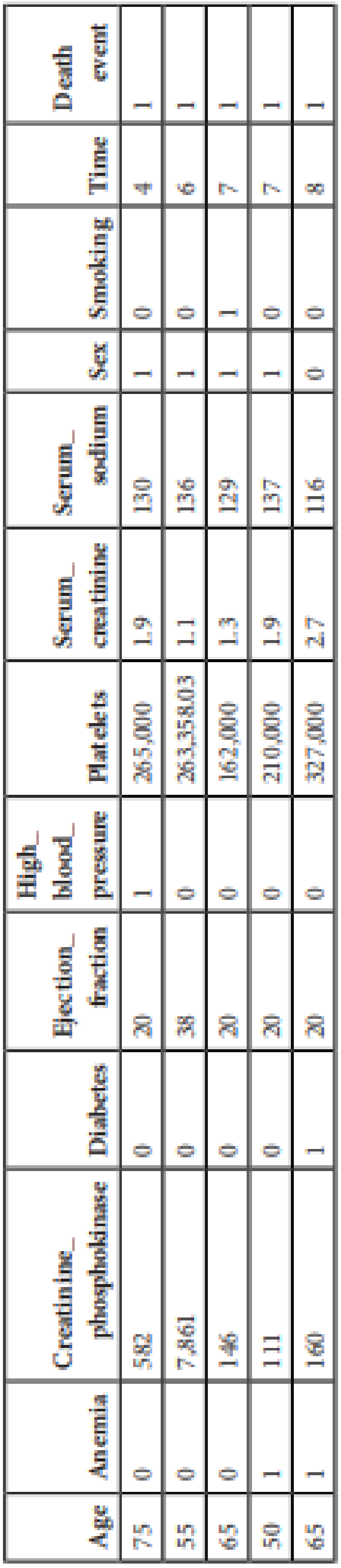
For this study, we used a heart failure clinical dataset containing medical records of patients who had left ventricular systolic dysfunction and with a history of heart failure. The dataset consists of 300 patient's records from Faisalabad Institute of Cardiology, from April to December 2015.

About 64% of the patients in the dataset are Male and the rest 36%are female patients, with their ages ranging between 40 to 95 years old. There are 13 features for each patient in the dataset, out of which 6 features are the clinical variables of the patient like the level of serum creatinine, serum sodium, creatinine phosphokinase, ejection fraction, blood platelets count, and the medical follow-up period of the patients. The target variable in this dataset is the death event feature with binary label, survived patient (death event = 0) and deceased patient (death event = 1). Figure 2.3 shows the attributes in the dataset, and Figure 2.4 shows the sample dataset taken for analysis.



Feature	Description	Measurement
Age	Age of the patient	Years
Anemia	Decrease of hemoglobin	Boolean
High blood pressure	If a patient is hypotensive	Boolean
Creatinine phosphokinase (CPK)	Level of CPK enzyme in the blood	mcg/L
Diabetes	If a patient is diabetic	Boolean
Ejection fraction (EF)	Percentage of blood leaving heart at each contraction	Percentage
Sex	Woman or man	Boolean
Platelets	Blood Platelets	Kiloplatelets/ mL
Serum creatinine	Level of creatinine in blood	mg/dL
Serum sodium	Level of sodium in blood	mEq/L
Smoking	If the patient smokes	Boolean
Time	Follow-up time of the patient	Days
Death event	If the patient died during the follow-up time	Boolean

Figure 2.3: Description of each feature in the dataset.



Age	Anemia	Creatinine_ phosphokinase	Diabetes	Ejection_ fraction	High_ blood_ pressure	Platelets	Serum_ creatinine	Serum_ sodium	Sex	Smoking	Time	Death event
75	0	582	0	20	1	265,000	1.9	130	1	0	4	1
55	0	7,861	0	38	0	263,358.03	1.1	136	1	0	6	1
65	0	146	0	20	0	162,000	1.3	129	1	1	7	1
50	1	111	0	20	0	210,000	1.9	137	1	0	7	1
65	1	160	1	20	0	327,000	2.7	116	0	0	8	1

Figure 2.4: Sample dataset.

2.6.1 Data Pre-Processing

In this study, the goal is to predict the survival rate of a patient with a history of heart failure and Figure 2.5 shows the complete architecture of the proposed system. There is a moderate class imbalance in these data; only 33% of the records contain information about the deceased patients as shown in Figure 2.6. This imbalance in classes will generate a biased model impacting the analysis. To address the class imbalance problem, we applied SMOTE to increase the number of records in the death event attribute.

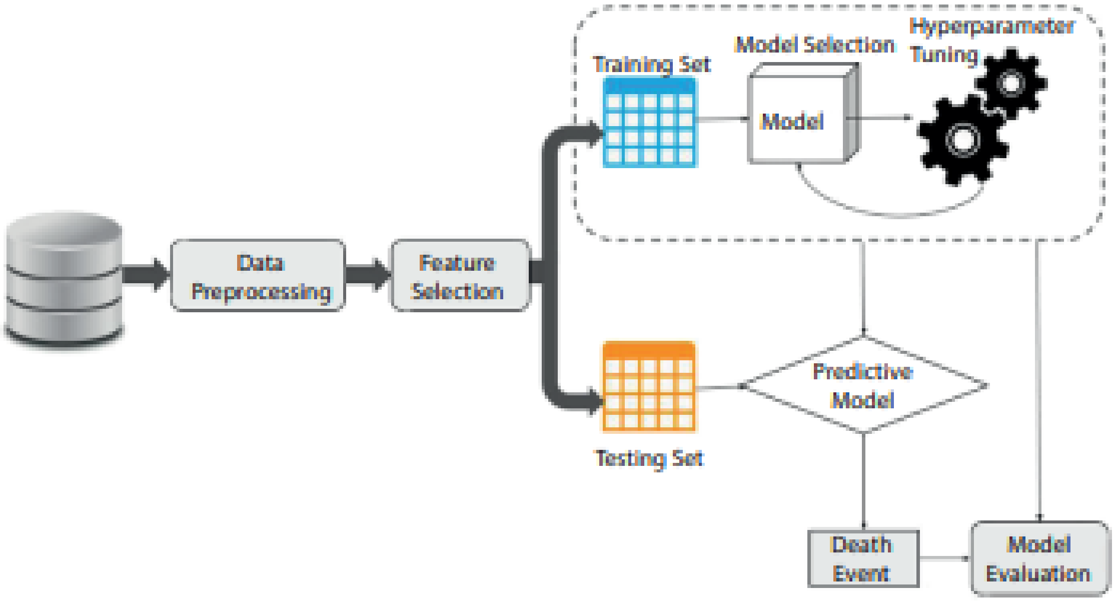


Figure 2.5: Architecture of proposed system.

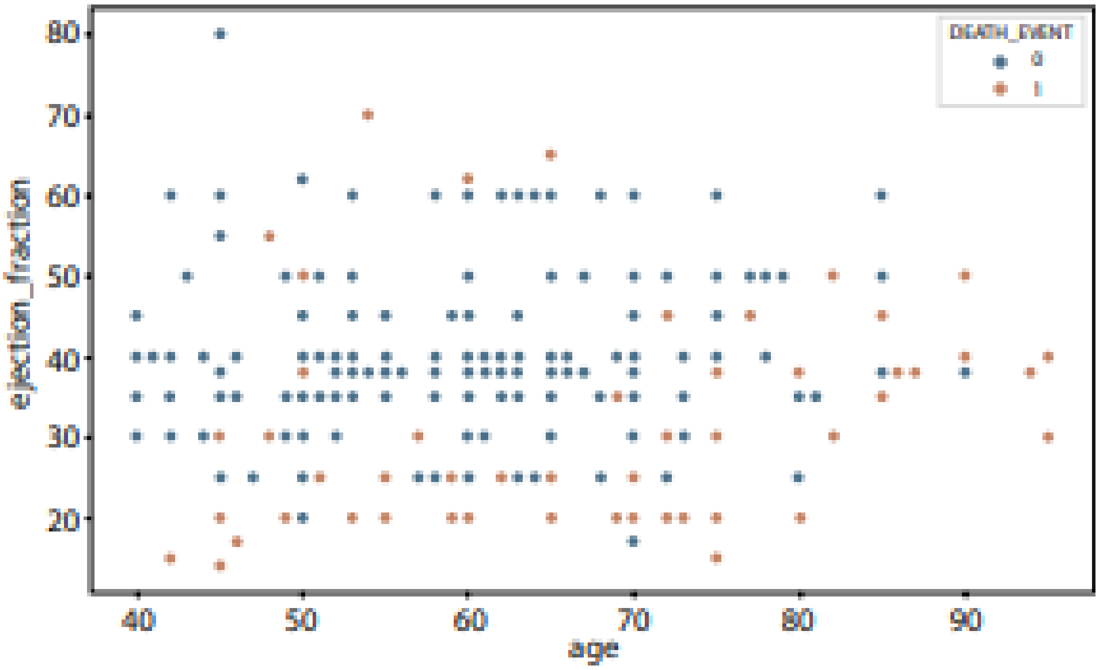


Figure 2.6: Original dataset distribution.

After applying sampling techniques, the number of records is increased to 406 with 203 numbers of entries for each class.

Figure 2.7 shows the distribution of target class (DEATH EVENT) and it is evident from the plot that class 0 is the majority class and class 1 is minority class. After applying SMOTE resampling technique the classes are balanced with 203 records under each class as shown in Figures 2.8.

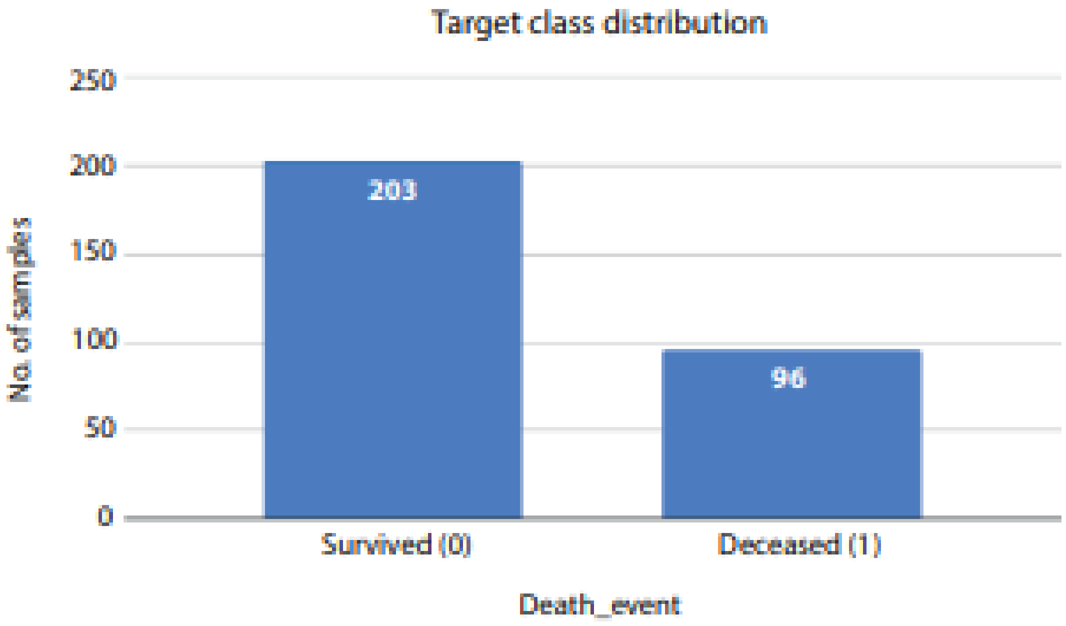


Figure 2.7: Figure 8.5 Target class distribution.

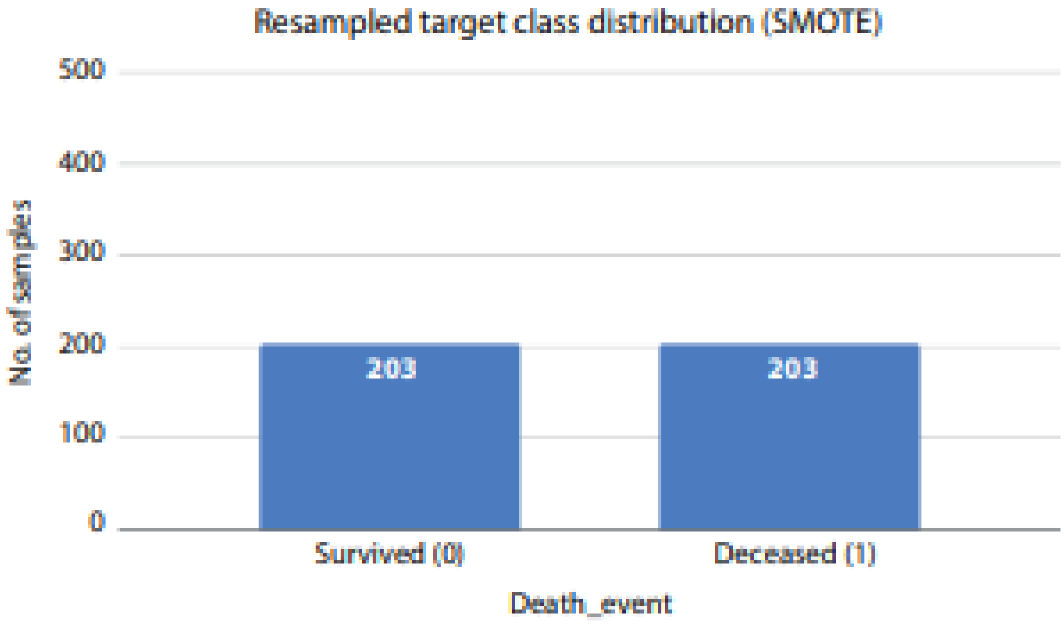


Figure 2.8: Figure 8.6 Resampled distribution applying SMOTE.

2.6.2 Feature Selection

The dataset consists of 13 attributes of which 3 most relevant features are selected using various feature selection algorithms, Extra Tree Classifier, Forward selection, Chi-square, Pearson Correlation, and Logit model. The Extra Tree Classifier generates the follow-up time, ejection fraction, and serum creatinine as the most important features that influence the survival rate of the heart failure patient as shown in Figure 2.9. Forward feature selection results show that age, ejection fraction, and serum creatinine as the top features.

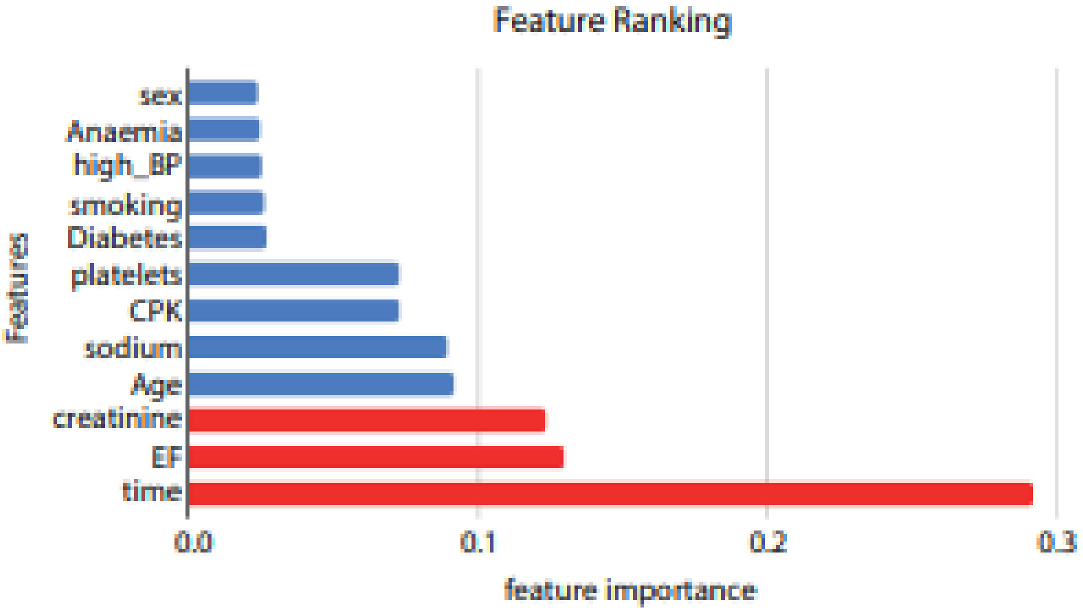


Figure 2.9: Feature ranking using Extra tree classifier.

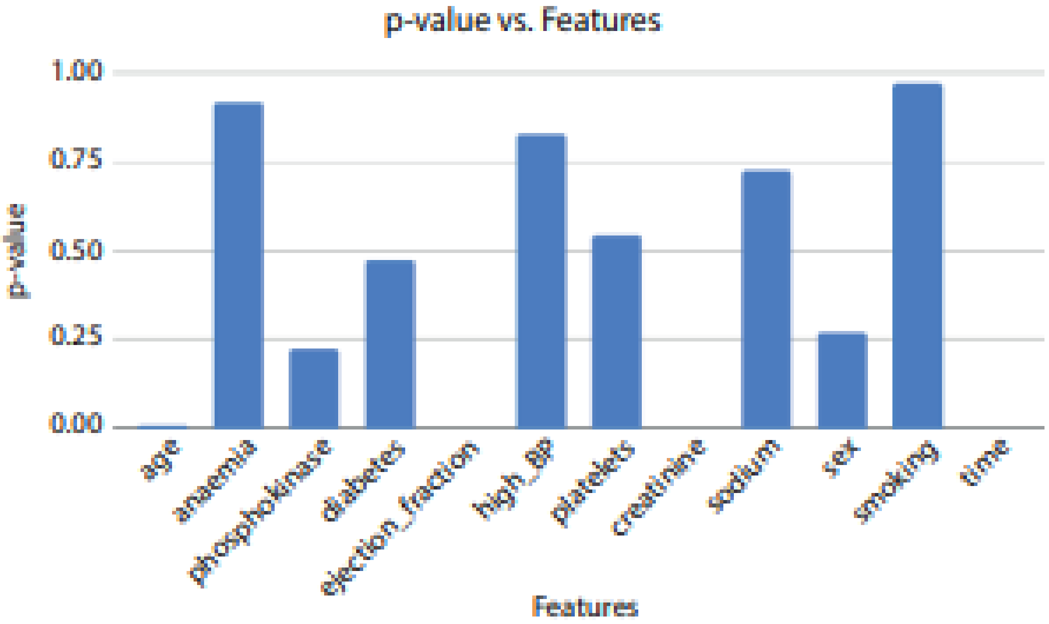


Figure 2.10: p-values of the features.

From Extra Tree Classifier, Pearson Correlation, and Logit model, the hypothesis suggests that time, ejection fraction, and serum creatinine as the top three most significant features. Whereas the other models, Forward Selection and Chi-Square generate that age, ejection fraction, and serum creatinine as the three most significant features. The p-values of ejection fraction, serum creatinine, and time are 0.0 (as shown in Figure 2.10) and are the most significant factors affecting the survival rate of the heart failure patient.

2.6.3 Model Selection

For the analysis, the dataset was split into $80\% and 20\%$ of the data as training and testing samples respectively. The model is trained for a two-class problem where the patient's survival rate is predicted given the medical parameters of the patient. The patient is classified as deceased or survived based on the most important medical features which affect the survival rate of the heart failure patient. The study is divided into three analyses based on the features selected by the various feature selection algorithms. The first study considers the important features predicted by the forward selection algorithm (age, serum creatinine, and ejection fraction). The second study predicts the survival rate of the patient using the attributes which are predicted most important by the Extra Tree Classifier and p-value (time, serum creatinine, and ejection fraction). The third study trains the ML model using all 13 features in the dataset to predict the death event of the patient. The dataset is trained on different ML models and neural networks, Logistic regression, Naive Bayes, SVM, Decision tree, Random forest, K-nearest neighbors, Bagging, AdaBoost, XGBoost, ANN, and Convolutional Neural Network. The hyperparameters of the models are tuned during the learning process using Grid Search and cross-validation techniques to optimize the model and minimize the loss by providing better results. For models such as SVM, neural networks, random forest, and decision trees where hyperparameter tuning is applied, the dataset is split into 70%of training samples, 15% of validation samples, and 15%as testing samples.

2.6.4 Model Evaluation

The classification model can be evaluated with the N × N confusion matrix, where N is the number of classes. The matrix compares the predicted variables and the actual variable to evaluate how well the model performs. The 2 × 2 matrix contains four values: True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN). The rows and columns of the matrix represent the actual and predicted values. The confusion matrix for the target variable (Death Event) is found, and the performance of the supervised models trained is quantified using classification evaluation metrics accuracy and F1-score.

2.7 Analysis

There is no universal model that performs best for the whole pool of datasets available for analysis. The prediction and the performance of the model depend on the nature and quality of the dataset on which the model is trained. From our analysis, it shows that the ensemble models XGBoost and random forest out performs the baseline ML algorithms and neural networks for the considered dataset. The tree-based model gives a decent result in a fast manner, whereas neural networks lack the one reason for this contusion that is the size of the data. Neural networks like CNN require a large and feature engineered dataset to train on and hypertuned parameters to outperform in the evaluation. With increasing dataset, the performance of CNN is also observed to be increasing, but XGBoost gives the desired results with 1,000 records itself and further increasing the dataset samples may cause the ML models to overfit even though neural networks may see an increasing trend. Hence, ensemble and ML classifier models are adequate for the prediction of survival rate of the heart failure patient with minimum resources and reduced time complexity.

2.8 Conclusion

Convolutional neural networks (CNNs) have accomplished astonishing achievements across a variety of domains, including medical research, and an increasing interest has emerged in facial expretion. Although deep learning has become a dominant method in a variety of complex tasks such as image classification and object detection, it is not a panacea. Being familiar with key concepts and advantages of CNN as well as limitations of deep learning is essential in order to leverage it in biometric research with the goal of improving image processing performance.

3 Experiments and Results

Contents

3.1 Introduction **53**

3.2 Overall architecture of the application **53**

3.3 Image acquisition **53**

3.4 Pretraitment **53**

3.5 Feature extraction **53**

3.6 Convolution neural network **54**

3.6.1 pooling 55

3.7 parameters **55**

3.7.1 Epoch 55

3.7.2 Batch 55

3.8 Database (CK+) **56**

3.9 Parameters tuning **56**

3.10 Resnet-34 . . . . **58**

3.11 Resnet-50 . . . . **61**

3.12 Densnet 121 . . . . **64**

3.13 Wideresnet50 . . . . **67**

3.14 Comparisons **70**

3.15 Real-time application **70**

3.15.1 Python 70

3.15.2 Pytorch 71

3.16 Visual Studio Code **72**

3.17 Application **72**

3.18 General conclusion **79**

3.1 Introduction

In this work, we used a framework for extracting facial expressions from a set of facial images. Experimental large-scale results were presented using a standard database: CK + (seven categories of expression) (fear, anger, happiness, disgust, surprise, sadness, contempt) The implementation of the proposed system has been divided into four components In the first component, a region of interest as face detection has been performed from the captured input image. For extracting more distinctive and discriminant features, in the second component, a deep learning-based convolutional neural network architecture has been proposed to perform feature learning tasks for classification purposes to recognize the types of expressions. in the third component, The system categorized the results. In the fourth component, Results are displayed Models that will apply are resnet-34

3.2 Overall architecture of the application

The overall architecture of our system is as follows:

3.3 Image acquisition

Image Acquisition is the first step in any image processing system. The general aim of any image acquisition is to transform an optical image (real-world data) into an array of numerical data which could be later manipulated on a computer. Image acquisition is achieved by suitable cameras. We use different cameras for different applications. If we need an X-ray image, we use a camera (film) that is sensitive to X-rays. If we want an infrared image, we use cameras that are sensitive to infrared radiation. For normal images (family pictures, etc.), we use cameras that are sensitive to the visual spectrum.

3.4 Pretraitment

we propose a new framework to integrate these two tasks using unified cascaded CNNs by multitask learning. The proposed CNNs consist of three stages. In the first stage, it produces candidate windows quickly through a shallow CNN. Then, it refines the windows by rejecting a large number of nonfaces windows through a more complex CNN. Finally, it uses a more powerful CNN to refine the result again and output five facial landmarks positions.

3.5 Feature extraction

A CNN is not only a deep neural network with many hidden layers but also a large network that simulates and understands stimuli as the visual cortex of the brain processes. CNN's output layer typically uses the neural network for multiclass classification. CNN uses the feature extractor in the training process instead of manually implementing it. CNN's feature extractor consists of special types of neural networks that decide the weights through the training process. CNN provides better image recognition when its neural network feature extraction becomes deeper (contains more layers), at the cost of the learning method complexities that had made CNN inefficient and neglected for some time. CNN is a neural network that extracts input image features and another neural network classifies the image features. The input image is used by the feature extraction network. The extracted feature signals are utilized by the neural network for classification. The neural network classification then works on the basis of the image features and produces the output. The neural network for feature extraction includes convolution layer piles and sets of pooling layers. As its name implies, the convolution layer transforms the image using the process of the convolution. It can be described as a series of digital filters. The layer of pooling transforms the neighboring pixels into a single pixel. The pooling layer then decreases the image dimension. As CNN's primary concern is the image, the convolution and pooling layers' procedures are intuitively in a two-dimensional plane. This is one of CNN's distinctions with other neural networks (Kim, 2017).

3.6 Convolution neural network

Convolution is a mathematical operation which involves a combination of two functions to produce a third function. In CNN the convolution is performed on the input data with the use of a filter to produce a feature map.

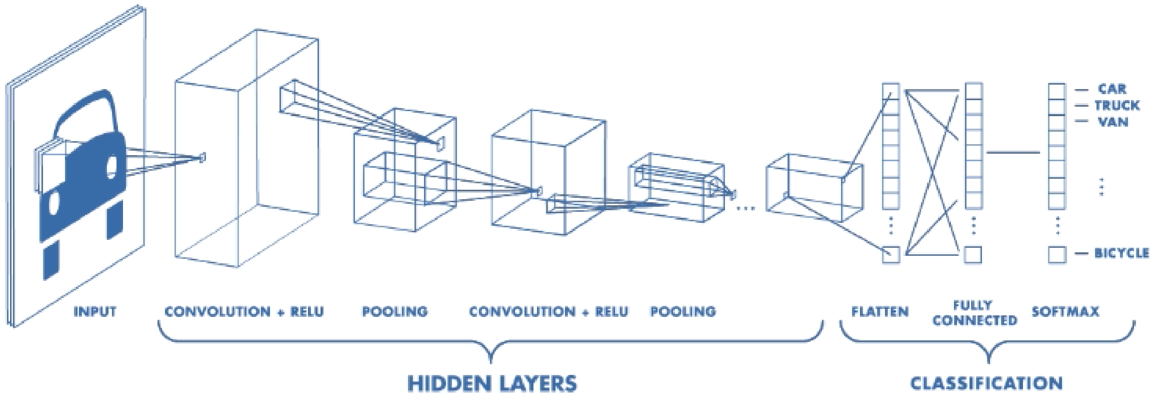


Figure 3.1: Architecture of a CNN. — Source https://www.mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks–1489512765771.html.

3.6.1 pooling

Pooling layer is added after a convolution layer. It performs continuous dimensionality reduction i.e reduces the number of parameters and computations thereby shortening training time and controlling overfitting. One such pooling technique is called max-pooling, which takes the maximum value in each window which decreases the feature map size while keeping the significant information.

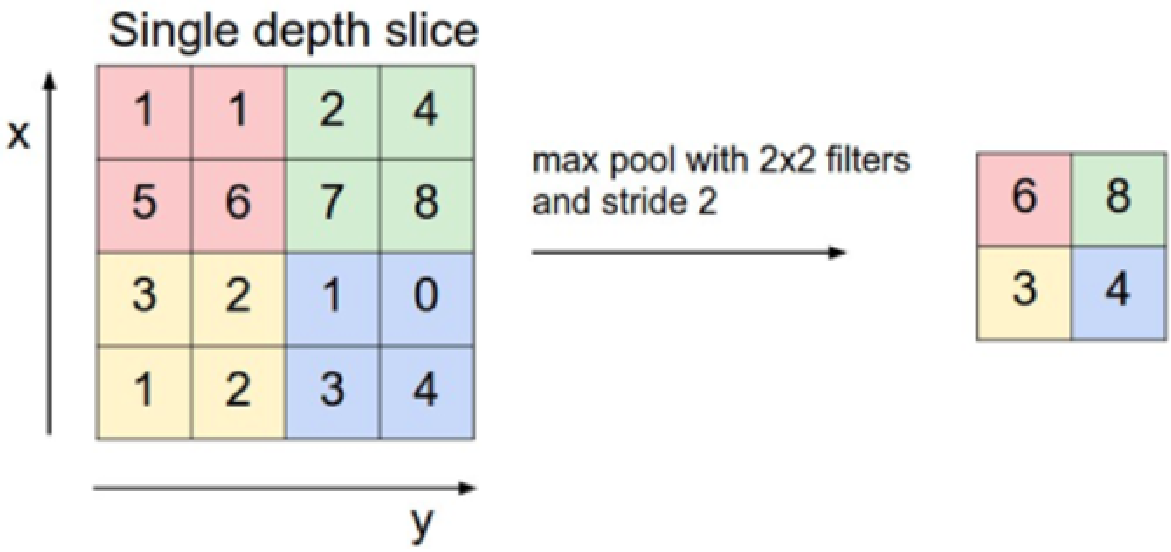


Figure 3.2: Max pooling takes the largest values. — Source: http://cs231n.github.io/convolutional-networks/

3.7 parameters

In this section, we will explain parameters what is used in the process of drawing conclusions.

3.7.1 Epoch

The number of epochs is a hyperparameter that defines the number times that the learning algorithm will work through the entire training dataset. One epoch means that each sample in the training dataset has had an opportunity to update the internal model parameters. An epoch is comprised of one or more batches. For example, as above, an epoch that has one batch is called the batch gradient descent learning algorithm.

3.7.2 Batch

The batch size is a hyperparameter that defines the number of samples to work through before updating the internal model parameters. Think of a batch as a for-loop iterating over one or more samples and making predictions. At the end of the batch, the predictions are compared to the expected output variables and an error is calculated. From this error, the update algorithm is used to improve the model. A training dataset can be divided into one or more batches. When all training samples are used to create one batch, the learning algorithm is called batch gradient descent. When the batch is the size of one sample, the learning algorithm is called stochastic gradient descent. When the batch size is more than one sample and less than the size of the training dataset, the learning algorithm is called mini-batch gradient descent.

3.8 Database (CK+)

The Extended Cohn-Kanade (CK+) dataset contains 593 video sequences from a total of 123 different subjects, ranging from 18 to 50 years of age with a variety of genders and heritage. Each video shows a facial shift from the neutral expression to a targeted peak expression, recorded at 30 frames per second (FPS) with a resolution of either 640x490 or 640x480 pixels. Out of these videos, 327 are labelled with one of seven expression classes: anger, contempt, disgust, fear, happiness, sadness, and surprise. The CK+ database is widely regarded as the most extensively used laboratory-controlled facial expression classification database available, and is used in the majority of facial expression classification methods.

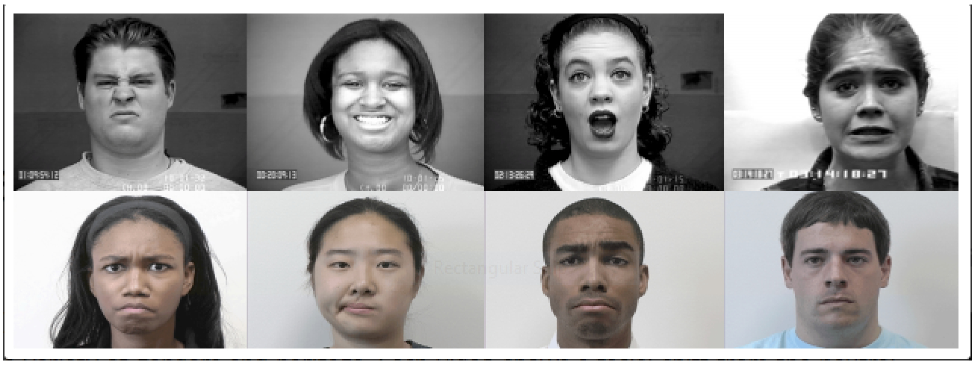


Figure 3.3: CK+ samples corresponding to facial expressions

3.9 Parameters tuning

In this work, we presented five modes to work on with four models, we explain modes as follows:

- MODE(1): size(88), epoch(50), batch(25)
- MODE(2): size(88), epoch(50), batch(50)
- MODE(3): size(88), epoch(100), batch(25)
- MODE(4): size(88), epoch(100), batch(50)
- MODE(5): size(112), epoch(100), batch(50)

The four models that we tested are as follows:

Densenet121. In a traditional feed-forward Convolutional Neural Network (CNN), each convolutional layer except the first one (which takes in the input), receives the output of the previous convolutional layer and produces an output feature map that is then passed on to the next convolutional layer. Therefore, for 'L' layers, there are 'L' direct connections; one between each layer and the next layer.

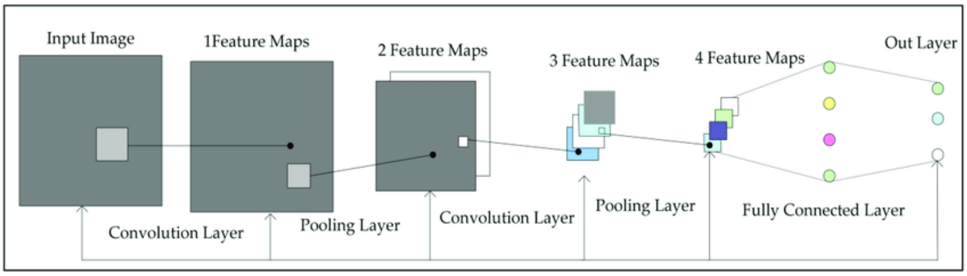


Figure 3.4: DenseNet architecture.

However, as the number of layers in the CNN increase, i.e. as they get deeper, the 'vanishing gradient' problem arises. This means that as the path for information from the input to the output layers increases, it can cause certain information to 'vanish' or get lost which reduces the ability of the network to train effectively.

DenseNets resolve this problem by modifying the standard CNN architecture and simplifying the connectivity pattern between layers. In a DenseNet architecture, each layer is connected directly with every other layer, hence the name Densely Connected Convolutional Network. For 'L' layers, there are L(L+1)/2 direct connections.

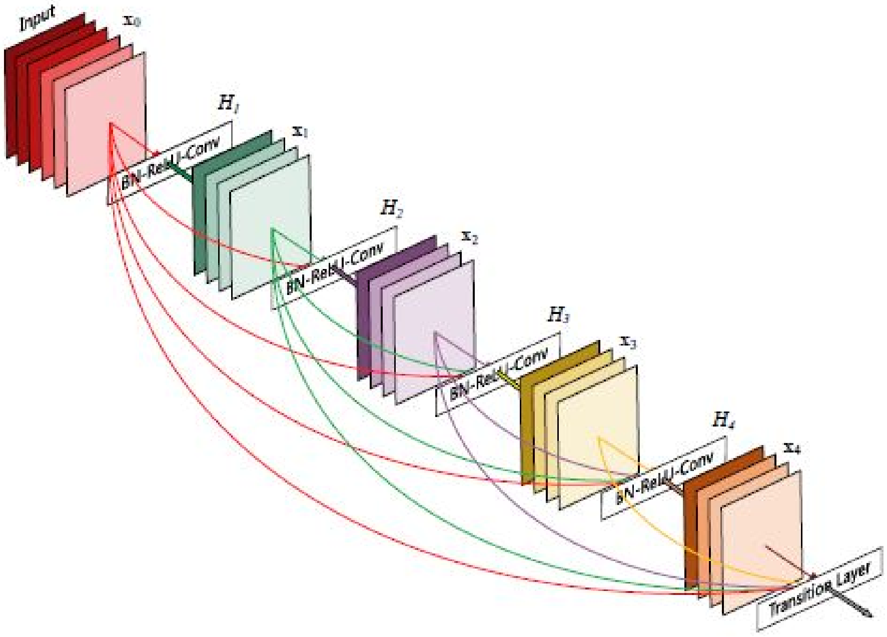


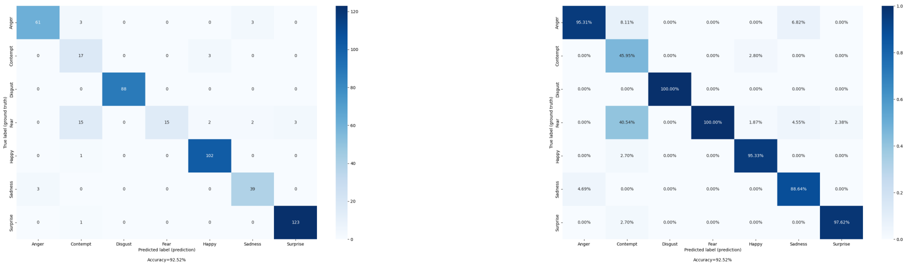
Figure 3.5: DenseNet architecture.

Resnet34. Resnet34 is a 34 layer convolutional neural network that can be utilized as a state-of-the-art image classification model. This is a model that has been pre-trained on the ImageNet dataset–a dataset that has 100,000+ images across 200 different classes. However, it is different from traditional neural networks in the sense that it takes residuals from each layer and uses them in the subsequent connected layers (similar to residual neural networks used for text prediction).

Resnet50. ResNet50 is a variant of ResNet model which has 48 Convolution layers along with 1 MaxPool and 1 Average Pool layer. It has 3.8×10^9 Floating points operations. It is a widely used ResNet model and we have explored ResNet50 architecture in depth.

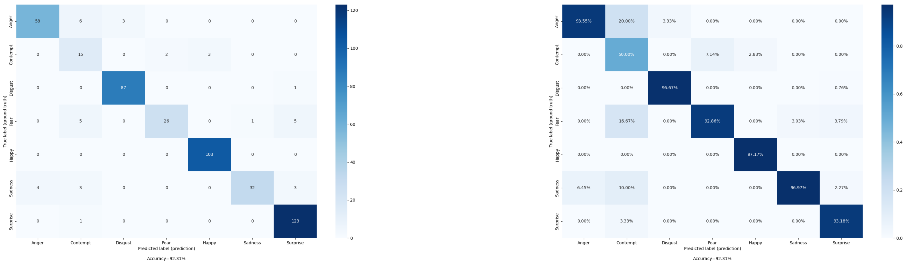
Wideresnet50. Released in 2017 by Sergey Zagoruyko and Nikos Komodakis, this model provides improvement on existing residual networks. By decreasing the depth of the architecture and increasing the width of the network, state-of-the-art accuracies and much faster training were achieved. ImageNet classes are mapped to Wolfram Language Entities through their unique WordNet IDs.

3.10 Resnet-34



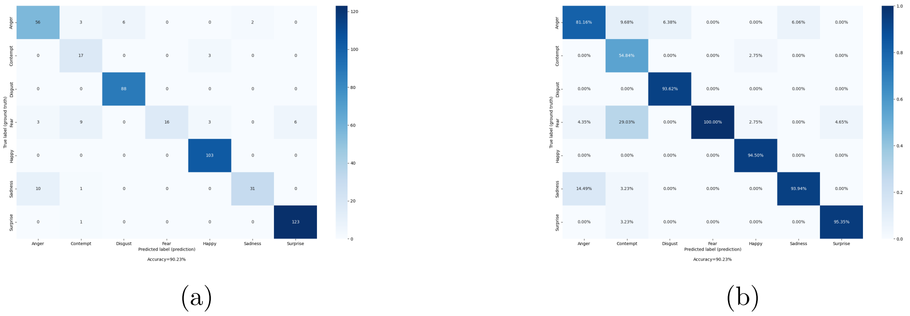
(a) (b)

Figure 3.6: Confusion matrices of mode(1) corresponds to Resnet-34 model.



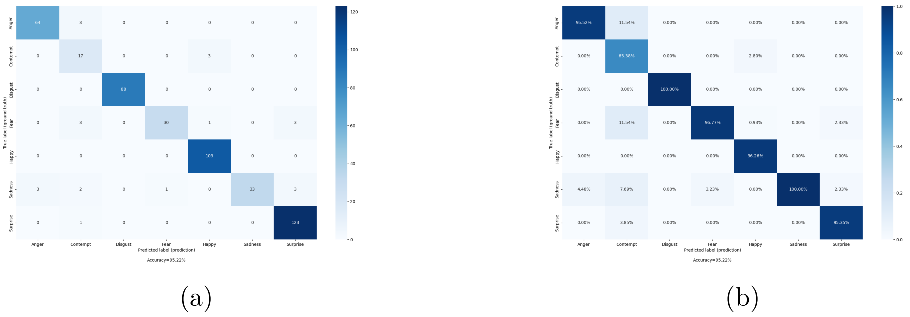
(a) (b)

Figure 3.7: Results of mode(2) for Resnet-34 model.



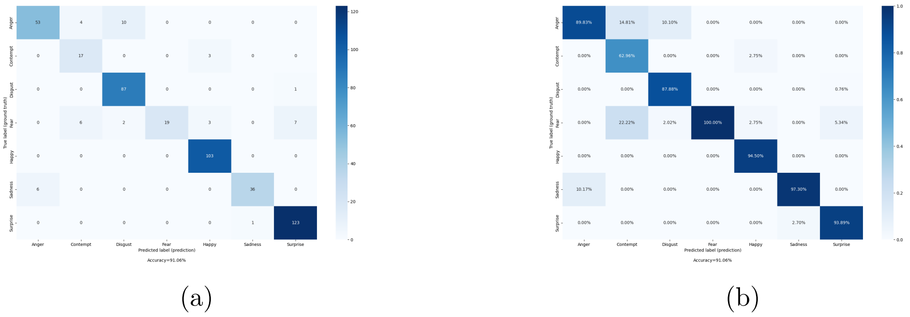
(a) (b)

Figure 3.8: Confusion matrices of mode(3) corresponds in Resnet-34 model.



(a) (b)

Figure 3.9: Confusion matrices of mode(4) corresponds to Resnet-34 model.

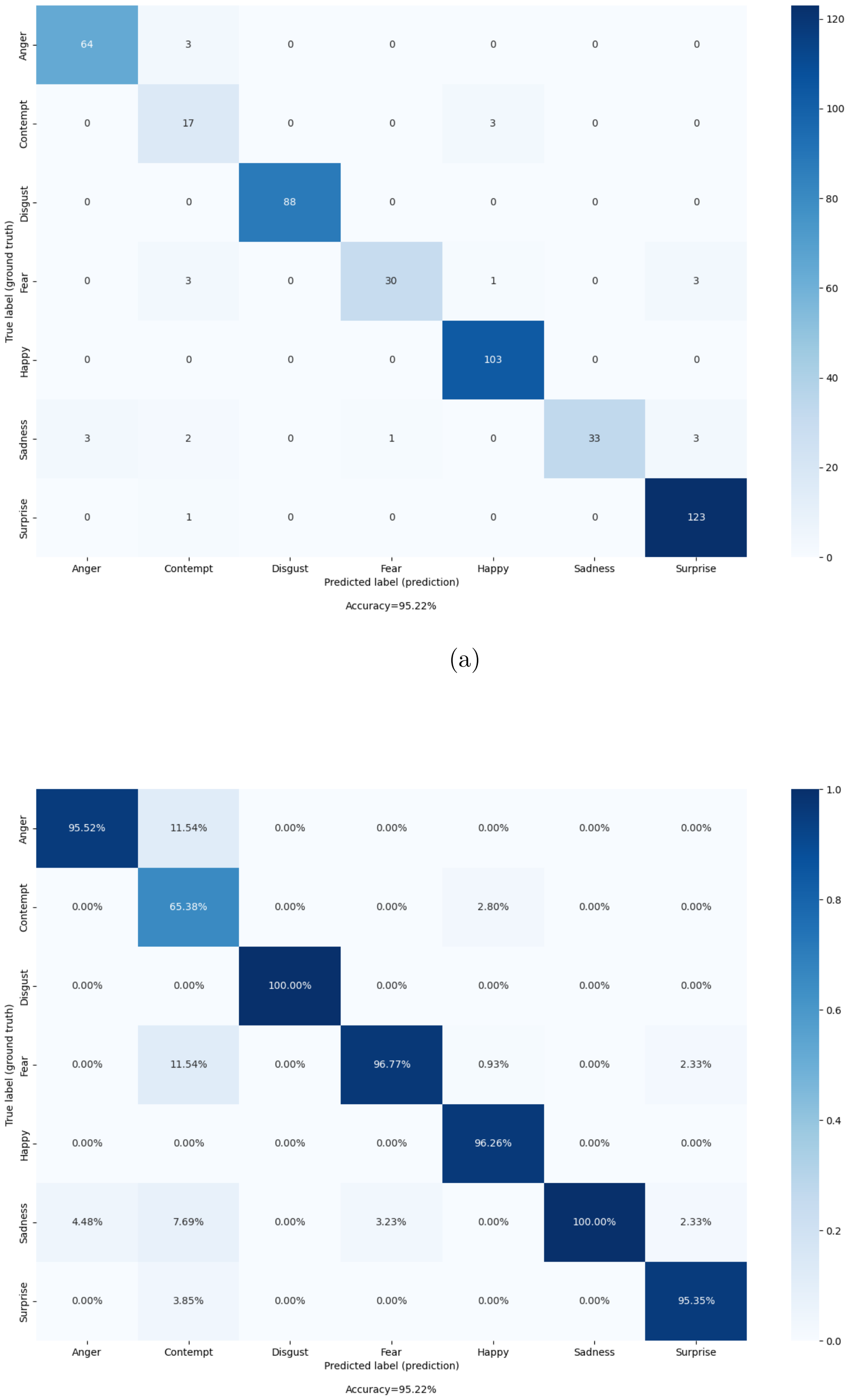


(a) (b)

Figure 3.10: Confusion matrices of mode(5) corresponds to Resnet-34 model.

Figures 3.6,3.7,3.8,3.11,3.10 show the results mode(1), mode(2), mode(3), mode(4),and mode(5), respectively. We can see how these figures that mode (4) show in Figure 3.11 has better result compared to the other modes for Resnet-34.

In this model, a good accuracy rate, regarding feelings of sadness, the percentage of matching 100%, we also notice some confusion in the classification where 5 pictures (3 sadness and 2 fear) are feelings of surprise and one picture belongs to its category on the basis of disgust.

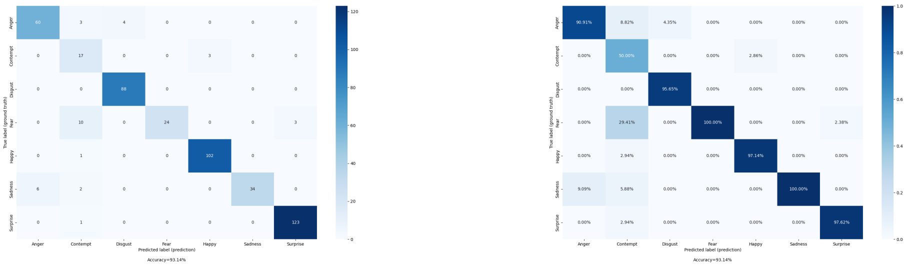


(a)

(b)

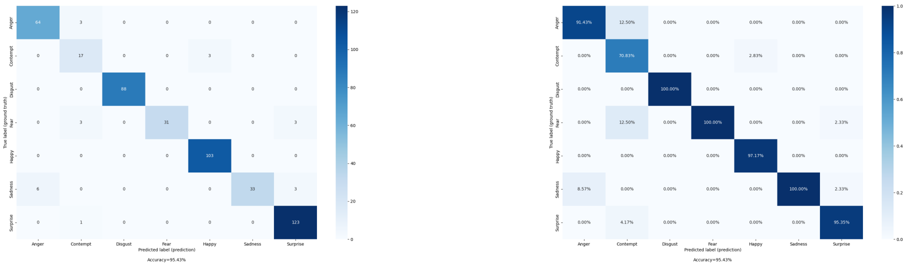
Figure 3.11: Better results for Resnet-34 model.

3.11 Resnet-50



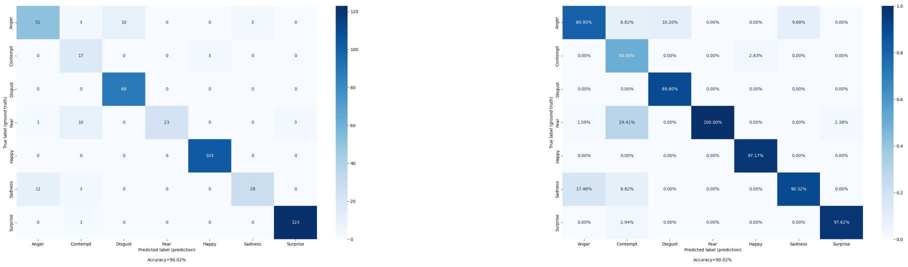
(a) (b)

Figure 3.12: Confusion matrices of mode(1) corresponds to Resnet-50 model.



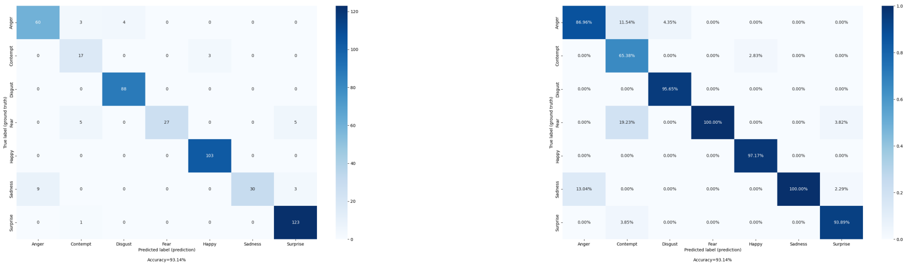
(a) (b)

Figure 3.13: Confusion matrices of mode(2) corresponds to Resnet-50 model.



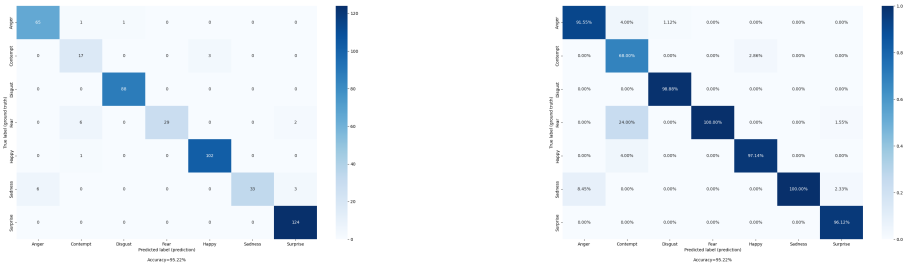
(a) (b)

Figure 3.14: Confusion matrices of mode(3) corresponds to Resnet-50 model.



(a) (b)

Figure 3.15: Confusion matrices of mode(4) corresponds to Resnet-50 model.

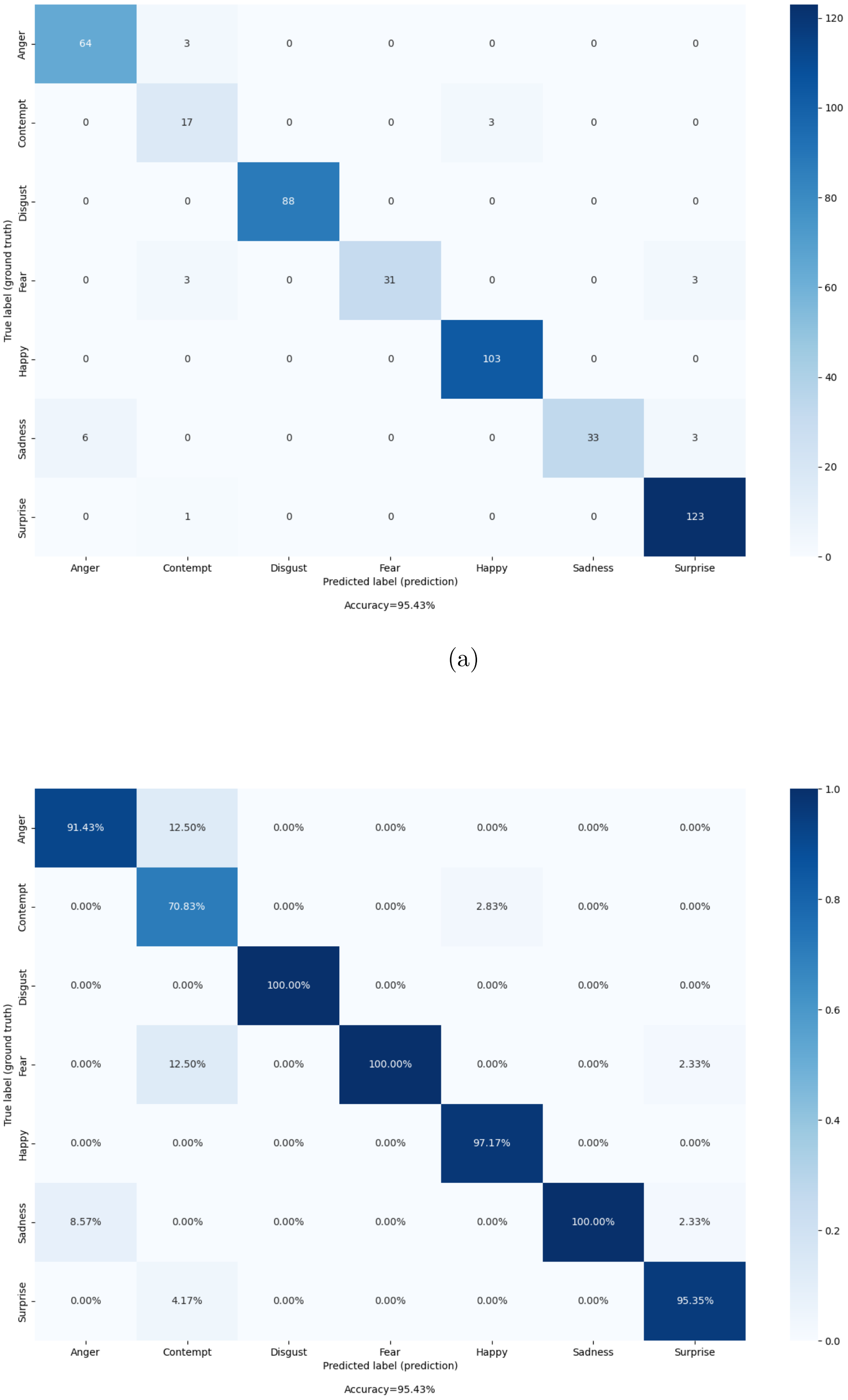


(a) (b)

Figure 3.16: Confusion matrices of mode(5) corresponds to Resnet-50 model.

Figures 3.12,3.17,3.14,3.15,3.16 show the results mode(1),mode(2),mode(3),mode(4),and mode(5),respectivly.we can see how these figures that mode (2) show in figure 3.17 has better result compared to the other modes for Resnet-50.

In this model, you get an accuracy of 94.80% and 100% in the feelings of sadness, fear, and disgust. We also notice 6 images of anger, which he classified as sadness, and 5 images of contempt for their expression as disgust. This confusion occurred due to the convergence of negative emotions (sadness, disgust, fear.....)



	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise
Anger	64	3	0	0	0	0	0
Contempt	0	17	0	0	3	0	0
Disgust	0	0	88	0	0	0	0
Fear	0	3	0	31	0	0	3
Happy	0	0	0	0	103	0	0
Sadness	6	0	0	0	0	33	3
Surprise	0	1	0	0	0	0	123

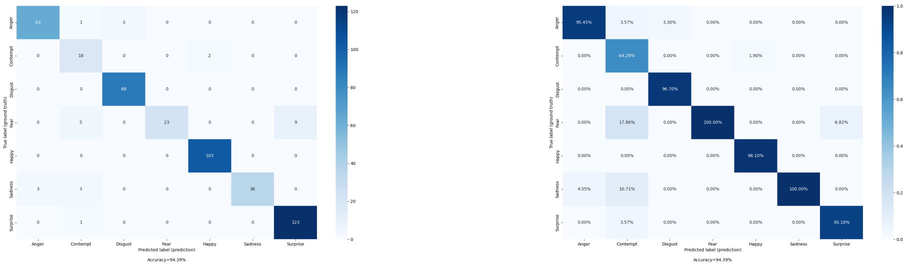
(a)

	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise
Anger	91.43%	12.50%	0.00%	0.00%	0.00%	0.00%	0.00%
Contempt	0.00%	70.83%	0.00%	0.00%	2.83%	0.00%	0.00%
Disgust	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
Fear	0.00%	12.50%	0.00%	100.00%	0.00%	0.00%	2.33%
Happy	0.00%	0.00%	0.00%	0.00%	97.17%	0.00%	0.00%
Sadness	8.57%	0.00%	0.00%	0.00%	0.00%	100.00%	2.33%
Surprise	0.00%	4.17%	0.00%	0.00%	0.00%	0.00%	95.35%

(b)

Figure 3.17: Better results for Resnet-50 model.

3.12 Densnet 121



	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise
Anger	63	1	3	0	0	0	0
Contempt	0	18	0	0	2	0	0
Disgust	0	0	88	0	0	0	0
Fear	0	5	0	23	0	0	9
Happy	0	0	0	0	103	0	0
Sadness	3	3	0	0	0	36	0
Surprise	0	1	0	0	0	0	123

Accuracy=94.39%

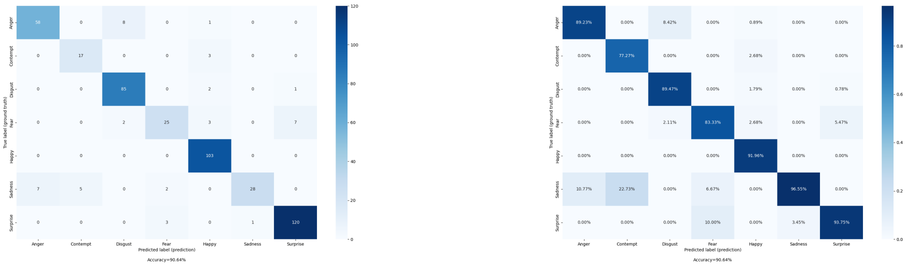
(a)

	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise
Anger	95.45%	3.57%	3.30%	0.00%	0.00%	0.00%	0.00%
Contempt	0.00%	64.29%	0.00%	0.00%	1.90%	0.00%	0.00%
Disgust	0.00%	0.00%	96.70%	0.00%	0.00%	0.00%	0.00%
Fear	0.00%	17.86%	0.00%	100.00%	0.00%	0.00%	6.82%
Happy	0.00%	0.00%	0.00%	0.00%	98.10%	0.00%	0.00%
Sadness	4.55%	10.71%	0.00%	0.00%	0.00%	100.00%	0.00%
Surprise	0.00%	3.57%	0.00%	0.00%	0.00%	0.00%	93.18%

Accuracy=94.39%

(b)

Figure 3.18: Confusion matrices of mode(1) corresponds to Densenet121 model.



	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise
Anger	58	0	8	0	1	0	0
Contempt	0	17	0	0	3	0	0
Disgust	0	0	85	0	2	0	1
Fear	0	0	2	25	3	0	7
Happy	0	0	0	0	103	0	0
Sadness	7	5	0	2	0	28	0
Surprise	0	0	0	3	0	1	120

Accuracy=90.64%

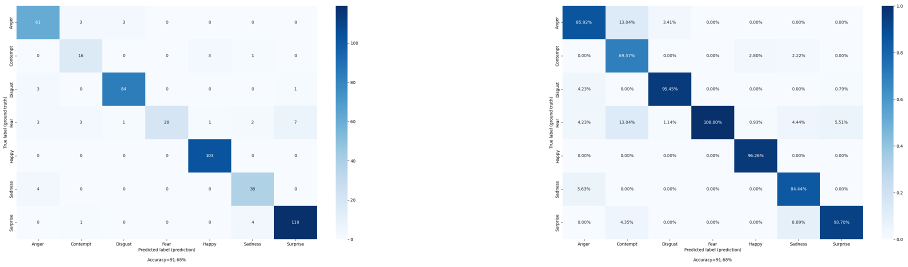
(a)

	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise
Anger	89.23%	0.00%	8.42%	0.00%	0.89%	0.00%	0.00%
Contempt	0.00%	77.27%	0.00%	0.00%	2.68%	0.00%	0.00%
Disgust	0.00%	0.00%	89.47%	0.00%	1.79%	0.00%	0.78%
Fear	0.00%	0.00%	2.11%	83.33%	2.68%	0.00%	5.47%
Happy	0.00%	0.00%	0.00%	0.00%	91.96%	0.00%	0.00%
Sadness	10.77%	22.73%	0.00%	6.67%	0.00%	96.55%	0.00%
Surprise	0.00%	0.00%	0.00%	10.00%	0.00%	3.45%	93.75%

Accuracy=90.64%

(b)

Figure 3.19: Confusion matrices of mode(2) corresponds to Densenet121 model.



	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise
Anger	61	3	3	0	0	0	0
Contempt	0	16	0	0	3	1	0
Disgust	3	0	84	0	0	0	1
Fear	3	3	1	20	1	2	7
Happy	0	0	0	0	103	0	0
Sadness	4	0	0	0	0	38	0
Surprise	0	1	0	0	0	4	119

Accuracy=91.68%

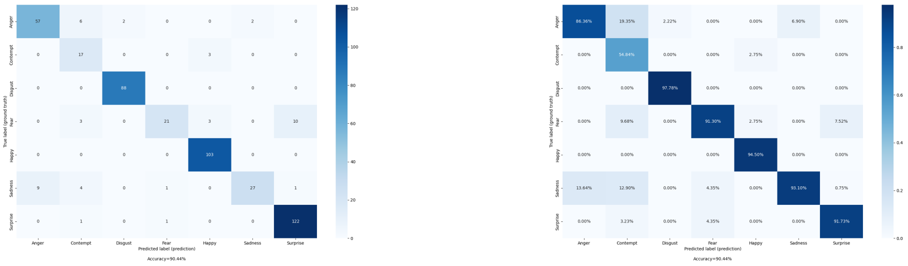
(a)

	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise
Anger	85.92%	13.04%	3.41%	0.00%	0.00%	0.00%	0.00%
Contempt	0.00%	69.57%	0.00%	0.00%	2.80%	2.22%	0.00%
Disgust	4.23%	0.00%	95.45%	0.00%	0.00%	0.00%	0.79%
Fear	4.23%	13.04%	1.14%	100.00%	0.93%	4.44%	5.51%
Happy	0.00%	0.00%	0.00%	0.00%	96.26%	0.00%	0.00%
Sadness	5.63%	0.00%	0.00%	0.00%	0.00%	84.44%	0.00%
Surprise	0.00%	4.35%	0.00%	0.00%	0.00%	8.89%	93.70%

Accuracy=91.68%

(b)

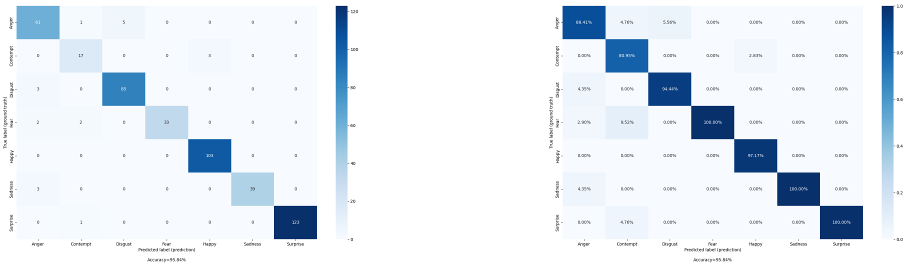
Figure 3.20: Confusion matrices of mode(3) corresponds to Densenet121 model.



(a)

(b)

Figure 3.21: Confusion matrices of mode(4) corresponds to Densenet121 model.



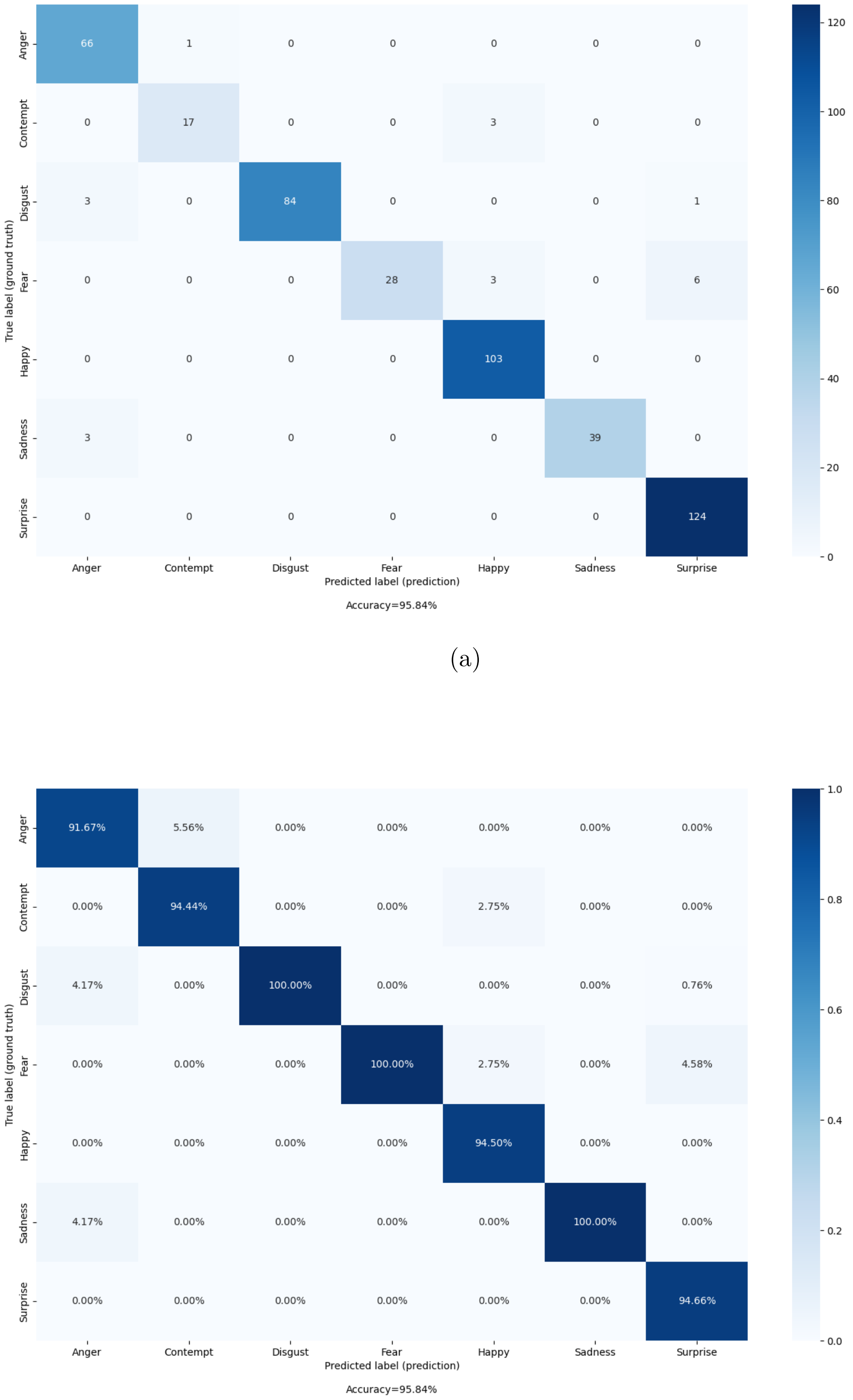
(a)

(b)

Figure 3.22: Confusion matrices of mode(5) corresponds to Densenet121 model.

figures 3.18,3.19,3.20,3.21,3.23 show the results mode(1), mode(2), mode(3), mode(4), and mode(5), respectively. We can see how these figures that mode (3) show in figure 3.23 has better result compared to the other modes for Densnet-121 model.

In this model, we achieved an accuracy of 93.14%, and a good score for feelings of contempt was 85% (the highest percentage compared to other models) and 100% for feelings of fear.

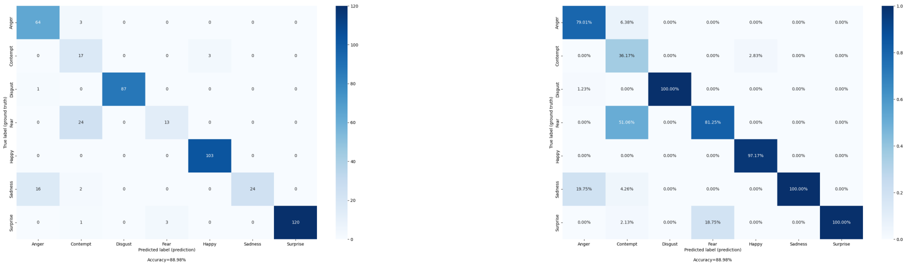


(a)

(b)

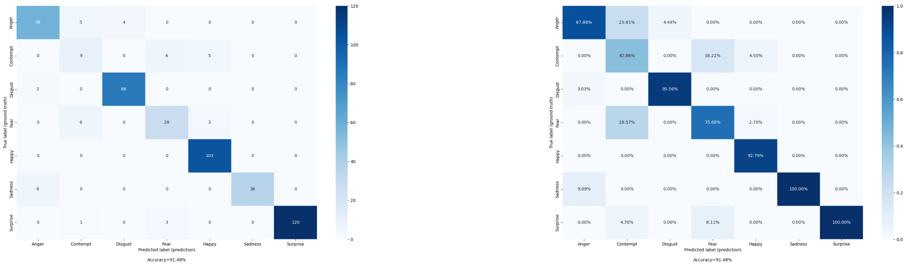
Figure 3.23: Better results for Densnet-121 model.

3.13 Wideresnet50



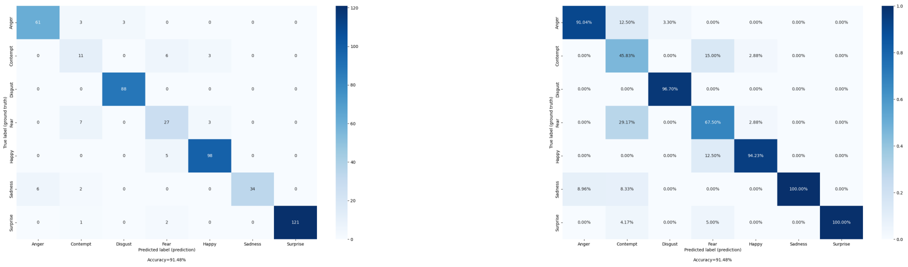
(a) (b)

Figure 3.24: Confusion matrices of mode(1) corresponds to Wideresnet50 model.



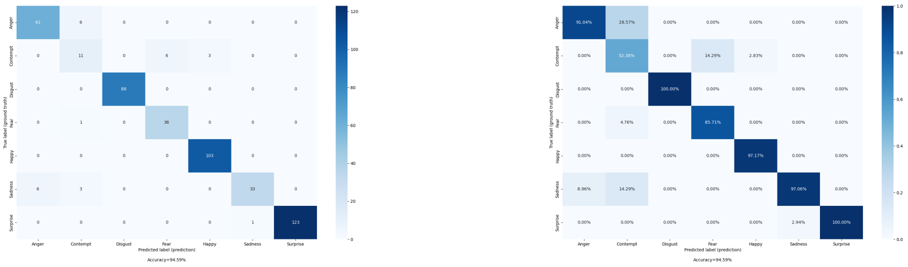
(a) (b)

Figure 3.25: Confusion matrices of mode(2) corresponds) to Wideresnet50 model.



(a) (b)

Figure 3.26: Confusion matrices of mode(3) corresponds to Wideresnet50 model.



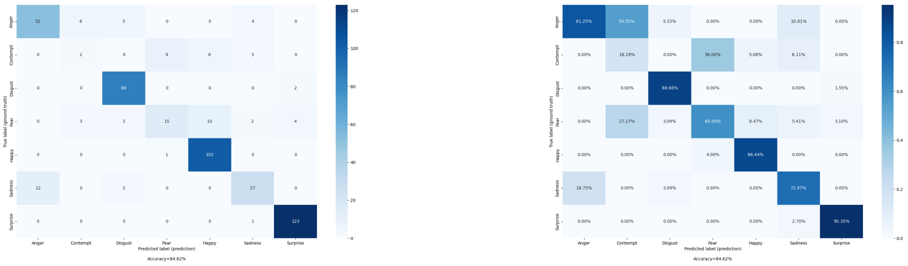
(a)

	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise
Anger	61	6	0	0	0	0	0
Contempt	0	11	0	6	3	0	0
Disgust	0	0	88	0	0	0	0
Fear	0	1	0	36	0	0	0
Happy	0	0	0	0	103	0	0
Sadness	6	3	0	0	0	33	0
Surprise	0	0	0	0	0	1	123

(b)

	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise
Anger	91.04%	28.57%	0.00%	0.00%	0.00%	0.00%	0.00%
Contempt	0.00%	52.38%	0.00%	14.29%	2.83%	0.00%	0.00%
Disgust	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
Fear	0.00%	4.76%	0.00%	85.71%	0.00%	0.00%	0.00%
Happy	0.00%	0.00%	0.00%	0.00%	97.17%	0.00%	0.00%
Sadness	8.96%	14.29%	0.00%	0.00%	0.00%	97.06%	0.00%
Surprise	0.00%	0.00%	0.00%	0.00%	0.00%	2.94%	100.00%

Figure 3.27: Confusion matrices of mode(4) corresponds to Wideresnet50 model.



(a)

	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise
Anger	52	6	5	0	0	4	0
Contempt	0	2	0	9	6	3	0
Disgust	0	0	86	0	0	0	2
Fear	0	3	3	15	10	2	4
Happy	0	0	0	1	102	0	0
Sadness	12	0	3	0	0	27	0
Surprise	0	0	0	0	0	1	123

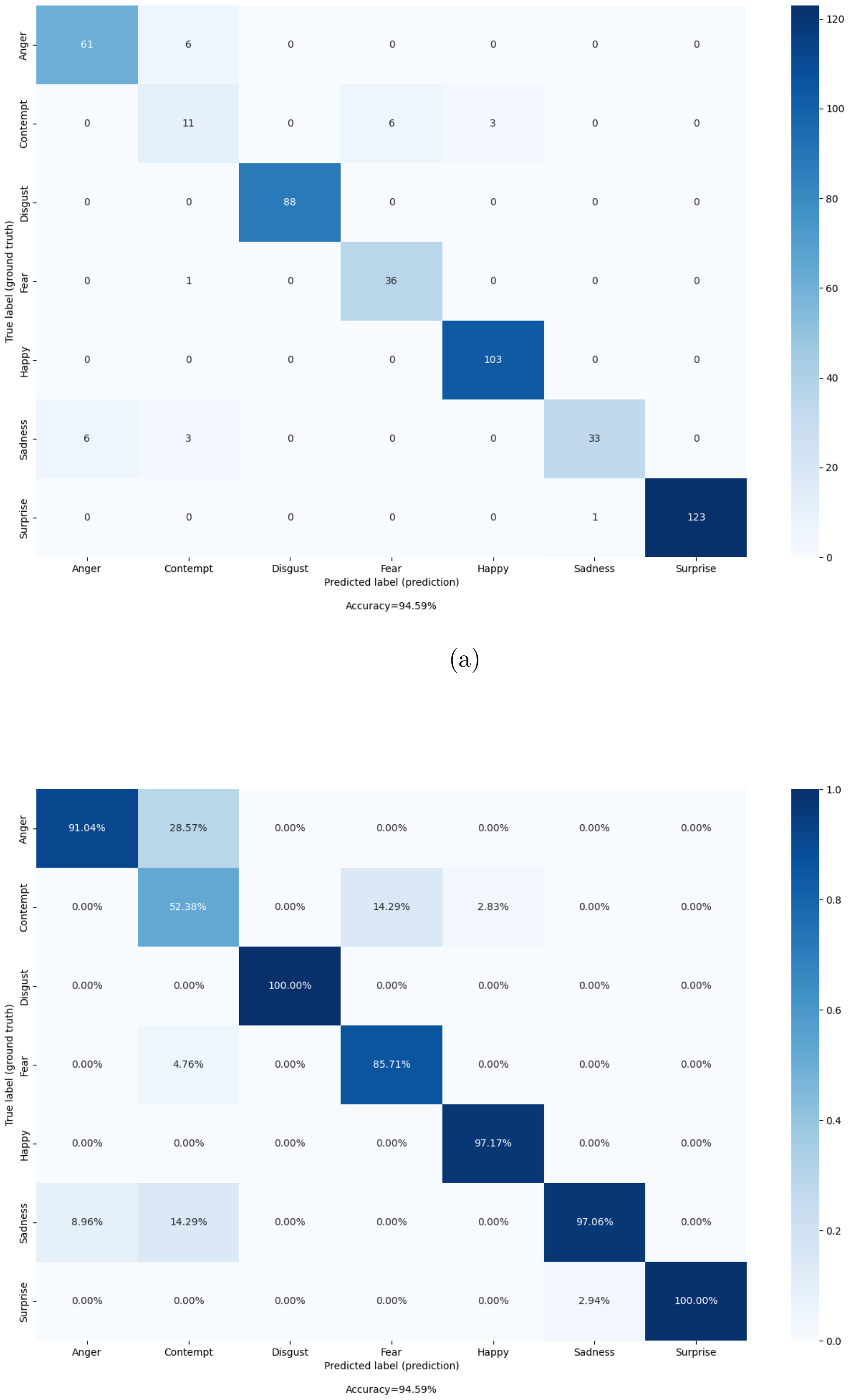
(b)

	Anger	Contempt	Disgust	Fear	Happy	Sadness	Surprise
Anger	81.25%	54.55%	5.15%	0.00%	0.00%	10.81%	0.00%
Contempt	0.00%	18.18%	0.00%	36.00%	5.08%	8.11%	0.00%
Disgust	0.00%	0.00%	88.66%	0.00%	0.00%	0.00%	1.55%
Fear	0.00%	27.27%	3.09%	60.00%	8.47%	5.41%	3.10%
Happy	0.00%	0.00%	0.00%	4.00%	86.44%	0.00%	0.00%
Sadness	18.75%	0.00%	3.09%	0.00%	0.00%	72.97%	0.00%
Surprise	0.00%	0.00%	0.00%	0.00%	0.00%	2.70%	95.35%

Figure 3.28: Confusion matrices of mode(5) corresponds to Wideresnet50 model.

Figures 3.24,3.25,3.26,3.29,3.28 show the results mode(1), mode(2), mode(3), mode(4), and mode(5), respectively. We can see how these figures that mode (5) show in Figure 3.28 has better result compared to the other modes for Wideresnet50 model.

In this model we got an accuracy of 95.63% (the highest of the four models) and 100% in feelings of fear and surprise.



(a)

(b)

Figure 3.29: Better results for Wideresnet50 model.

3.14 Comparisons

Figure shows comparisons between the four models and its corresponding modes. This figure confirms the previous remarks and clearly depict that Wideresnet50 has the best result in CK+ database. As we know, Wideresnet50 is widely used for action recognition systems, and that is why we got best results when we applied it in facial emotion recognition topic.

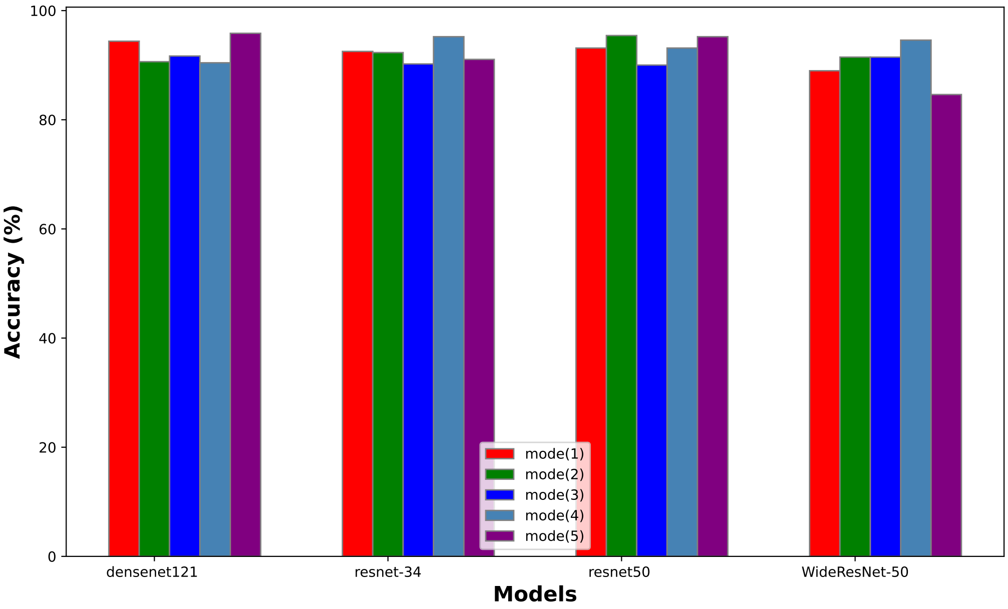


Figure 3.30: Comparisons between the four models and its corresponding modes.

3.15 Real-time application

After completing an empirical study, we tried to transform our knowledge and results After completing an empirical study, we tried to transform our knowledge and results into a realistic practical application that reflects our aspirations and our passion for seeing our study on the ground. In this part, we will try to explain the stages of development of the application and its characteristicsinto a realistic practical application that reflects our aspirations and our passion for seeing our study on the ground. In this part, we will try to explain the stages of development of the application and its characteristics.

3.15.1 Python

Python is a high-level, interpreted, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation.

Python is dynamically-typed and garbage-collected. It supports multiple programming paradigms, including structured (particularly procedural), object-oriented and functional programming. It is often described as a "batteries included" language due to its comprehensive standard library.

Guido van Rossum began working on Python in the late 1980s as a successor to the ABC programming language and first released it in 1991 as Python 0.9.0. Python 2.0 was released in 2000 and introduced new features such as list comprehensions, cycle-detecting garbage collection, reference counting, and Unicode support. Python 3.0, released in 2008, was a major revision that is not completely backward-compatible with earlier versions. Python 2 was discontinued with version 2.7.18 in 2020.

Python consistently ranks as one of the most popular programming languages.

3.15.2 Pytorch

PyTorch is an open source machine learning (ML) framework based on the Python programming language and the Torch library. It is one of the preferred platforms for deep learning research. The framework is built to speed up the process between research prototyping and deployment.

PyTorch is similar to NumPy and computes using tensors that are accelerated by graphics processing units (GPU). Tensors are arrays, a type of multidimensional data structure, that can be operated on and manipulated with APIs. The PyTorch framework supports over 200 different mathematical operations.

The popularity of PyTorch continues to rise as it simplifies the creation of artificial neural network (ANN) models. PyTorch is mainly used for applications of research, data science and artificial intelligence (AI).

Key Features of PyTorch

- TorchScript- This is the production environment of PyTorch that enables users to seamlessly transition between modes. TorchScript optimizes functionality, speed, ease-of-use and flexibility. `\item` .
- Dynamic graph computation- This feature allows users to change network behavior on the fly, rather than waiting for the entire code to be executed.
- Automatic differentiation- This technique numerically computes the derivative of a function by making backward passes in neural networks.
- Python support- Because PyTorch is based on Python, it can be used with popular libraries and packages such as NumPy, SciPy, Numba and Cynthon.

PyTorch Benefits

Using PyTorch can provide developers with the following benefits:

- Is based on Python, making it easy to learn and simple to code.
- Allows for easy debugging with popular Python tools.
- Is well supported on major cloud platforms, making it easy to scale.
- Has a small community of focused on open source
- Can export learning models to the Open Neural Network Exchange (ONNX) standard format.
- Has a user-friendly interface.
- Offers a C++ front end interface option.

3.16 Visual Studio Code

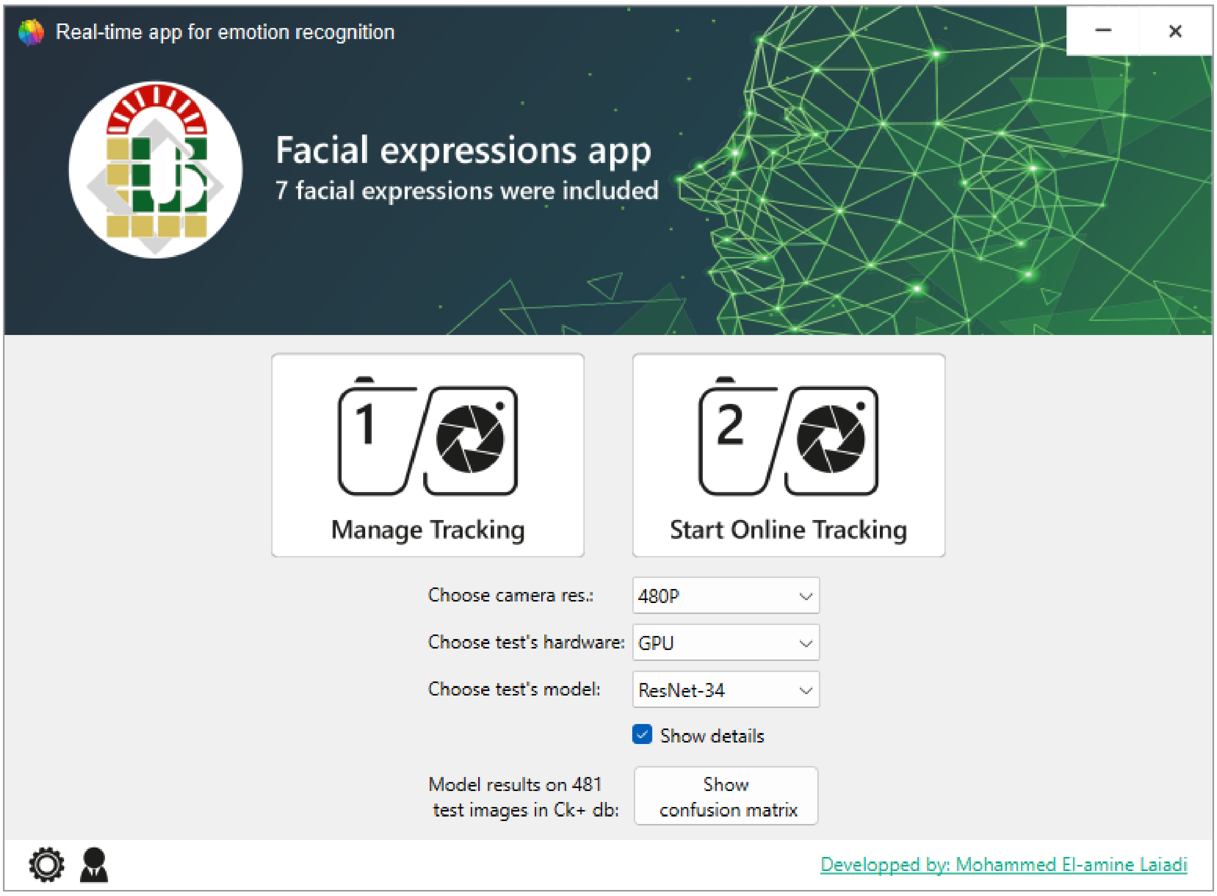
Visual Studio Code, also commonly referred to as VS Code, is a source-code editor made by Microsoft for Windows, Linux and macOS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git. Users can change the theme, keyboard shortcuts, preferences, and install extensions that add additional functionality.

In the Stack Overflow 2021 Developer Survey, Visual Studio Code was ranked the most popular developer environment tool, with 70% of 82,000 respondents reporting that they use it.

3.17 Application

In this section, we will see our real-time application for facial expression recognition. Our training is made by 500 facial images of CK+ database using data augmentation. The four used models are pretrained on ImageNet database (has 1000 classes) and we fine-tune these pretained models for facial emotion recognition using seven emotions (anger, contempt, disgust, fear, happiness, sadness, and surprise). For the evaluation stage, we evaluated the provided 481 facial images on the four trained models. Finally, in the real test stage we used our real-time application for real-life scenarios. In the following we will show step by step our main application.

Figures 3.31 and 3.32, show the main form/widget of our real time application. Our application contains two buttons (for Manage tracking and for Start online tracking), three combo-boxes, one for choosing the resolution of the camera (480P, 720P and 1080P), one for choosing the test's hardware (GPU or CPU) and the last one for choosing the test's model (we used for models as mentioned earlier in this chapter).



Real-time app for emotion recognition

Facial expressions app

7 facial expressions were included

Choose camera res.: 480P

Choose test's hardware: GPU

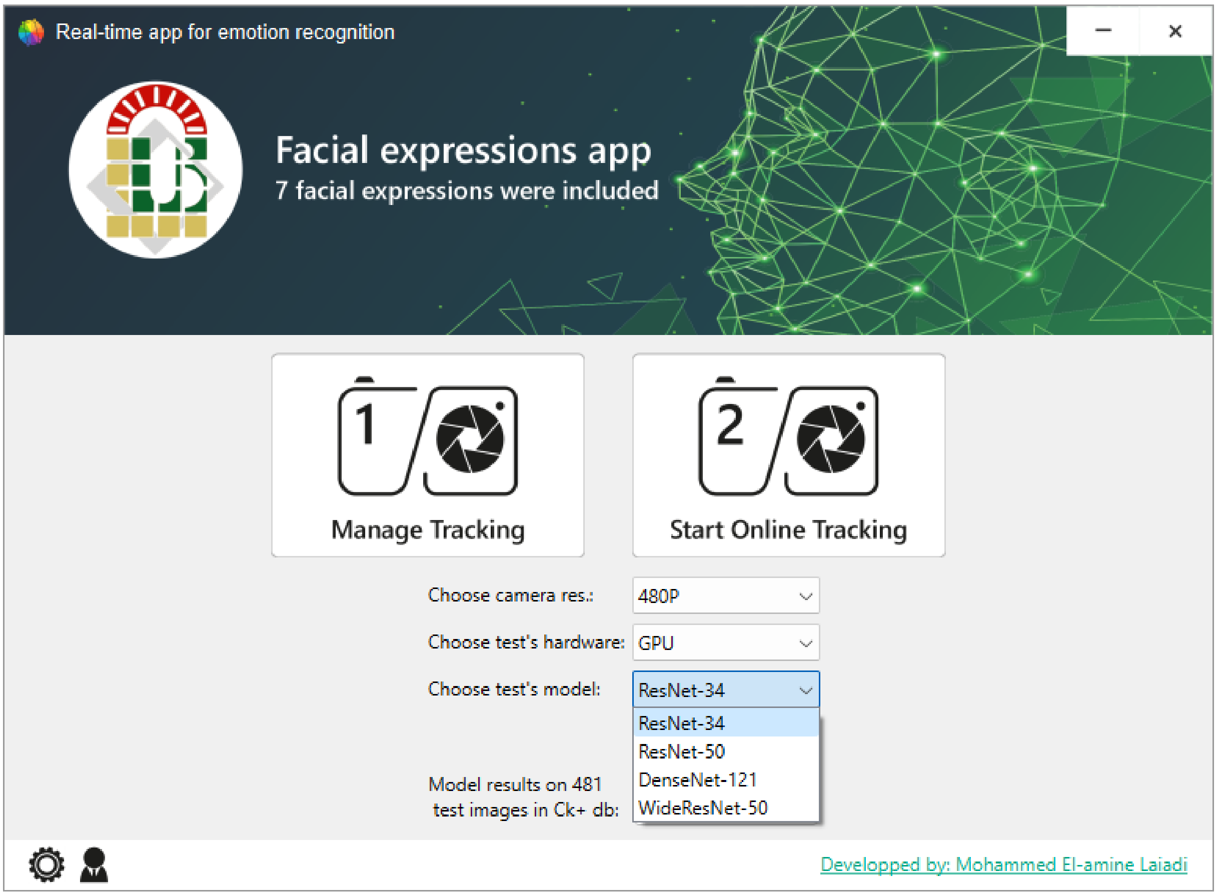
Choose test's model: ResNet-34

☑ Show details

Model results on 481 test images in Ck+ db:

Developped by: Mohammed El-amine Laiadi

Figure 3.31: The main interface of the real-time application.


Real-time app for emotion recognition
Facial expressions app
7 facial expressions were included
Manage Tracking
Start Online Tracking
ResNet-34
ResNet-50
DenseNet-121
WideResNet-50
Developped by: Mohammed El-amine Laiadi

Choose camera res.: 480P

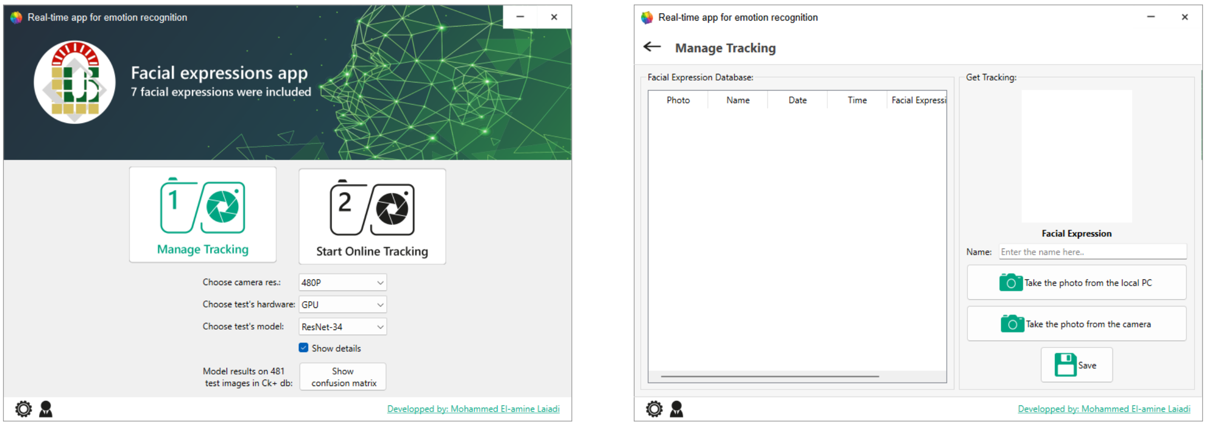
Choose test's hardware: GPU

Choose test's model: ResNet-34

Model results on 481 test images in Ck+ db:

Figure 3.32: The four used models in our application.

Figure 3.33 shows action of clicking of "Manage Tracking" button and its response and resulting form. The "Manage Tracking" button can save the facial image (photo), the name, the date of when capturing, time when capturing and the corresponding facial expression. These data is inserted in predefined SQLITE database. Furthermore, the "Manage Tracking" option can let us save the previous data from the real-time camera or simply by uploading a facial image.


Real-time app for emotion recognition
Facial expressions app
7 facial expressions were included
Manage Tracking
Start Online Tracking
Show confusion matrix
Developped by: Mohammed El-amine Laiadi
Real-time app for emotion recognition
Manage Tracking
Take the photo from the local PC
Take the photo from the camera
Save
Developped by: Mohammed El-amine Laiadi

Choose camera res.: 480P

Choose test's hardware: GPU

Choose test's model: ResNet-34

☑ Show details

Model results on 481 test images in Ck+ db:

(a)

Facial Expression Database:

Photo	Name	Date	Time	Facial Expressi

Get Tracking:

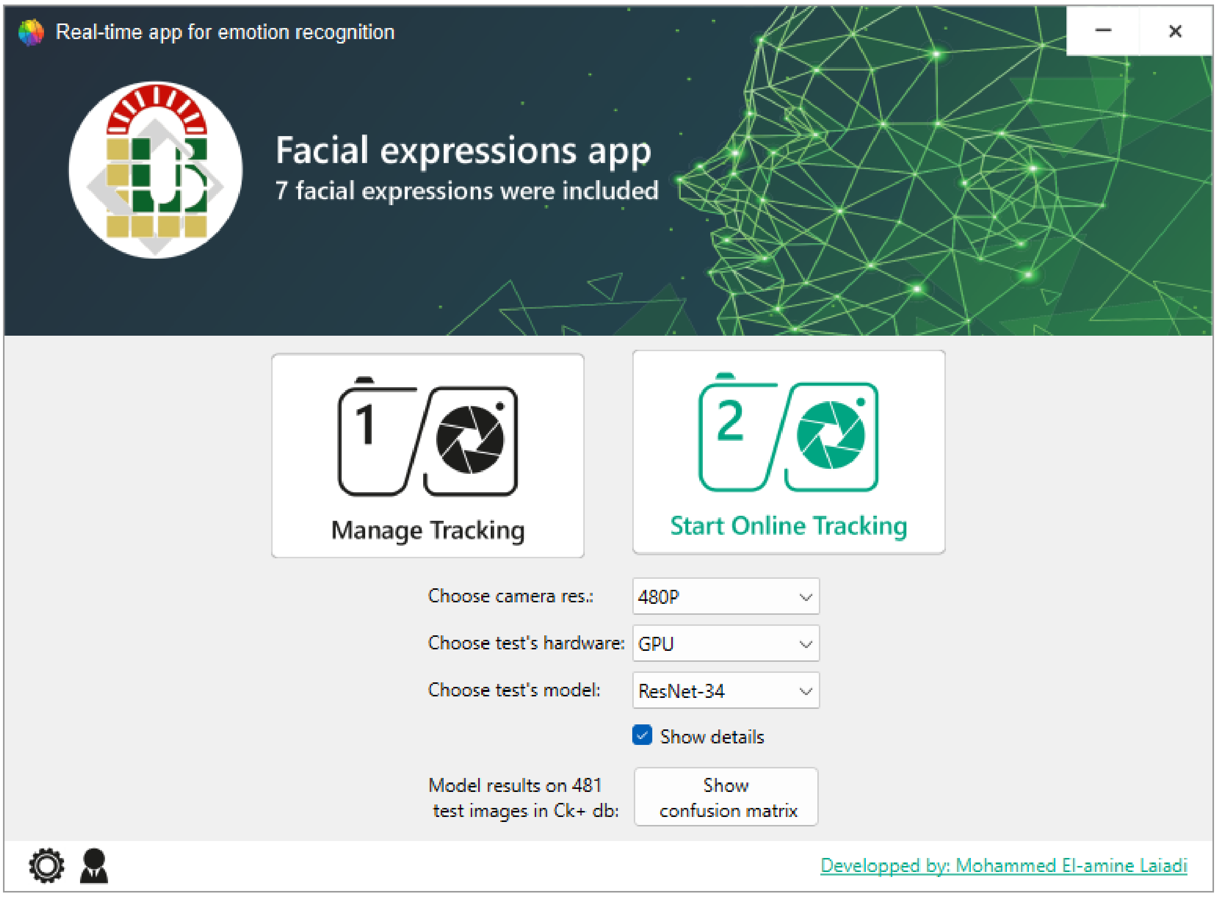
Facial Expression

Name: Enter the name here..

(b)

Figure 3.33: Action of manage tracking button (a) Click on Manage Tracking button, (b) Resulting form of the clicked button.

Figure 3.34 shows action of clicking of "Start Online Tracking" button and its response and resulting form. The "Start Online Tracking" button activate the webcam (camera) installed on the running PC and start track faces introduced in the capture vision (captured in the scene/vision of the camera). Our real-time application can capture mono-introduced face and multi-introduced faces (monocular and multiple inputs of individuals).


Real-time app for emotion recognition
Facial expressions app
7 facial expressions were included
Manage Tracking
Start Online Tracking

Choose camera res.: 480P

Choose test's hardware: GPU

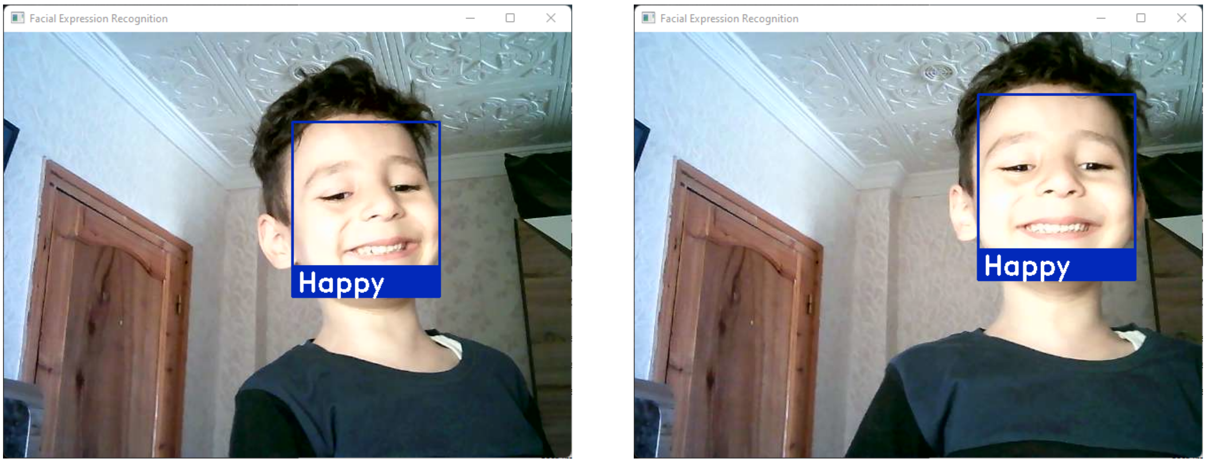
Choose test's model: ResNet-34

☑ Show details

Model results on 481 test images in Ck+ db: Show confusion matrix

Developped by: Mohammed El-amine Laiadi

Figure 3.34: Action of start online tracking.



(a)

(b)

Figure 3.35: Detection feeling of happiness.

Figure 3.35 shows the operation of detecting the face and recognizing the feeling by our real-time application using a simple webcam installed on the PC. Furthermore, our application which is based on the four CNN models shows its robustness against blur-motion, lighting, and face position, which is a very good indication that our application is robust for real scenarios.

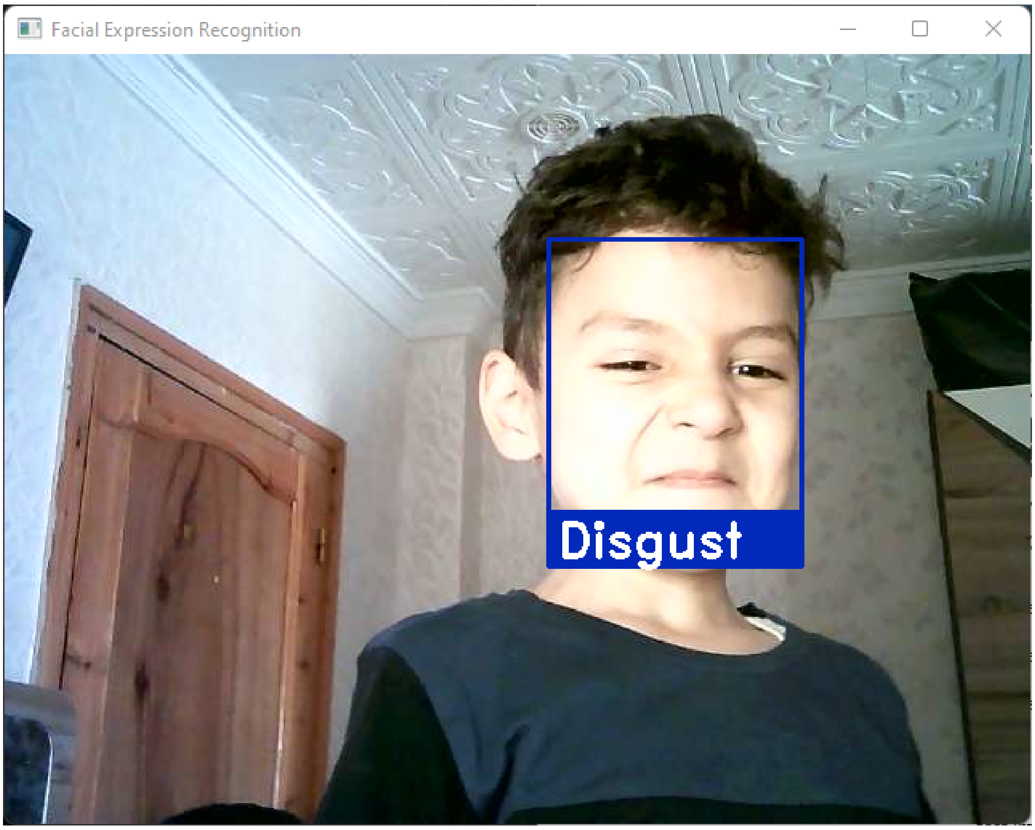


Figure 3.36: Detection feeling of disgust.

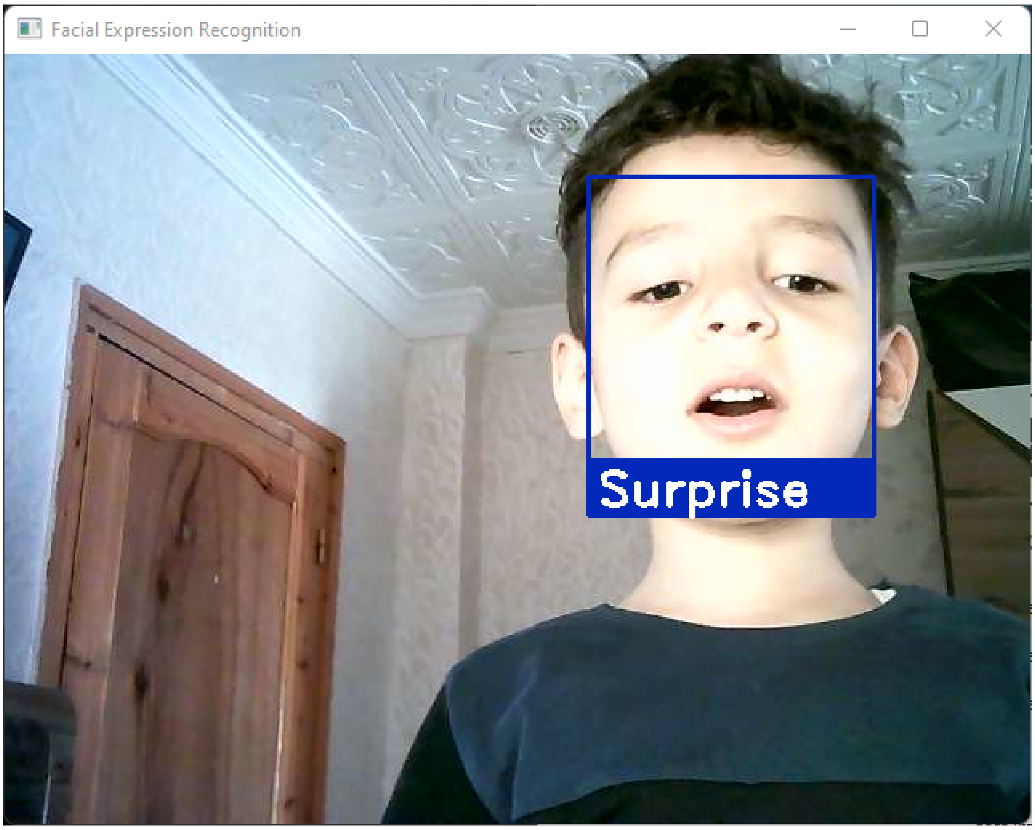


Figure 3.37: Detection feeling of surprise.

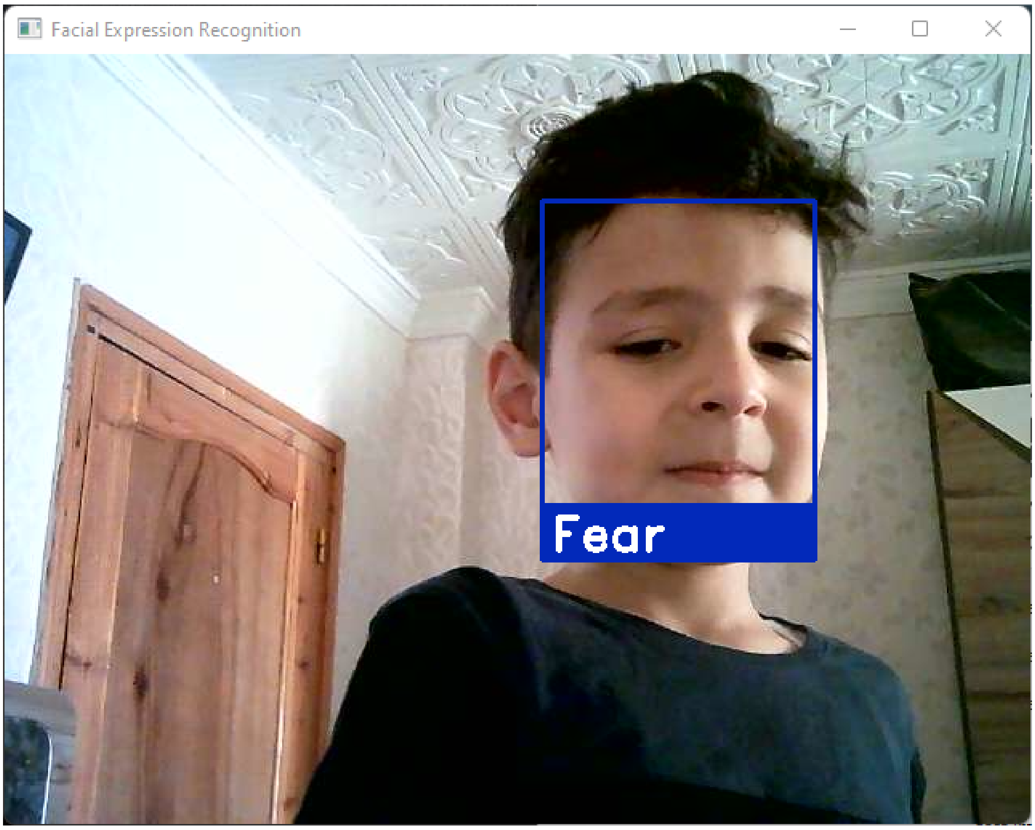


Figure 3.38: Detection feeling of fear.

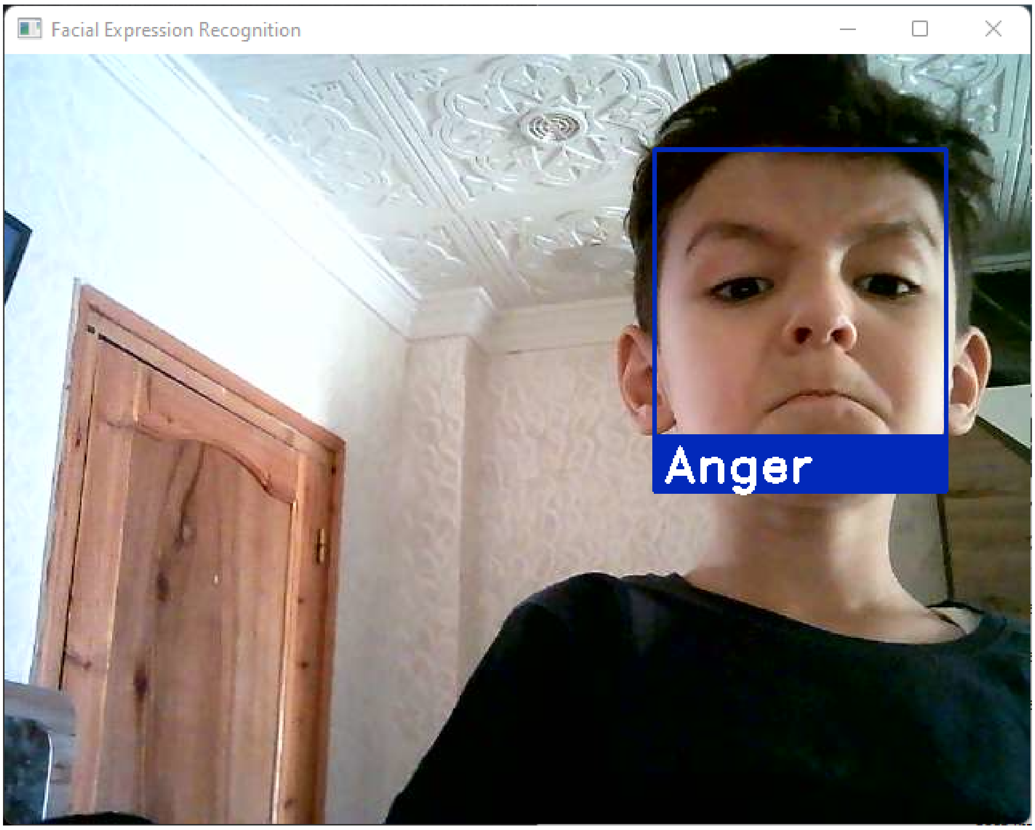


Figure 3.39: Detection feeling of anger.

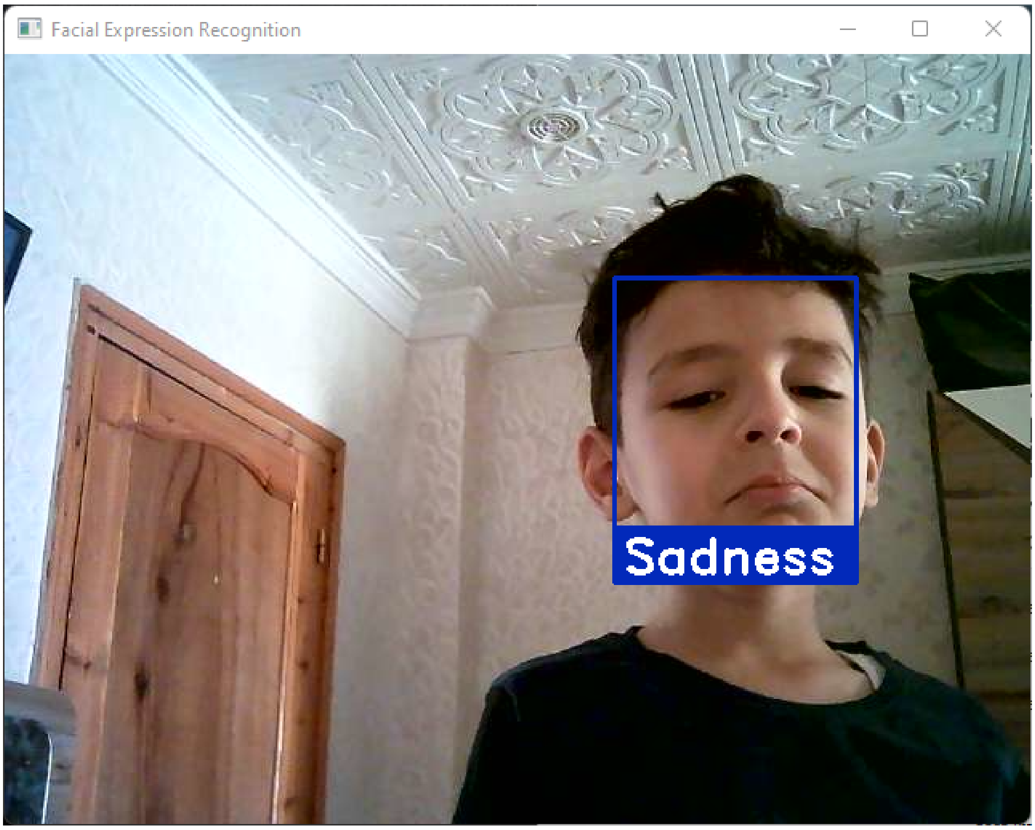


Figure 3.40: Detection feeling of sadness.

Figures 3.36, 3.37, 3.38, 3.39, and 3.40 show different emotions captured by our real time application which are: Disgust, Surprise, Fear, Anger, and Sadness, respectively. Furthermore, our application show a good performances against age and we can see that the models are robust against age variation. Our application capture the emotion from faces even if these faces are belong to children or adults.

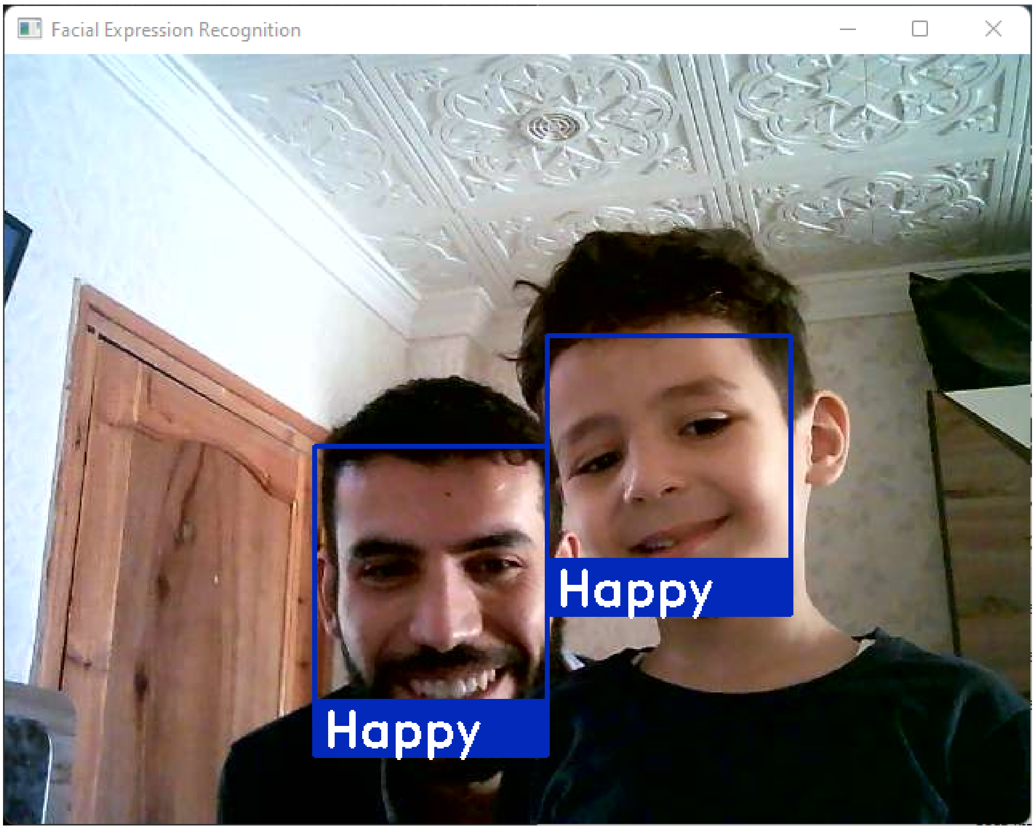


Figure 3.41: Detection for the feeling of happiness belonging to two individuals at the same scene (same time).

Figure 3.41 shows the detection for the feeling of happiness belonging to two individuals at the same scene (same time). These two different individuals are with different ages and with/without beard and moustache. This figure confirms our claims about robustness, performance and speed of detection. Furthermore, confirm that our application work well when more than one individual (one face) is introduced to the camera vision. Also, our application has an awareness of humans faces and other object in the background (background of the camera vision).

3.18 General conclusion

The ability to understand facial expressions is an important part of nonverbal communication. If you only listen to what a person says and ignore what their face is telling you, then you really won't get the whole story. Often, words do not match emotions, and the face betrays what a person is actually feeling.

We convey a lot of nonverbal information in our faces, and we tend to focus on different areas of the face when we try to interpret what each expression might mean. We look at the eyes to determine if someone is sad or angry, for example, and at the mouth to check if someone is happy.

Eyebrows. Eyebrows can show distinctive emotional signals (and they're potentially as important as the eyes for facial recognition). Eyebrows can be:

- Raised and arched (showing surprise).
- Lowered and knit together (often meaning anger, sadness, or fear).
- Drawn up in the inner corners (which could convey sadness).

Eyes. The eyes are often described as "windows to the soul," and we often look to them to determine what someone else may be feeling. The eyes might be:

- Blinking quickly (meaning distress or discomfort) or blinking too little (which may mean that a person is trying to control their eyes).
- Dilated (showing interest or even arousal).
- Staring intensely (which could show attention or anger) or looking away (showing discomfort or distraction).

Mouth. The mouth can convey more than just a smile. People often use their mouths to mask other emotions their face is conveying—for example, a forced smile might cover up an eye micro-expression showing someone's true feelings.

Look out for:

- A dropped jaw (which signals surprise) Open mouth (showing fear).
- One side of the mouth raised (which could indicate hate or contempt).
- Raised corners (meaning happiness).
- Corners that are drawn down (conveying sadness).

Other signals to look for are:

- Lip biting (which may be a sign of anxiety).
- Pursed lips (showing distaste).
- Covering the mouth (which could mean they are hiding something).

Emotional facial expressions can inform researchers about an individual's emotional state. Recent technological advances open up new avenues for automatic Facial Expression Recognition (FER). Based on machine learning, such technology can tremendously increase the amount of processed data. FER is now easily accessible and has been validated for the classification of standardized prototypical facial expressions.

Facial Emotion Recognition is a technology used for analysing sentiments by different sources, such as pictures and videos. It belongs to the family of technologies often referred to as 'affective computing', a multidisciplinary field of research on computer's capabilities to recognise and interpret human emotions and affective states and it often builds on Artificial Intelligence technologies.

Convolutional Neural Net is a popular deep learning technique for current visual recognition tasks. Like all deep learning techniques, CNN is very dependent on the size and quality of the training data. Given a well prepared dataset, CNNs are capable of surpassing humans at visual recognition tasks. However, they are still not robust to visual artifacts such as glare and noise, which humans are able to cope. The theory of CNN is still being developed and researchers are working to endow it with properties such as active attention and online memory, allowing CNNs to evaluate new items that are vastly different from what they were trained on. This better emulates the mammalian visual system, thus moving towards a smarter artificial visual recognition system.

Bibliography

[1] Abate, A.F., Nappi, M., Riccio, D., Sabatino, G.: 2d and 3d face recognition: A survey. Pattern recognition letters **28**(14), 1885–1906 (2007)

[2] Ahmad, T., Munir, A., Bhatti, S.H., Aftab, M., Raza, M.A.: Survival analysis of heart failure patients: A case study. PloS one **12**(7), e0181,001 (2017)

[3] Alasadi, S.A., Bhaya, W.S.: Review of data preprocessing techniques in data mining. Journal of Engineering and Applied Sciences **12**(16), 4102–4107 (2017)

[4] Bennett, J.A., Riegel, B., Bittner, V., Nichols, J.: Validity and reliability of the nyha classes for measuring research outcomes in patients with cardiac disease. Heart & Lung **31**(4), 262–270 (2002)

[5] Borlaug, B.A.: Evaluation and management of heart failure with preserved ejection fraction. Nature Reviews Cardiology **17**(9), 559–573 (2020)

[6] Bredy, C., Ministeri, M., Kempny, A., Alonso-Gonzalez, R., Swan, L., Uebing, A., Diller, G.P., Gatzoulis, M.A., Dimopoulos, K.: New york heart association (nyha) classification in adults with congenital heart disease: relation to objective measures of exercise and outcome. European Heart Journal-Quality of Care and Clinical Outcomes **4**(1), 51–58 (2018)

[7] Chicco, D., Jurman, G.: Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. BMC medical informatics and decision making **20**(1), 1–16 (2020)

[8] Fayed, H.A., Atiya, A.F.: Speed up grid-search for parameter selection of support vector machines. Applied Soft Computing **80**, 202–210 (2019)

[9] Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition ieee transactions on circuits and systems for video technology. Special Issue on Image-and Video-Based Biometrics **14**(1) (2004)

[10] Javeed, A., Rizvi, S.S., Zhou, S., Riaz, R., Khan, S.U., Kwon, S.J.: Heart risk failure prediction using a novel feature selection method for feature refinement and neural network for classification. Mobile Information Systems **2020** (2020)

[11] Mhatre, A., Palla, S., Chikkerur, S., Govindaraju, V.: Efficient search and retrieval in biometric databases", spie defense and security. In: Symposium, March-2005. Citeseer (2001)

[12] Nithya, R., Santhi, B.: Comparative study on feature extraction method for breast cancer classification. Journal of Theoretical and Applied Information Technology **33**(2), 220–226 (2011)

[13] Nithya, R., Santhi, B.: Mammogram classification using maximum difference feature selection method. Journal of Theoretical and Applied Information Technology **33**(2), 197–204 (2011)

[14] Nithya, R., Santhi, B.: Breast cancer diagnosis in digital mammogram using statistical features and neural network. Research Journal of Applied Sciences, Engineering and Technology **4**(24), 5480–5483 (2012)

[15] Nithya, R., Santhi, B.: Application of texture analysis method for mammogram density classification. Journal of Instrumentation **12**(07), P07,009 (2017)

[16] Penny, W.: Biometrics: A double edged sword-security and privacy. SANS Institute (2002)

[17] Pfeffer, M.A., Shah, A.M., Borlaug, B.A.: Heart failure with preserved ejection fraction in perspective. Circulation research **124**(11), 1598–1617 (2019)

[18] Ratha, N.K., Connell, J.H., Bolle, R.M.: Enhancing security and privacy in biometrics-based authentication systems. IBM systems Journal **40**(3), 614–634 (2001)

[19] Rishiikeshwer, B., Shriram, T.A., Raju, J.S., Hari, M., Santhi, B., Brindha, G.: Farmer-friendly mobile application for automated leaf disease detection of real-time augmented data set using convolution neural networks. Journal of Computer Science (2019)

[20] Robert, B.M., Brindha, G., Santhi, B., Kanimozhi, G., Prasad, N.R.: Computational models for predicting anticancer drug efficacy: A multi linear regression analysis based on molecular, cellular and clinical data of oral squamous cell carcinoma cohort. Computer methods and programs in biomedicine **178**, 105–112 (2019)

[21] Saeys, Y., Abeel, T., Peer, Y.V.d.: Robust feature selection using ensemble feature selection techniques. In: Joint European conference on machine learning and knowledge discovery in databases, pp. 313–325. Springer (2008)

[22] Samir, N., Michael, T., Raj, N.: Biometrics: Identity verification in a networked world (2002)

[23] Shakthi, K.A., Brindha, G., Bharathi, N.: Enhanced classification through improved feature selection

[24] Spinella, E.: Biometric scanning technologies: Finger, facial and retinal scanning. SANS Institute, San Francisco, CA **28** (2003)

[25] Taha, K., Ross, H.J., Peikari, M., Mueller, B., Fan, C.P.S., Crowdy, E., Manlhiot, C.: An ensemble-based approach to the development of clinical prediction models for future-onset heart failure and coronary artery disease using machine learning. Journal of the American College of Cardiology **75**(11_Supplement_1), 2046–2046 (2020)

[26] Timmis, A., Townsend, N., Gale, C.P., Torbica, A., Lettino, M., Petersen, S.E., Mossialos, E.A., Maggioni, A.P., Kazakiewicz, D., May, H.T., De Smedt, D., Flather, M., Zuhlke, L., Beltrame, J.F., Huculeci, R., Tavazzi, L., Hindricks, G., Bax, J., Casadei, B., Achenbach, S., Wright, L., Vardas, P., of Cardiology, E.S.: European Society of Cardiology: Cardiovascular Disease Statistics 2019. European Heart Journal **41**(1), 12–85 (2019). DOI 10.1093/eurheartj/ehz859. URL `https://doi.org/10.1093/eurheartj/ehz859`

[27] Wang, Z., Zhu, Y., Li, D., Yin, Y., Zhang, J.: Feature rearrangement based deep learning system for predicting heart failure mortality. Computer Methods and Programs in Biomedicine **191**, 105,383 (2020). DOI 10.1016/j.cmpb.2020.105383

[28] Yan, Y., Liu, R., Ding, Z., Du, X., Chen, J., Zhang, Y.: A parameter-free cleaning method for smote in imbalanced classification. IEEE Access **PP**, 1–1 (2019). DOI 10.1109/ACCESS.2019.2899467

[29] Zimmerman, M.: Biometrics and user authentication. sans institute. m. zimmerman. Biometrics and User Authentication, SANS Institute (2002)

[30] Zunkel, R.L.: Hand geometry based verification. In: Biometrics, pp. 87–101. Springer (1996)