

Article

Asian Affective and Emotional State (A2ES) Dataset of ECG and PPG for Affective Computing Research

Nor Azlina Ab. Aziz ^{1,*}, Tawsif K. ¹, Sharifah Noor Masidayu Sayed Ismail ², Muhammad Anas Hasnul ¹ , Kamarulzaman Ab. Aziz ³ , Siti Zainab Ibrahim ⁴, Azlan Abd. Aziz ¹ and J. Emerson Raja ¹ 

¹ Faculty of Engineering & Technology, Multimedia University, Bukit Beruang 75450, Melaka, Malaysia

² Faculty of Information Science & Technology, Multimedia University, Bukit Beruang 75450, Melaka, Malaysia

³ Faculty of Business, Multimedia University, Bukit Beruang 75450, Melaka, Malaysia

⁴ School of Computing and Informatics, Albukhary International University, Jalan Tun Abdul Razak, Alor Setar 05200, Kedah, Malaysia

* Correspondence: azlina.aziz@mmu.edu.my

Abstract: Affective computing focuses on instilling emotion awareness in machines. This area has attracted many researchers globally. However, the lack of an affective database based on physiological signals from the Asian continent has been reported. This is an important issue for ensuring inclusiveness and avoiding bias in this field. This paper introduces an emotion recognition database, the Asian Affective and Emotional State (A2ES) dataset, for affective computing research. The database comprises electrocardiogram (ECG) and photoplethysmography (PPG) recordings from 47 Asian participants of various ethnicities. The subjects were exposed to 25 carefully selected audio–visual stimuli to elicit specific targeted emotions. An analysis of the participants’ self-assessment and a list of the 25 stimuli utilised are also presented in this work. Emotion recognition systems are built using ECG and PPG data; five machine learning algorithms: support vector machine (SVM), k-nearest neighbour (KNN), naive Bayes (NB), decision tree (DT), and random forest (RF); and deep learning techniques. The performance of the systems built are presented and compared. The SVM was found to be the best learning algorithm for the ECG data, while RF was the best for the PPG data. The proposed database is available to other researchers.

Keywords: affective computing; emotion recognition system; physiological signals



Citation: Ab. Aziz, N.A.; K., T.; Ismail, S.N.M.S.; Hasnul, M.A.; Ab. Aziz, K.; Ibrahim, S.Z.; Abd. Aziz, A.; Raja, J.E. Asian Affective and Emotional State (A2ES) Dataset of ECG and PPG for Affective Computing Research. *Algorithms* **2023**, *16*, 130. <https://doi.org/10.3390/a16030130>

Academic Editor: Fuyi Li

Received: 19 January 2023

Revised: 15 February 2023

Accepted: 20 February 2023

Published: 27 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As the World Health Organization’s Director-General, Tedros Adhanom Ghebreyesus remarked, in 2020, that mental health is essential for overall health and well-being [1]. The outbreak of the COVID-19 pandemic brought new challenges to the issue of mental health. According to the Kaiser Family Foundation’s investigation into the effect of COVID-19 on American life, the respondents were concerned about losing income due to the fact of job loss, workplace closure, or reduced job hours during the pandemic [2]. Six out of ten adults were concerned about getting an infection or exposing themselves or their family to the virus while working. All of these concerns have negative effects on mental health and emotions. Additionally, according to a survey conducted by Changwon Son’s team [3], 71% of students in the United States claimed that their anxiety and stress levels increased as a result of the pandemic. A report from the University of Saskatchewan, Canada [4], focusing on the university’s medical students, also showed a similar result. These findings demonstrate the seriousness of the COVID-19 pandemic’s impacts on mental health. Therefore, in this challenging era of COVID-19, research on intelligent systems that monitor for symptoms of unpleasant emotions building up in a person is becoming more pressing.

An emotion recognition system (ERS) can recognise human emotions and can be used in many different fields. For example, a stress detector that assesses employees’ stress

levels using electrocardiogram (ECG) and galvanic skin response (GSR) is proposed in [5]. A cardiac-based ERS is proposed in [6–8] to assess driver stress levels and drowsiness detection in [9]. Additionally, ERS have been proposed for various uses in the education industry. In [10], voice-based emotion identification for affective e-learning is proposed. A facial ERS that enables teachers to monitor students' moods throughout class [11] and physiological signal-based ERS adoption in an intelligent tutoring system (ITS) [12] are also found among the works that reported the usage of ERS for education.

From the works discussed above, it can be observed that an ERS can be built using multiple modalities: ECG, GSR, and voice and facial images. Notably, physiological signals are commonly used. Among the physiological signals that are often utilised as ERS modalities are electroencephalogram (EEG) [13,14] and ECG [15–18]. Some works integrate several modalities for their ERS [19–21], while others use a single modality [17,22,23]. Due to the high demand, the number of works on physiological-based ERS utilising wearable devices and noninvasive sensors has also increased. Physiological-based ERS are good for social masking avoidance [24] and are less prone to fake emotions and manipulation [25]. The utilisation of wearable devices is supported by the popularity of their usage among consumers. Rock Health surveyed digital health adoption and discovered that wearable device usage has increased significantly, from 24% in 2018 to 33% in 2019 [26]. Additionally, Statista, a German-based online statistics source predicted that the number of smartwatch users is expected to reach 1.2 million by 2024 [27]. According to this statistic, the endeavour to build ERS using wearable devices represents a path towards a proper future with significant advancements.

Many labelled emotion databases have been produced in recent years that comprise various modalities [28], for example, a database for emotion analysis using physiological signals (DEAP) [13]; a database for affect, personality, and mood research on individuals and groups (AMIGOS) [18]; a database for decoding affective physiological responses (DECAF) [20]; and a multimodal physiological emotion database for discrete emotion recognition (MPED) [16]. Several databases are composed of data collected using nonportable devices and expensive technology; meanwhile, the databases for emotion recognition through EEG and ECG (DREAMER) [14], wearable stress and affect detection (WESAD) [29], and emotion recognition smartwatches (ERS) [30] are compilations of signals collected from wireless, low-cost, and off-the-shelf devices. These databases have been utilised in studies by researchers with different levels of success [13,18,20].

In past research, the issues of racial inequities and bias toward wearable technology, particularly for those with darker skin tones, have been raised [31,32]. Those with darker skin tones, tattoos, or arm hair have lower accuracy when using wearable devices that track their health activity or monitor their heart conditions. Noseworthy et al. [32] recommended that researchers should be aware of racial bias and disseminate study results across demographic subgroups to minimise bias. To the best of our knowledge, there are no existing physiological affective datasets collected from wearable devices that look at this issue and include multi-Asian ethnicities. For example, the DEAP dataset comprises data from European participants [13], while the MPED dataset consists of data from Chinese participants only.

Thus, this paper introduces the Asian Affective and Emotional State (A2ES) Database consisting of ECG and PPG recordings of 47 participants from various Asian ethnicities. Both ECG and PPG recordings have been reported to be affected by skin colour [31,33]. An ECG is used to detect the heart's electrical activity, which starts from the sinoatrial node to contract the heart muscles for continuing the blood pumping action in the body [34]. As illustrated in Figure 1, the ECG comprises three primary components: P wave, QRS wave, and T wave. On the other hand, PPG is a low-cost and noninvasive way to measure blood volume changes in a human during heart activity. PPG has two main components: incoherent light source and photoreceiver [35]. A typical PPG signal element is shown in Figure 2, complete with the systolic period associated with blood in-rush, the diastolic period associated with relaxation, and the dicrotic notch associated with pulse reflection [36].

The subjects that participated in the data collection were exposed to 25 audio–visual stimuli to elicit specific emotions. The self-assessment ratings from the participants and the list of the 25 stimuli are also presented here.

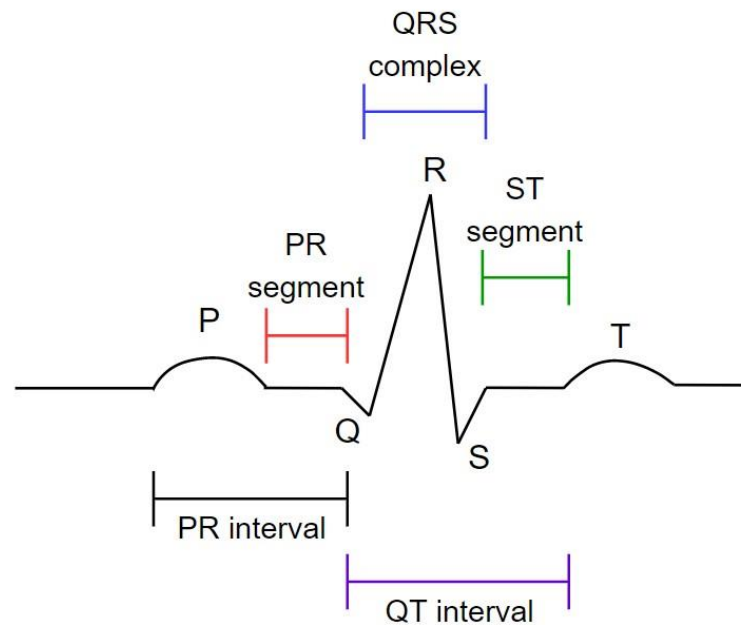


Figure 1. The P, QRS, and T waves in a single cycle of a standard ECG reading.

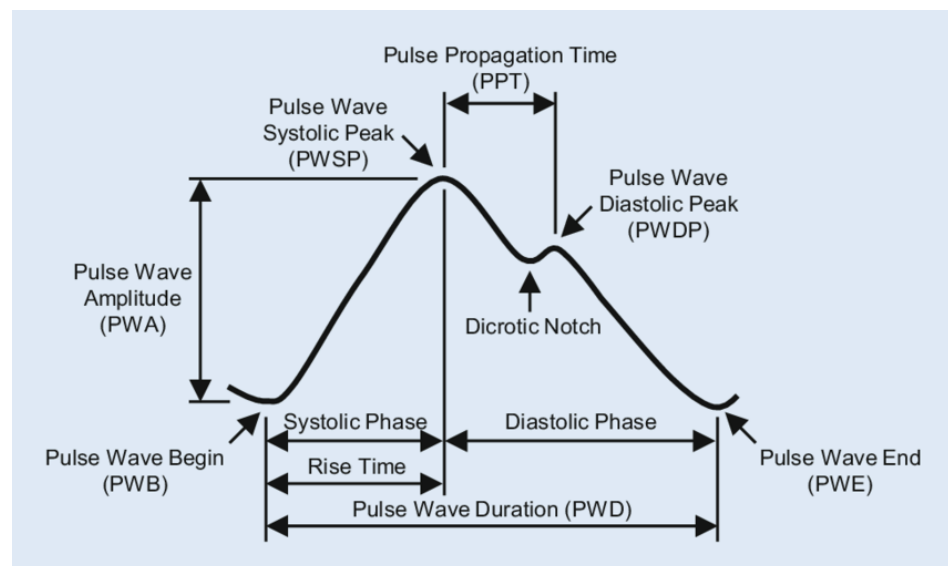


Figure 2. Typical transmission of a PPG signal with the systolic period. Reproduced with permission from [37].

The applicability of the A2ES's ECG and PPG data for building an ERS was tested using machine learning and deep learning approaches. Five machine learning algorithms, namely, support vector machine (SVM), naive Bayes (NB), K-nearest neighbours (KNN), decision tree classifier (DT), and random forest (RF), were applied. The ECG-based ERS built using SVM and the PPG-based ERS built using RF were found to be the best. The small data size did not suit deep learning, and poor performances were reported.

The rest of this paper is organised as follows. In Section 2, related works, including ECG- and PPG-based ERS, as well as ECG- and PPG-based databases, are described. The

experiment protocol is covered in Section 3, which includes the stimuli selection procedure, participants' details, and data collection setting and protocol. Section 4 describes the data preprocessing and feature extraction process. In Section 5, an evaluation of the ECG- and PPG-based ERS performances are presented. A concluding discussion and future work directions are provided in Section 6.

2. Related Works

ECG and PPG are popular modalities for ERS. Many studies using these modalities have achieved promising results in representing human emotions. Bagirathan et al. [22] utilised ECG signals to recognise positive and negative valence states in children with autism spectrum disorder (ASD). The proposed system successfully obtained an accuracy of 81%. Meanwhile, a PPG-based ERS with a convolutional neural network (CNN) is proposed in [38] for the fast emotional recognition of valence and arousal. The system achieved a 75.3% and 76.2% valence and arousal accuracy within 1.1 s for short-term emotion recognition. In 2021, Preethi et al. developed a real-time ERS to automate a music selection system using emotion recognized based on PPG signals [39]. An accuracy of 91.81% was achieved utilising features extracted from phase-space geometry (Poincare's analysis). For binary classification and multiclass classification, maximum accuracy rates of 96.67% and 91.11% were achieved, respectively. Hasnul et al. evaluated the performance of an ECG-based ERS with the features extracted using two distinct feature extraction toolboxes, TEAP and AUBT, and achieved an accuracy of up to 65% [17].

ECG and PPG are also commonly integrated with other physiological signals as a strategy to improve ERS performance. In [40], ECG was used together with temperature (TEMP), galvanic skin response (GSR), electromyography (EMG), respiration (RESP), accelerometer signals, and facial expressions to recognise dimensional emotional states (high arousal and high valence (HAHV), high arousal and low valence (HALV), low arousal and high valence (LAHV), and low arousal and low valence (LALV)), arousal, and valence. The accuracy obtained was in the range of 40 to 70%. Zainudin et al. [41] proposed stress detection using ECG and GSR signals and categorised them using two approaches: machine learning and deep learning. Their work successfully achieved the best accuracy of 95%. Tian Chen et al. proposed a multimodal fusion ERS that includes EEG and ECG [42]. The fusion ERS was better than the single-modality ERS, with an accuracy for valence of 85.38% and for arousal of 77.52%.

In [43], another emotion-based music recommendation engine system was built using a combination of PPG and GSR signals from wearables. The emotional information from PPG and GSR was fed to a collaborative and content-based recommendation engine, and the best accuracy rate obtained exceeded 70%. Domínguez-Jiménez et al. [44] also proposed an ERS using PPG and GSR from wearable devices. The ERS recognises three emotions: amusement, sadness, and neutral. The system successfully recognised all three emotions, with a testing accuracy of up to 100%. In [45], a deep physiological affect network, which is a robust physiological model that recognises human emotions using PPG and EEG signals, is presented. The proposed system achieved 78.72% and 79.03% overall accuracy for recognising valence and arousal emotions, respectively.

Although both ECG and PPG signals can be used independently or integrated with other physiological signals, they can also be fused together to upgrade the robustness and improve an ERS's performance. For example, Li et al. [46] proposed a group-based individual response specificity (IRS) to improve the emotion recognition performance by fusing the statistical features from ECG and PPG with GSR. The highest performance achieved was 78.06% using the MLP classifier. The authors of [47] also proposed an automatic ERS with the fusion of the ECG and PPG features and successfully achieved the best performance of 85.70%. Additionally, the fusion of the ECG and PPG features was also used in [48]. They classified three emotions, positive, neutral, and negative, using a CNN and achieved an accuracy of 75.40%.

In affective computing, existing datasets that collect data from either a single modality or a multimodality using physiological and physical signals are important for the advancement of this field. Existing datasets and their size (number of participants and number of stimuli), type of stimuli, modalities used, devices, and labels are tabulated in Table 1. Most of the listed datasets contain ECG signals. The ECGs were collected using various devices, namely, Shimmer, Biosemi Active System, Biopac System, FlexComp, Procomp Infinity, and Mobi. Only two of the datasets contain PPG only without ECG: DEAP [13] and DEAR-MULSEMEDIA [49]. In these works, the PPG signals were recorded using Biosemi ActiveTwo and Shimmer devices. Four works have both cardio-based physiological signals (ECG and PPG): CASE [50], CLAS [51], ECSMP [52], and K-EmoCon [19]. The datasets CASE and CLAS contain ECG and PPG signals measured using Thought Technology and Shimmer3, respectively. ECSMP and K-EmoCon used AECG-100 and Polar H7 for ECG and Empatica E4 for PPG. The ECSMP [52] dataset has the greatest number of subjects, and EMDC [53] has the greatest number of stimuli compared to the other datasets. Twelve datasets used audio–visual stimuli, making it the most common type of stimuli to elicit emotions. Additionally, most of the datasets used valence and arousal as emotion annotations in addition to basic emotions, such as joy, anger, sadness, fear, disgust, stress, or neutral. A review paper [25] discusses in detail most of these datasets.

Table 1. Summary of the existing cardiological-based ERS datasets.

Dataset	Data	Stimuli Used	Modalities Used	Cardiac-Based Devices	Emotion Label
AMIGOS [18]	40 subjects, 24 stimuli	Audio–visual	ECG, EEG, GSR	Shimmer	Valence, arousal, dominance
ASCERTAIN [54]	58 subjects, 36 stimuli	Audio–visual	ECG, EEG, GSR	NA	Valence, arousal
AuBT [55]	1 subject, 4 stimuli	Audio	ECG, EMG, RESP, GSR	NA	Joy, anger, sadness, pleasure
CASE [50]	30 subject, 20 stimuli	Audio–visual	ECG, PPG, EMG, GSR	Thought Technology	Valence, arousal
CLAS [51]	62 subjects, 32 stimuli	Audio–visual, visual	ECG, PPG, GSR	Shimmer3	Valence, arousal
DEAP [13]	32 subjects, 40 stimuli	Audio–visual	PPG, EEG, GSR, EOG	Biosemi ActiveTwo	Valence, arousal, liking
DEAR-MULSEMEDIA [49]	18 subjects, 4 stimuli	Audio–visual, tactile, olfaction, haptic	PPG, EEG, GSR	Shimmer	Valence, arousal
DECAF [20]	30 subjects, 76 stimuli	Audio–visual	ECG, EOG, EMG, MEG, facial expression	NA	Valence, arousal, dominance
DREAMER [14]	23 subjects, 18 stimuli	Audio–visual	ECG, EEG	Shimmer	Valence, arousal, dominance
DSDRWDT [56]	24 subjects	Driving task	ECG, EMG, GSR, RESP	FlexComp	Stress
ECSMP [52]	89 subjects, 6 stimuli	Audio–visual, cognitive assessment task	ECG, PPG, EEG, GSR, TEMP, ACC	AECG-100A (ECG), Empatica E4 (PPG)	Neutral, fear, sad, happy, anger, disgust, fatigue

Table 1. Cont.

Dataset	Data	Stimuli Used	Modalities Used	Cardiac-Based Devices	Emotion Label
EMDC [53]	3 subjects, 360 stimuli	Audio	ECG, EMG, GSR, RESP	Procomp ² Infiniti	Valence, arousal
K-EmoCon [19]	32 subjects	Naturalistic conversation	ECG, PPG, EEG, GSR, TEMP	Polar H7 (ECG), Empatica E4 (PPG)	Valence, arousal
MAHNOB-HCI [57]	27 subjects, 20 stimuli	Audio–visual	ECG, EEG, GSR, RESP, TEMP	Biosemi ActiveTwo	Valence, arousal, dominance
MPED [16]	23 subjects, 28 stimuli	Audio–visual	ECG, EEG, GSR, RESP	Biopac System	Joy, funny, anger, fear, disgust, sad, neutral
RECOLA [58]	46 subjects	Spontaneous and naturalistic interactions	ECG, GSR, voice, facial expression	Biopac MP36	Valence, arousal
SWELL [5]	25 subjects, 4 working conditions with stressors	Writing, presenting, reading, searching task	ECG, GSR, facial expression, body posture	Mobi (TMSi)	Valence, arousal, stress
WESAD [29]	15 subjects, 10 stimuli	Audio–visual, public speaking, mental arithmetic task	ECG, PPG, GSR, EMG, TEMP, RESP	RespiBan Professional ² (ECG), Empatica E4 (PPG)	Neutral, stress, amusement

3. Data Collection Protocol

3.1. Emotion Annotation and Stimuli Selection

The labelling and annotating of the A2ES data were conducted based on the discrete emotional model (DEM), also known as basic emotions [16]. The seven selected basic emotions were happy, sad, anger, fear, disgust, surprise, and neutral. The self-labelling process was conducted directly after the subjects watched each video. A self-assessment form, as shown in Figure 3, was prepared and a brief description for the user on how to comply was written for the first video assessment. The subjects were encouraged to be truthful concerning their introspective emotions instead of thinking of what is expected from the videos. The first part identified the emotion experienced from watching the video, and the second part requested the subjects to rate the intensity of the emerged feeling on a scale, where one was the lowest and five was the highest. The combination of these two parts allowed us to map the emotion to the valence and arousal scale.

The experiment was designed to collect data with an equal distribution of the six emotions to promote variation and reduce bias. Prior to the data collection, a pilot study on stimuli selection with respect to the targeted emotion was conducted. The findings of the pilot study are presented in [59]. Based on the outcome of the pilot study, the stimuli selection was refined. The stimuli were suited to the targeted participant's background. All videos were procured from YouTube, with a duration ranging from one to five minutes. The total duration of all of the video clips was 1 h 15 min. The subjects were presented with one neutral video before and after three consecutive videos with a similar targeted emotion. The targeted emotional sequence was happy > surprise > fear > disgust > sad > anger, with an interlude of a neutral video. The list of the 25 selected videos used to elicit an emotional response for the data collection can be found in [60].

Section 5 of 30

Video 1

Direction: Please complete the questions after you watch each video. You were encouraged to rate the video according to what you had actually felt at the specific time you were watching the video and not what you believed people should feel in reaction to the videos or your general mood of the day.

Based on Video #1, what did you feel when you were watching the video? *

☐ Happy
☐ Sad
☐ Anger
☐ Fear
☐ Disgust
☐ Surprise
☐ Neutral

Based on your choice above, please rate your feelings to the stimulus on the following scales. *

Lowest 1 2 3 4 5 Highest
☐ ☐ ☐ ☐ ☐

Figure 3. The self-assessment form prepared in Google Forms.

The ECG signals were recorded using a KardiaMobile (KM) device. KM is a one-lead ECG device by AliveCor. It works by placing two fingers of each hand on the electrodes (Figure 4). It can capture 30, 60, or 300 s of raw and filtered ECG data and transfer them to the connected smartphone via an ultrasonic audio-based wireless communication protocol. Several studies have assessed and validated the KM device and its algorithm. The sensitivity and specificity obtained varied between 55–100% and 84–99%, respectively, depending on the patient population and reference technique [61–66]. The PPG was collected using a Maxim Band. It is a wrist-worn activity and heart-rate monitor that makes use of a maxim analogue front-end (AFE), accelerometer, optical sensors, and an internal algorithm. The band is shown in Figure 5. When collecting the signals from the ECGs, the subject was instructed to start the recording only when they began to feel the emotion. This was because an ECG has a limitation of one minute as a maximum length. As for PPG, the signals were recorded continuously while the video played.



Figure 4. AliveCor Kardia Mobile device and its application.



Figure 5. Maxim PPG Band.

Figure 6 shows the percentage of the data distribution between the targeted emotion and real samples. The targeted happy, sad, anger, fear, disgust, and surprise data were equally distributed at 12%. Neutral was considered as the absence of any particular emotion, and the targeted sample size percentage was larger at 28%. The real samples based on the participants' feedback had a size for neutral of 27%. In the real data distribution, happiness had the highest sample compared to the other five discrete emotions, with an extra 8%, while anger had the least, with only 8% of the total data. The reason for such an imbalance is that some emotions, such as anger and sadness, are typically harder to trigger by only watching videos. Another reason for the slight imbalance is because different persons have different perspectives when dealing with stimuli.

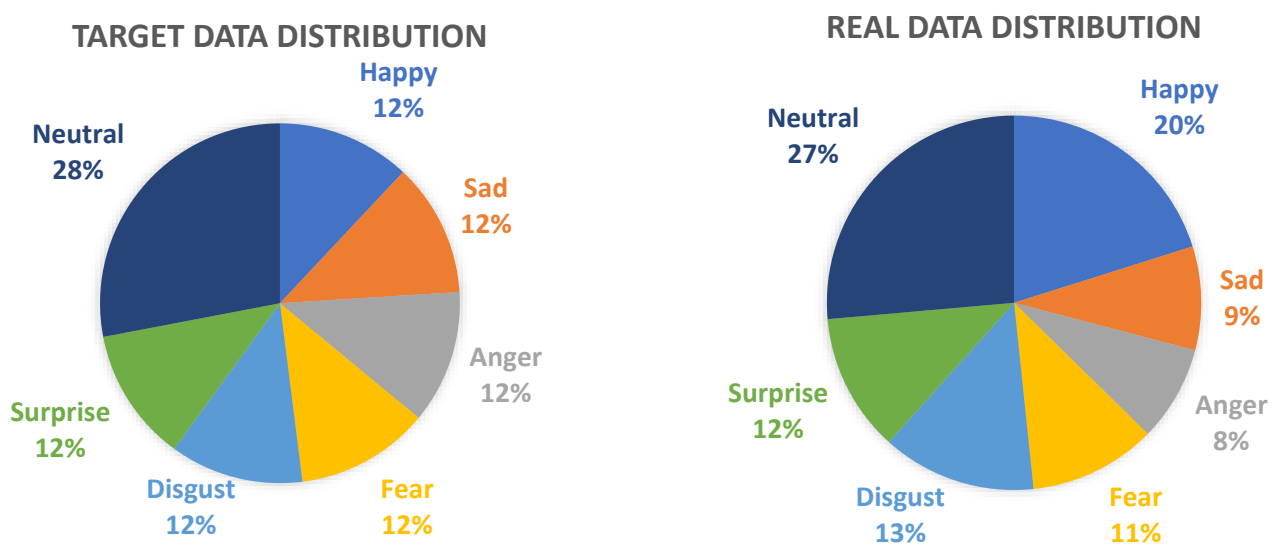


Figure 6. Data distribution of the emotion labels.

Table 2 shows the number of the subjects' responses towards each video. As can be observed, perspective differences existed among the participants. Additionally, some participants experienced contradictory emotions that were opposed to the targeted emotion. Video 4, for example, failed to induce the targeted emotion of happiness in the majority of the participants. The video concerns the anticipation of a happy ending in a well-known fictional movie; the popularity of this movie, as well as the suspense element of the clip, contributed to a greater number of participants selecting neutral and surprise. In another case involving videos 18 and 20, they were supposed to conjure sadness in the viewers, and the majority of the subjects reported feeling happy instead. Video 18 portrays a pitiful situation of a young man who encounters difficulties, but the act of kindness shown by the strangers towards him may have caused the viewers to feel happy instead of sad. Video 20

shows a collection of heart-warming father–daughter relationships in the context of the latter’s wedding day. Although the scene is touching and sad, most viewers perceived it as happiness, because the situations take place on a wedding day, which is typically perceived as a happy occasion. In occasions where the targeted emotion was different than the participants’ feedback, this study considered the individual self-assessments as the data labels.

Meanwhile, disgust had the most similar majority votes, with all three of the videos managing to obtain more than 40 subjects (out of 48) experiencing the targeted emotion. Video 15 recorded the highest number of participants experiencing the targeted disgust emotion (46), and only two participants experienced other emotions, with one participant reporting being not affected by the video (neutral), while another participant felt fear after watching the video. This video depicted a girl eating frogs. For the surprise, fear, and anger videos, the majority of the votes in the self-assessments were similar to the targeted emotions. Videos 1, 5, 9, 13, 17, 21, and 25 were considered neutral, and they were played between other targeted emotion videos to regulate the participants’ emotions and lower the intensity of the emotion felt from the previous videos. Hence, some of the subjects might have experienced a residual emotion from the previous videos, causing them to label the presence of emotions instead of neutrality. Meanwhile, after a sequence of unpleasant videos, such as anger-inducing videos, a neutral video might provide a pleasant and happy feeling in the participants, where 15 reported feeling happy for video 25. The deviance of the reported emotion by the participants and the targeted emotion was not a major problem, as the objective of the experiment was to record the ECG and PPG signals of the subjects when they experienced a specific emotion. Nonetheless, this contributed to an imbalance of the real data distribution in contrast to the targeted distribution, as shown in Figure 6. In the second part of the self-assessment form, the intensity of an emotion was captured. This was used to analyse whether the intensity of an emotion was low or high when experienced by the participants after exposure to the stimulus. The intensity of the emotions experienced by the 48 participants when watching the videos are shown in Figure 7. The three videos targeting the same emotion are clustered together in the figure. The blue with a darker shade on the left side shows that the subjects selected a low intensity for the felt emotion, while the blue with a brighter shade on the right side shows the opposite. Video 15 managed to elicit the highest intensity of emotion (disgust), as it shows an act of eating foods that are beyond social norms. The second highest intensity count was for video 16, also in the disgust category, where rotten foods are shown in the clip. Video 8 also had a high count of subjects rating it as high intensity. The surprising factor that contributed to such a degree of scale came from the extraordinary human capacity to solve difficult tasks. A low intensity of an emotion towards the stimulus might be because the subjects were already familiar with the videos shown. No subject rated videos 2, 19, 20, and 23 as one in the scale.

Table 2. Number of subjects that labelled the videos according to the discrete emotional model.

		Videos (Number of Subjects That Chose the Labelled Emotion)																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Labelled Emotion	Happy	6	37	29	4	7	7	2	15	14	2	1		9					38	6	33	15	1			15
	Sad		5	1	2					3	1	5	17	2					7	37	10	2	2	5	8	1
	Anger				1													1					36	24	37	
	Fear	2		2	9	2	9	2		1	37	37	24		2	1		1		2			1			
	Disgust				1	1	2	1			5	2	1	1	42	46	41						2	14	1	
	Surprise		4	7	17	1	27	35	29		2	3	2		1		6	4		2	1				3	
Neutral		40	2	9	14	37	3	8	4	30	1		4	36	3	1	1	42	3	1	4	31	6	2	2	32
Targeted Emotion		Neutral	Happy	Happy	Happy	Neutral	Surprise	Surprise	Surprise	Neutral	Fear	Fear	Fear	Neutral	Disgust	Disgust	Disgust	Neutral	Sad	Sad	Sad	Neutral	Anger	Anger	Anger	Neutral
Target Reached?		Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes

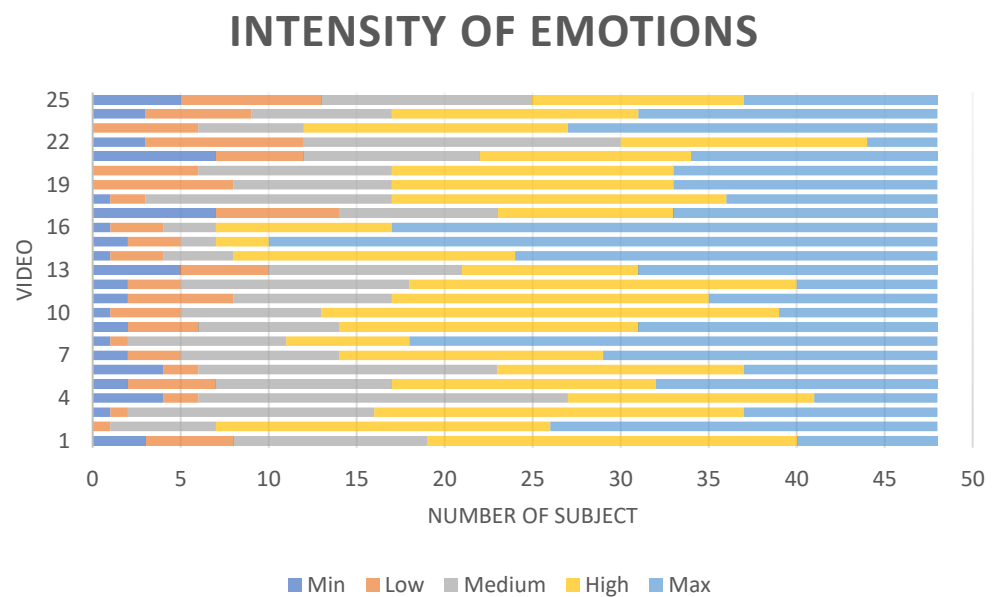


Figure 7. Intensity of the emotions rated by the subjects for each video, except neutral.

3.2. Participants

A total of 47 people took part in this study. Participation was on a volunteer basis. Each session started with the participants filling out a self-report of any psychological problems and cardiovascular disease. Additionally, since the data collection was conducted during the height of the COVID-19 pandemic, the participants were asked about COVID-19 symptoms. The data collection session for a participant proceeded only if the participant answered “no” for all screening questions.

Among the 47 participants, there were 29 men and 18 women (refer to Figure 8). As shown in Figure 9, the ages ranged from 19 to 47 years old (mean = 27.81 years). Out of the 47 participants, 20 were between 18 and 24 years old, 19 were between 25 and 36 years old, and 8 were older than 36 years, with the oldest being 47 years old. The ethnic diversity included Malay (=18), Bangladesh (=10), Arab (=7), Chinese (=4), Indian (=3), Myanmar (=2), Pakistani (=2), and others (=1). Almost 75% of the participants were students (=35) of our university, and the rest were either academicians (=7) or from the community (=5) near our institution. Figures 10 and 11 provide a demographic chart of the participants’ race and occupation.

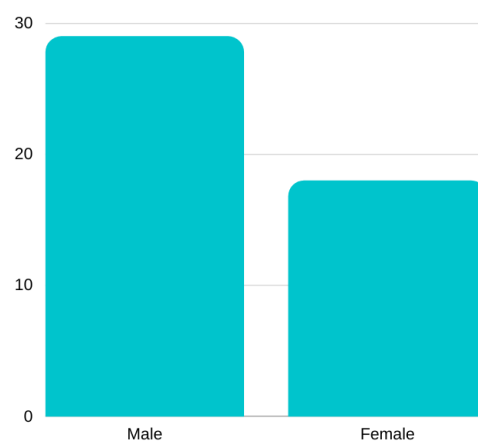


Figure 8. The demography of the participants’ gender.

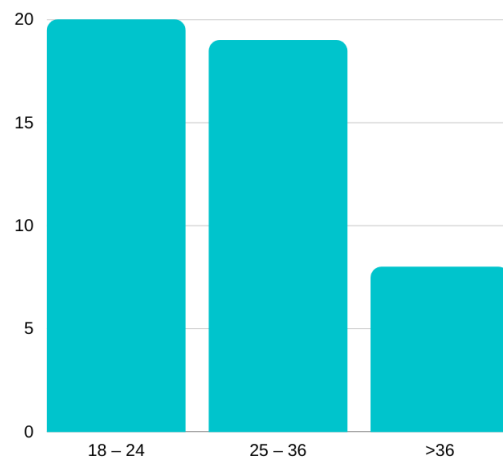


Figure 9. The demography of the participants' age.

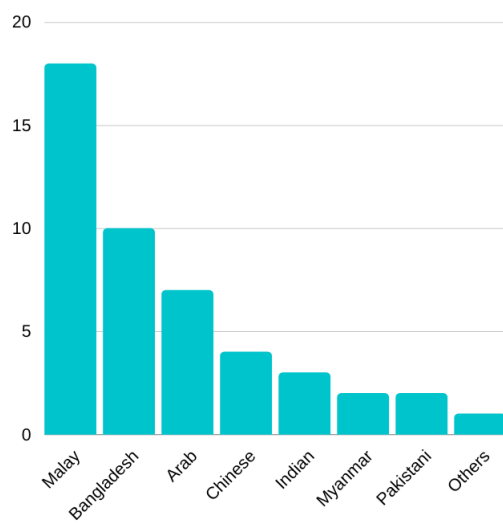


Figure 10. The demography of the participants' race.

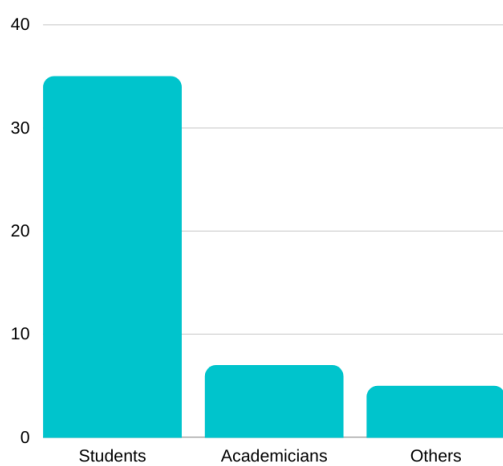


Figure 11. The demography of the participants' occupation.

There was a maximum number of two data collection sessions per day, totalling 47 sessions altogether. Once the participant arrived at the data collection lab, the attendee

explained the data collection procedure and devices used and obtained the participant's consent via the provided form. The KM device was positioned in front of the participant at an arms-level height. The participants had to place two fingers from each hand on the electrodes to record a 60 s ECG whenever they experienced an intense emotion. The MaximBand was worn on the participant's left hand, and the duration of the PPG recordings varied depending on the length of the videos.

Two computers were used for the data collection, as depicted in Figure 12. The setup of the left computer was for the participant to watch the videos and self-evaluate their emotional states, while the setup of the right computer was for the attendee to control the video displayed, as well as the participant's system. A divider was placed between the attendee and the participant to allow them to maintain their full concentration on the videos. The room temperature was set to 22 degrees Celsius. After the participants grasped the data collection procedure, the sensors status was verified, and the attendee started to play the videos without any further interaction with the participant. After each session, the PPG and ECG readings were gathered and stored on a secure drive. The KM and MaximBand applications were then reset for the next session.

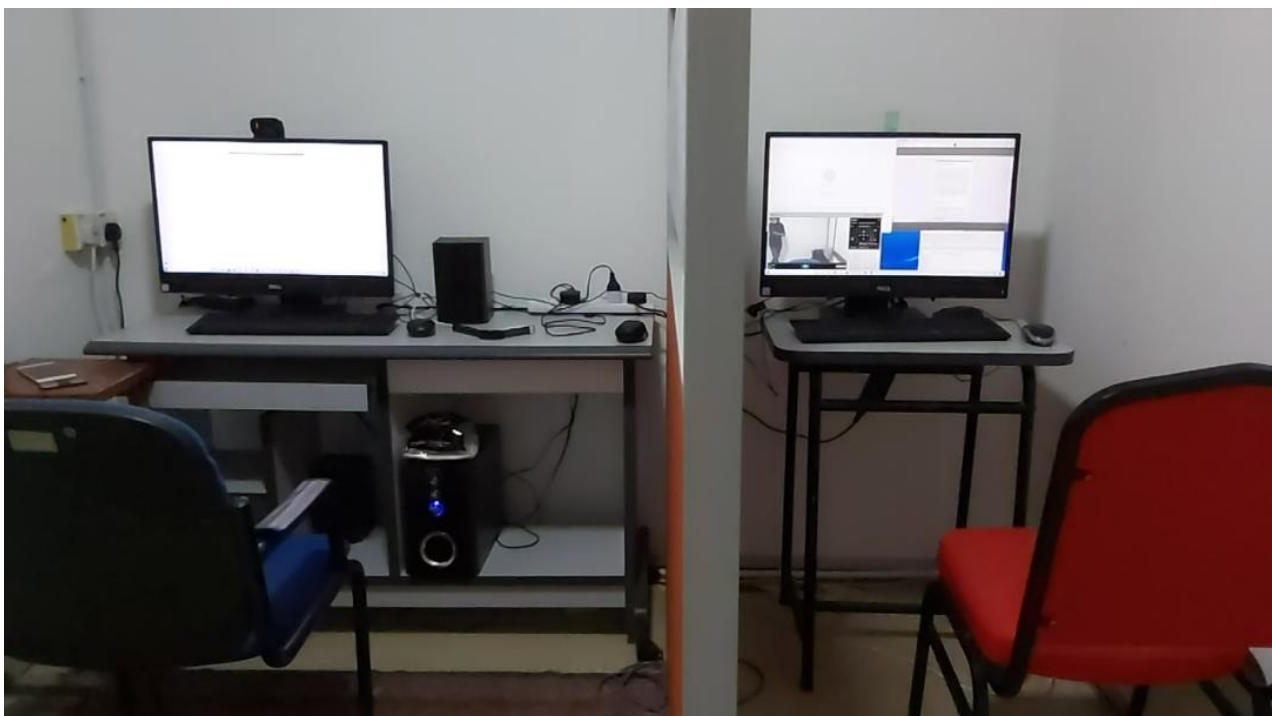


Figure 12. Data collection lab setup.

4. Data Preprocessing and Features Extraction

4.1. ECG

Augsburg Biosignal Toolbox (AuBT) is a MATLAB-based emotion recognition toolbox developed by a team of researchers at the University of Augsburg, Germany [55]. The functionality of the toolbox includes a comprehensive graphical user interface (GUI) with ECG preprocessing, feature extraction, feature combination, feature selection, and classification. Additionally, the toolbox also has the capability to process EMG, SC, and RSP. This toolbox was adopted for the A2ES ECG data preprocessing and feature extraction. The toolbox has also been adopted in prior research, such as in [14].

Prior to the extraction of the HR and HRV features from the ECG signals, lowpass filtering and normalization were applied during the preprocessing. Next, each P, Q, R, S, and T peak was detected. There were nine feature types, with a total of 81 mixed variations that could be extracted using the AuBT. The list of statistical features is shown in Table 3.

The mean, median, standard deviation (Stdev), max, min, and range (max–min) of an interval refer to the amplitude characteristics of the time series. For the heart rate variability (HRV), the RR interval of the time series was taken for measurement. Both the heart rate (HR) and HRV features could be used to detect emotions and stress. The HRV feature pNN50 is the number of adjacent R-to-R intervals, also known as the normal-to-normal, and the percentage was greater than 50 ms. The mean of the frequency spectrum (specRange) of the HRV was calculated based on the calculated range. The triangular index (TriInd) represents the sum of all normal-to-normal intervals divided by the height of the histogram of all RR intervals restrained on a distinct scale with bins of 7.8125 ms. In short, 66 HR statistical features from the ECG intervals and selected amplitudes with 15 time domain HRV features were extracted and combined, leading to a total of 81 features extracted by the AuBT.

Table 3. AuBT's ECG features.

Underlying Features	Statistical Features	Number of Features
RR, PP, QQ, SS, TT, PQ, QS, ST interval	Mean, Median, Stdev, Min, Max, Range	48
P, R, S amplitude	Mean, Median, Stdev, Min, Max, Range	18
HRV	Mean, Median, Stdev, Min, Max, Range, pNN50, specRange	8
HRV distribution	Mean, Median, Stdev, Min, Max, Range, TriInd	7
Total number of features:		81

4.2. PPG

The Toolbox for Emotional feAture extraction from Physiological signals (TEAP) is an open-source toolbox developed in [67] to process multimodal physiological signals for emotion detection. The TEAP can preprocess and extract features from EEG, GSR, ECG, PPG, EMG, RSP, and ST. This study focused only on applying the TEAP for PPG, which was implemented in MATLAB.

Upon implementing the TEAP on the raw PPG signals, the preprocessing is performed automatically. A low-pass median filter with a window equal to the sample rate cleaned the signals. Then, 17 features were extracted from the time and frequency domains of the clean PPG signals, and the list of features are shown in Table 4. The inter-beat-interval (IBI) is the time interval calculated between individual heartbeats. Based on the IBI, the HRV can be calculated by applying the standard deviation of all normal-to-normal intervals contained in each segment. The mean square error (MSE) features were calculated from the multiscale entropy at five levels, and they provide an insight into the complexity of the PPG signal fluctuations over the range of the time scale. The low, medium, and high frequencies of the tachogram power were also calculated as features. Four frequency ranges of the power spectral density (PSD) along with the statistical features of the mean and the mean IBI were also extracted. The last two features were the energy ratio for the spectral power density and tachogram power.

Table 4. PG features extracted using TEAP.

Features	Abbreviation	Description	Number of Features
IBI	meanIBI, HRV	Mean IBI, HRV (std(IBI))	2
MSE	MSE1, MSE2, MSE3, MSE4, MSE5	Multiscale entropy at 5 levels	5
Tachogram power	Tachogram_LF, Tachogram_MF, Tachogram_HF	$\log(PXLF(f))$, $\log(PXMF(f))$, $\log(PXHF(f))$, $\log(PXMF(f))/\log(PXLF(f)) + \log(PXHF(f))$, where LF: $f \in [0.01, 0.08]$ Hz, MF: $[0.08, 0.15]$ Hz, HF: $[0.15, 0.4]$ Hz	3
PSD	sp0001, sp0102, sp0203, sp0304	$\text{Log}(PX(f))$, $f \in \{[0, 0.01], [0.1, 0.2], [0.2, 0.3], [0.3, 0.4]\}$ Hz $\log(PXLF(f)PXHF(f))$, where LF: $f \in [0.0, 0.08]$ Hz and HF: $f \in [0.15, 5.0]$ Hz	4
Statistical moments	mean_	Mean	1
Energy ratio	sp_energyRatio, tachogram_energy_ratio	Spectral power ratio between the 0.0–0.08 Hz and 0.15–0.5 Hz bands, energy ratio $\text{tachogram_MFSP}/(\text{tachogram_HSP} + \text{tachogram_LFSP})$	2
Total number of Features:			17

5. Experimental Results and Discussion

In order to validate the proposed dataset, an emotion recognition experiment was conducted using machine learning and deep learning algorithms. The emotion recognition was performed using each modality: ECG and PPG. From the emotion types and intensities labelled by the participants, the emotions were reclassified according to the arousal classes of high and low, valence classes of high and low, and the dimensional classes of high arousal and high valence (HAHV), high arousal and low valence (HALV), low arousal and high valence (LAHV), and low arousal and low valence (LALV). Classifying emotions according to arousal and valence is commonly adopted in emotion recognition works, as seen in [68]. Before the classification, the data were split into subsets for training and testing. The ratio of the training and testing was 70 to 30%.

5.1. Machine Learning

In this study, five machine learning (ML) algorithms were used. The algorithms were, namely, support vector machine (SVM), naive bayes (NB), K-near neighbours (KNN), decision tree classifier (DT), and random forest (RF). These algorithms have been observed to be popularly chosen among researchers of affective computing [14,20,21,54,69]. Grid-SearchCV [70] was utilized to tune the hyperparameters of the algorithms, and the model was fit with the optimal parameters. This study also utilised the KFold cross-validation technique, with 10 folds. The accuracy of each ML in classifying the emotions were then compared.

5.1.1. ECG

Table 5 displays the classification performance of the ERS based on the ECG signals according to the arousal, valence, and dimensional model. The classification was conducted using the features extracted by AuBT. The results demonstrate that SVM was the best classifier among the five ML algorithms for the arousal and valence classification, with 68.75% and 58.81%, respectively. Other affective datasets have also reported an accuracy within same range [14]. KNN provided the highest accuracy for the dimensional values, with 32.10%. The classification of the dimensional multiclass was more complex than the

binary classification of arousal and valence. Thus, a lower accuracy was expected. Overall, the findings indicate that SVM and KNN are suitable for predicting arousal, valence, and dimensional emotions. On the other hand, the NB classifier was not able to provide a good performance in the classification of the ECG signals according to the three types of classification problems.

Table 5. Machine learning models' performance for ECG.

	SVM	DT	NB	KNN	RF
Arousal	68.75	66.19	40.06	66.76	65.91
Valence	58.81	55.4	54.55	54.83	54.55
Dimensional	29.26	30.11	17.33	32.1	31.82

5.1.2. PPG

The accuracy of the emotion classification utilising PPG signals is shown in Table 6. The RF had the highest classification accuracy for the arousal and dimensional emotions, with 67.30% and 40.00%, respectively. Whereas the SVM obtained the highest PPG-based ERS classification accuracy for valence at 64.94%. Table 7 demonstrates that compared to other algorithms, SVM and RF performed comparably well in identifying emotions based on the overall results, where they were either the best or the second best algorithm. Interestingly, although the features extracted were fewer, it was observed that the accuracy of the PPG-based ERS built for the valence and dimensional emotional model were better than the ECG-based ERS. Meanwhile, for arousal the difference was marginal.

Table 6. ML models' performance for PPG.

	SVM	DT	NB	KNN	RF
Arousal	64.61	63.64	63.31	60.06	67.30
Valence	64.94	53.25	54.22	54.55	57.07
Dimensional	37.01	36.69	25.00	34.42	40.00

Table 7. DL parameter settings.

Parameter	Parameter Setting
Input nodes	17,920
Hidden layers	33
Activation function (hidden layers)	Relu
Activation function (output layer)	Sigmoid

5.2. Deep Learning

In addition to ML, deep learning (DL) was also used in this study to assess the usability of the proposed ECG- and PPG-based ERS datasets. The DL network implemented here is depicted in Figure 13. The architecture consists of 33 convolutional layers, followed by a fully connected layer with a SoftMax activation function. Table 7 provides a summary of the DL parameter settings.

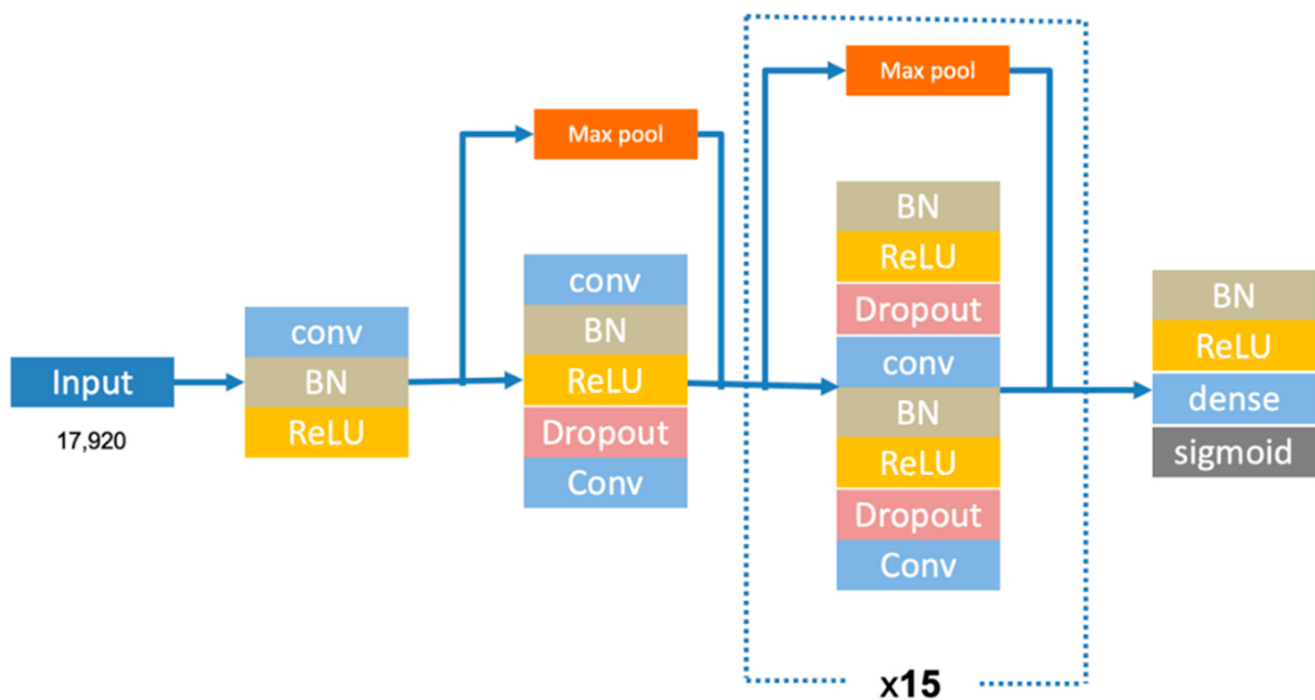


Figure 13. Deep learning model architecture.

5.2.1. ECG

The DL achieved a testing accuracy of 63.50% for arousal, 53.26% for valence, and 57.50% for the dimensional classes. The results are tabulated in Table 8. Compared to the results obtained by ML, DL had a poorer performance in classifying arousal and valence. Specifically, the SVM achieved a better performance, whereas in the dimensional classification, DL had the best performance, outperforming ML. It is worth mentioning that the size of A2ES is relatively small. Techniques to increase the size of the data such as data augmentation were not applied here. Future work should focus on this aspect before the adoption of DL for ERS development using the A2ES dataset.

Table 8. DL model's performance for ECG.

	DL
Arousal	63.50
Valence	53.26
Dimensional	57.50

5.2.2. PPG

The performance of the PPG-based ERS is tabulated in Table 9. DL obtained 34.63% for arousal, 56.8% for valence, and 24.35% for dimensional. These results are lacking in comparison to what was obtained by ML. The disparity in the results can be explained by the fact that the A2ES's small dataset may not be ideal for deep learning, which requires a larger dataset to train effectively. Additionally, the number of features for PPG was also less.

Table 9. DL model's performance for PPG.

	DL
Arousal	34.63
Valence	56.8
Dimensional	24.35

6. Discussion and Conclusions

In this era of COVID-19 and many other challenges, developing an emotion aware system is beneficial for society's mental health. Therefore, an affective research dataset, A2ES, was proposed in this paper. The dataset consists of ECG and PPG recordings collected from 47 Asian participants from various ethnicities using wearables and off-the-shelf devices. This was conducted to address the lack of such datasets for affective computing research and bias avoidance in future research. The participants were exposed to 25 audio–visual stimuli to elicit specific targeted emotions. The self-assessment ratings from the participants and a list of the 25 stimuli used were included, along with the ECG and PPG performance evaluations using ML and DL approaches. The findings prove the usability of the A2ES for emotion recognition. The performance of ML in classifying emotions using the A2ES with ECG and PPG was better than DL. This was because the size of the data was limited due to the small sample size of the A2ES dataset. The A2ES data are available upon request for other researchers and noncommercial purposes. Although, the data are labelled according to the seven basic emotions of neutral, happy, surprise, fear, disgust, sad, and anger, as well as their intensity, the data can be relabelled to arousal and valence. The data are not tagged to the participants. It is suggested that future research adopting the A2ES should consider different methods of feature extraction and feature selection and reduction to ensure only informative features are applied for more accurate classification, enhanced classification algorithms, and ensemble classifiers, as well as for addressing the imbalance in the data of the different classes. Additionally, to benefit from the strength of DL, the prospective focus should be to enhance the ERS by increasing the size of the data, such as by applying a data augmentation technique to expand the size of the data. The inclusion of the A2ES dataset with other affective computing datasets in building an ERS is expected to lead to an unbiased ERS.

Author Contributions: Conceptualization, N.A.A.A., T.K., S.N.M.S.I. and M.A.H.; methodology, T.K., S.N.M.S.I. and M.A.H.; software, T.K. and S.N.M.S.I.; validation, N.A.A.A. and K.A.A.; formal analysis, N.A.A.A., T.K. and S.N.M.S.I.; investigation, T.K., S.N.M.S.I. and M.A.H.; resources, N.A.A.A.; writing—original draft preparation, T.K., S.N.M.S.I. and M.A.H.; writing—review and editing, N.A.A.A., K.A.A., S.Z.I., A.A.A. and J.E.R.; visualization, T.K., S.N.M.S.I. and M.A.H.; supervision, N.A.A.A., S.Z.I., A.A.A. and J.E.R.; project administration, N.A.A.A. and K.A.A.; funding acquisition, N.A.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a TM Research and Development Grant (RDTC/190988) that was awarded to the Multimedia University.

Institutional Review Board Statement: The Multimedia University Research Ethics Committee (REC) has approved the data collection of A2ES (approval number: EA0282021).

Data Availability Statement: The A2ES dataset is available upon request for other researchers and noncommercial purposes at azlina.aziz@mmu.edu.my.

Acknowledgments: Special thanks to all participants of the A2ES data collection exercise.

Conflicts of Interest: The authors declare no conflict of interest.

References

- World Health Organisation. COVID-19 Disrupting Mental Health Services in Most Countries, WHO Survey; World Health Organization: Geneva, Switzerland, 2020; pp. 2–5.
- Findings, P.; Impact, T.; Coronavirus, O.; Lifeamerica, O. KFF Health Tracking Poll—Early April 2020: The Impact of Coronavirus on Life In America Poll Findings: The Impact of Coronavirus on Life In America Americans See Major Disruptions to Their Own Lives, Report No End in Sight; KFF: San Francisco, CA, USA, 2020; pp. 1–22.
- Son, C.; Hegde, S.; Smith, A.; Wang, X.; Sasangohar, F. Effects of COVID-19 on college students' mental health in the United States: Interview survey study. *J. Med. Internet Res.* **2020**, *22*, 14. [\[CrossRef\]](#) [\[PubMed\]](#)
- Stacey, A.; D'Eon, M.; Madojemu, G. Medical student stress and burnout: Before and after COVID-19. *Can. Med. Educ. J.* **2020**, *11*, e204. [\[CrossRef\]](#) [\[PubMed\]](#)
- Koldijk, S.; Sappelli, M.; Verberne, S.; Neerinx, M.A.; Kraaij, W. The Swell knowledge work dataset for stress and user modeling research. In Proceedings of the 2014 International Conference on Multimodal Interaction, Istanbul, Turkey, 12–16 November 2014.
- Rastgoo, M.N.; Nakisa, B.; Maire, F.; Rakotonirainy, A.; Chandran, V. Automatic driver stress level classification using multimodal deep learning. *Expert Syst. Appl.* **2019**, *138*, 112793. [\[CrossRef\]](#)
- Lee, D.S.; Chong, T.W.; Lee, B.G. Stress Events Detection of Driver by Wearable Glove System. *IEEE Sens. J.* **2017**, *17*, 194–204. [\[CrossRef\]](#)
- Spencer, C.; Koc, I.A.; Suga, C.; Lee, A.; Dhareshwar, A.M.; Franzén, E.; Iozzo, M.; Morrison, G.; McKeown, G.J. A Comparison of Unimodal and Multimodal Measurements of Driver Stress in Real-World Driving Conditions. *arXiv* **2020**. [\[CrossRef\]](#)
- Lee, J.; Kim, J.; Shin, M. Correlation Analysis between Electrocardiography (ECG) and Photoplethysmogram (PPG) Data for Driver's Drowsiness Detection Using Noise Replacement Method. *Procedia Comput. Sci.* **2017**, *116*, 421–426. [\[CrossRef\]](#)
- Bahreini, K.; Nadolski, R.; Westera, W. Towards real-time speech emotion recognition for affective e-learning. *Educ. Inf. Technol.* **2016**, *21*, 1367–1386. [\[CrossRef\]](#)
- Wang, W.; Xu, K.; Niu, H.; Miao, X. Emotion Recognition of Students Based on Facial Expressions in Online Education Based on the Perspective of Computer Simulation. *Complexity* **2020**, *2020*, 4065207. [\[CrossRef\]](#)
- Alqahtani, F.; Katsigiannis, S.; Ramzan, N. Using Wearable Physiological Sensors for Affect-Aware Intelligent Tutoring Systems. *IEEE Sens. J.* **2021**, *21*, 3366–3378. [\[CrossRef\]](#)
- Koelstra, S.; Muhl, C.; Soleymani, M.; Lee, J.S.; Yazdani, A.; Ebrahimi, T.; Pun, T.; Nijholt, A.; Patras, I. DEAP: A database for emotion analysis using physiological signals. *IEEE Trans. Affect. Comput.* **2012**, *3*, 18–31. [\[CrossRef\]](#)
- Katsigiannis, S.; Ramzan, N. DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals from Wireless Low-cost Off-the-Shelf Devices. *IEEE J. Biomed. Health Inform.* **2018**, *22*, 98–107. [\[CrossRef\]](#)
- Minhad, K.N.; Ali, S.H.M.; Reaz, M.B.I. Happy-anger emotions classifications from electrocardiogram signal for automobile driving safety and awareness. *J. Transp. Health* **2017**, *7*, 75–89. [\[CrossRef\]](#)
- Song, T.; Zheng, W.; Lu, C.; Zong, Y.; Zhang, X.; Cui, Z. MPED: A multi-modal physiological emotion database for discrete emotion recognition. *IEEE Access* **2019**, *7*, 12177–12191. [\[CrossRef\]](#)
- Hasnul, M.A.; Ab Aziz, N.A.; Aziz, A.A. Evaluation of TEAP and AuBT as ECG's Feature Extraction Toolbox for Emotion Recognition System. In Proceedings of the 2021 IEEE 9th Conference on System, Process and Control, ICSPC 2021, Malacca, Malaysia, 10–11 December 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 52–57.
- Miranda Correa, J.A.; Abadi, M.K.; Sebe, N.; Patras, I. AMIGOS: A Dataset for Affect, Personality and Mood Research on Individuals and Groups. *IEEE Trans. Affect. Comput.* **2018**, *12*, 479–493. [\[CrossRef\]](#)
- Park, C.Y.; Cha, N.; Kang, S.; Kim, A.; Khandoker, A.H.; Hadjileontiadis, L.; Oh, A.; Jeong, Y.; Lee, U. K-EmoCon, a multimodal sensor dataset for continuous emotion recognition in naturalistic conversations. *Sci. Data* **2020**, *7*, 293. [\[CrossRef\]](#)
- Abadi, M.K.; Subramanian, R.; Kia, S.M.; Avesani, P.; Patras, I.; Sebe, N. DECAF: MEG-Based Multimodal Database for Decoding Affective Physiological Responses. *IEEE Trans. Affect. Comput.* **2015**, *6*, 209–222. [\[CrossRef\]](#)
- Udovičić, G.; Derek, J.; Russo, M.; Sikora, M. Wearable Emotion Recognition System based on GSR and PPG Signals. In Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care, Mountain View, CA, USA, 13 October 2017; pp. 53–59. [\[CrossRef\]](#)
- Bagirathan, A.; Selvaraj, J.; Gurusamy, A.; Das, H. Recognition of positive and negative valence states in children with autism spectrum disorder (ASD) using discrete wavelet transform (DWT) analysis of electrocardiogram signals (ECG). *J. Ambient Intell. Humaniz. Comput.* **2021**, *12*, 405–416. [\[CrossRef\]](#)
- Hsu, Y.L.; Wang, J.S.; Chiang, W.C.; Hung, C.H. Automatic ECG-Based Emotion Recognition in Music Listening. *IEEE Trans. Affect. Comput.* **2017**, *11*, 85–99. [\[CrossRef\]](#)
- Mand, A.A.; Wen, J.S.J.; Sayeed, M.S.; Swee, S.K. Robust stress classifier using adaptive neuro-fuzzy classifier-linguistic hedges. In Proceedings of the 2017 International Conference on Robotics, Automation and Sciences, ICORAS 2017, Melaka, Malaysia, 27–29 November 2017; pp. 1–5.
- Hasnul, M.A.; Aziz, N.A.A.; Alelyani, S.; Mohana, M.; Aziz, A.A. Electrocardiogram-based emotion recognition systems and their applications in healthcare—A review. *Sensors* **2021**, *21*, 5015. [\[CrossRef\]](#)
- Rock Health. Stanford Medicine Center for Digital Health Digital Health Consumer Adoption Report 2019; Rock Health: San Francisco, CA, USA, 2019; pp. 1–6.

27. Laricchia, F. Smartwatches—Statistics and Facts. 2020. Available online: <https://www.statista.com/topics/4762/smartwatches/#editorsPicks> (accessed on 30 November 2022).
28. Jemioło, P.; Storman, D.; Mamica, M.; Szymkowski, M.; Żabicka, W.; Wojtaszek-Główka, M.; Ligeza, A. Datasets for Automated Affect and Emotion Recognition from Cardiovascular Signals Using Artificial Intelligence—A Systematic Review. *Sensors* **2022**, *22*, 2538. [\[CrossRef\]](#)
29. Schmidt, P.; Reiss, A.; Duerichen, R.; Van Laerhoven, K. Introducing WeSAD, a multimodal dataset for wearable stress and affect detection. In Proceedings of the 2018 International Conference Multimodal Interaction, Boulder, CO, USA, 16–20 October 2018; pp. 400–408. [\[CrossRef\]](#)
30. Quiroz, J.C.; Geangu, E.; Yong, M.H. Emotion recognition using smart watch sensor data: Mixed-design study. *JMIR Ment. Health* **2018**, *5*, e10153. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Colvonen, P.J.; DeYoung, P.N.; Bosompra, N.O.A.; Owens, R.L. Limiting racial disparities and bias for wearable devices in health science research. *Sleep* **2020**, *43*, zsaa159. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Noseworthy, P.A.; Attia, Z.I.; Brewer, L.P.C.; Hayes, S.N.; Yao, X.; Kapa, S.; Friedman, P.A.; Lopez-Jimenez, F. Assessing and Mitigating Bias in Medical Artificial Intelligence: The Effects of Race and Ethnicity on a Deep Learning Model for ECG Analysis. *Circ. Arrhythmia Electrophysiol.* **2020**, *13*, 208–214. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Koerber, D.; Khan, S.; Shamsheri, T.; Kirubakaran, A.; Mehta, S. The Effect of Skin Tone on Accuracy of Heart Rate Measurement in Wearable Devices: A Systematic Review. *J. Am. Coll. Cardiol.* **2022**, *79*, 1990. [\[CrossRef\]](#)
34. Rizal, A.; Hidayat, R.; Nugroho, H.A. Signal Domain in Respiratory Sound Analysis: Methods, Application and Future Development. *J. Comput. Sci.* **2016**, *11*, 1005–1016. [\[CrossRef\]](#)
35. Kamshilin, A.A.; Margaryants, N.B. Origin of Photoplethysmographic Waveform at Green Light. *Phys. Procedia* **2017**, *86*, 72–80. [\[CrossRef\]](#)
36. Renesas. *OB1203 Heart Rate, Blood Oxygen Concentration, Pulse Oximetry, Proximity, Light and Color Sensor: Signal to Noise Ratio*; Renesas: Tokyo, Japan, 2020; pp. 1–12.
37. Fischer, C.; Glos, M.; Penzel, T.; Fietze, I. Extended algorithm for real-time pulse waveform segmentation and artifact detection in photoplethysmograms. *Somnologie* **2017**, *21*, 110–120. [\[CrossRef\]](#)
38. Lee, M.S.; Lee, Y.K.; Pae, D.S.; Lim, M.T.; Kim, D.W.; Kang, T.K. Fast emotion recognition based on single pulse PPG signal with convolutional neural network. *Appl. Sci.* **2019**, *9*, 3355. [\[CrossRef\]](#)
39. Preethi, M.; Nagaraj, S.; Madhan Mohan, P. Emotion based Media Playback System using PPG Signal. In Proceedings of the 2021 International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET 2021, Chennai, India, 25–27 March 2021; pp. 426–430.
40. Yang, W.; Rifqi, M.; Marsala, C.; Pinna, A. Physiological-Based Emotion Detection and Recognition in a Video Game Context. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 July 2018. [\[CrossRef\]](#)
41. Zainudin, Z.; Hasan, S.; Shamsuddin, S.M.; Argawal, S. Stress Detection using Machine Learning and Deep Learning. *J. Phys. Conf. Ser.* **2021**, *1997*, 012019. [\[CrossRef\]](#)
42. Chen, T.; Yin, H.; Yuan, X.; Gu, Y.; Ren, F.; Sun, X. Emotion recognition based on fusion of long short-term memory networks and SVMs. *Digit. Signal Process. A Rev. J.* **2021**, *117*, 103153. [\[CrossRef\]](#)
43. Ayata, D.; Yaslan, Y.; Kamasak, M.E. Emotion Recognition from Multimodal Physiological Signals for Emotion Aware Healthcare Systems. *J. Med. Biol. Eng.* **2020**, *40*, 149–157. [\[CrossRef\]](#)
44. Domínguez-Jiménez, J.A.; Campo-Landines, K.C.; Martínez-Santos, J.C.; Delahoz, E.J.; Contreras-Ortiz, S.H. A machine learning model for emotion recognition from physiological signals. *Biomed. Signal Process. Control.* **2019**, *55*, 101646. [\[CrossRef\]](#)
45. Kim, B.H.; Jo, S. Deep Physiological Affect Network for the Recognition of Human Emotions. *IEEE Trans. Affect. Comput.* **2020**, *11*, 230–243. [\[CrossRef\]](#)
46. Li, C.; Xu, C.; Feng, Z. Analysis of physiological for emotion recognition with the IRS model. *Neurocomputing* **2016**, *178*, 103–111. [\[CrossRef\]](#)
47. Shahid, H.; Butt, A.; Aziz, S.; Khan, M.U.; Hassan Naqvi, S.Z. Emotion Recognition System featuring a fusion of Electrocardiogram and Photoplethysmogram Features. In Proceedings of the 2020 14th International Conference on Open Source Systems and Technologies (ICOSST), Lahore, Pakistan, 16–17 December 2020; pp. 1–6. [\[CrossRef\]](#)
48. Yang, C.J.; Fahier, N.; Li, W.C.; Fang, W.C. A Convolution Neural Network Based Emotion Recognition System using Multimodal Physiological Signals. In Proceedings of the 2020 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan 2020), Taoyuan, Taiwan, 28–30 September 2020; pp. 2020–2021. [\[CrossRef\]](#)
49. Raheel, A.; Majid, M.; Anwar, S.M. DEAR-MULSEMEDIA: Dataset for emotion analysis and recognition in response to multiple sensorial media. *Inf. Fusion* **2021**, *65*, 37–49. [\[CrossRef\]](#)
50. Sharma, K.; Castellini, C.; van den Broek, E.L.; Albu-Schaeffer, A.; Schwenker, F. A dataset of continuous affect annotations and physiological signals for emotion analysis. *Sci. Data* **2019**, *6*, 196. [\[CrossRef\]](#)
51. Markova, V.; Ganchev, T.; Kalinkov, K. CLAS: A Database for Cognitive Load, Affect and Stress Recognition. In Proceedings of the International Conference on Biomedical Innovations and Applications (BIA 2019), Varna, Bulgaria, 8–9 November 2019.
52. Gao, Z.; Cui, X.; Wan, W.; Zheng, W.; Gu, Z. ECSMP: A dataset on emotion, cognition, sleep, and multi-model physiological signals. *Data Br.* **2021**, *39*, 107660. [\[CrossRef\]](#)

53. Kim, J.; André, E. Emotion recognition based on physiological changes in music listening. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 2067–2083. [\[CrossRef\]](#)
54. Subramanian, R.; Wache, J.; Abadi, M.K.; Vieriu, R.L.; Winkler, S.; Sebe, N. Ascertain: Emotion and personality recognition using commercial sensors. *IEEE Trans. Affect. Comput.* **2018**, *9*, 147–160. [\[CrossRef\]](#)
55. Wagner, J. *Augsburg Biosignal Toolbox (Aubt)*; University of Augsburg: Augsburg, Germany, 2014.
56. Healey, J.A.; Picard, R.W. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intell. Transp. Syst.* **2005**, *6*, 156–166. [\[CrossRef\]](#)
57. Soleymani, M.; Lichtenauer, J.; Pun, T.; Pantic, M. A multimodal database for affect recognition and implicit tagging. *IEEE Trans. Affect. Comput.* **2012**, *3*, 42–55. [\[CrossRef\]](#)
58. Ringeval, F.; Sonderegger, A.; Sauer, J.; Lalanne, D. Introducing the RECOLA multimodal corpus of remote collaborative and affective interactions. In Proceedings of the 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Shanghai, China, 22–26 April 2013. [\[CrossRef\]](#)
59. Ismail, S.N.M.S.; Aziz, N.A.A.; Ibrahim, S.Z.; Khan, C.T.; Rahman, M.A. Selecting Video Stimuli for Emotion Elicitation via Online Survey. *Hum.-Cent. Comput. Inf. Sci.* **2021**, *11*, 19. [\[CrossRef\]](#)
60. Sayed Ismail, S.N.M.; Nor, N.A.; Ibrahim, S.Z. A comparison of emotion recognition system using electrocardiogram (ECG) and photoplethysmogram (PPG). *J. King Saud Univ.-Comput. Inf. Sci.* **2022**, *34*, 3539–3558. [\[CrossRef\]](#)
61. Lau, J.K.; Lowres, N.; Neubeck, L.; Brieger, D.B.; Sy, R.W.; Galloway, C.D.; Albert, D.E.; Freedman, S.B. iPhone ECG application for community screening to detect silent atrial fibrillation: A novel technology to prevent stroke. *Int. J. Cardiol.* **2013**, *165*, 193–194. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Haberman, Z.C.; Jahn, R.T.; Bose, R.; Tun, H.; Shinbane, J.S.; Doshi, R.N.; Chang, P.M.; Saxon, L.A. Wireless Smartphone ECG Enables Large-Scale Screening in Diverse Populations. *J. Cardiovasc. Electrophysiol.* **2015**, *26*, 520–526. [\[CrossRef\]](#)
63. Tarakji, K.G.; Wazni, O.M.; Callahan, T.; Kanj, M.; Hakim, A.H.; Wolski, K.; Wilkoff, B.L.; Saliba, W.; Lindsay, B.D. Using a novel wireless system for monitoring patients after the atrial fibrillation ablation procedure: The iTransmit study. *Heart Rhythm* **2015**, *12*, 554–559. [\[CrossRef\]](#)
64. Lowres, N.; Mulcahy, G.; Gallagher, R.; Ben Freedman, S.; Marshman, D.; Kirkness, A.; Orchard, J.; Neubeck, L. Self-monitoring for atrial fibrillation recurrence in the discharge period post-cardiac surgery using an iPhone electrocardiogram. *Eur. J. Cardiothorac. Surg.* **2016**, *50*, 44–51. [\[CrossRef\]](#)
65. Desteghe, L.; Raymaekers, Z.; Lutin, M.; Vijgen, J.; Dilling-Boer, D.; Koopman, P.; Schurmans, J.; Vanduyndhoven, P.; Dendale, P.; Heidbuchel, H. Performance of handheld electrocardiogram devices to detect atrial fibrillation in a cardiology and geriatric ward setting. *Europace* **2017**, *19*, 29–39. [\[CrossRef\]](#)
66. Bumgarner, J.M.; Lambert, C.T.; Hussein, A.A.; Cantillon, D.J.; Baranowski, B.; Wolski, K.; Lindsay, B.D.; Wazni, O.M.; Tarakji, K.G. Automated Atrial Fibrillation Detection Algorithm Using Smartwatch Technology. *J. Am. Coll. Cardiol. Autom. Atr. Fibrillation Detect. Algorithm Using Smartwatch Technol.* **2018**, *71*, 2381–2388. [\[CrossRef\]](#)
67. Soleymani, M.; Villaro-Dixon, F.; Pun, T.; Chanel, G. Toolbox for Emotional feature extraction from Physiological signals (TEAP). *Front. ICT* **2017**, *4*, 1–7. [\[CrossRef\]](#)
68. Shu, L.; Xie, J.; Yang, M.; Li, Z.; Liao, D.; Xu, X.; Yang, X. A review of emotion recognition using physiological signals. *Sensors* **2018**, *18*, 2074. [\[CrossRef\]](#)
69. Mano, L.Y. Emotional condition in the Health Smart Homes environment: Emotion recognition using ensemble of classifiers. In Proceedings of the 2018 IEEE (SMC) International Conference on Innovations in Intelligent Systems and Applications, INISTA 2018, Roma, Italy, 3–5 July 2018.
70. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B. Scikit-learn: Machine Learning in Python. *J. Ofmachine Learn. Res.* **2011**, *12*, 2825–2830. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.