ACS presents: The Annual
Dennis Moore Oration & 1962 awards

Building AI Tools

People & Preserving Privacy

with Professor
Tom Gedeon

Proceedings

Opening and welcome	Dr David Cook FACS CP
Introduction of opening speaker	Mr Jerome Chiew MACS (Snr) CP
Opening address	Adam Edwards, Office of Digital Government
Introduction	1962 Prize and Medal Finalists Professor Terry Woodings FACS
Presentation of 1962 Awards	Professor Alex Reid FACS CP
Introduction 1962 Educators	Professor Terry Woodings FACS
Presentation of Educator Awards	Dr Brian von Konsky FACS CP
Welcome by event partner	DC Alliance
Introduction of 2023 Orator	Dr David Cook FACS CP
Oration delivered	Professor Tom Gedeon
Vote of thanks	Dr Brian von Konsky FACS CP

Oration Organising Committee

Dr David Cook FACS CP

Dr Bob Cross FACS CP

Dr Brian von Konsky FACS CP

Professor Terry Woodings FACS

1962 Awards Judges

Dr Bob Cross FACS CP

Professor Tanya McGill FACS

Professor Tony Watson FACS

Professor Terry Woodings FACS

The Australian Computer Society Annual Dennis Moore Oration Dinner. The University Club of Western Australia, 24 November 2023.

Orator: **Professor Tom Gedeon**

We wish to thank our exclusive event sponsor

Annual Dennis Moore Oration Past Orators

Since 2012, to commemorate fifty years of digital computing in Western Australia, the WA Branch of the ACS has invited a distinguished scholar and researcher with a connection to WA to present a lecture on the leading edge of an important and emerging area of information and computer technology.

2023

Year	Orator
2012	Professor Andrew Rohl
2013	Professor Ian Reid
2014	Professor Craig Valli
2015	Professor Svetha Venkatesh
2016	Dr Adrian Boeing
2017	Professor Matthew Bellgard
2018	Professor Jinbo Wang
2019	Associate Professor Rachel Cardwell-Oliver
2020	No Oration held due to Covid restrictions
2021	Associate Professor Doina Olaru
2022	Associate Professor Vidy Potdar

1962 Prize

From a suggestion of Dennis Moore (and with his strong support) 2012 also saw the setting up of an annual prize for the best graduating student in ICT from a WA university. Although the primary criteria are based on academic performance, the candidates are also judged on their ability to promote their ideas in computing and contribution so far.

Previous winners of the 1962 Prize are:

Year	Winner
2022	David Adams & Yuval Berman – University of Western Australia
2021	Alistair Martin – Murdoch University
2020	Samual Heath – University of Western Australia
2019	Jarryd Wimbridge – Edith Cowan University
2018	Taaqif Peck – University of Western Australia
2017	Mark Shelton – University of Western Australia
2016	Dalibor Borkovic - Murdoch University
2015	Michael Martis - University of Western Australia
2014	Anthony Long - Curtin University
2013	Laurence Da Luz - Edith Cowan University
2012	Kevin Adnan - Curtin University

The 1962 Prize finalists for 2023 in alphabetical order are:

Year	Winner
Max Barker	Curtin University
Erik Martin Estevez	Murdoch University
Benjamin Gray	Curtin University
Cory Gyarmathy	Curtin University
Shuang Li	Murdoch University
Charles Mein	Curtin University

1962 Medal

In 2019 the 1962 Awards was expanded to include a new award for the most outstanding candidate who completed Doctoral research (eg PhD) in Western Australia in the field of Information Technology and Computer Science.

Previous winners of the 1962 Medals are:

Year	Winner
2022	Dr Uzair Nadeem, University of Western Australia
2021	Dr Naeha Sharif, University of Western Australia
2020	Dr Anupiya Nugaliyadde, Murdoch University
2019	Dr Qiuhong Ke, University of Western Australia

2023

The 1962 Medal finalists for 2023 in alphabetical order:

Dr Nur Al Hasan Haldar	The University of Western Australia
Dr Md Shamim Hossein	Edith Cowan University
Dr Sirui Li	Murdoch University
Dr Hira Maqbool	Edith Cowan University
Dr Mohammad Mursalin	Edith Cowan University
Dr Manou Malin Rosenberg	The University of Western Australia

1962 Educator Recognition

New in 2023, the 1962 awards will recognise teachers and lecturers that have received awards from other organisations. These are the dedicated people who make the other awards possible.

The 1962 Educators for 2023 that being recongised in alphabetical order are:

Donna Buckley	John Curtin College of the Arts
Michelle Chomiak	St Marks Community School
Dr Michelle Ellis	Edith Cowan University School of Science
Vinicius Madeiros	National Institute of Technology
Dr Chris McDonald	CSSE, University of Western Australia

Dennis Moore AM MA (Cantab) FACS:

Dennis Moore was born in NSW in 1937. He was educated on scholarships at The King's School, Parramatta where he was captain and dux of the school, and at Queens' College Cambridge where he graduated in 1958 in mathematics.

After a period with commerce and industry in computing and operations research in NSW, he pioneered computing in Western Australia, installing the first computer at UWA in 1962. He introduced WA's first computing qualification – the DipNAAC – at UWA.

In 1965, he was responsible for the purchase and installation of the DEC PDP-6. This was the world's first commercial installation of a time-shared computer and Australia's first high precision graphics device.

He was foundation president of the WA Computer Society, which later merged with the Australian Computer Society, becoming the first WA Branch Chairman. He was Director of the Western Australian Regional Computing Centre in the sixties and seventies. This provided computing services to CSIRO and State Government Departments as well as the University.

He was executive director of Government Computing for WA from 1978 to 1984. During this period he promoted the development of inter-departmental systems and was closely associated with the development of the WA Land Information System and the WA Technology Park. This was followed by a two year stint managing a computer company in Malaysia, including a consultancy to the Sarawak Government.

He then undertook research in RAN DATA, an encryption company which he had helped establish, and was appointed foundation Head of School of Computing at Curtin University of Technology in 1987. From 1998 to 2002 he was Director of Academic Planning at Curtin. From 1995 to 1999 he was Chair of the State Government's Information Policy Council.

Dennis Moore was elected a Fellow of the Australian Computer Society in 1970 and was made a Member of the Order of Australia for services to Information Technology in 1997. He retired in 2002 and was made an Honorary Life Member of the ACS in 2014.

Privacy—preserving AI tools to respond to and understand people

Prof. Tom Gedeon
Curtin University
tom.gedeon@curtin.edu.au

Abstract

With the availability of low-cost sensors in today's environment, we are increasingly capturing data directly from individuals' behaviours via wearables and cameras. This enables us to create AI tools capable of reading human actions and reactions to the outcomes created by our AI systems, allowing us to enhance or adjust their outputs, and be responsive to the human. This closely resembles the nonverbal cues used by people during a conversation.

Responsive AI refers to cutting-edge AI that responds to human actions and reactions to predict subtle emotional states. In practice today, these sensors include wearable devices that detect skin conductance, heart rate, muscular activation, and skin temperature, among other signals. Cameras are used to track eye gaze behaviours, capture videos, record thermal images, or utilise hyperspectral technology, and with a long term additional goal of replacing wearable sensors. The deployment of AI to discern nuanced human internal states introduces new privacy concerns in addition to the expected privacy implications associated with video cameras. These risks can be alleviated through the adoption of privacy-preserving techniques referred to as Responsible AI. This is a privacy-by-design method for managing the use of private and personal data. In practice, this means using adversarial generative algorithms to remove personally identifying data from sensor data streams and videos. We will discuss previous work in these areas to show how fully utilising Responsive AI needs the incorporation of Responsible AI principles.

Biography

Tom Gedeon holds the Human-Centric Advancements Chair in Artificial Intelligence at Curtin University. Prior to this, he was Optus Chair in AI. He was Professor of Computer Science and former Deputy Dean of the College of Engineering and Computer Science at ANU. He gained his BSc (Hons) and PhD from the University of Western Australia. Professor Gedeon's main research area is Responsive and Responsible AI. His focus is on the development of automated systems for information extraction, from eye gaze and physiological data, as well as textual, camera and other data, and for the synthesis of the extracted information into humanly useful information resources, primarily using neural/deep networks and fuzzy logic methods, and delivered in real, augmented and virtual environments. Professor Gedeon has over 400 publications and has run multiple international conferences. He is a former president of the Asia-Pacific Neural Network Assembly and former President of the Computing Research and Education Association of Australasia. He has been General Chair for the International Conference on Neural Information Processing (ICONIP) three times. He has been nominated for VC's awards for postgraduate supervision at three Universities. He was recently a member of the Australian Research Council's College of Experts, and continues in that role from 2024. He is an associate editor of the IEEE Transactions on Fuzzy Systems and the INNS/Elsevier Neural Networks journal.

1. Introduction

Building responsive AI tools while preserving privacy is a critical challenge in the field of artificial intelligence. It involves maintaining a balance between delivering personalised and effective AI-driven experiences based on context information. An AI-driven responsive experience is either *intrapersonal* or *interpersonal*. Here we will consider *intrapersonal* aspects to cover different emotional responses of a subject primarily from an observer's point of view and *interpersonal* aspects cover interactions among two or multiple individuals.

Responsible AI, also known as Ethical AI or AI Ethics, refers to the practice of developing and deploying artificial intelligence (AI) systems in a way that aligns with ethical principles, human values, and societal norms. The goal of responsible AI is to ensure that AI technologies are developed and used in a manner that is fair, transparent, accountable, and respects the rights and well-being of individuals and communities.

Key principles and considerations of responsible AI include but not limited to *fairness*, *transparency*, *accessibility*, *privacy*, *security*. Any AI system should be designed to avoid bias and discrimination, ensuring that it does not unfairly advantage or disadvantage particular groups of people. The AI system should follow *Transparency*, where the decision-making processes and algorithms behind an AI system should be transparent and explainable, allowing users and stakeholders to understand how decisions are made. Developers and organizations responsible for AI systems should be accountable for the outcomes of these systems. This includes addressing errors, biases, and harm caused by their AI system. AI systems should respect individuals' privacy rights and protect sensitive data. Data collection and usage should be transparent and consent-based. AI systems should be built with robust security measures to protect against malicious attacks and unauthorized access. AI technologies should be designed to be accessible to all individuals, including those with disabilities, to ensure inclusivity. AI systems should not be used to harm, deceive, or infringe upon human rights. Developers and users of AI should consider the ethical implications of their applications. Stakeholders from diverse backgrounds, including ethicists, policymakers, technologists, and affected communities, should collaborate in the development and deployment of AI systems to ensure a well-rounded perspective.

Additionally, any AI system requires continuous monitoring and evaluation. Ongoing monitoring and evaluation of AI systems are essential to identify and address ethical and social issues that may arise over time. Promoting ethical AI requires a multidisciplinary approach that involves not only technical experts but also ethicists, policymakers, and society as a whole. Many organizations and initiatives, both public and private, are working to establish guidelines, standards, and best practices for the ethical development and deployment of AI technologies to ensure that they benefit humanity while minimizing risks and harms. Finally, all AI systems should be used responsibly.

Responsible AI is a component of ethical AI, and focuses on alleviating privacy risks through the adoption of privacy-preserving techniques referred to as Responsible AI. This is a privacy-by-design method for managing the use of private and personal data. In practice, this means using adversarial generative algorithms to remove personally identifying data from sensor data streams and videos.

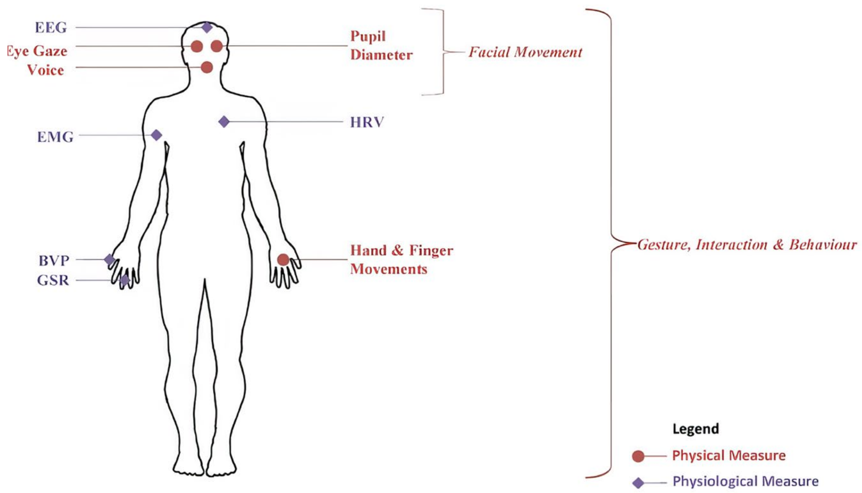


Figure 1: Intrapersonal Responsive AI

2. Intrapersonal Responsive AI

In the context of human-centred AI and computing research, we aim to place sensors on humans to extract useful information. This information can be used to enhance AI-based software or to understand human attributes and reactions. We discuss a few examples ranging from stress to empathy. The exploration space for collecting sensory information is shown in Figure 1.

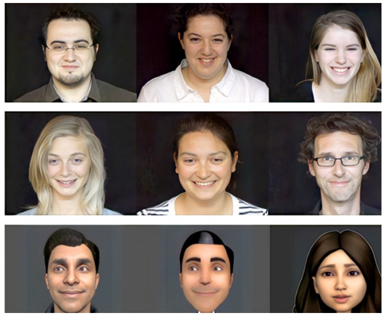
2.1 Stress

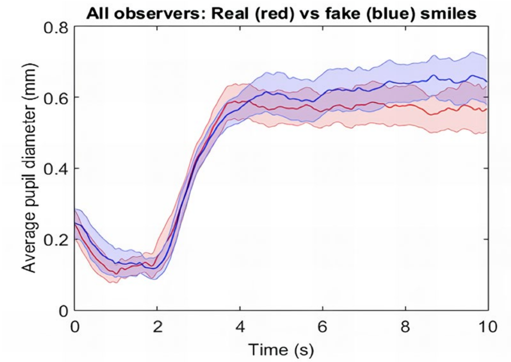
Stress is a major problem in our society today and poses major concerns for the future [1]. It is important to gain an objective understanding of how average individuals respond to events they observe in typical environments they encounter. We developed a computational model of stress based on objective human responses collected from human observers of environments. In the process, we investigated whether a computational model can be developed to recognise observer stress in abstract virtual environments (text), virtual environments (films) and real environments (real-life settings) using physiological and physical response sensor signals [2]. Our work proposes an architecture for a computational observer stress model. The architecture was used to implement models for the different types of environments. Sensors appropriate to the different types of environments were investigated where the aims were to achieve unobtrusive methods for stress response signal collection, reduce encumbrance and hence, enhance methods to capture natural observer behaviours and produce stress models that recognised stress more robustly. We discuss the motivations for each investigation and detail the experiments we conducted to collect stress data sets for observers of the different types of environments. We describe individual-independent artificial neural networks and support vector machine-based models that were developed to recognise stress patterns from observer response signals. The classifiers were extended to include a genetic algorithm, which was used to select features that were better for stress recognition and reduce the use of redundant features. The outcomes of this research provide a possible future extension on managing stress objectively.

2.2 Emotion Veracity

Interactions with emotionally expressive faces has a fascinating charm that has the potential to deeply capture human attention. When a face shows an expression, it not only reflects the individual's inner mental state, but it also appears to seek to influence the emotional condition of those who observe this display. In our study, we discovered that physiological responses can more accurately recognise the smiles and anger displays in people than using verbal replies from the same people [5], [6]. The ability to perceive facial expressions is critical to human social interaction and is present in the majority of social communications. Understanding genuine facial expressions is seen as a valuable social skill. A smile is regarded as a compliment in this context. It is an easily recognisable yet can be a confusing facial display where real smiles can be labelled as both *feeling* and *showing* happiness, but posed smiles are labelled as *showing* rather than *feeling* happiness.

An initial pupil data timeline analysis indicated a common trend for each stimulus. Figure 2b depicts the time point average of pupil diameter across all observers while watching all video stimuli. The pupil constricted from stimulus initiation and reached a minimum in 1-2 seconds, after which a rapid dilatation began and lasted 3-4 seconds. Then either a smooth dilation or constriction began and maintained in a consistent range until the end of our analysis time-frame. It is worth noting that the trends were divided based on real and posed smiling stimuli starting around the 3s. The paired sample permutation tests revealed that pupil dilation differed significantly between real and posed smiling stimuli ($t = 4.56$). In the context, we proposed an *independent approach* (also called leave-one-smile-and-observer- out) and compared with the *smiler independent* (leave-one-smile-out, which means that the classifiers have seen no physiological features from any of the observers on that smiler) and the *observer independent* (leave-one-observer-out, which means that the classifiers have detected no physiological features from any of the smilers on that observer) approaches (see Figure 3) [7]. Our approach is fully unbiased from both smile videos and observers.





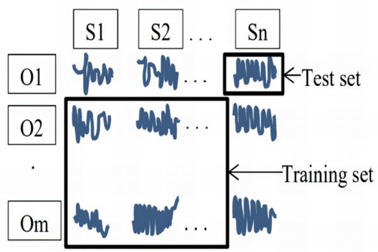
(a) Real and Posed Smiles [3].

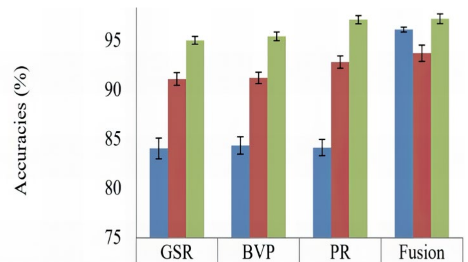
(b) Variations of pupilary responses [4].

Figure 2: Smiles and observed pupillary responses.

Figure 3b shows that when using PR features, the observer-independent approach achieves better accuracies, specifically 97.1%. The data of the test observer is eliminated from training in this strategy, but information from comparable smiler videos watched by other observers remains in the training set. The smiler-independent strategy, on the other hand, obtains a good accuracy of 92.8% utilising PR features, where information from similar observers is maintained in the training set but data from the test smilers' movies is not used during training. Our independent approach, on the other hand, has a lower accuracy of 84.1% due to greater robustness during training. When a classifier is evaluated with a dataset that contains similar data to that encountered during training, it naturally performs better. The classifier in the observer-independent approach is biased with data from smile videos during training, making it less accurate. Similarly, information from observers biases the smiler-independent approach. Both approaches deal with biasing data added during classifier training, which results in lower accuracy when applied to wholly fresh data. The independent approach, on the other hand, is trained without bias from either observers or smilers. As a result, its results are actually more accurate (as opposed to measured accuracy) and free of bias, with the reported accuracy predicted to hold on completely new data.
We note that fusion of multiple physiological signals improves the independent approach more than other approaches.

Given that the smiler-independent and observer-independent approaches are not completely free of bias, we pri- oritised the development and use of our independent approach. To improve the performances using our independent approach, we incorporate feature level fusion by combining features from PR, BVP, and GSR. Final results demonstrate that participants are verbally 52.7% (on average) to 68.4% (by voting) correct, but physiologically 96.1% correct when employing our independent approach with ensemble technique. In comparison, computer vision approaches reached an accuracy of 92% [8] using an end-to-end fully automated approach and 95% [9] using a feature engineering-based pipeline. The comparison supports our view that computer vision approaches attempt to use all of the information accessible in the images/videos, which is the same information available to the human observers. A somewhat superior performance by human physiological signals could be attributed to the human observers having more training from their lives prior to the experiments.





	GSR	BVP	PR	Fusion
Independent Approach	84	84.3	84.1	96.1
Smiler Independent	91.1	91.2	92.8	93.7
Observer Independent	95	95.4	97.1	97.2

(a) An independent approach.

(b) Performance comparison using ensemble classifier.

Figure 3: An independent approach for classifying smiles, S = smiler, O = observer, n = number of smile videos, m = number of observers, PR = Pupilary Responses, BVP = Blood Volume Pulses, GRS = Galvanic Skin Responses [7].

Extending our work on smiles, we found that physiological responses can predict the veracity of **anger** better than their own verbal responses. For example, acted anger can be expressed when the stimulus is not genuinely angry with an aim to manipulate the observer. Here, we aim to examine if the veracity of anger can be recognised from an observer's pupillary data with computational approaches [6], [10]. In [10], we use a genetic algorithm-based feature selection method to select time series pupillary features of observers who see acted and genuine anger as video stimuli. We then use the selected features to train a simple fully connected neural network and a two-stream neural network. Our results show that the two-stream architecture is able to achieve a promising recognition result with an accuracy of 93.6% when the pupillary responses from both eyes are available. It also shows that the genetic algorithm-based feature selection method can effectively improve the classification accuracy by 3.1%. We hope our work could help current research such as human-machine interaction and psychology studies that require emotion recognition.

The previous research revealed that people are generally poor at consciously distinguishing genuine and acted-anger facial expressions, with a mere 65% accuracy of verbal answers. This statistic motivates us to investigate whether a group of feed-forward neural networks can perform better using raw pupillary dilation signals from individuals [11]. Our results show that a single neural network cannot accurately discern the veracity of an emotion based on raw physiological signals, with an accuracy of 50.5%. Nonetheless, distinct neural networks using pupillary dilation signals from different individuals display a variety of genuineness for discerning the anger emotion, from 27.8% to 83.3%. By leveraging these differences, our novel Misaka neural networks can compose predictions using different individuals' pupillary dilation signals to provide a more accurate overall prediction than from even the highest performing single individual, reaching an accuracy of 88.9%. Further research will involve the investigation of the correlation between two groups of high-performing predictors using verbal answers and pupillary responses.

Given the significant variations in predicted behavioural reactions on smiles and anger and some initial work on fear and surprise, the unexpectedly parallel results suggest the possible application of this method for identifying the sincerity of diverse emotions. We propose that, using virtual reality or screen avatars to convey these emotions for the purpose of eliciting cooperation, persuasion, or engagement in contexts such as chronic condition management, caring for autistic individuals, or caring for elderly people could benefit significantly from the ability to measure the perceived authenticity of these emotional displays by human observers.

2.3 Doubt, Deception and Lies

Another aspect in this context is doubt, deceit recognition from observer's point of view [12], [13]. We live in a world surrounded by 'fake news' and manipulated information, so a system assisting people with knowing what information to trust would be beneficial. Our research investigates situations where the presenters themselves have doubts about the information they are delivering, and we detect this via advanced affective computing techniques [12].

First of all, we investigated the physiological underpinnings to detect the 'doubt effect' – where a presenter's subjective belief in some information has been manipulated. We constructed stimulus videos in which presenters delivered information that in some cases they were led to doubt, but asked to 'present anyway'. We then showed these stimuli to observers and measured their physiological signals (pupillary responses). Neural networks trained with two statistical features reached a higher accuracy in differentiating the doubt/ manipulated-belief compared to the observers' own veracity judgments, which is overall at chance level. We also trained confirmatory neural networks for the predictability of specific stimuli and extracted significant information on those stimulus presenters. We further showed that a semi-unsupervised training regime can use subjective class labels to achieve similar results to using the ground truth labels, opening the door to much wider applicability of these techniques, as expensive ground truth labels (provenance) of stimuli data can be replaced by crowd source evaluations (subjective labels). Overall, we showed that neural networks can be used on subjective data, which includes observer perceptions of the doubt felt by the presenters of information.

Our ability to detect this doubt effect is due to our observers' underlying emotional reactions to what they see, reflected in their physiological signals, and learnt by our neural networks.

This kind of technology using physiological signals collected in real time from observers could be used to reflect audience distrust, and perhaps could lead to increased truthfulness in statements presented via the Media.

We showed that a generalised neural network trained on a population base reached a higher accuracy in differen- tiating doubting and trusting information compared with the conscious veracity judgments from the same observers (See Figure 4). This recognition was significantly improved when multi-task learning (MTL) was used to account for individual differences or group differences. This is attributable to the ability of MTL which can both allow each individual to have a model customised for them and share the data of other people through the shared hidden layers in the neural network.

2.4 Behavioural Modification

2.4.1 Pavlok for breaking bad habits

This study proposes a feedback mechanism to 'break bad habits' using the Pavlok device. Pavlok utilises beeps, vibration and shocks as a mode of aversion technique to help individuals with behaviour modification [14]. While the device can be useful in certain periodic daily life situations, like alarms and exercise notifications, the device relies on manual operations that limit its usefulness.
To this end, we design a user interface to generate an automatic feedback mechanism that integrates Pavlok and a deep learning based model to detect certain behaviours via an integrated user interface i.e. mobile or desktop application. Our proposed solution is implemented and verified in the context of snoring, which first detects audio from the environment following a prediction of whether the audio content is a snore or not. Based on the prediction of the deep learning model, we use Pavlok to alert users for preventive measures. We believe that this simple solution can help people to change their atomic habits, which may lead to long-term benefits.

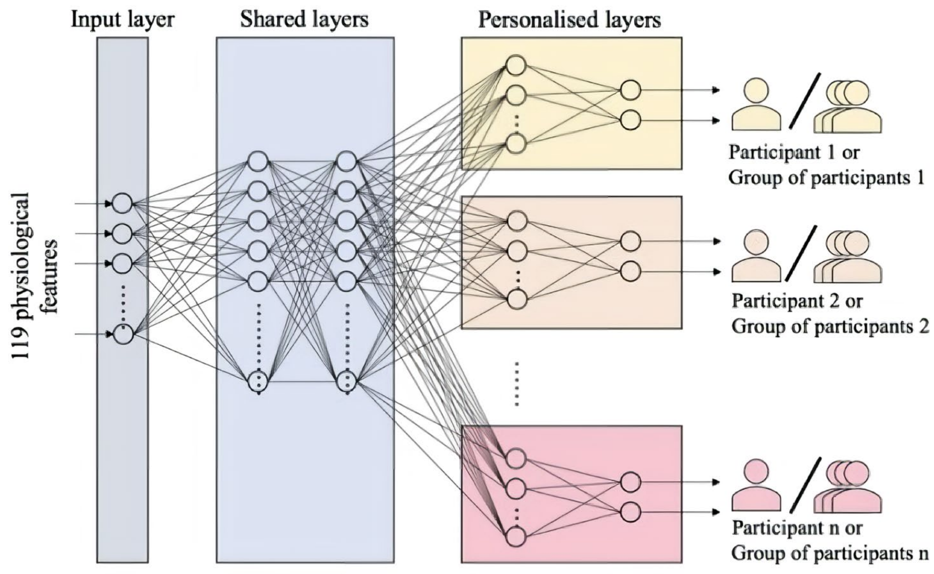


Figure 4: Deceit recognition framework.

2.5 Static Facial Expression

To advance responsive and responsible AI in human face-related research, quality data from real-world settings is crucial. Most existing datasets for analyzing human facial expressions come from controlled lab environments, and therefore, we introduce a new database, Static Facial Expressions in the Wild (SFEW), derived from temporal ex- pressions in movies [15]. Prior research has proposed various effective methods, but inconsistencies in databases and protocols hinder meaningful comparisons. To address this, we suggest a person-independent training and testing protocol for expression recognition.

2.6 Empathy, Distress and Personality Traits

On a very broad aspect, empathy refers to the capacity to comprehend and express appropriate emotions in response to the emotions, perspectives, and beliefs of others. The development of interpersonal relationships and the reduction of stress and unhappiness among people in our society are fundamentally dependent on empathy. Think about a scenario where a family member becomes unwell and causes you to experience emotional discomfort. Genuine support from a coworker after revealing your emotions about your family member may help you feel much better and strengthen your relationship with your coworker. Empathy is important in a variety of real-world human interactions, such as the relationships between patients and doctors [16], teachers and students [17], and even humans and robots [18]. In order to increase communication and patient outcomes, empathic clinicians are better able to comprehend their patients' problems [16]. Empathy plays a crucial role in education – especially with the move towards online learning brought on by the COVID-19 pandemic – in helping teachers understand their students' emotional states and foster a happy learning environment [17]. Socially assistive robots that recognise and respond empathically to the emotional states of the elderly, stroke survivors, and patients with autism spectrum disorder or Alzheimer's disease can improve assistance and care.

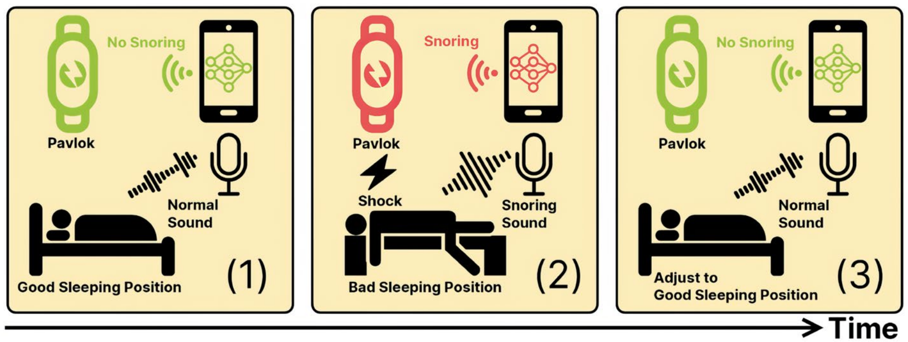


Figure 5: Pavlok-Nudge is an interactive, nearly real-time software application aiming to aid individuals for atomic behaviour modification. Here, we present the use case for snoring behaviour modification. Our proposed solution consists of the wrist borne Pavlok and a smartphone (see 1, 2, 3). The microphone on the mobile device captures audio which is further passed through a deep learning-based framework to detect snoring sounds. Upon snoring detection, the mobile device provides a signal to the Pavlok to nudge the user for posture change to stop snoring (see 2). The scenarios 1, 2 and 3 are happening in sequence over time.

In [19], we propose to detect people's empathy, distress and personality traits (Conscientiousness, Openness, Ex- traversion, Agreeableness and Stability) from their written essays and conversations in response to newspaper articles involving harm to individuals, nature or organisations (Figure 6). Given that empathy and these personality traits are subjective and depend on demographic factors [20], we leveraged demographic information after applying appropriate templates to transform numerical data into meaningful text information. The newspaper articles in the dataset were lengthy, and so we summarised lengthy articles using a pre-trained language model (PLM). We further leveraged another PLM for rephrasing textual contents (essays, articles and demographic sentences) as a means of data augmen- tation. Finally, to predict empathy and personality traits, we experimented with three PLMs (BERT, DistilBERT and ALBERT) with necessary hyperparameter tuning by the Optuna hyperparameter optimisation framework. Our system achieved its best Pearson correlation coefficient of 0.187 and 0.750 between the ground truth and predicted empathy scores from written essays and conversations, respectively. We make the codes and evaluation accessible to everyone at ***https://github.com/hasan-rakibul/WASSA23-empathy-emotion***.

Given that the above performance on the written essay was sub-optimal compared to the empathy prediction on conversations (0.187 vs 0.750), we investigated further. Apart from the primary text data used for empathy prediction, a significant distinction between these two tasks (empathy prediction on essays ver-sus conversations) lies in the annotation protocol employed: self-assessment by all participants (essays) as opposed to the controlled annotation of all samples by three external anno-tators (conversations). While crowdsource annotation offers a quick and straightforward means of obtaining data, it is sus-ceptible to inaccuracies and misinformation, as highlighted by Sheehan [21]. Consequently, addressing the issue of annotation noise becomes a crucial concern. To address this issue, we intro-duce a large language model (LLM)-guided empathy prediction workflow (Figure 7).

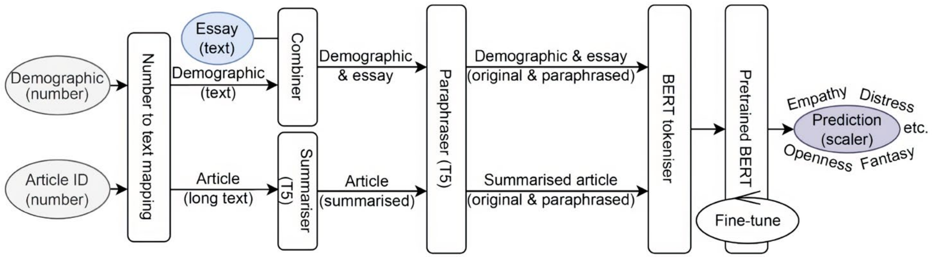


Figure 6: Our overall system for predicting empathy and personality traits, as we propose in [19]. First, we convert numerical features into semantically meaningful text. Next, we harness a pre-trained language model (PLM) based on T5 to summarise lengthy articles. As a method of data enrichment, we rephrase the textual contents. Finally, we fine-tune a BERT-based PLM to predict empathy levels and various other personality traits.

This system first corrects annotation errors, ensuring the reliability of the annotations for conventional empathy prediction models like BERT-based PLMs. Our approach systematically chooses between LLM and self-assessed annotations on a per-sample basis. In our earlier study [19], we convert numerical demographic information to fixed sentences. We hypothesise that naturally varying sentences would benefit the prediction and hence, we utilised the GPT-3.5 LLM to transform numerical demographic data into semantically meaningful textual sequences, enabling seamless integration. Through extensive experimentation on three datasets, which involve predicting individuals' empathy levels towards newspaper articles, we demonstrate the effectiveness of the method we propose compared to other approaches. We construct a RoBERTa-based model and train it using a combination of LLM and self-assessed annotations, achieving state-of-the-art performance.

Apart from text sequences, empathy prediction in real-life video interactions is important and has real-life impli- cations in evaluating people's empathic capabilities in various contexts, such as patient-doctor and teacher-student interactions. Despite its importance, computational empathy in videos remains relatively unexplored compared to computational emotions and facial expressions [25]. Prior works on computational empathy from videos [26], [27] have predominantly relied on scripted or semi-scripted interactions, falling short in capturing the intricacies and subtleties of real-life human interactions. These studies have also overlooked the potential insights offered by physiological signals, which are recognised for carrying crucial emotional information beyond what is consciously available [28].

We address these gaps by introducing a multimodal deep-learning architecture to predict empathy in videos (Fig- ure 8). Our architecture leverages a combination of multimodal signals, including audio signals, audio transcripts, facial expressions and physiological indicators such as heart rate and skin conductance. To bolster existing datasets, we collected real-life interaction videos, along with audio and physiological signals. Additionally, we employ a neural architecture search approach to evaluate multiple architectural alternatives and determine the most effective one, primarily targeting edge devices. The successful execution of this model holds substantial implications for improving the quality of social interactions, as it enables the measurement of empathy in a systematic and automated manner.

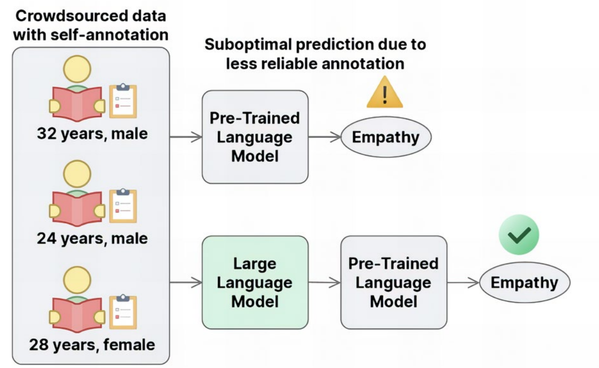


Figure 7: Comparison between a standard empathy prediction process utilising a pre-trained language model, as seen in previous works [22]–[24] and our novel LLM-guided approach. The conventional workflow, which relies on crowdsource data, often suffers from the inherent noise in the annotations, leading to less-than-ideal predictions. In contrast, our proposed workflow leverages the power of LLM to automatically enhance or redefine noisy annotations, resulting in superior performance when we compare to the standard approach.

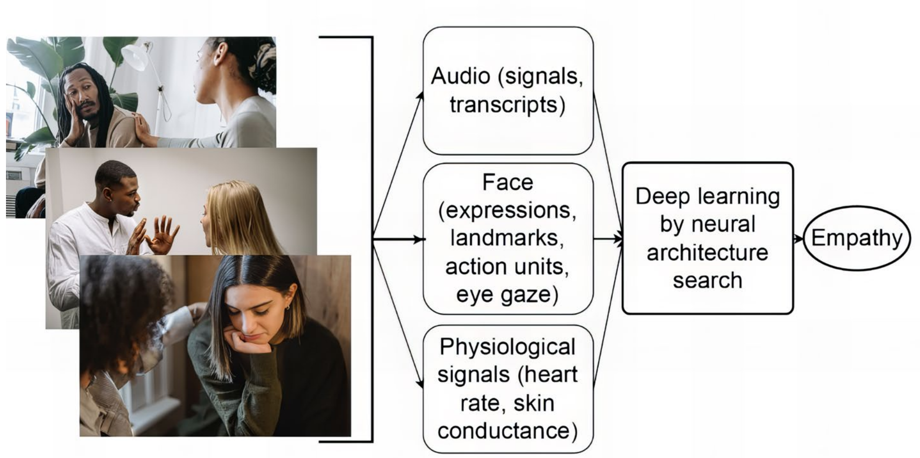


Figure 8: Overall framework of empathy detection from multimodal signals.

3. Interpersonal Responsive AI

In the case of *interpersonal* interactions, we investigate the holistic aspect of two and more people interacting with each other from emotion, cohesion and influential/dominant person perspectives in images and videos [29]–[32] as shown in Figure. 10. There are several attributes such as facial expression [15], gaze direction [33], headpose [34], and social context which influences subjective perceptions.

3.1 Group Emotion Estimation

The recent advancement of social media has given users a platform to socially engage and interact with a larger population. Millions of images and videos are being uploaded everyday by users on the web from different events and social gatherings. There is an increasing interest in designing systems capable of understanding human manifestations of emotional attributes and affective displays. As images and videos from social events generally contain multiple subjects, it is an essential step to study these groups of people. To this end, we study the problem of happiness intensity analysis of a group of people in an image using facial expression analysis [35]–[37]. A user perception study is conducted to understand various attributes, which affect a person's perception of the happiness intensity of a group. We identify the challenges in developing an automatic mood analysis system and propose three models based on the attributes in the study. An 'in the wild' image-based database is collected. To validate the methods, both quantitative and qualitative experiments are performed and applied to the problem of shot selection, event summarisation and album creation. The experiments show that the global and local attributes defined in the paper provide useful information for theme expression analysis, with results close to human perception results (See Figure 9).

Figure 9: Left: The overview of HAPPEI database [35]. Right: Group happiness intensity estimation using min- span tree. Different aspects such as social context, relative position of people, event etc play crucial role in a Group's Mood [35].

3.2 Group Cohesion Estimation

The cohesiveness of a group is an essential indicator of the emotional state, structure and success of the group. We study the factors that influence the perception of group-level cohesion and propose methods for estimating the human- perceived cohesion on the group cohesiveness scale [29], [38] (See Figure 10). To identify the visual cues (attributes) for cohesion, we conducted a user survey. Image analysis is performed at a group level via a multi-task convolutional neural network. A capsule network is explored for analysing the contribution of facial expressions of the group members on predicting the Group Cohesion Score (GCS). We add the GCS to the Group Affect database and propose the 'GAF- Cohesion database'. The model we propose performs well on the database and achieves near human-level performance in predicting a group's cohesion score. It is interesting to note that group cohesion as an attribute, when jointly trained for group-level emotion prediction, helps in increasing the performance for the later task. This suggests that group-level emotion and cohesion are correlated. Further, we investigate the effect of face-level similarity, body pose and subset of a group on the task of automatic cohesion perception.

Figure 10: Frames from VGAF dataset [39] across different emotion and cohesion categories. This is a third person subjective point of view.

3.3 Most Influential/Dominant Person

Estimating the most dominant person in a social interaction setting is a challenging feat even with the advancement of deep learning techniques due to the problem's complexity, limited availability of labelled data and subjective biases in annotations [32] (See Figure 11). To this end, we aim to reformulate the problem of detecting the Most Dominant Person (MDP) as a person ranking problem by utilizing person-specific attributes such as facial gestures, eye gaze, vi- sual attention and speaking patterns. The framework we propose, attributed *Graph*-based dominant person ranking in social *I* n*T* erac*TI* on videos, *GraphITTI*, learns generic and robust person rankings on top of context level features [40]. To inject domain knowledge into the *GraphITTI* framework, we consider inter-personal and intra-personal aspects along with spatiotemporal context patterns. Our extensive quantitative analysis suggests that *GraphITTI* framework performs favourably over the current state-of-the-art for dominant person detection and ranking.

Figure 11: Most Influential/Dominant Person perception [32]: The left image is a group where the baby in the centre is the most 'important'. In the centre image, although it is a friend circle, still we can see that the person holding the phone is the 'important' one. Finally, in the rightmost image it is about socially prominent people but without this info, an AI model model tries to predict 'Who is the most important/dominant person in a conversation?'

4. Physiological Sensors

We also investigate the connection between music and human physiological responses, which has the potential to improve music therapy techniques and potentially contribute to addressing mental disorders and epileptic seizures triggered by specific musical stimuli. In [41], we examine the impact of distinct music genres on participants' electro- dermal activity.

4.1 Feature extraction from Physiological Sensors

In our study involving physiological signals collected from participants, it's common practice to reduce the dimen- sionality and complexity of the data by extracting relevant features. This process helps in making the data more manageable for analysis and reduces computational costs. The most common features used in emotion related domain are as follows:

1. *Mean of Filtered and Normalized Signals:* This feature calculates the average value of the filtered and normalized physiological signals, providing an overall measure of signal intensity.
2. *Maximum and Minimum of Filtered Signals:* These features capture the highest and lowest values within the filtered signals, indicating signal extremes.
3. *Standard Deviation and Interquartile Range of Filtered Signals:* Standard deviation measures the dispersion or spread of the data, while interquartile range quantifies the

spread between the upper and lower quartiles, providing insights into signal variability.

4. *Variance and Kurtosis of Filtered Signals:* Variance measures the average of the squared differences from the mean, representing signal variability. Kurtosis measures the "tailedness" or shape of the signal distribution.

5. *Number of Peaks for Periodic Signals:* For periodic signals, this feature counts the number of peaks, which can be indicative of certain physiological patterns.

6. *Means of the Absolute Values of the First and Second Differences of the Normalized and Filtered Signals:* These features assess the rate of change in the signals, helping to identify abrupt changes or trends.

7. *Mean of the First 10 Points Derived Using Welch Power Spectral Density:* Welch Power Spectral Density analysis calculates the distribution of signal power across different frequencies. Taking the mean of the first 10 points can provide a summary of the signal's dominant frequency components.

These features offer a comprehensive view of the physiological signals' characteristics, including their central tendency, variability, shape, frequency components, and rate of change. By extracting and analyzing these features, one can gain insights into how physiological signals vary across different conditions or participants, which may be useful for tasks like emotion recognition or other applications in physiological data analysis.

4.2 Gingerbread Animation

This approach provides a creative and engaging way to visualize physiological data and it is interactive with multiple signals over time duration. The Gingerbread Animation is useful for conveying complex information to a wider audience or for gaining visual insights into the data's patterns and trends. For this purpose, physiological signals, including EDA (Electrodermal Activity), BVP (Blood Volume Pulse), and PD (Pupillary Dilation), were analyzed over a time duration. Then, data for each participant were segmented based on the length of audio stimuli. In Figure 12, EDA, BVP, and PD values were represented in red, blue, and green colors, respectively, and ST (skin temperature) is displayed in grey. The Gingerbread Animation uses a stylized 2D representation of a human body to visualize time series of physiological signals, i.e., BVP, PD (Pupil Diameter), EDA, and ST. The locations on the gingerbread figure correspond to the locations where these signals are generated. For example, PD is displayed in the right eye, BVP in the heart, EDA in the left wrist, and ST in the right foot area. These colors can mix and create various shades on the gingerbread figure's surface. The result is a sequence of images that can be compiled to create a video that visually represent each experimental trial while retaining a representation of the physiological signals' dynamics.

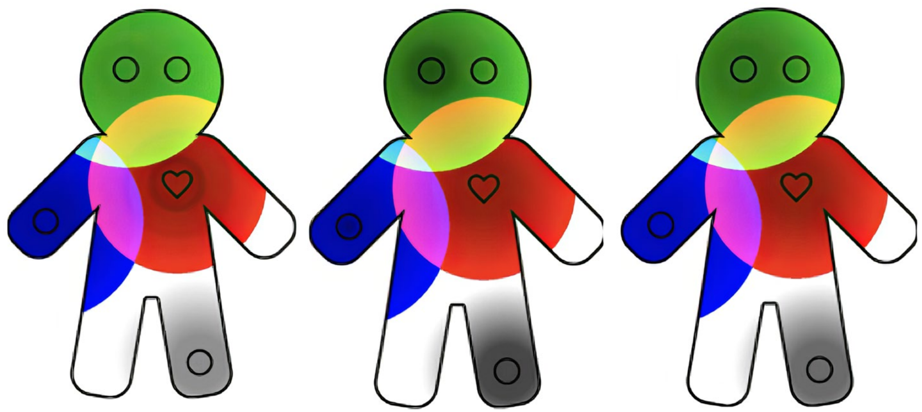


Figure 12: Physiological Signals Representation in an Animation (Red = BVP, Blue = EDA, Green =PD, Gray = ST)

5. Responsible AI

Whether it is *intrapersonal* or *interpersonal* interaction, there is always a need to protect personal identity information. However, it is challenging to remove such information that has infinite classes (open set).

5.1 Privacy of Physiological Signals

Physiological signals all contain personal identity information. We will use the example of electroencephalogram (EEG) signals. EEG data is sensitive to noise and rich with diverse brain-related information. In [42], we focus on filtering out information-preserving features from EEG signals. We present two key components: (1) a discriminative feature extractor capable of classifying multiple labels from short-term EEG signals and (2) an approach to eliminate undesirable features from EEG data guided by our feature extractor. This process resembles a person doing selective listening at a noisy gathering, analogous to the cocktail party problem in machine learning.

We leverage image-wise autoencoder to extract distinctive and resilient features from EEG images. In the process of training the autoencoder, it effectively minimizes the reconstruction loss, even achieving a very low level for the test dataset, which allows it to serve as a generator within the GAN framework. We leverage Convolutional Neural Networks (CNNs) to extract relevant features (Figure 13).

Leveraging the success of our discriminative model for short-term EEG data, the feature filter (Figure 14) is designed as an end-to-end framework. It is trained to transform EEG signals containing unwanted features into signals devoid of those unwanted elements.

Our experiments with an alcoholism dataset (UCI EEG dataset) demonstrate that our proposed model can elim- inate over 90% of alcoholism-related information from EEG signals while losing only an average of 4.2% in useful feature accuracy.

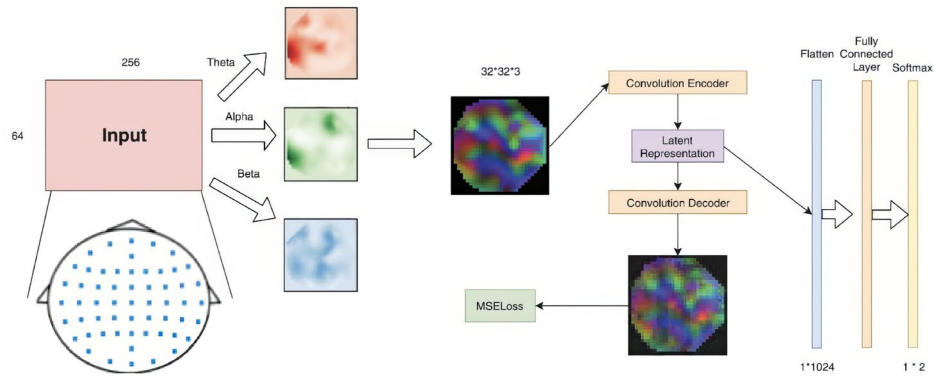


Figure 13: Architecture of the image-wise autoencoder [42]. It comprises three frequency bands used to create EEG images. These EEG images serve as input for the classification model, which simultaneously conducts image reconstruction and classification tasks. The model's training is guided by a combined cross-entropy and mean squared error (MSE) loss components. We convert EEG traces to images to be able to leverage the advanced computer vision models developed in that community.

In [44], we propose an approach to disguise the identity information in EEG signals with dummy identities while preserving the key features. The dummy identities are obtained by applying a grand average on EEG spectra across the subjects within a group that has common attributes. The personal identity information in original EEGs is transformed into disguised ones with a CycleGAN-based EEG disguising model. With the constraints added to the model, the features of interest in EEG signals can be preserved. We evaluate the model by performing classification tasks on both the original and the disguised EEG and compare the results. For evaluation, we also experiment with ResNet classifiers, which perform well, especially on the identity recognition task, with an accuracy of 98.4%. The results show that our EEG disguising model can hide about 90% of personal identity information and can preserve most of the other key features.

5.2 Privacy of Images and Videos

Images and videos also contain useful personal attributes that, when mishandled or exploited, can potentially violate people's privacy. These attributes may encompass facial features, body language and identifiable surroundings. The most important of these is generally facial features, as this is how people primarily recognise each other.

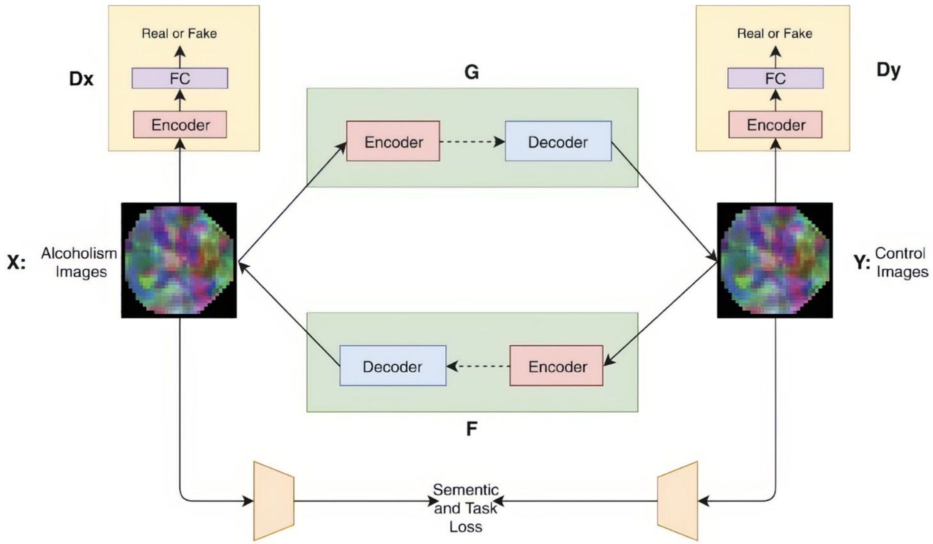


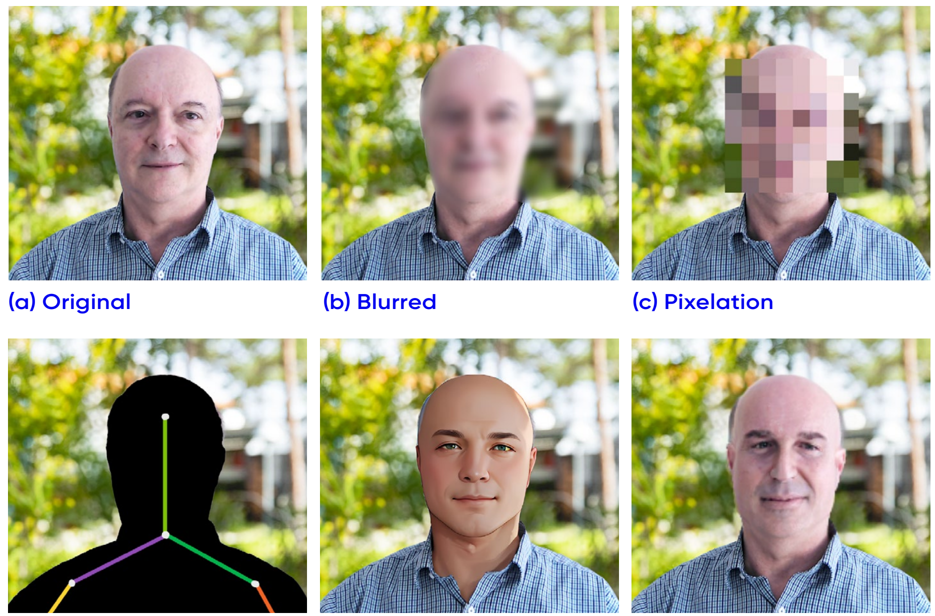
Figure 14: Architecture of the feature filter [42]. It involves training it to directly transform EEG images containing unwanted features (alcoholism-related information) into EEG signals without those unwanted attributes, often repre- sented as control images. Adapted from the structure of CycleGAN [43], this feature filter incorporates an additional classifier (C), which contributes semantic and task-related loss functions. These additional components help ensure that essential features are retained throughout the feature filtering process.

5.2.1 Blurring, Pixelation, Skeletonisation or Avatarisation

Common strategies of anonymisation include blurring, pixelation, skeletonisation and avatarisation (see Figure 15b to Figure 15e). This could lead to a significant loss of information if we blur, pixelate, skeletonise or avatarise the images/videos. Such loss of information is detrimental when small facial gestures or other non-verbal cues are important. Avatarisation may contain facial expressions but often looks unnatural. If sufficient video is available, it is possible to de-blur images and retrieve the original face. With even more video, pixelation can, in principle, also be reversed, though with the large pixels used in our example, this would be impractical.

If we intend to collect data for some purpose, anonymisation, which obscures identity too well, can be detrimental. For example, in an interpersonal interaction, with two pixellated or skeletonised heads, it may be difficult for an algorithm to consistently track which person is which. This applied to a lesser extent for blurring and avatarisation.

The final deepfake face looks like a normal image, just the identity is changed. Please note that the facial expression is retained, which may be significant for later re-analysis of the stored deepfake version. We assume the original face image is not stored to preserve privacy.



(a) Original (b) Blurred (c) Pixelation

(d) Skeletonisation (e) Avatar (f) Deepfake

Figure 15: An original image and its anonymised version using different techniques.

5.2.2 Disguising personal identity via Deepfake faces (deepfaces)

For better anonymisation, we open up the avenue for 'deep fake faces for good', where we can replace the original face information anonymously with the same approximate age, ethnicity and gender in real time, with the transformation ideally performed in-camera or on the edge [45] (Figure 15f). Most deepfake detection methods focus on detecting spatial and/or spatio-temporal changes in facial attributes. This is because available benchmark datasets contain mostly visual-only modifications. However, a sophisticated deepfake may include small segments of audio or audio- visual manipulations that can completely change the meaning of the content. We believe that this content-driven deepfake can be a powerful tool for anonymization [45].

We need to consider the potential applications of our deep fake faces, as this will impact on the actual approach used. Let us consider a potential use case of realistic deepfaces in aged care facilities for pre-fall wobbliness detection. It would be useful for the aged care staff to partially recognise the person walking in a wobbly way. Maybe we can tune our deepface model to make them look similar enough to look like a sibling? So, Mary's face is not shown, so her privacy is protected, yet to the staff who are monitoring the CCTV footage at the nursing station, the wobbly person looks a bit like Mary, so it is Mary – which is important when they jump up to help the wobbly person. Further, Mary's deepface should look the same today as it will tomorrow. Since we would assume there is no central database (privacy issues), the camera needs to retain enough differentiating information between the relatively small number of faces it sees in an aged care facility to be able to consistently generate the same deepface for Mary. This also suggests we need alternative approaches for public settings where there will be many more indentities – but we may there not insist on reproducibility of deepfaces.

Of course, we need to discuss perceptual anonymisation versus true anonymisation. In our aged care scenario, the individual and their family probably mostly care that Mary's face is not visible, whereas the aged care facility probably needs to be able to show they took appropriate action when a deep-fake-person is wobbly so they need to know or be able to find out who is wobbly. Another issue arises from the notion of consistency of deepface as mentioned above. To fully anonymise the video feed, we would need to go beyond deepfaces and change clothing appearances, skin texture to recognisable wrinkle patterns or tattoos. We would also need to change the dynamics of movement: body language, gait and so on. In today's world where we share and leak so much information to Google, Facebook and so on, we believe human perceptual privacy (faces) should be the key privacy goal.

6. Conclusion and Future Work

We have described a longitudinal body of work in the construction of Responsive AI, increasingly focused on appli- cations to signals. The necessary complementary approach of Responsible AI is a critical and evolving concept that encompasses various principles, practices, and ethical considerations to ensure that artificial intelligence technologies are developed, deployed, and used in a manner that benefits humanity while minimizing potential harm.

In conclusion, responsible AI is a holistic approach that requires ethical considerations, fairness, transparency, accountability, and ongoing commitment to addressing the societal impacts of AI technologies. Achieving responsible AI involves collaboration among various stakeholders and the recognition that AI should benefit all of humanity while minimizing harm. It is an ongoing journey that requires continuous vigilance, education, and adaptation to ensure that AI aligns with our ethical and societal values.

The future of responsible AI holds several exciting possibilities and challenges as technology continues to advance. Here are some key aspects to consider when thinking about the future of responsible AI:

1. **Enhanced Ethical Frameworks:** As AI technologies become more sophisticated, there will be a growing need for more robust and adaptable ethical frameworks. These frameworks will need to address complex ethical dilemmas arising from AI's increasing autonomy, decision-making power, and interaction with human society.

2. **AI Regulation:** Governments and regulatory bodies are starting to play a more prominent role in shaping the responsible use of AI. We can see the beginning of the development of comprehensive AI regulations that set standards for fairness, transparency, privacy, and accountability. These regulations will aim to ensure that AI benefits society while minimizing risks.

3. **Bias and Fairness Mitigation:** Efforts to mitigate bias in AI will continue to evolve. Advanced techniques, such as adversarial debiasing and fairness-aware machine learning, will be developed to reduce bias in AI systems and ensure fairness in decision-making processes.

4. **Explainability and Transparency:** The demand for AI explainability and transparency will intensify. Re- searchers will work on making AI models more interpretable and understandable, helping users and stakeholders trust AI systems and better understand their decision-making processes.

5. **Human-AI Collaboration:** The future of responsive AI will involve closer collaboration between humans and AI systems. Human-AI partnerships will become more integrated into various domains, enhancing productivity, decision-making, and problem-solving capabilities.

6. **Privacy-Preserving AI:** Techniques for privacy-preserving AI will continue to advance, allowing AI to work with sensitive data while protecting individuals' privacy rights. Federated learning, differential privacy, and secure multiparty computation are

examples of privacy-preserving approaches. A new approach is the generation of synthetic data which matches important distributional properties of private data, and can be used to train models which are then tested on the private data which therefore does not need to 'leave the building'. We consider privary-preserving AI as the keystone of responsible AI.

7. **Ethical AI Research:** There will be a growing emphasis on conducting research specifically focused on the ethical aspects of AI, including the exploration of novel ethical theories, frameworks, and guidelines tailored to AI systems.

8. **Corporate Responsibility:** Companies and organizations will be expected to demonstrate a strong commit- ment to responsible AI practices. Ethical AI development and deployment will become a competitive advantage, and consumers may prioritize products and services that align with their values.

9. **Ethical AI in Emerging Technologies:** As AI intersects with emerging technologies like quantum computing, biotechnology, and autonomous vehicles, ethical considerations will become even more critical. Responsible AI frameworks will need to adapt to these evolving technologies.

In summary, the future of responsible AI holds the promise of more ethical, transparent, and accountable AI systems. However, it also poses challenges related to the evolving nature of technology, the need for adaptive ethical frameworks, and the importance of global collaboration. Responsible AI will continue to be a dynamic field, evolving to meet the ethical and societal demands of AI in an ever-changing world.

References

1. N. Sharma and T. Gedeon, "Objective measures, sensors and computational techniques for stress recognition and classification: A survey," *Computer methods and programs in biomedicine*, vol. 108, no. 3, pp. 1287–1301, 2012.

2. N. Sharma and T. Gedeon, "Modeling observer stress: A computational approach," *Intelligent Decision Tech- nologies*, vol. 9, no. 2, pp. 191–207, 2015.

3. M. Z. Hossain and T. Gedeon, "Discriminating real and posed smiles: Human and avatar smiles," in *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, 2017, pp. 581–586.

4. M. Hossain, T. Gedeon, R. Sankaranarayana, D. Apthorp, A. Dawel, *et al.*, "Pupillary responses of asian observers in discriminating real from fake smiles: A preliminary study," in *Measuring Behavior*, 2016, pp. 170–176.

5. M. Z. Hossain, T. Gedeon, and R. Sankaranarayana, "Using temporal features of observers' physiological mea- sures to distinguish between genuine and fake smiles," *IEEE Transactions on Affective Computing*, vol. 11, no. 1, pp. 163–173, 2020. doi: ***10.1109/TAFFC.2018.2878029***.

6. L. Chen, T. Gedeon, M. Z. Hossain, and S. Caldwell, "Are you really angry? detecting emotion veracity as a proposed tool for interaction," in *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, 2017, pp. 412–416.

7. M. Z. Hossain and T. D. Gedeon, "An independent approach to training classifiers on physiological data: An example using smiles," in *Neural Information Processing: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, December 13–16, 2018, Proceedings, Part II 25*, Springer, 2018, pp. 603–613.

8. Y. Yang, M. Z. Hossain, T. Gedeon, and S. Rahman, "Realsmilenet: A deep end-to-end network for spontaneous and posed smile recognition," in *Proceedings of the Asian Conference on Computer Vision*, 2020.

9. H. Dibeklio˘glu, A. A. Salah, and T. Gevers, "Recognition of genuine smiles," *IEEE Transactions on Multimedia*, vol. 17, no. 3, pp. 279–294, 2015.

10. C. Huang, X. Zhu, and T. Gedeon, “A genetic feature selection based two-stream neural network for anger veracity recognition,” in *Neural Information Processing: 27th International Conference, ICONIP 2020, Bangkok, Thailand, November 23–27, 2020, Proceedings, Part I 27*, Springer, 2020, pp. 3–11.

11. Z. Qin, T. Gedeon, L. Chen, X. Zhu, and M. Z. Hossain, “Artificial neural networks can distinguish genuine and acted anger by synthesizing pupillary dilation signals from different participants,” in *Neural Information Process- ing: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, December 13–16, 2018, Proceedings, Part V 25*, Springer, 2018, pp. 299–310.

12. X. Zhu, Z. Qin, T. Gedeon, R. Jones, M. Z. Hossain, and S. Caldwell, “Detecting the doubt effect and subjec- tive beliefs using neural networks and observers’ pupillary responses,” in *Neural Information Processing: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, December 13-16, 2018, Proceedings, Part IV 25*, Springer, 2018, pp. 610–621.

13. X. Zhu, T. Gedeon, S. Caldwell, R. Jones, and X. Gu, “Deceit detection: Identification of presenter’s subjective doubt using affective observation neural network analysis,” in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, IEEE, 2020, pp. 3174–3181.

14. S. Ghosh, M. R. Hasan, P. Agrawal, *et al.*, “Pavlok-nudge: A feedback mechanism for atomic behaviour modifi- cation,” *arXiv preprint arXiv:2305.06110*, 2023.

15. A. Dhall, R. Goecke, S. Lucey, and T. Gedeon, “Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark,” in *2011 IEEE international conference on computer vision workshops (ICCV workshops)*, IEEE, 2011, pp. 2106–2112.

16. B. D. Jani, D. N. Blane, and S. W. Mercer, “The role of empathy in therapy and the physician-patient relation- ship,” *Complementary Medicine Research*, vol. 19, no. 5, pp. 252–257, 2012. doi: ***10.1159/000342998***.

17. K. Aldrup, B. Carstensen, and U. Klusmann, “Is empathy the key to effective teaching? a systematic review of its association with teacher-student interactions and student outcomes,” *Educational Psychology Review*, vol. 34, no. 3, pp. 1177–1216, 2022. doi: ***10.1007/s10648-021-09649-y***.

18. M. Spitale, S. Okamoto, M. Gupta, H. Xi, and M. J. Matari´c, “Socially assistive robots as storytellers that elicit empathy,” *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 11, no. 4, pp. 1–29, 2022. doi: ***10.1145/3538409***.

19. M. R. Hasan, M. Z. Hossain, T. Gedeon, S. Soon, and S. Rahman, “Curtin OCAI at WASSA 2023 empathy, emo- tion and personality shared task: Demographic-aware prediction using multiple transformers,” in *Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, Toronto, Canada: Association for Computational Linguistics, Jul. 2023, pp. 536–541. doi: ***10.18653/v1/2023.wassa- 1.47***.

20. R. A. Borracci, H. C. Doval, L. Celano, A. Ciancio, D. Manente, and J. G. Calder´on, “Patients’ perceptions of Argentine physicians’ empathy based on the Jefferson scale of patient’s perceptions of physician empathy: Psychometric data and demographic differences,” *Education for Health*, vol. 30, no. 1, pp. 19–25, 2017. doi: ***10.4103/1357-6283.210513***.

21. K. B. Sheehan, “Crowdsourcing research: Data collection with amazon’s mechanical turk,” *Communication Monographs*, vol. 85, no. 1, pp. 140–156, 2018. doi: ***10.1080/03637751.2017.1342043***.

22. S. Tafreshi, O. De Clercq, V. Barriere, S. Buechel, J. Sedoc, and A. Balahur, “WASSA 2021 shared task: Predict- ing empathy and emotion in reaction to news stories,” in *Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Online: Association for Computational Lin- guistics, Apr. 2021, pp. 92–104. [Online]. Available: ***https://aclanthology.org/2021.wassa-1.10***.

23. V. Barriere, S. Tafreshi, J. Sedoc, and S. Alqahtani, “WASSA 2022 shared task: Predicting empathy, emotion and personality in reaction to news stories,” in *Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis*, Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 214–227. doi: ***10.18653/v1/2022.wassa-1.20***.

24. V. Barriere, J. Sedoc, S. Tafreshi, and S. Giorgi, “Findings of WASSA 2023 shared task on empathy, emotion and personality detection in conversation and reactions to news articles,” in *Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, Toronto, Canada: Association for Computational Linguistics, Jul. 2023, pp. 511–525. [Online]. Available: ***https://aclanthology.org/2023. wassa-1.44***.

25. A. Kaklauskas, A. Abraham, I. Ubarte, *et al.*, “A review of ai cloud and edge sensors, methods, and applications for the recognition of emotional, affective and physiological states,” *Sensors*, vol. 22, no. 20, 2022, issn: 1424-8220. doi: ***10.3390/s22207824***.

26. P. Barros, N. Churamani, A. Lim, and S. Wermter, “The omg-empathy dataset: Evaluating the impact of affective behavior in storytelling,” in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2019, pp. 1–7. doi: ***10.1109/ACII.2019.8925530***.

27. L. Mathur, M. Spitale, H. Xi, J. Li, and M. J. Matarí c, “Modeling user empathy elicited by a robot storyteller,” in *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2021, pp. 1–8. doi: ***10.1109/ACII52823.2021.9597416***.

28. M. Z. Hossain and T. Gedeon, “Classifying posed and real smiles from observers’ peripheral physiology,” in *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 2017, pp. 460–463. doi: ***10.1145/3154862.3154893***.

29. S. Ghosh, A. Dhall, N. Sebe, and T. Gedeon, “Automatic prediction of group cohesiveness in images,” *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1677–1690, 2020.

30. S. Ghosh, Z. Cai, P. Gupta, *et al.*, “Emolysis: A multimodal open-source group emotion analysis and visualization toolkit,” *arXiv preprint arXiv:2305.05255*, 2023.

31. S. Ghosh, A. Dhall, and N. Sebe, “Automatic group affect analysis in images via visual attribute and feature networks,” in *2018 25th IEEE International Conference on Image Processing (ICIP)*, IEEE, 2018, pp. 1967– 1971.

32. S. Ghosh and A. Dhall, “Role of group level affect to find the most influential person in images,” in *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018, pp. 0–0.

33. S. Ghosh, M. Hayat, A. Dhall, and J. Knibbe, “Mtgls: Multi-task gaze estimation with limited supervision,” in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2022, pp. 3223–3234.

34. S. Ghosh, A. Dhall, M. Hayat, and J. Knibbe, “Av-gaze: A study on the effectiveness of audio guided visual attention estimation for non-profilic faces,” in *2022 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2022, pp. 2921–2925.

35. A. Dhall, R. Goecke, and T. Gedeon, “Automatic group happiness intensity analysis,” *IEEE Transactions on Affective Computing*, vol. 6, no. 1, pp. 13–26, 2015.

36. A. Dhall, R. Goecke, S. Ghosh, J. Joshi, J. Hoey, and T. Gedeon, “From individual to group-level emotion recognition: Emotiw 5.0,” in *Proceedings of the 19th ACM international conference on multimodal interaction*, 2017, pp. 524–528.

37. A. Dhall, “Emotiw 2019: Automatic emotion, engagement and cohesion prediction tasks,” in *2019 International Conference on Multimodal Interaction*, 2019, pp. 546–550.

38. S. Ghosh, A. Dhall, N. Sebe, and T. Gedeon, “Predicting group cohesiveness in images,” in *2019 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2019, pp. 1–8.

39. G. Sharma, S. Ghosh, and A. Dhall, “Automatic group level affect and cohesion prediction in videos,” in *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, IEEE, 2019, pp. 161–167.

40. G. Sharma, S. Ghosh, A. Dhall, M. Hayat, J. Cai, and T. Gedeon, “Graphitti: Attributed graph-based dominance ranking in social interaction videos,” in *ACM ICMI 2023 workshop*, ACM, 2023, pp. –.

41. J. S. Rahman, T. Gedeon, S. Caldwell, R. Jones, M. Z. Hossain, and X. Zhu, “Melodious micro-frissons: Detecting music genres from skin response,” in *2019 International Joint Conference on Neural Networks (IJCNN)*, 2019, pp. 1–8. doi: ***10.1109/IJCNN.2019.8852318***.

42. Y. Yao, J. Plested, and T. Gedeon, “Information-preserving feature filter for short-term eeg signals,” *Neurocom- puting*, vol. 408, pp. 91–99, 2020. doi: ***10.1016/j.neucom.2019.11.106***.

43. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223– 2232.

44. S. Liu, Y. Yao, C. Xing, and T. Gedeon, “Disguising personal identity information in eeg signals,” in *Neural Information Processing: 27th International Conference, ICONIP 2020, Bangkok, Thailand, November 18–22, 2020, Proceedings, Part V 27*, Springer, 2020, pp. 87–95.

45. Z. Cai, S. Ghosh, T. Gedeon, A. Dhall, K. Stefanov, and M. Hayat, “‘glitch in the matrix!’: A large scale benchmark for content driven audio-visual forgery detection and localization,” *Computer Vision and Image Understanding*, 2023.

Powering Australia's technology brilliance

SUPPORTING LOCAL INDUSTRY

Private Cloud

It's Secured.

Perimeter DDOS Protection Powered by Fortinet
Hosted in Tier III Certified Data Centre ISO 27001
Multi-Factor Authentication MIFARE DESFire EVI encoded access card

It's Connected.

Connected to TPG Fabric NaaS, Choice of alternative Telco, ISP
Multi-Cloud Connectivity Megaport, IX, Telco
Comms Diversity Dual pathway from 'Meet Me' room

It's Uniquely Yours.

Tailored to Suit BUaaS, IaaS, NaaS
White Labelling MSP wholesale
Remote Hands In-house expertise, LI/L2 support

AI

Ready for your AI workload

dcalliance.com.au