University of Essex

Department of Computer Science and Electronic Engineering

CE901:DISSERTATION

AI and Machine Learning tool and web application for emotion recognition used in the rehabilitation of neurological patients with emotion recognition deficits

MADIHA FATHIMA

PG2112262

Supervisor: **Dr. Vito De Feo**

December 13, 2022

Colchester

ABSTRACT

The use of visual information to recognize emotion in humans has been shown to be effective. This has proven to be quite challenging and is an active area of research. The study is based on an attempted production of a desktop application for carrying out neuroscience experiments of various natures that have been recognized as a common denominator of the emotions of the participant in order to allow a rehabilitation training adequate.

Using techniques such as EEG, Eye-tracker and GSR made it possible to research. First, data is collected in experiments, they are then processed by a machine learning algorithm whose study and development are followed, which attempts to predict correctly the emotion experienced by the participant. Particular emphasis in this study is given to the Facial coding that, through the registration of the face, provides a prediction about the observed emotion.

ACKNOWLEDGEMENT

I take this opportunity to sincerely thank my Supervisor, Dr. Vito De Feo, for constantly assisting and supporting me throughout my research, which lasted this summer.
I also want to express my gratitude to everyone on the team who helped make this project a success.

MadihaFathima

ABBREVIATIONS

- AI - Artificial Intelligence
- GSR - Galvanic Skin Response
- FER - Facial Emotion Recognition
- HCI - Human-Computer Interaction
- CNN - Convolutional Neural Network
- SVM - Support Vector Machine
- EEG - Electroencephalogram
- TBI - Traumatic brain injury
- HC - Healthy Controls
- VA - Valance-Arousal
- BCI - Brain-Computer Interfaces
- AFEA - Automatic Facial Expression Analysis
- FACS - Facial Action Coding System

Contents

List of Figures

CHAPTER 1

Introduction

Humans, as emotional beings, are unable to effectively Without emotions, they communicate and interact with one another. Emotions are required to convey our feelings and intentions. [1] Similarly, our interactions with machines can be enhanced. significantly if they can comprehend our message completely, and that is only conceivable if they can determine the emotional state of the person delivering communication and Visual information is used to recognize emotions. In the age of technology, the way we do things has changed as well. Our lives are generally made easier by technology. Whether we use our phones to unlock by recognizing our faces, track food delivery couriers as they arrive at our doorstep, or use apps for identifying leaves and birds we spot on hikes, there is no end to the possibilities. This has a far greater role to play in human-computer interaction (HCI). There are three major steps in Automatic Facial Emotion Recognition (FER)[2]. which include feature extraction, preprocessing, and classification. Geometric and appearance features are the two broad categories of facial features. The geometric features include information about specific face portions such as the eyes, mouth, brows, and corner points. However, geometric feature-based methods are typically difficult to implement because they require precise and dependable facial feature detection, which can cause issues in a variety of situations[3]. The appearance features, on the other hand, deal with the entire face or a specific region of it, rather than focusing on a specific portion of the face. One of the most popular methods for obtaining such features is to use texture filters, such as Gabor filters, which focus on basic textures and are sensitive to wrinkles and bulges in the skin. Computer systems that attempt to automatically analyze and recognize facial motions and facial feature changes from visual information are referred to as facial expression analysis systems. In the computer vision domain, facial expression analysis is sometimes confused with emotion analysis. A higher level of knowledge is required for emotion analysis. Facial expressions, for example, can convey emotion as well as intention, cognitive processes, physical effort, or other intra or interpersonal meanings. Context, body gestures, voice, individual differences, and cultural factors, as well as facial configuration and timing, all help with interpretation. Computer facial expression analysis systems must be able to analyze facial actions regardless of context, culture, gender, or other factors.

The purpose of this thesis is to develop an application for conducting neuro scientific experiments, as well as the first steps in the use of specialised technologies for data collection and processing in sector research organisations. The path taken while collaborating on a small project is shown, but it connects people from various research disciplines who are working toward a common goal. The project's time horizon will also become clear as the discussion progresses in the following chapters. I worked with Professor Vito De Feo's group, knowing that I was taking part in something that would not have been completed by the end of my thesis work, but that it would have been beneficial to be a part of it

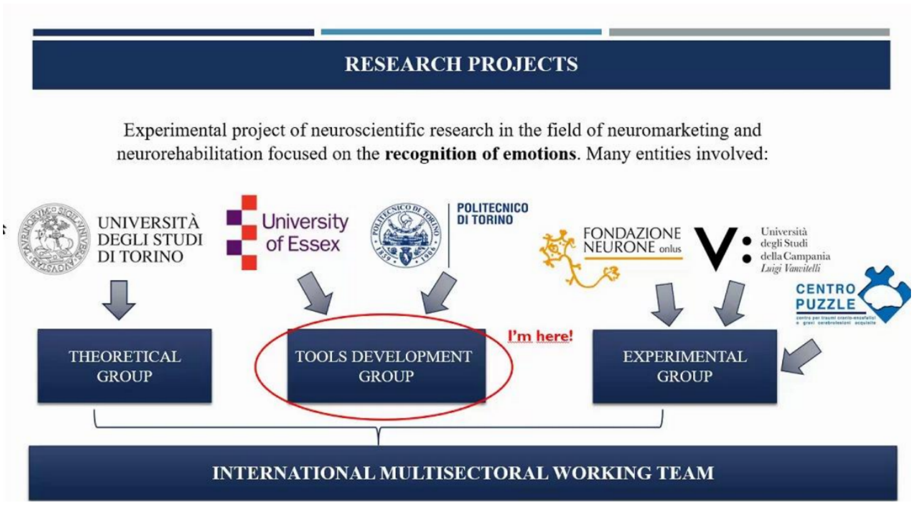


Figure 1.1: The project overview

The presence of colleagues from this foreign university, which is known for its multiculturalism, has allowed me to interact with other mentalities and approaches to common issues, which has assisted me in confronting diverse mentalities and approaches, even those from very far away. These opportunities, in my opinion, are extremely beneficial to a professional's personal development in a field like ours, information technology, which is becoming less and less bound by national borders and which underpins constant interaction with foreign countries that are becoming closer by the day. Not only that, but the great thematic diversity of students and professionals present in the various research subgroups allowed me not only to approach topics that other minds would never have had the opportunity to consider, but also to develop my own ideas.

Specifically this thesis has been structured in 8 chapters, in addition to this brief introduction, each dealing with an aspect or a phase of what my work has been within this project.

The first chapter is a brief introduction of the project followed with next chapter where I've explained Neurorehabiliation and conditions that can benefit from neurological rehab. In the third chapter I've explained the technology devices used to perform the experiment which includes GSR, Eye Tracker and EEG. Next, the fourth chapter includes the explanation and snap shots of the lab application in which I talk about the programming part and the process. This chapter also has a brief description about the programming language used and the advantages of the languages. In the fifth chapter I've explained the iMotions software and the working of the application. This chapter also includes the working of external stimulus software and Pyschopy application. It also seemed necessary to give a brief description about artificial Intelligence and Machine Learning concepts in the further chapters. The data extracted from performing the experiment is also mentioned to explain the analysis of the emotion recognition.

CHAPTER 2

Neurorehabilitation

2.1 What is Neurological rehabilitation?

Neurological rehabilitation (rehab) is a doctor-supervised program designed for people with diseases, injury, or disorders of the nervous system. Neurological rehab can often improve function, reduce symptoms, and improve the well-being of the patient[4]. Traumatic brain injury (TBI) refers to head injuries that impair normal brain function. [5]Â TBI is frequently associated with a wide range of potential psychosocial functional deficits. Although many factors influence psychosocial function after TBI, increasing evidence indicates that social cognitive skills are critical contributors. One of the higher-level skills of social cognition is the ability to perceive and recognise others' emotional states based on their facial expressions. Numerous studies have evaluated facial emotion recognition performance in adult TBI patients. However, the results have been inconsistent. The purpose of this study is to perform a meta-analysis to characterise facial emotion recognition in adult TBI patients.

A colleague majoring in psychology at the University of the Turin Studies has in fact developed a neuro rehabilitation experiment who can help people who have lost or lost or reduced ability to recognize one's emotions. Going to observe in a more general way the topic, we can define neuro rehabilitation as "a branch of medicine specialized in healing and rehabilitating the body following an injury to the nervous system and to minimize or compensate any resulting functional alterations" [4]. It is a relatively discipline.new, but that is finding great success and importance given the large number of discomforts, and their variety, that a problem with the nervous system may involve. This damage can be caused by a large number of Reasons, in our study we have roughly grouped the cause of discomfort in three categories: vascular damage, on co logical damage and head trauma.The problems due to the onset of diseases that cause an injury to the nervous system or more simply a head injury can belong to more than one domain: they can be linguistic, visual,motor, perceptual, or even heavily worsen the capacity of the individual to reason lucidly and, precisely, to recognize in a way clear one's own or others' emotions.

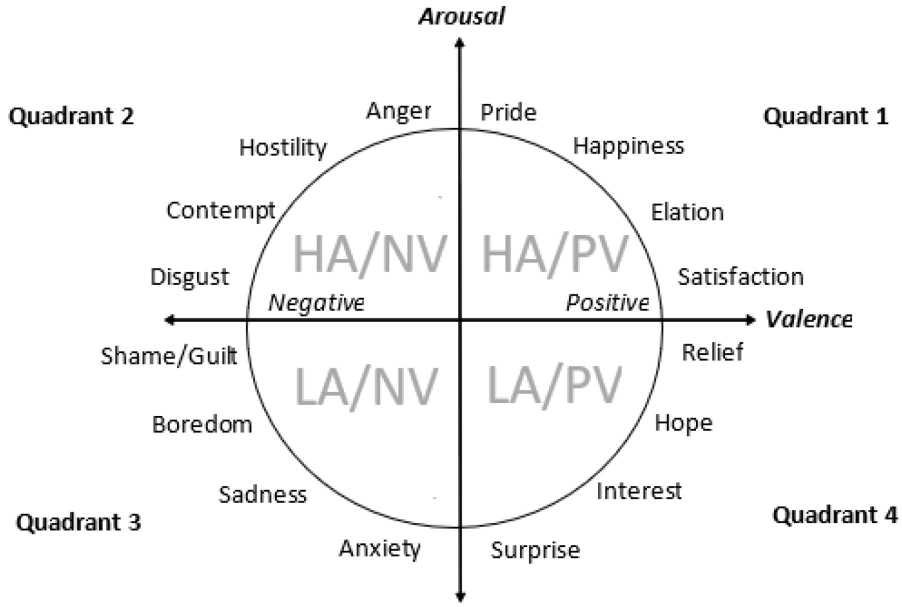


Figure 2.1: Arousal valence Space (AVS) Model.

2.2 What conditions can benefit from neurological rehab?

Injuries, infections, degenerative diseases, structural defects, tumours, and circulatory system disorders can all impair the nervous system. Some of the conditions that may benefit from neurological rehabilitation are as follows:

- Vascular disorders, such as ischemic strokes (caused by blood clots), hemorrhagic strokes (caused by bleeding in the brain), subdural hematoma, and transient ischemic attacks (TIAs)
- Infections, such as meningitis, encephalitis, polio, and brain abscesses
- Trauma, such as brain and spinal cord injury

- Structural or neuromuscular disorders, such as Bell palsy, cervical spondylosis, carpal tunnel syndrome, brain or spinal cord tumors, peripheral neuropathy, muscular dystrophy, myasthenia gravis, and Guillain-BarrÃ© syndrome
- Functional disorders, such as headache, seizure disorder, dizziness, and neuralgia
- Degenerative disorders, such as Parkinson disease, multiple sclerosis, amyotrophic lateral sclerosis (ALS), Alzheimer disease, and Huntington chorea

The ability to recognise emotional expression is critical for social interactions, environmental adaptation, and overall quality of life. People with Alzheimer's disease (AD) have impaired emotion recognition, so rehabilitating these skills has the potential to yield significant benefits. Each year, it is estimated that approximately 10 million traumatic brain injuries (TBI) occur worldwide, posing a global health problem (Dewan et al., 2018). TBI causes a slew of deficits in cognitive and socio-emotional domains (Kelly et al., 2017). Importantly, when compared to other deficits, socio-emotional deficits produce worse outcomes, negatively affecting quality of life, relationships, employment prospects, independence, and autonomy post-injury (Ponsford et al., 2014).

2.3 Measures

2.3.1 Visual emotion recognition

The Ekman 60 Faces Test, which is part of the Facial Expressions of Emotion: Stimuli and Tests, was used to assess visual emotion recognition (FEEST). On a computer screen, patients were shown 60 static images of faces with emotional expressions. They chose the word that best described the emotion expressed by the image shown from a list of six emotions (happiness, sadness, anger, fear, surprise, and disgust). The new Dutch standards were used to interpret the scores. A normal score in the fifth percentile or lower was considered impaired.

2.3.2 Cognition

To screen for cognitive deficits, the Montreal Cognitive Assessment (MoCA) was used. The MoCA was designed to screen for mild cognitive impairments. It assesses the following cognitive domains: visual and executive functions, denomination, memory, attention, mental flexibility, abstraction, language, and orientation. The score is calculated by adding the points earned for each successfully completed task and ranges from 0 to 30.

2.3.3 Mood

The Patient Health Questionnaire-4 (PHQ-4) was used to assess mood, which is an ultrashort self-report questionnaire designed to measure depressive symptoms. The PHQ-4 uses four core symptoms of depression and anxiety: loss of interest, depressed mood, feeling anxious, and difficulty stopping or controlling worrying. The PHQ-4 total scores range from 0 to 6 for each of the two subscales, with a maximum of 6. with a cutoff score of 3 indicating depression or an anxiety disorder. All measures had good reliability and validity.

CHAPTER 3

Technology

Facial Emotion Recognition (FER) is a technology that analyses facial expressions in static images and videos to reveal information about an individual's emotional state. The complexity of facial expressions, the technology's potential use in any context, and the involvement of new technologies such as artificial intelligence all raise significant privacy concerns.

3.1 GSR

The galvanic skin response (GSR, also known as electrodermal activity, or EDA) refers to changes in sweat gland activity that reflect the intensity of our emotional state, also known as emotional arousal. Our level of emotional arousal changes in response to our surroundings; if something is frightening, threatening, joyful, or otherwise emotionally relevant, the subsequent change in emotional response increases eccrine sweat gland activity. This has been linked to emotional arousal in studies. It is worth noting that both positive ("happy" or "joyful") and negative ("threatening" or "saddening") stimuli can increase arousal - and thus skin conductance. As a result, the GSR signal is not representative of theÂ type of emotion, but the intensity of it.

3.1.1 The Background of GSR signals

Vigouroux was the first to discover a link between mental state and GSR activity, discovering a link with patient sedation and skin resistance. In the 120 years since this seminal discovery,

Figure 3.1: The Galvanic Skin Response

thousands of articles have been written about the relationship between emotional response and GSR signal. While sweat secretion is important for thermoregulation and sensory discrimination, emotional stimulation also causes changes in skin conductance; the higher the arousal, the higher the skin conductance.

Skin conductance is unconsciously controlled. [20]Instead, sympathetic activity, which drives aspects of human behaviour as well as cognitive and emotional states, modulates it autonomously. As a result, skin conductance provides direct insights into autonomous emotional regulation.It can be used as an additional source of information to validate self-reports, surveys, or interviews conducted with study participants.

Sweat gland density varies across the human body, but is greatest in the hand and foot regions (200-600 sweat glands per cm squatre), where the GSR signal is typically collected. The signal's time course is thought to be the result of two additive processes: a tonic base level driver that fluctuates slowly (seconds to minutes) and a faster-varying phasic component (fluctuating within seconds).The continuous data stream can detect changes in phasic activity because these bursts have a steep incline to a distinct peak and a slow decline relative to the baseline level. When studying GSR signal changes in response to sensory stimuli, researchers focus on the latency and amplitudes of the phasic bursts with respect to stimulus onset (images, videos, sounds).

3.2 Eye-Tracker

Eye tracking is a practise that has been around for over a century. Since the first follow-the-speck-on-the-eye methodology, much has changed, and most setups now involve simply sitting comfortably in front of a candy-bar-sized tracker. While it is obvious that hardware has become easier to use, faster, and better, software has also changed significantly.

The history of eye tracking software is not as long as that of hardware, but it follows a similar path, beginning with rudimentary but functional solutions and progressing to where we are today - easier-to-use, faster, and better. Whereas the software previously only provided the essential raw data, it is now possible to go much further with a few clicks. Eye tracking uses various technological processes to measure eye movements, positions, and points of gaze. Eye tracking, in other words, identifies and tracks a person's visual attention in terms of location, objects, and duration.

Eye tracking technology is frequently mentioned in relation to:

- Blinking pattern measurement
- Identifying what a subject ignores or does not look at
- Evaluation of pupil reaction to visual and emotional stimuli
- Human-computer interaction and machine learning are made easier.

Images are processed by the brain via light-sensing cells in the retina. These rod and cone cells detect light passing through the pupil and transmit it to the brain. While there are far fewer cone cells in the retina than rod cells, the former allow humans to see in high resolution and full colour. The eye performs a variety of movements, such as vergence and torsional. Fixations, saccades, and smooth pursuit are the most important ones measured by eye tracking. Fixations happen when the eye stops collecting visual data. Although the duration of a fixation varies greatly, the longer it lasts, the more visual information is processed.

Saccades are quick "jumps" performed by the eye between fixations in a static environment. The eye moves from one object of interest to another in order to acquire new high-resolution visual data. Vision is suppressed due to the extreme speed of saccades. As a result, they are not as important in gaze tracking as fixations are. They do, however, reveal information

Figure 3.2: Eye Tracker

about the trajectory of fixations and visual attention. Smooth pursuit is the eye movement that occurs when looking at and following an object in motion. Because smooth pursuit allows for visual intake, movement is important for tracking eye movements.

3.2.1 Is eye tracking used for qualitative or quantitative research?

While standard limitations of eye tracking face a series of limitations, 3D eye tracking for depth sensors solves the challenges with the following solutions:

- Wide range
- Multi-person tracking
- 3D line of sight
- No calibration
- Software only (if eye tracking device integrates 3D camera)

Eye tracking works by passively tracking the position and movement of the eyes. The pupil is illuminated by an invisible source of near-infrared or infrared light. As a result, a reflection forms on the cornea. An infrared camera will then record that reflection, defining the pupil's centre, determining eye rotation, and determining gaze direction. Pupil Center Corneal Reflection (PCCR) is the formal term for the main method of eye tracking that uses optical monitoring of the pupil and corneal reflections. Advanced mathematical algorithms are used to calculate eye position, point of gaze, and eye movements.

Following that, the recorded data is converted into raw data, which is then processed by eye tracking software.

3.3 EEG

The EEG is a noninvasive electrophysiological technique for recording electrical activity from the human brain. Hans Berger, a German psychiatrist, presented the first report on the application of this technique in humans in 1924. [18] EEG signals are typically collected using a specialised device known as an electroencephalogram. The main components of this device are special metal plate electrodes that should be placed on the human scalp, while alternative needle electrodes can be inserted directly into the scalp in special cases. In most cases, eight, sixteen, or thirty-two pairs of electrodes are placed on four standard locations on the head: the nasion, inion, and right and left preauricular points in the figure. Electrodes

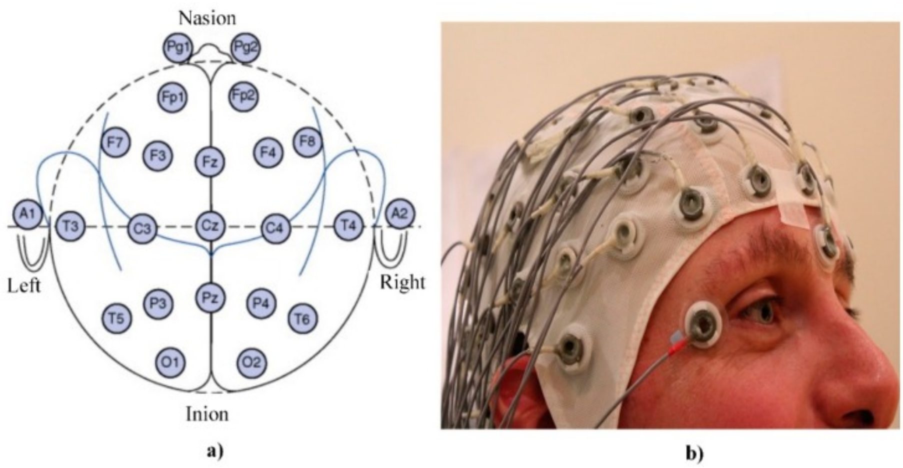


Figure 3.3: Electroencephalography (EEG) measurements: (a) distribution of EEG electrodes on human scalp; (b) special headset with installed electrodes

can be applied to the human scalp using adhesive-conducting gel or special headsets with electrodes already installed. The EEG signal is a voltage fluctuation between two paired electrodes over time, and the signal amplitude is typically measured using the peak to peak technique. The response of the brain to various stimuli is usually measured and analysed by five frequency ranges from EEG signals, namely: delta, theta, alpha, beta, and gamma.

These band waves are found throughout the brain and are associated with various emotional states. A variety of methods for processing and analysing EEG signals can be implemented depending on the object of interest. If the goal is to assess an average level of valence and arousal or to detect the efficiency of applied stimulus, a fast Fourier transformation or latency test can be used. If the purpose is to identify a specific emotion and its strength, statistical methods or machine learning techniques can be implemented. A review of related works based on only EEG signals.

CHAPTER 4

Application

4.1 Introduction to Python

Python appears to be the most plausible answer after considering the previous questions. [24]. Guido Van Rossum created Python, a high-level programming language, in the early 1990s. The language is pseudo-interpreted, which means that an interpreter turns on a virtual machine and runs one line of code at a time, converting it into a machine language that is platform agnostic. His programmer's goal was to make approaching it extremely simple because of the similarity to pseudocodes and the very streamlined structure, which is often taught as a first language thanks to the indentation of several code lines and have not left any room for ambiguity like the option to group the different logical blocks.

Python supports a variety of programmes such as procedure, imperative, functional, and objective-oriented programming objects, allowing for reasonably flexible language use that can be tailored to the programmer's needs and skills. It can also be used to create sophisticated, object-oriented platforms that are completely cooperative, as opposed to orders that only share the running sequence (scripting). Another important feature of the language is its strong dynamic typing, which allows the assignment of any type of value to variables, allowing for greater flexibility in programme design but with the risk of errors being detected in runtimes when everything appeared to work properly in advance.

Finally, Python employs a waste collector based on the theory of production and object barriers, which indicate the number of objects in a given generation when a 'collection process' is initiated in order to free the memory of components not already in use in the application. When memory is not freed or is freed at an early stage, the programmer can modify this technique to avoid errors, sometimes at the expense of performance. It's a powerful part of the Python ecosystem; it's a large standard library with a wide range of tools for most different jobs, as well as an extensive third-party library called Python Pack. The ability to change and redistribute code is central to the ecosystem.

SDKs for all other tools, starting with the face and progressing through the Eyes-Tracker, coupled with the GSR, to the EEG, are all consistent with Python, which is even the language in some that the majority of materials are contained in the applications. Because Python is platform-independent, there are no issues with the environment in which it is used, so we can start working on Microsoft without worrying about compatibility.

4.2 Advantages of Python

Not only was the language choice confirmed because we responded to the fundamental constraints of our project, which were required to get started, but it also satisfied us of some other problems that are constantly part of this approach to growth and look forward to the project's upcoming. First and foremost, it was agreed that the machine learning component in Python would be built in light of its overwhelming control and widespread usage in literature. In terms of client and server structure, the web app has no significant limitations.

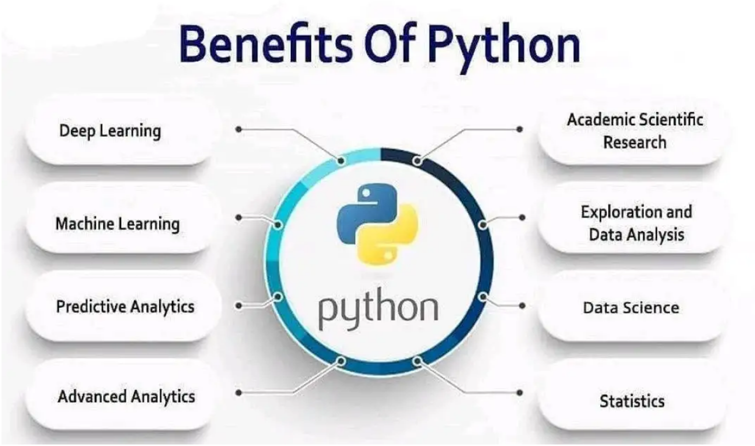


Figure 4.1: Benefits of Python Programming

There was no significant constraint in selecting a language on the customer to host the web app and another on the server to prepare and send responses. An independent structure was chosen for the desktop app. As a result, it would clearly be more convenient for all of its parts to use the same language, even if, as previously stated, it would not be limited to communicating two modules in different languages. As previously stated, versatility was one of the project's top priorities; if we had worked in the technical-scientific field, we could have considered using tools like Matlab, which allows us to perform any math-related task in a dense and instantaneous manner. This is significant even though arrays can be obtained from components in the C, C++, and Fortran language families, so that you can optimise certain processes that in Pythons turn out to be impediments to a particular process task through the use of languages which work at bottom elevation or simply through the use of certain librariesÂ which are not implemented in Python.

4.3 Tkinter

Following the selection of a language, the next step was to select the extra pieces that comprise the application environment for the various experiences. The next step was taken. Given the requirement for an interface to begin studies and to minimise participant data, it was clear that a library was required to govern the user interface. Because the software system will only be installed in the laboratories where the study will be conducted, it has been decided to avoid focusing on customer happenings outside of the test, and thus not to invest time in developing particularly sophisticated. We decided to build a functional and minimal programme with a few years ago's standard Laboratory Application modelsÂ withÂ few icons and list menu, minimum structure and look, without too much consideration to the aesthetic component.

We chose Tkinter because of the limited criteria and the desire to preserve the code, the project, and several examples of existing literature, tutorials, and other types of reports. Tkinter is a Python binding for the Tk graphical user interface, which is used in conjunction with the Tcl high-level computer language. This complementarity is provided by the development of graphical user interfaces in a variety of languages and technologies. There are several alternatives, some of which are used within the program's libraries to provide additional capabilities but necessitate more research and source code in order to achieve more sophisticated and visible results.

The debate can be continued by identifying positive and negative aspects of Tkinter; I adore its characteristics, so let us begin with the positive: It has excellent portability in both the Windows and Mac environments, producing native widgets on all platforms; its binding is undeniably more robust. WxPython, for example, is more flexible and consistent than many other libraries. It is quite mature and stable, having been the Python standard for nearly 30 years. It has simple syntax and APIs. For example, the textual and canvas components are both simple and powerful, and the geometry managers (paket, location, and grid) allow you to easily control the page layout.

4.4 Image Experiment

PyCharm is a dedicated Python Integrated Development Environment (IDE) providing a wide range of essential tools for Python developers, tightly integrated to create a convenient environment for productive Python, web, and data science development. ...

To comprehend 'What is PyCharm?' and 'What is PyCharm used for?' we should be able to answer 'What is an IDE?' first. An integrated development environment (IDE) is made up of an editor and a compiler that we use to write and compile programmes. It possesses a set of features required for software development.

The presence of an IDE greatly simplifies the development and programming processes. It interprets what we type and suggests the appropriate keyword to insert. We can tell the difference between a class and a method because the IDE gives them different colours. The IDE also uses different colours to indicate which keywords are correct and which are incorrect. If we type the wrong keyword, it attempts to predict the keyword that we intend to type and auto completes it.

- User Window :

  This application is developed in PyCharm with the help of python language. Prior the experiment, we need to login to the application with the credentials and select the participant in order to start the experiment. The user window allows us to add and delete the participant as well. It also has the option to search the participant. Once the participant is selected, the experiment can be proceeded.

  Next, the window displays the option to select the image experiment. There is also

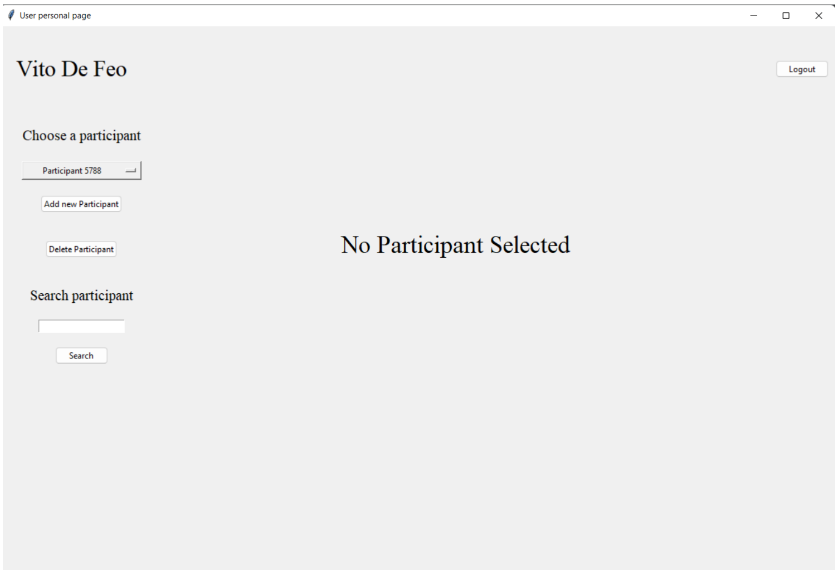


Figure 4.2: User Window

an option to export data, delete data and option to record. Select the setting of the application and you can proceed with the experiment.

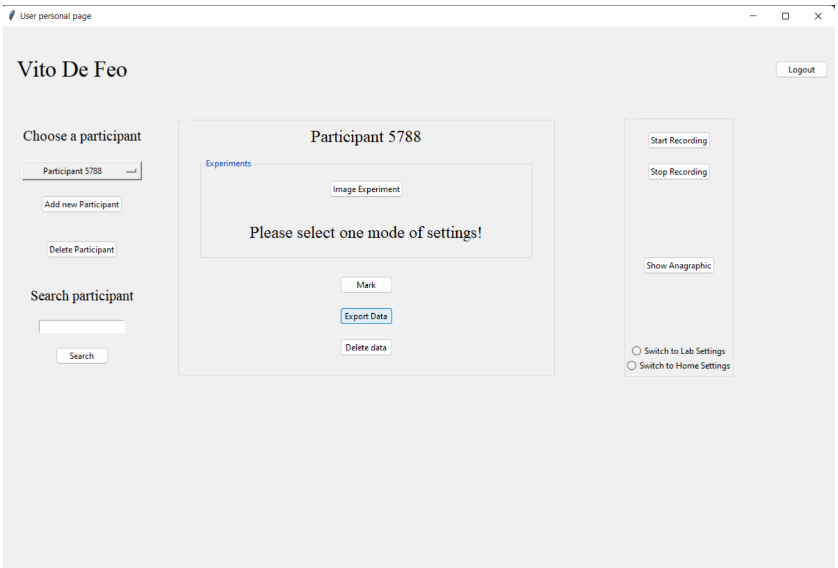


Figure 4.3: User Personal Page

- Start Image experiment and Survey :
  By clicking on the start image experiment button, a window displays with images along with the counter on top. The image window has a skip button to skip the stimulus or wait till the counter runs and displays next image automatically. After the image window is skipped, or when the counter reaches 5 seconds, the survey page displays

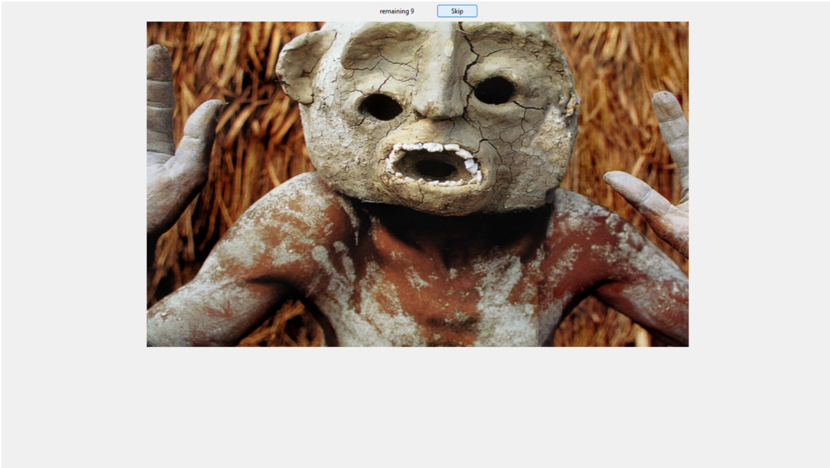


Figure 4.4: Image Window

where the participant is allowed to respond to the survey options by entering the emotions responses in the range 1-9 according to participant.

Figure 4.5: Survey Window

Tkinter in Python os used to create Graphical User interfaces (GUIs) and is included in all standard Python Distributions. In fact, it's the only framework built into the Python standard library. It provides various response types which makes it easier to design GUI's. This Python framework interfaces with the Tk toolkit and functions as a thin object-oriented layer on top of Tk. The Tk toolkit is a cross-platform collection of 'graphical control elements,' also known as widgets, that can be used to create application interfaces. This framework allows Python users to easily create GUI elements by utilising the widgets found in the Tk toolkit. In a Python application, Tk widgets can be used to create buttons, menus, data fields, and so on. These graphical elements, once created, can be associated with or interact with features, functionality, methods, data, or even other widgets. A button widget, for example, can accept mouse clicks and be programmed to perform an action, such as exiting the application.

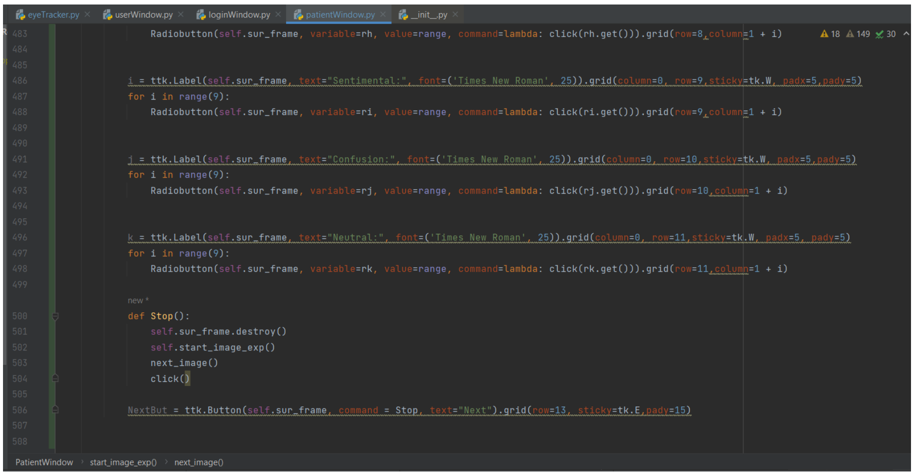


Figure 4.6: Tkinter code to build survey

CHAPTER 5

iMotions

5.1 Introduction to iMotions

iMotions has built quite an ecosystem of software offerings for a tech company that is fast approaching 20 years in the market, and in this blog, we will provide an overview of how far we have come in almost two decades. In 2005, the answer to the question of what iMotions does was fairly simple: it is an eye tracking company with big ambitions. Sixteen years later, the company, expertise, and offerings have evolved alongside and beyond eye tracking to now include a variety of research pathways. With that in mind, we'd like to provide an updated answer to that question and walk you through all of the options for behavioural research available with iMotions today.

iMotions is, at its core, a software company that provides solutions for the present and future of behavioural research. We enable research to keep up with the world's rapid changes. We have the technology to help you understand how humans behave in each environment, whether your research takes place in a carefully controlled lab, on a ship at sea, or in the local supermarket. We understand that the challenges that researchers face today are not the same as they have always been - the pace of research necessitates flexibility. That is why we have focused on providing a system that is adaptable, not only in terms of sensor configurations, but also in terms of where research can take place: anywhere.

5.2 Working with iMotions

In 2021, the iMotions ecosystem will consist of a few components, each designed in a unique way to address today's human behaviour research challenges. We provide two types of software products to help with research: lab-based and remote. Each offers a unique approach to understanding human behaviour through technology, and is complemented by our tier of training and services. This includes the following:

Lab-based	Remote	Research Enablement
• iMotions Core (includes surveys + API)	• iMotions Online Data Collection (ODC)	• iMotions Onboarding & Customer Success Program
• iMotions modules (sensor-based)	• Mobile Research Platform (Wearable sensors)	• iMotions Academy
• R Notebooks		• iMotions Services

Figure 5.1: iMotions Services

The iMotions modules connect to Core and allow you to conduct research in any way you want - each module is a piece of software that offers different data collection and analysis options for various biosensors. For example, you can collect facial emotion data using the Facial Expression Analysis module, which also allows you to collect facial coding data from participants. This can be seamlessly supplemented with data from over 50 different biosensors, such as eye tracking, electrodermal activity, EEG, and others.

- Eye tracking â screen-based
- Eye tracking â glasses
- Eye tracking â VR
- Eye tracking â webcam

- Facial expression analysis
- Electrodermal activity (known as EDA, or GSR)
- Electroencephalography (EEG)
- Electrocardiography (ECG)*
- Electromyography (EMG)*

5.3 External Stimulus Softwares

The ability to present stimuli to your respondents while measuring their psychophysiological responses is one of the software's pillars. Users can easily design sophisticated studies using various types of stimuli such as videos, audio, images, and websites thanks to our built-in stimulus presentation engine. The software keeps the biometric data in sync with the stimulus presentation, making it simple to extract various types of metrics from eye-tracking on a stimulus-by-stimulus basis.

Different dedicated stimuli presentation tools are frequently used in academic research in experimental psychology. E-prime is a popular tool that is used in many publications. Its graphical user interface allows users to design studies as well as perform inline scripting. Users can also create extremely complex study designs. Other stimulus presentation tools, such as Psychtoolbox, PsychoPy, and Presentation, are also used, but the user must have coding experience to use them, even though OpenSesame provides a graphical user interface as an overlay to PsycoPy.

5.4 Image - Video Experiment

5.4.1 PsychoPy

There are two types of triggers/markers that can be sent to iMotions via the event receiving API. The first option is to use a point/discrete marker, which creates a single marker on your iMotions data at a specific point in time. This might be a good option if you just want to know when something happened. However, when interacting with third-party applications, I prefer the more sophisticated range/scene marker function.

The psychopy gives options to both build and code. The builder window has multiple options such as stimulus, responses and coder where we can add images, buttons and import files.

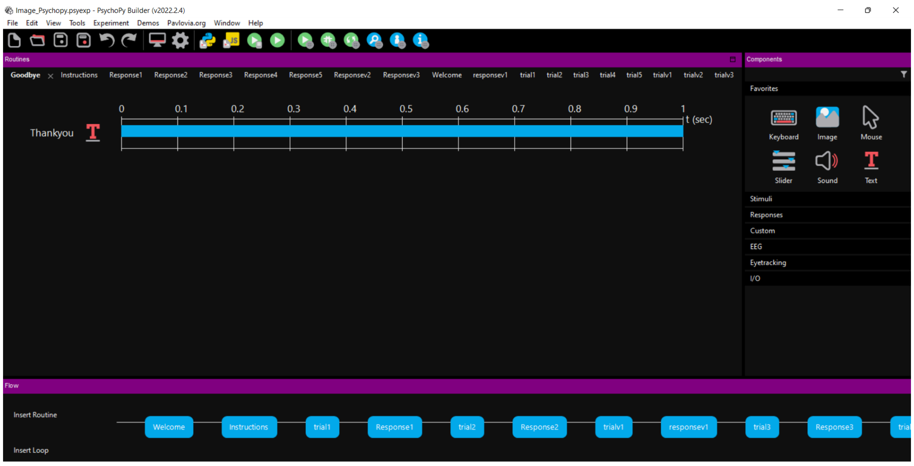


Figure 5.2: Psychopy Builder View

The builder view allows us to add images in loop or manually. Once we run the experiment, it takes quite a few seconds to start the experiment which displays a welcome page. By clicking on space bar, the experiment starts. But, prior to the experiment, an instruction page displays where the participant can get the instructions on how the experiment is to be performed.

INSTRUCTIONS

1. Do not press anything untill you get instructions from the supervisor.
2. Keep your eyes on the screen while performing the experiment.
3. The timer will start once you click the spacebar. So, you will have limited time to finish this experiment.
4. Once you finish the experiment, inform the supervisor.
5. Press ‘Space’ to start the experiment.

Figure 5.3: Instructions Page

After going through the instructions, the image-video experiment starts. Every stimuli displays for 5 seconds and then survey page is displayed where the user is allowed to respond to the survey. This page displays multiple emotions with radio-buttons for the participant to respond according to the range of emotions he/she is feeling

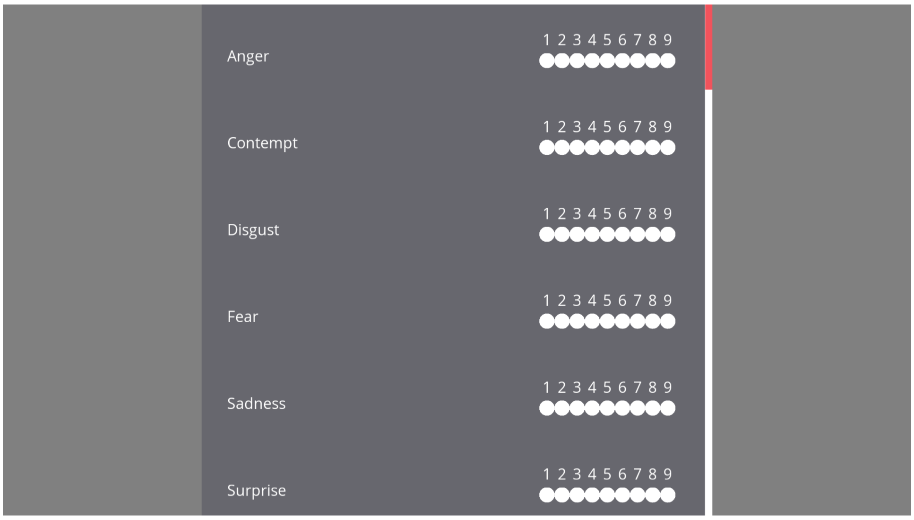


Figure 5.4: Survey

By clicking the responses, the results are saved in an excel sheet which can be later used for data analysis to conclude results. The Coder view allows us to code application which also allows various packages to design the stimulus application by importing images and excel sheets to produce forms. We can also set colours and fonts to style the application accordingly.

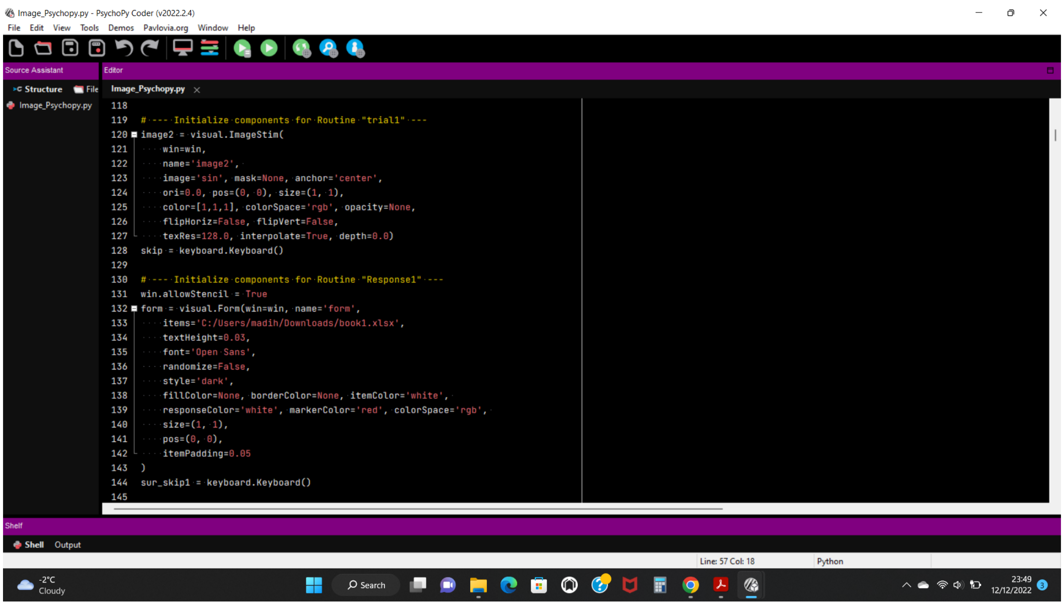


Figure 5.5: Psychopy Coder View

CHAPTER 6

Machine Learning

Machine learning (ML) is a subset of artificial intelligence (AI) that enables software applications to become more accurate at predicting outcomes without explicitly programming them to do so. Machine learning algorithms predict new output values by using historical data as input. Machine learning is important because it provides enterprises with insights into trends in customer behaviour and business operational patterns, as well as assisting in the development of new products. Machine learning is central to the operations of many of today's leading companies, including Facebook, Google, and Uber. For many businesses, machine learning has become a significant competitive differentiator.

6.1 What are the different types of machine learning?

Classical machine learning is frequently classified by how an algorithm learns to improve its prediction accuracy. There are four fundamental approaches to learning: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. The algorithm that data scientists use is determined by the type of data they want to predict.

6.1.1 Supervised learning

In this type of machine learning, data scientists provide labelled training data to algorithms and specify which variables they want the algorithm to look for correlations between. The algorithm's input and output are both specified. Supervised machine learning necessitates

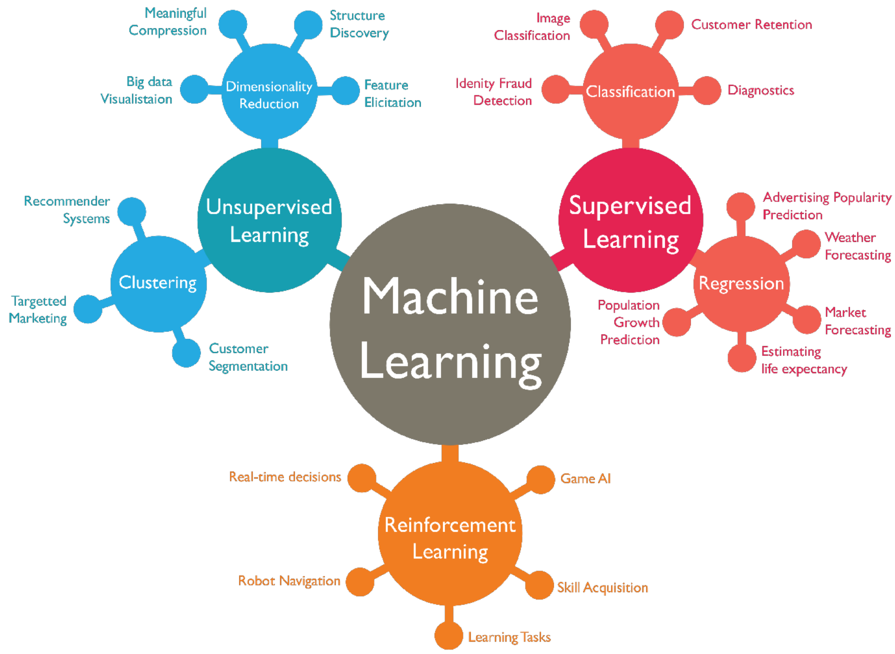


Figure 6.1: Different types of machine learning

the data scientist training the algorithm with both labelled inputs and desired outputs. Supervised learning algorithms are useful for the following tasks:

- Binary classification is the division of data into two categories.
- Multi-class classification: Choosing between more than two types of answers.
- Predicting continuous values with regression modelling.
- Ensembling: The process of combining the predictions of multiple machine learning models to produce an accurate prediction.

6.1.2 Unsupervised learning

Algorithms that train on unlabeled data are used in this type of machine learning. The algorithm scans data sets for any meaningful connections. The data used to train algorithms, as well as the predictions or recommendations they produce are predetermined. Unsupervised machine learning algorithms do not need labelled data. They sift through unlabeled data for patterns that can be used to classify data points into subsets. The vast majority of deep learning algorithms, including neural networks, are unsupervised. The following tasks are well suited to unsupervised learning algorithms:

- Clustering is the process of dividing a dataset into groups based on similarities.
- Anomaly detection is the detection of unusual data points in a data set.
- Association mining is the process of identifying groups of items in a data set that frequently occur together.
- Dimensionality reduction refers to the process of reducing the number of variables in a data set.

6.1.3 Semi-supervised learning

This approach to machine learning combines the two preceding types. Data scientists may feed an algorithm mostly labelled training data, but the model is free to explore the data on its own and develop its own understanding of the data set. Semi-supervised learning works by data scientists feeding a small amount of labelled training data to an algorithm. The algorithm learns the dimensions of the data set from this, which it can then apply to new, unlabeled data. When algorithms are trained on labelled data sets, their performance typically improves. However, labelling data can be time-consuming and costly. Semi-supervised learning falls somewhere between the performance of supervised learning and the efficiency of unsupervised learning. Semi-supervised learning is used in the following contexts:

- Machine translation is the process of teaching algorithms to translate languages using less than a complete dictionary of words.
- Fraud detection: Detecting cases of fraud when you only have a few positive examples.
- Data labelling: Algorithms trained on small data sets can learnÂ to apply data labels to larger sets automatically.

6.1.4 Reinforcement learning

Data scientists typically use reinforcement learning to teach a machine to complete a multi-step process with clearly defined rules. Data scientists program an algorithm to complete a task and give it positive or negative cues as it works out how to complete a task. But, for the most part, the algorithm decides what steps to take along the way. Reinforcement learning operates by programming an algorithm with a specific goal and a set of rules for achieving that goal. Data scientists also programme the algorithm to seek positive rewards, which it receives when it performs an action that benefits the ultimate goal, and to avoid punishments, which it receives when it performs an action that moves it away from the ultimate goal. Reinforcement learning is frequently used in situations such as:

- Robotics: Using this technique, robots can learn to perform tasks in the physical world.
- Reinforcement learning has been used to teach bots to play a variety of video games.
- Video gameplay: Reinforcement learning has been used to teach bots to play a number of video games.
- Resource management: Given finite resources and a defined goal, reinforcement learning can help enterprises plan out how to allocate resources.

6.1.5 Multi-Task Learning

"Multi-Task Learning (MTL) is a machine learning subfield in which a shared model learns multiple tasks at the same time." The benefits of such approaches include improved data economy, less over-appropriate representation, and rapid learning using auxiliary data information.â [22] For the development of a successful model, ML, which uses knowledge to improve the interpretation of later data and leverages useful information from conventional details, generally requires a large number of labialised details. Deep models with multiple hidden layers and various parameters in the form of neural networks are a common type of machine learning model. To identify the correct parameters, these models typically require millions of data samples. This need may, however, not be met in other fields, such as medical image analysis, as the labelling of samples requires extra handling. Multi-task learning (MTL), including significant information from other relevant learning, is an efficient technique for decreasing data dispersion in these situations [21].

6.2 What are the advantages and disadvantages of machine learning?

Machine learning has been used in a variety of applications, from predicting customer behaviour to developing the operating system for self-driving cars. In terms of benefits, machine learning can assist businesses in better understanding their customers. Machine learning algorithms can learn associations and help teams tailor product development and marketing initiatives to customer demand by collecting customer data and correlating it with behaviours over time.

Machine learning is a primary driver in some businesses' business models. Uber, for example, matches drivers with riders using algorithms. Machine learning is used by Google to surface ride advertisements in searches. However, machine learning has drawbacks. For starters, it can be expensive. Data scientists, who command high salaries, are typically in charge of machine learning projects. These projects also require software infrastructure that can be expensive.There is also the issue of bias in machine learning. Algorithms trained on data sets that exclude certain populations or contain errors can result in inaccurate models of the world that fail at best and are discriminatory at worst. When an organisation bases its core business processes on biassed models, it risks regulatory and reputational harm.

CHAPTER 7

Artificial Intelligence

Artificial intelligence uses computers and machines to simulate the human mind's problem-solving and decision-making abilities. While there have been several definitions of artificial intelligence (AI) over the last few decades, John McCarthy offers the following definitio, "It is the science and engineering behind the creation of intelligent machines, particularly intelligent computer programs. It is similar to the similar task of using computers to understand human intelligence, but AI does not have to limit itself to biologically observable methods." [23] However, the birth of the artificial intelligence debate was marked decades before this definition by Alan Turing's seminal work, "Computing Machinery and Intelligence," which was published in 1950. Stuart Russell and Peter Norvig then went on to publish Artificial Intelligence: A Modern Approach, which quickly became one of the leading textbooks in the field of AI research. In it, they delve into four potential AI goals or definitions, distinguishing computer systems based on rationality and thinking vs. acting:
Human-centered approach:

- Human-like computer systems
- System that behaves like humans

The best approach is:

- Rational-thinking systems
- Systems that behave logically

Alan Turing's definition would have fallen under the category of "human-like systems." In its most basic form, artificial intelligence is a field that combines computer science and large datasets to solve problems. It also includes the subfields of machine learning and deep learning, which are frequently mentioned in the context of artificial intelligence. AI algorithms are used in these disciplines to create expert systems that make predictions or classifications based on input data.

7.1 Types of artificial intelligence

Weak AI, also known as Narrow AI or Artificial Narrow Intelligence (ANI), is AI that has been trained and focused on performing specific tasks. The majority of the AI that surrounds us today is driven by weak AI. This type of AI is anything but weak; it powers applications like Apple's Siri, Amazon's Alexa, IBM Watson, and self-driving cars.

Artificial General Intelligence (AGI) and Artificial Super Intelligence (ASIC) are components of strong AI (ASI). Artificial general intelligence (AGI), also known as general AI, is a theoretical form of AI in which a machine has an intelligence comparable to humans; it has a self-aware consciousness capable of solving problems, learning, and planning for the future. Artificial Superintelligence (ASI)âalso known as superintelligence. would outperform the human brain's intelligence and capability While strong AI is still entirely theoretical, with no practical examples in use today, AI researchers are still investigating its development. Meanwhile, the best examples of ASI may come from science fiction, like HAL, the superhuman, rogue computer assistant in 2001: A Space Odyssey.

7.2 Deep learning vs. machine learning

Because deep learning and machine learning are often used interchangeably, it's important to understand the differences between the two. As previously stated, deep learning and machine learning are both subfields of artificial intelligence, with deep learning being a subfield of machine learning. Neural networks are the building blocks of deep learning. Deep learning algorithms can be defined as neural networks with more than three layers (including inputs and outputs). The difference between deep learning and machine learning is in how each algorithm learns. Deep learning automates much of the feature extraction

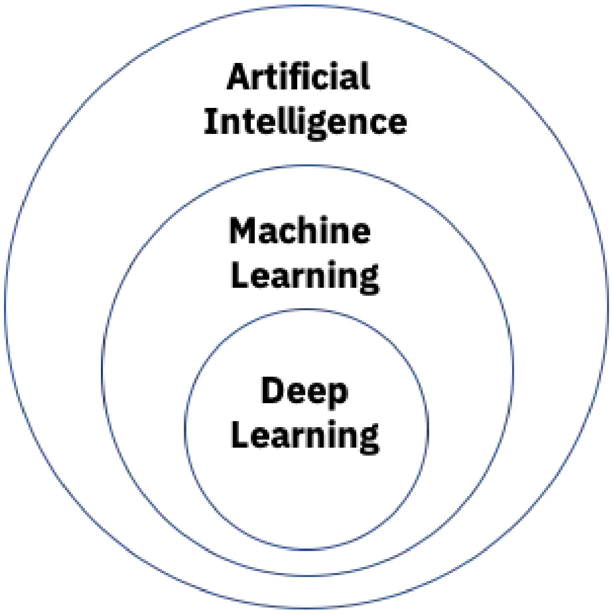


Figure 7.1: Difference between Machine learning and Deep learning

process, removing some of the manual human intervention and allowing for the use of larger data sets.

7.3 Applications

There are numerous, real-world applications of AI systems today. Below are some of the most common examples:

- Speech Recognition is a capability that uses natural language processing (NLP) to convert human speech into a written format. It is also known as automatic speech recognition (ASR), computer speech recognition, or speech-to-text. Many mobile devices incorporate speech recognition into their systems to conduct voice searches (e.g., Siri) or to improve texting accessibility.

- Customer service: Throughout the customer journey, online chatbots are replacing human agents. They respond to frequently asked questions (FAQs) about topics such as shipping or provide personalised advice, such as cross-selling products or recommending sizes for users, altering the way we think about customer engagement across websites and social media platforms. Examples include messaging bots on e-commerce sites with virtual agents, messaging apps like Slack and Facebook Messenger, and tasks that are typically performed by virtual assistants and voice assistants.

- Computer Vision: This AI technology allows computers and systems to derive meaningful information from digital images, videos, and other visual inputs, and then act on that information. This ability to make recommendations sets it apart from image recognition tasks. Computer vision, which is powered by convolutional neural networks, has applications in photo tagging in social media, radiology imaging in healthcare, and self-driving cars in the automotive industry.

- AI algorithms can help discover data trends that can be used to develop more effective cross-selling strategies by using past consumption behaviour data. This is used by online retailers to make relevant add-on recommendations to customers during the checkout process.

- Automated stock trading: Designed to optimize stock portfolios, AI-driven high-frequency trading platforms make thousands or even millions of trades per day without human intervention.

CHAPTER 8

Emotion Recognition

8.1 Facial Emotion Recognition

A number of studies have recently examined facial emotion recognition deficits in adult TBI patients. However, research on specific emotion recognition deficits in TBI has yielded inconclusive results. Wearne et. al [7] discovered that, when compared to healthy controls (HC), adult patients with TBI have higher overall accuracy in recognising emotion, specifically happy and sad emotions, whereas Byom et. al [6] discovered no difference between TBI patients and HC in happy, but a significant difference in sad expressions. Furthermore, the sample sizes in these studies were small, and the magnitude of group differences between adult patients with TBI and HC varied significantly. A meta-analysis can help increase statistical power and clarify inconsistent findings in individual studies. So in this study, we will conduct a meta-analysis to investigate the magnitude of emotion recognition deficits in adult patients with TBI in comparison to HC. Furthermore, we will conduct subgroup meta-analyses to determine whether difficulties in recognising negative emotions (anger, disgust, fear, and sadness) or positive emotions (happy and surprise) or any specific basic emotion (for example, anger or happiness) can be a more distinguishing feature of TBI. In addition, meta-regression analyses will be carried out to investigate the effects of potential co-founders on emotion recognition deficits, such as age, gender, education level, and disease duration. Our meta-analysis will aid in understanding the patterns of emotion recognition function in adult TBI patients, which may be useful in identifying targets for affect recognition interventions and developing useful training intervention programs.

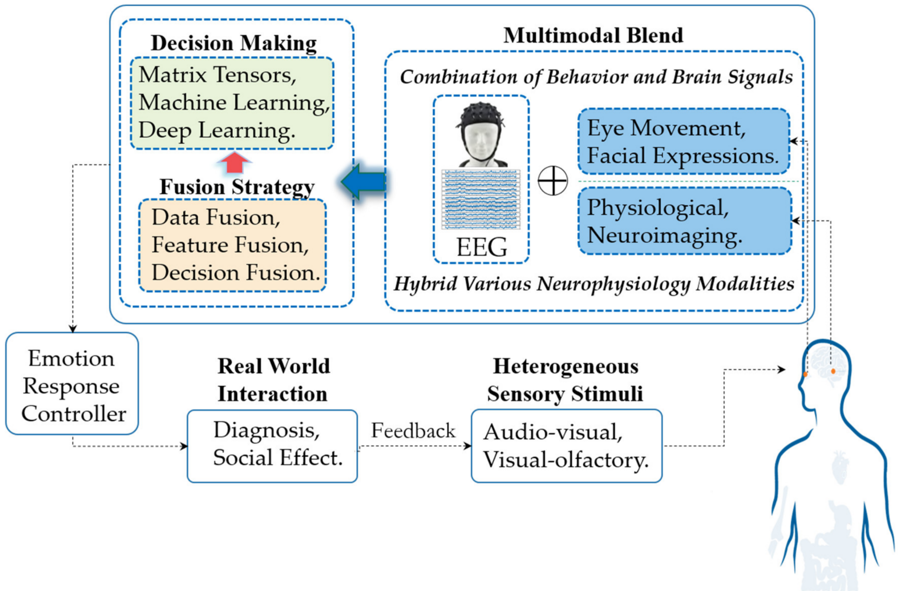


Figure 8.1: Flowchart of multimodal emotional involvement for electroencephalogram (EEG)-based affective brainâcomputer interfaces (BCI)

- In the signal detection and input stage, the system's signal acquisition devices capture modality data from combined brain signals and behaviour modalities (such as expression and eye movement) or various hybrid neurophysiology modalities (such as functional Near-infrared spectroscopy (fNIRS) and fMRI) induced by heterogeneous sensory stimuli (such as audio-visual stimulation) during interaction with other groups or by certain events. At this point, multimodal aBCI can be classified into three types: aBCI based on a combination of behavioural and brain signals, aBCI based on various hybrid neurophysiology modalities, and aBCI based on heterogeneous sensory stimuli.

- Obtaining pure EEG signals containing emotional signals is critical in signal processing. The preprocessing process for neurophysiological signals is to denoise and remove artefacts from the collected original signals. When it comes to image signals like expressions, irrelevant information in images is removed while information representing emotional state is restored and enhanced. We can use various fusion strategies during the modalities fusion stage, such as data fusion, feature fusion, and decision fusion. Emotional state decision output can be used to select spatial tensors, machine learning, or deep learning methods as the final classification decision-maker in decision-making.

- The complete modality data are the source domain for the modality problem, and the missing modality data are the target domain. The goal is to transfer knowledge between the various modality signals. The authors proposed a novel semisupervised multiview deep generative framework for multimodal emotion recognition in the presence of incomplete data. Each modality of emotional data is treated as a separate view in this framework, and the importance of each modality is inferred automatically by learning a nonuniformly weighted Gaussian mixture posterior approximation for the shared latent variable. Our framework naturally addresses the labeled-data-scarcity problem by recasting the semisupervised classification problem as a specialised missing data imputation task. The incomplete-data issue is elegantly avoided by treating missing views as latent variables and integrating them.

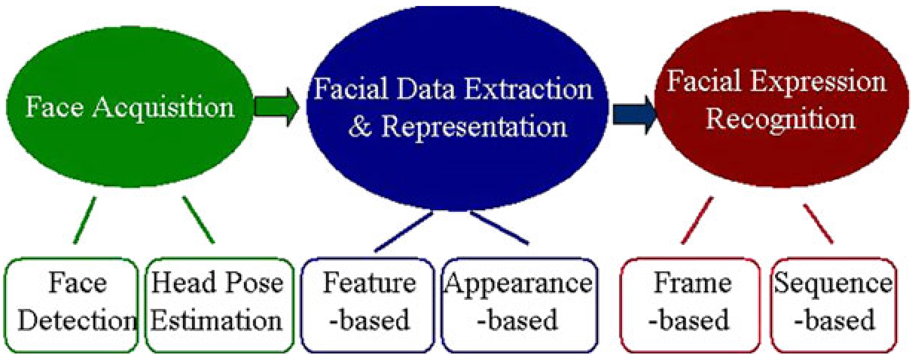


Figure 8.2: Basic structure of facial expression analysis systems

Achievements in related fields such as psychology, human movement analysis, face detection, face tracking, and recognition enable automatic facial expression analysis. Automatic facial expression analysis can be used in a variety of fields, including clinical psychology, psychiatry, neurology, pain assessment, lie detection, intelligent environments, and multimodal human computer interfaces (HCI).

8.2 Facial Action Coding System (FACS)

The Facial Action Coding System (FACS) is the most widely recognised, comprehensive, and credible method of measuring visually discernible facial movements. It was created by psychologists Paul Ekman and Wallace V. Friesen to objectively measure facial expressions for behavioural science research (Ekman, Friesen, 1978; Ekman, Friesen, Hager, 2002)[11].

Prior to the development of FACS, there was no standard system for measuring the face, which made collecting reliable scientific data in the field extremely difficult. FACS is anatomically based, with 44 distinct Action Units (AUs) that correspond to each motion of the face, as well as several sets of head and eye movements and positions. Each AU has a numeric code that identifies the muscle or muscle groups being contracted or relaxed, which create the changes in appearance that are observed, providing the ability to analyze minuet facial motions. The intensity of each AU facial action is recorded on a five-point scale. Events are individual facial expressions that can be coded in their entirety after being deconstructed into an AU-based description, which can include a combination of Action Units or a single Action Unit. Despite the fact that FACS is based on muscle movement, there is no direct relationship between muscles and AUs. Each AU is associated with a specific muscle or groups of muscles; however, because a single muscle can act differently and produce a wide range of facial changes, AU codes refer to how the muscle moves to produce the specific facial change.

While the number of AUs used in FACS is not large, over 7,000 distinct AU combinations have been observed (Ekman Scherer, 1982). FACS analyses video or photographed facial events frame by frame to determine facial movement. A trained FACS coder can thoroughly identify and categorise each detail of an expression to reveal the specific AUs that produced the expression using FACS' descriptive power. The FACS codes for each universal facial expression are listed below. The * symbol represents the most reliable muscle movements, implying that they are the most difficult to flex voluntarily or to fake an expression of. Each number is the Action Unit (AU) code; these numbers represent muscle movements.

Action Unit		**Description**
AU 4		Brows are lowered and drawn together
AU 1+4		Inner parts of the brows are raised and drawn medially
AU 1+2		Entire brow is raised

Figure 8.3: Description of action units in the brow region

8.3 Principles of Facial Expression Analysis

Facial expression analysis includes both facial motion measurement and expression recognition. Face acquisition, facial data extraction and representation, and facial expression recognition are the three steps in the general approach to automatic facial expression analysis (AFEA) as shown in the figure 8.2. Face acquisition is a stage of processing that finds the face region in input images or sequences automatically. It can be a detector that detects faces in each frame or one that detects faces in the first frame and then tracks them throughout the video sequence. A facial expression analysis system can use a head finder, head tracking, and pose estimation to handle large amounts of head motion.

The next step is to extract and represent the facial changes caused by facial expressions after the face has been located. There are two approaches to facial feature extraction for expression analysis: geometric feature-based methods and appearance-based methods. Geometric facial features depict the shape and placement of facial components (including mouth, eyes, brows, nose, etc.). The facial components or feature points are extracted and combined to form a feature vector that represents the face geometry. Image filters, such as Gabor wavelets, are applied to either the entire face or specific regions in a face image to extract a feature vector in appearance-based methods. The effects of in-plane head rotation and different scales of the faces can vary depending on the different facial feature extraction methods be eliminated by face normalization before the feature extraction or by feature representation before the step of expression recognition.

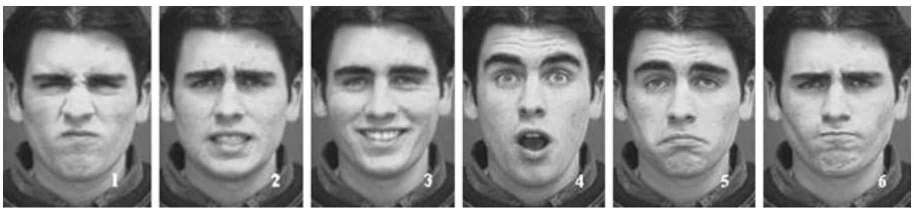


Figure 8.4: Emotion-specified facial expression

8.4 Problem Space for Facial Expression Analysis

8.4.1 Level of Description

1, disgust; 2, fear; 3, joy; 4, surprise; 5, sadness; 6, anger With few exceptions, most AFEA systems attempt to recognise a small set of prototypic emotional expressions (e.g., disgust, fear, joy, surprise, sadness, anger), as shown in Fig. 8.3. This practise may stem from Darwin's [9] work, as well as the work of Ekman, Friesen, and Izard et al., [10] who proposed that emotion-specific expressions have corresponding prototypic facial expressions. However, such prototypic expressions are relatively rare in everyday life. Instead, subtle changes in one or a few discrete facial features, such as tightening of the lips in anger or obliquely lowering the lip corners in sadness, are used to communicate emotion. Changes in isolated features, particularly the brows or eyelids, are typical of paralinguistic displays; for example, raising the brows signals greeting.

Automated recognition of fine-grained changes in facial expression is required to capture the subtlety of human emotion and paralinguistic communication. The facial action coding system is a system that uses human observers to detect subtle changes in facial features. Trained observers can manually FACS [12] code all possible facial displays, which are referred to as action units and can occur individually or in combinations, while viewing videotaped facial behaviour in slow motion. FACS is made up of 44 action units. Thirty are anatomically linked to the contraction of a particular set of facial muscles. The anatomic basis for the remaining 14 is unknown. FACS refers to these 14 as miscellaneous actions. Many action units can be classified as either symmetrical or asymmetrical. A 5-point scale is used for action units that vary in intensity.

8.4.2 Individual Differences in Subjects

Face shape, texture, colour, and facial and scalp hair differ depending on gender, ethnic background, and age [13]. For example, infants have smoother, less textured skin and frequently lack facial hair in the brows or scalp. The difference in eye opening and contrast between iris and sclera between Asians and Northern Europeans may affect the robustness of eye tracking and facial feature analysis in general. Face features may be obscured by beards, eyeglasses, or jewellery. Individual differences in appearance may have significant implications for face analysis. There have been few attempts to investigate their impact. A study by Zlochower et al [14]. found that algorithms for optical flow and high-gradient component detection optimised for young adults performed worse when used in infants.Â The reduced texture of infantsâ skin, their increased fatty tissue, juvenile facial conformation, and lack of transient furrows may all have contributed to the differences observed in face analysis between infants and adults.

Individual differences in expressiveness, in addition to differences in appearance, refer to the degree of facial plasticity, morphology, frequency of intense expression, and overall rate of expression. Individual differences in these characteristics are well established and play an important role in determining one's identity (these individual differences in expressiveness and in biases for particular facial actions are sufficiently strong that they may be used as a biometric to augment the accuracy of face recognition algorithms [19]). Individuals who have suffered facial nerve or central nervous system damage exhibit extreme variability in expressiveness. A large sample of people is required to develop algorithms that are resistant to individual differences in facial features and behaviour of varying ethnic background, age, and sex, which includes people who have facial hair and wear jewelry or eyeglasses and both normal and clinically impaired individuals.

8.4.3 Transitions Among Expressions

The assumption that expressions are singular and begin and end in a neutral position simplifies facial expression analysis. In reality, facial expression is more complex, particularly at the action unit level. Combinations of action units are possible, as is serial dependence. Transitions from one action unit or combination of actions to another may not involve a neutral state in between. A robust facial analysis system must be capable of parsing the stream of behaviour, and training data must include dynamic combinations of action units that can be additive or non-additive. Smiling is an example of an additive combination, as shown in the figure. (AU 12) with mouth open is coded as AU 12 + 25, AU 12 + 26 or AU system would have to detect transitions between all three levels of mouth opening while continuing to recognise AU 12, which could be changing in intensity at the same time. Non-additive combinations add to the complexity. Following its application in These interactions are known as co-articulation effects in speech science. A common example is the combination AU 12+15, which occurs frequently during embarrassment. Although AU

12 raises the cheeks and lip corners, the downward action of AU 15 modifies its action on the lip corners. The resulting change in appearance is highly dependent on timing. Lip corner downward action can occur simultaneously or sequentially. The latter appears to be the more common option. A database should include individual action units as well as

Figure 8.5: Some examples of combination of FACS action units

additive and nonadditive combinations, particularly those involving co articulation effects, to be comprehensive. A classifier trained solely on single action units may underperform in combinations with co-articulation effects.

8.4.4 Intensity of Facial Expression

The intensity of facial actions can vary. Manual FACS coding, for example, describes the intensity variation of action units using a 3- or, more recently, a 5-point intensity scale. Furthermore, some related action units function as sets to represent intensity variation. Action units 41, 42, 43, or 45 in the eye region can represent intensity variation from slightly drooped to closed eyes. Several computer vision researchers proposed methods for automatically representing intensity variation. Essa and Pentland [15] used optical flow to represent intensity variation in smiling. In addition, quantified intensity variation in emotion-specified expression and action units. However, these authors did not attempt the more difficult step of distinguishing intensity variation within types of facial actions.

Instead, intensity measures were used for the more limited purpose of distinguishing between different types of facial actions. The researchers compared manual and automatic intensity variation coding. They discriminated intensity variation in eye closure as reliably as human coders did using Gabor features and an artificial neural network. Bartlett and colleagues recently examined action unit intensity by analysing facial expression dynamics. They used a correlation analysis to explicitly measure the relationship between the learned classifiers' output margin and expression intensity. They used the output ranking score to estimate intensity. These findings imply that it is possible to recognise intensity variation within different types of facial actions automatically. Whether or not researchers attempt to distinguish intensity variation within facial actions, it is critical that theÂ range of variation be described adequately. Methods that work for intense expressions may generalize poorly to ones of low intensity.

8.4.5 Deliberate Versus Spontaneous Expression

The majority of face expression data has been gathered by having subjects perform a series of expressions. The appearance and timing of these directed facial action tasks may differ from spontaneously occurring behaviour. Separate motor pathways, the pyramidal and extrapyramidal motor tracks, mediate deliberate and spontaneous facial behaviour, respectively. As a result, fine-motor control of deliberate facial actions is frequently inferior and less symmetrical than what happens naturally. Many people, for example, can raise their outer brows spontaneously while leaving their inner brows at rest; few can do so voluntarily. Sadness is characterised by spontaneous depression of the lip corners (AU 15) and raising and narrowing of the inner corners of the brow (AU 1 + 4).

Few people can perform these actions deliberately without training, which helps with lie detection [30]. Differences in the temporal organisation of spontaneous and deliberate facial actions are especially significant because many pattern recognition methods, such as hidden Markov modelling, are highly dependent on the timing of the appearance change. Unless a database contains both deliberate and spontaneous facial actions, it will likely be insufficient for developing face expression methods that are resistant to these differences.

8.4.6 Head Orientation and Scene Complexity

The presence and actions of other people, as well as background conditions, may all have an impact on face analysis. Face orientation has received deliberate attention in the face recognition literature. The FERET database, [17] for example, includes both frontal and oblique views, and several specialised databases have been gathered to try to develop face recognition methods that are invariant to moderate changes in face orientation. Multiple perspectives are rarely used in the face expression literature, and pose invariance has received relatively little attention. [16] Most researchers believe that face orientation is only affected in-plane or that out-of-plane rotation is minor. In reality, large out-of-plane rotations in head position are common and frequently accompany expression changes.

In reality, large out-of-plane rotations in head position are common and frequently accompany expression changes. Kraut and Johnson discovered that smiling usually occurs when one turns toward another person. Researchers discovered that when an infant pitches her head back, she expresses surprise. Image data with facial expression changes in combination with significant nonplanar change in pose are required to develop pose invariant methods of face expression analysis. There have been some attempts to handle large out-of-plane rotations in head position. Scene complexity, such as the presence of other people and the presence of a background, may influence the accuracy of face detection, feature tracking, and expression recognition. Most databases rely on image data in which the background is neutral or has a consistent pattern and there is only one person in the scene. Multiple people interacting with each other are likely in natural environments, and their effects must be understood. It will be difficult to develop and test algorithms that are robust to such variation unless this variation is represented in training data.

8.4.7 Image Acquisition and Resolution

Several factors influence the image acquisition procedure, including the properties and number of video cameras and digitizers, the size of the face image in relation to the total image dimensions, and the ambient lighting. All of these variables may have an impact on facial expression analysis. Images taken in low light or at a low resolution may contain less information about facial features. Similarly, when the size of the face image is small in comparison to the total image size, less information is available. The implications of downsampling from this rate are unknown for NTSC cameras, which record images at 30 frames per second. Many optical flow algorithms assume that pixel displacement between adjacent frames is small. Unless they are tested at a variety of sampling rates, the robustness to sampling rate and resolution will be compromised cannot be assessed.

Changes in head position relative to the light source and variations in ambient lighting can have a significant impact on face expression analysis within an image sequence. A light source above the subject's head casts shadows below the brows, which can obscure the eyes, particularly in subjects with prominent bone structure or hair. When the angle of lighting changes across an image sequence, methods that work well in studio lighting may perform poorly in more natural lighting (e.g., through an exterior window). When a frontal orientation is not required, most investigators use single-camera setups, which is problematic. Out-of-plane rotation may be difficult to standardise when using image data from a single camera. Multiple cameras may be required for large out-of-plane rotations. Another issue to consider is image resolution. Professional-grade PAL cameras, for example, produce images with extremely high resolution. Security cameras, on the other hand, produce severely degraded images. Although postprocessing may improve image resolution, the amount of improvement is likely to be limited. Also unknown are the effects of post-processing on expression recognition. Table 19.4 depicts a face at various resolutions. For a 69 93 pixel image, most automated face processing tasks should be possible. Facial features such as the corners of the eyes and the mouth become difficult to detect at 48 64 pixels. Facial expressions can be recognised at 4864 pixels but not at 2432 pixels. Algorithms that perform well at optimal full-face frontal image resolutions and studio lightingÂ can be expected to perform poorly when recording conditions are degraded or images are compressed.

CHAPTER 9

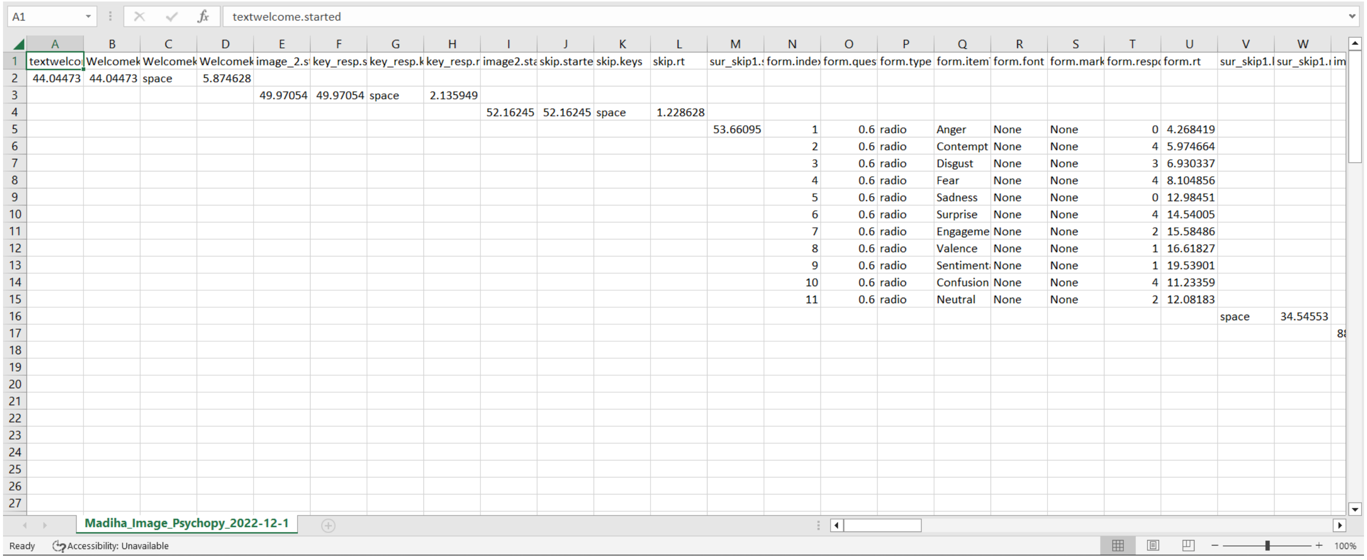
Result

The experiment was performed using multiple stimuli and as a result we've obtained a csv file with all the responses the participant has provided. Once we start the experiment, stimulus presentation starts and user is allowed to respon to the question in the survey.

Figure 9.1: Stimuli Presentation in Psychopy

The following results were obtained by performing the experiment on above image mentioned. The columns include time when we skipped the image, fixation time of the stimuli, various emotions and range the participant entered as a response. This data as a csv file is downloaded automatically once the experiment is finished. This is one of the features in psychopy which makes it easier to use the experiments for data analysis.

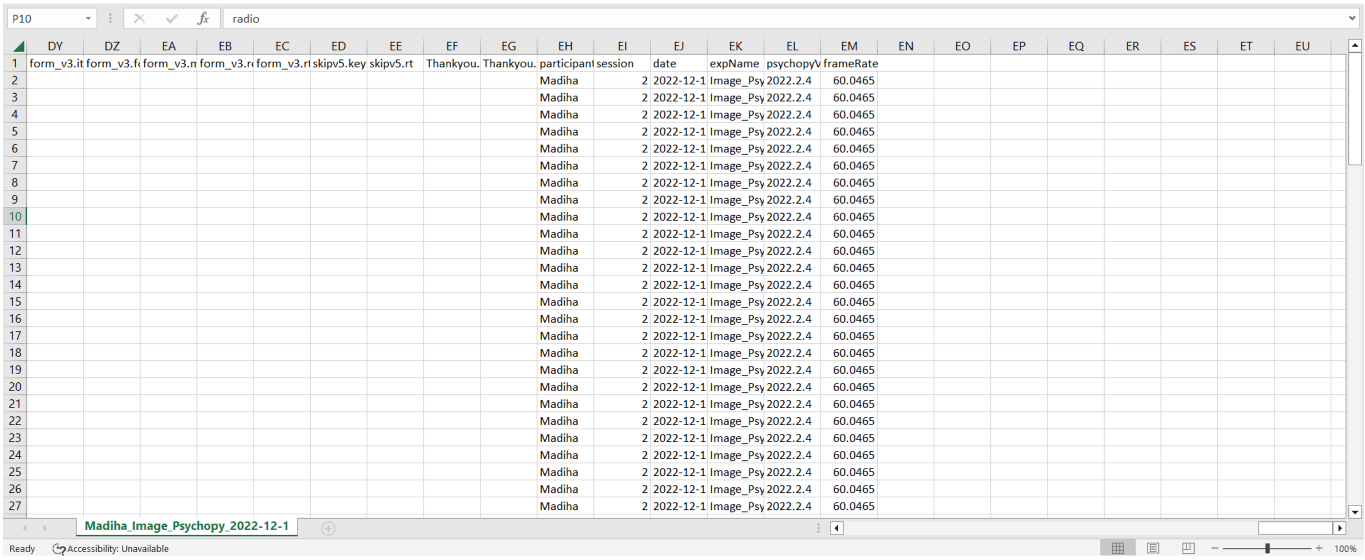
Figure 9.3 includes details of the participant and the session in which the experiment is



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	textwelco	Welcomek	Welcomek	Welcomek	image_2.s	key_resp.s	key_resp.k	key_resp.r	image2.sta	skip.starte	skip.keys	skip.rt	sur_skip1.s	form.index	form.ques	form.type	form.item	form.font	form.mark	form.respo	form.rt	sur_skip1.l	sur_skip1.r
2	44.04473	44.04473	space	5.874628																			
3					49.97054	49.97054	space	2.135949															
4									52.16245	52.16245	space	1.228628											
5													53.66095	1	0.6	radio	Anger	None	None	0	4.268419		
6														2	0.6	radio	Contempt	None	None	4	5.974664		
7														3	0.6	radio	Disgust	None	None	3	6.930337		
8														4	0.6	radio	Fear	None	None	4	8.104856		
9														5	0.6	radio	Sadness	None	None	0	12.98451		
10														6	0.6	radio	Surprise	None	None	4	14.54005		
11														7	0.6	radio	Engageme	None	None	2	15.58486		
12														8	0.6	radio	Valence	None	None	1	16.61827		
13														9	0.6	radio	Sentiment	None	None	1	19.53901		
14														10	0.6	radio	Confusion	None	None	4	11.23359		
15														11	0.6	radio	Neutral	None	None	2	12.08183		
16																						space	34.54553

Figure 9.2: Response data exported

performed. This helps us segregate the data based on person to person in order to analyse the different emotions of the participants on different stimuli.



	DY	DZ	EA	EB	EC	ED	EE	EF	EG	EH	EI	EJ	EK	EL	EM
1	form_v3.it	form_v3.f	form_v3.n	form_v3.r	form_v3.rt	skipv5.key	skipv5.rt	Thankyou.	Thankyou.	participant	session	date	expName	psychopyV	frameRate
2										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
3										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
4										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
5										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
6										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
7										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
8										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
9										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
10										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
11										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
12										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
13										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
14										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
15										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
16										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
17										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
18										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
19										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
20										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
21										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
22										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
23										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
24										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
25										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
26										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465
27										Madiha	2	2022-12-1	Image_Psy	2022.2.4	60.0465

Figure 9.3: Participant data

CHAPTER 10

Conclusion

Finally, I'd like to point out that the application's structure greatly assists us in simplifying the process. Everything is in its proper place, so that participants can easily understand what to choose and how the flow works. In the Emotions Application, we've removed the video experiment option and browser experiment option to make it purely an image experiment for this thesis. I've also added and export data option from which we can export the saved responses from the survey window to a csv file which can be later used in the data analysis.

The psychopy application works perfectly to display stimulus and take survey responses. I learnt so mauch doing this project, but there is still work need to be done which is to communicate with the iMotions API to connect the external software to the iMotions so that we get customised stimulus display for the participants to perform experiments.

Bibliography

[1] Zahid, Zaki and Shaukat, Arslan and Khan, Rizwan Ahmed and Akram, Usman and Byun, Yung-Cheo *2019 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC Asia-Pacific)*, title=Emotion Recognition in Video Clips Using Simple and Extended Center Symmetric LBP, year=2019 pages=1-6, doi=10.1109/ITEC-AP.2019.8903878

[2] Yingli Tian, Takeo Kanade, and Jeffrey F. Cohn *Facial Expression recognition, published on 26 January 2017.*

[3] Y. Wang L. Guan, *"Recognizing Human Emotional State From Audiovisual Signals*" in IEEE Transactions on Multimedia, vol. 10, no. 5, pp. 936-946, Aug. 2008, doi: 10.1109/TMM.2008.927665.*

[4] Johns Hopkins Medicine, Health *Neurological Rehabilitation*

[5] Lin X, Zhang X, Liu Q, Zhao P, Zhang H, Wang H, Yi Z. *Facial emotion recognition in adult with traumatic brain injury: A protocol for systematic review and meta-analysis. Medicine (Baltimore). 2020 Jul 17;99(29):e21154. doi: 10.1097/MD.0000000000021154. PMID: 32702870; PMCID: PMC7373508.*

[6] Byom L, Duff M, Mutlu B, et al. *Facial emotion recognition of older adults with traumatic brain injury.*

[7] Wearne TA, Osborne-Crowley K, Logan JA, et al. *Understanding how others feel: evaluating the relationship between empathy and various aspects of emotion recognition following severe traumatic brain injury.*

[8] Cohn, J., Zlochower, A., Lien, J., Kanade, T.: *Automated face analysis by feature point tracking has high concurrent validity with manual facs coding. Psychophysiology 36, 35â 43 (1999)*

[9] Darwin, C.: *The Expression of Emotions in Man and Animals. Murray, London (1872), reprinted by University of Chicago Press, 1965*

[10] Izard, C., Dougherty, L., Hembree, E.A.: *A system for identifying affect expressions by holistic judgments. Unpublished Manuscript, University of Delaware (1983)*

[11] Lien1998AutomatedFE, Jenn-Jier James Lien and Takeo Kanade and Jeffrey F. Cohn and Ching-Chung Li *Automated facial expression recognition based on FACS action units journal=Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition year=1998, pages=390-395*

[12] Ekman, P., Friesen,W.: *The Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, San Francisco (1978)*

[13] Farkas, L., Munro, I.: *Anthropometric Facial Proportions in Medicine. Charles C Thomas, Springfield (1987)*

[14] Zlochower, A., Cohn, J., Lien, J., Kanade, T.: *A computer vision based method of facial expression analysis in parent-infant interaction. In: International Conference on Infant Studies, Atlanta (1998)*

[15] Essa, I., Pentland, A.: *Coding, analysis, interpretation, and recognition of facial expressions. IEEE Trans. Pattern Anal. Mach. Intell. 19(7), 757â763 (1997)*

[16] Rizvi, S., Phillips, P., Moon, H.: *The Feret verification testing protocol for face recognition algorithms. In: Proceedings of the Third International Conference on Automatic Face and Gesture Recognition, pp. 48â55 (1998)*

[17] Vetter, T.: *Learning novels views to a single face image. In: Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, pp. 22â29 (1995)*

[18] EEG, ElectroencephalographyâiMotions Software and EEG Headsets. *[(accessed on 29 October 2019)]; Available online*

[19] GarcÃa, Imizcoz M, Perea MV, Soto F, Smith SJ, Calvo Simal S *The Efficacy of Emotion Recognition Rehabilitation for People with Alzheimer's Disease. J Alzheimers*

[20] Salimpoor, V.N., Benovoy, M., Longo, G., Cooperstock, J.R. Zatorre, R.J. (2009). *The rewarding aspects of music listening are related to degree of emotional arousal. PLoS ONE 4, e7487*

[21] M. Crawshaw, *Multi-Task Learning with Deep Neural Networks: A, sept 2020.*

[22] R. Caruana, *âMultitask Learning,â Machine Learning , p. 9, 1997.*

[23] IBM Learning hub *IBM Cloud Education 3 June 2020*

[24] Huenerfauth and G. Rossun, *Introduction of python*