Proceedings

Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies

18-20 June 2024
Alicante, Spain

visuAAL

goodbrother

Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies

Acknowledgments

This conference has been supported by:

- The visuAAL project on Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 861091 - https://cordis.europa.eu/project/id/861091.
- The COST Action GoodBrother - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living (CA19121), supported by COST (European Cooperation in Science and Technology) - https://www.cost.eu/actions/CA19121.
  COST (European Cooperation in Science and Technology) is a funding agency for research and innovation networks. Our Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation.
  www.cost.eu

Preface

The *Joint visuAAL-GoodBrother Conference on Trustworthy Video- and Audio-Based Assistive Technologies*, held from 18th to 20th June 2024 at the University of Alicante, Spain, brought together experts to explore the future of privacy-aware technologies for Active and Assisted Living (AAL). Organised by the *GoodBrother* COST Action (CA19121) and the *visuAAL* Marie Skłodowska-Curie Innovative Training Network, the event provided a platform for interdisciplinary discussions on the ethical, legal, and privacy challenges associated with audio and video monitoring in care environments.

The conference featured diverse presentations that covered innovative approaches to assistive technologies, such as enhancing robots' ability to understand natural language for home care, as well as the use of behavioural insights to improve the acceptance of camera-based systems by older adults. Researchers also discussed the legal frameworks surrounding digital health, comparing regulatory approaches in the EU and China, and examined how technologies such as digital twins could ensure compliance with privacy laws. These discussions reflected the conference's focus on balancing technological advancements with ethical responsibility and legal compliance.

A critical theme throughout the conference was the challenge of protecting user privacy without compromising safety. Presenters offered perspectives on how privacy-by-design principles can be integrated into AAL technologies, ensuring that they remain trustworthy and widely accepted by the public. Ethical issues, such as balancing autonomy and protection in assistive technologies for older people, were also explored in depth, with solutions that emphasised the need for responsible research and development.

By the end of the event, it was clear that advancing AAL technologies requires a collaborative, interdisciplinary approach, combining insights from law, healthcare, engineering, and social sciences. The presentations provided valuable contributions to the ongoing development of privacy-aware and user-centric assistive technologies, ensuring that future systems can support the well-being of older adults while respecting their privacy and dignity.

We extend our sincere thanks to all the participants, and we are confident that the ideas and solutions presented during the conference will continue to shape the future of AAL technologies.

Suggested citation

Francisco Florez-Revuelta (ed.) (2024): Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies. Alicante, Spain, 18-20 June 2024. DOI: 10.5281/zenodo.13969847.

Table of Contents

Irene Ballester and Martin Kampel (2024): Measuring dementia behaviours through depth sensors*. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Measuring dementia behaviors through depth sensors

Irene Ballester, Martin Kampel
Computer Vision Lab, TU Wien
irene.ballester@tuwien.ac.at, martin.kampel@tuwien.ac.at

Abstract

Human Behavior Modelling aims to analyze human behaviors using visual or 3D data. In dementia, behavioral changes are strongly correlated with cognitive decline, making their monitoring essential. To address privacy concerns, especially important for vulnerable populations, this study uses depth maps instead of RGB images. We present a novel Human Pose Estimation (HPE) method using point cloud sequences derived from depth data and discuss domain gaps between benchmark and real-world datasets.

Introduction

Dementia affects millions, manifesting not just as memory loss but through behavioral changes (WHO, 2021). Monitoring these changes is crucial for tracking progression and tailoring care (Ballester et al., 2024c). To address privacy concerns associated with RGB cameras, we utilize depth sensors that capture 3D data without identifiable features. Our aim is to develop a robust framework for measuring dementia behaviors. In this work, we present a novel HPE method that processes point cloud sequences, and discuss domain adaptation techniques to improve performance in real-world settings.

Methodology

HPE from Point Cloud Sequences (SPiKE)

SPiKE takes point cloud sequences and predicts the 3D coordinates of the joints (Ballester et al., 2024a). Unlike previous works that process timestamps independently (Zhou et al., 2020), SPiKE leverages temporal information by using a Transformer to encode the spatio-temporal structure of the sequence.

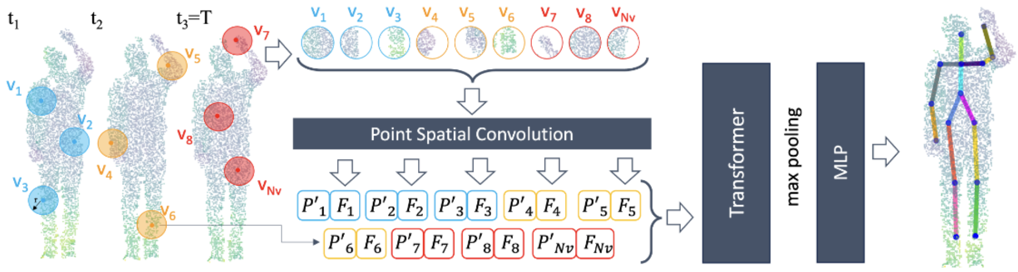


Figure 1. SPiKE pipeline: each point cloud is divided into local volumes for feature extraction using point spatial convolution, ensuring efficient processing while preserving spatial integrity.

In Ballester et al. (2024a), we show that SPiKE outperforms existing state-of-the-art methods on the ITOP benchmark with an mAP of 89.19%, Ablation studies highlight the benefits of using sequence information and maintaining spatial structure. Future work will assess SPiKE on real-world data.

BAD Dataset and domain gaps

The Bathroom Activities Dataset (BAD) consists of depth sequences from 19 dementia patients performing various activities and presents real-world challenges like occlusions, variable sensor angles, and unbalanced class distributions, making it ideal for testing robust models. Benchmarks often fail to reflect these complexities, resulting in significant performance gaps when models trained on controlled data are applied to real-world environments (Ballester and Kampel, 2024b; Fan et al., 2021).

Domain adaptation techniques

To bridge this gap, we explore Test-Time Training (TTT) (Sun et al., 2020) for point cloud sequences. The model's encoder is trained in the source domain with two heads: one for the primary task (e.g., HAR or HPE) and another for a self-supervised task, optimizing a combined loss. In the target domain, the encoder is fine-tuned while both heads remain frozen, allowing adaptation to target data using self-supervised learning. The challenge is finding an self-supervised task that effectively adapts the feature extractor to the target domain.

Conclusions

Robust and effective measurement of dementia behavior from depth maps requires not only the development of HAR and HPE models, but also domain adaptation strategies for real-world applicability. This study presents a new HPE model and discusses the application of TTT to point cloud sequences, with the ultimate goal of achieving robust performance in practical settings.

Acknowledgments

This publication is based upon work from COST Action GoodBrother - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living (CA19121), supported by COST (European Cooperation in Science and Technology). This publication is part of the visuAAL project on Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 861091.

References

Ballester, I., Peterka, O., & Kampel, M. (2024a). SPiKE: 3D Human Pose from Point Cloud Sequences. arXiv preprint arXiv:2409.01879

Ballester, I., & Kampel, M. (2024b). Action recognition from 4D point clouds for privacy-sensitive scenarios in assistive contexts. In International Conference on Computers Helping People with Special Needs (pp. 359-364). Cham: Springer Nature Switzerland.

Ballester, I., Gall, M., Münzer, T., & Kampel, M. (2024c). Vision-Based Toilet Assistant for People with Dementia in Real-Life Situations. In 2024 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops) (pp. 141-147). IEEE.

Fan, H., Yang, Y., & Kankanhalli, M. (2021). Point 4D Transformer Networks for Spatio-Temporal Modeling in Point Cloud Videos. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 14204-14213).

Sun, Y., Wang, X., Liu, Z., Miller, J., Efros, A., & Hardt, M. (2020). Test-time training with self-supervision for generalization under distribution shifts. In International Conference on Machine Learning (pp. 9229-9248). PMLR.

WHO (2021) Global status report on the public health response to dementia.

Zhou, Y., Dong, H., & El Saddik, A. (2020). Learning to estimate 3D human pose from point cloud. IEEE Sensors Journal, 20(20), 12334-12342.

***Wiktor Mucha, and Martin Kampel (2024): Understanding Human Behaviour With Wearable Cameras Based on Information From the Human Hand**. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Understanding Human Behaviour With Wearable Cameras Based on Information From the Human Hand

Wiktor Mucha, Martin Kampel
Computer Vision Lab, TU Wien, Vienna, Austria
wiktor.mucha@tuwien.ac.at, martin.kampel@tuwien.ac.at

Abstract

Hand pose estimation is one of the critical tasks in egocentric vision, as humans interact with thousands of objects with their hands every day. We discuss key factors for egocentric hand pose estimation for ma single camera and how retrieved hand skeletons can be used in applications for modelling human behaviour.

Introduction

Wearable cameras provide a unique perspective by capturing daily activities from an individual's viewpoint, an egocentric view. Placing a camera on the human body provides broad information about the wearer's life and health. Understanding human behaviour through wearable cameras can be approached in two main ways: long-term and short-term behaviour analysis. Considering an example of food intake, long-term behaviour analysis focuses on identifying individuals' routines and recurring activities, such as where and when they eat. In contrast, short-term behaviour analysis examines specific movements, determines the types of food consumed, and estimates their nutritional content. Hands play a crucial role in most Activities of Daily Living (ADLs), such as eating, drinking, or other essential tasks and are at the centre of the egocentric image in most cases (Núñez-Marcos, Azkune, & Arganda-Carreras, 2022). These activities often involve various hand movements in relation to objects, such as rotating, picking up, or manipulating them. In short-term analysis, examining these hand-object interactions through egocentric vision provides a detailed understanding of the actions performed, offering valuable insights into both specific tasks and broader patterns of behaviour. This makes egocentric hand pose estimation a critical challenge. We consider hand pose estimation a task of localising 21 key points that describe the finger joints, palm, and wrist positions form a single camera.

Methodologies for Egocentric Hand Pose Estimation

Hand pose estimation can be understood in 2D and 3D space, each with distinct advantages. While 2D models involve less complex networks and potentially faster processing, 3D models offer a richer understanding by capturing depth. One potential solution is a single-hand approach named *EffHandNet* (Mucha & Kampel, 2024), which begins with a pre-trained hand detector that identifies regions, $R_{1,2}$ of hands in the egocentric input image. It then predicts the hand pose within $R_{1,2}$. The feature extraction is handled by Convolutional Neural Network (CNN), followed by transposed convolutions to produce heatmaps representing joints' locations. Although this method benefits from extensive datasets, its drawbacks include increased computational load due to multiple passes through the backbone network. *EffHandEgoNet* (Mucha & Kampel, 2024) addresses these limitations using an egocentric approach that hand presence in the scene , $H_{L,R}$ and employs two up-sampling heads to refine the resolution. This method improves the modelling of hand-object interactions and outputs the hand pose as 2D coordinates. Its efficiency stems from a single forward pass through the backbone, and it excels in learning interactions between hands, though the limited availability of public egocentric datasets constrains it. Extending *EffHandEgoNet* to 3D (Mucha, Wray, & Kampel, 2024) incorporates a regression module to estimate the z coordinates. This results in 2.5D coordinates in the image space. 2.5D coordinates are then transformed into 3D representation using the pinhole camera model, enabling an understanding of hand poses in three-dimensional space.

Egocentric Applications of Hand Pose

Egocentric hand pose estimation has emerged as a valuable tool across various applications, e.g., action recognition, struggle determination, and hand rehabilitation. Following (Mucha & Kampel, 2024), In action recognition, this technique involves analysing sequences of frames $f_1, \ldots, f_n$. Actions shorter than *n* frames are zero-padded, while longer ones undergo uniform subsampling. The YOLOv7 model is employed for object detection, describing objects with bounding boxes and labels. Further, the hand pose with object information is embedded with a transformer encoder and action classification layer. Experiments show a strong correlation between the accuracy of action recognition and the hand pose. In the realm of struggle determination, egocentric hand pose estimation facilitates categorising struggle levels into binary classifications. Accurate recognition of struggle is crucial for providing effective support to individuals. Recent approaches have achieved an 89% accuracy in binary struggle determination by integrating hand pose information through a transformer encoder alongside spatial-temporal features from video. Furthermore, in hand rehabilitation for stroke patients who experience hand dysfunction, egocentric vision can address challenges such as exercise recognition, repetition counting, exercise detection, and form evaluation, employing hand pose estimation (Mucha, Tanaka, & Kampel, 2024).

Conclusion

Precise pose description is fundamental for multiple applications in egocentric image processing. 2D pose information is, at this stage, more accurate as it eliminates the challenging estimation of depth and can be an alternative to 3D pose estimation when inference time is crucial. While methods based on hand detection struggle with hand pose estimation in egocentric perspectives due to occlusions, alternative approaches show that the highest precision is reached when considering both hands in the input for learning complex interactions.

Acknowledgments

This publication is part of the visuAAL project on Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 861091.

References

Mucha, W., & Kampel, M. (2024). In My Perspective, in My Hands: Accurate Egocentric 2D Hand Pose and Action Recognition. *2024 IEEE 18th International Conference on Automatic Face and Gesture Recognition (FG)* (pp. 1-8). Istanbul, Turkiye: IEEE.

Mucha, W., Tanaka, K., & Kampel, M. (2024). REST–HANDS: Rehabilitation with Egocentric Vision using Smartglasses for Treatment of Hands after Surviving Stroke. *European Conference on Computer Vision Workshops.* Milan, Italy: Springer.

Mucha, W., Wray, M., & Kampel, M. (2024). SHARP: Segmentation of Hands and Arms by Range using Pseudo-Depth for Enhanced Egocentric 3D Hand Pose Estimation and

Action Recognition. *27th International Conference on Pattern Recognition (ICPR)*. Kolkata, India: Springer.
Núñez-Marcos, A., Azkune, G., & Arganda-Carreras, I. (2022). Egocentric Vision-based Action Recognition: A survey. *Neurocomputing*, 175-197.

***Siddharth Ravi, and Francisco Florez-Revuelta (2024): Contextual Privacy Preservation in Active and Assisted Living Using Omnidirectional Cameras**. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Contextual Privacy Preservation in Active and Assisted Living Using Omnidirectional Cameras

Siddharth Ravi, and Francisco Florez-Revuelta
University of Alicante, Spain
siddharth.ravi@ua.es, francisco.florez@ua.es

Abstract

This paper explores contextual privacy preservation using top-view omnidirectional dioptric cameras, particularly within Active and Assisted Living (AAL) environments. These technologies hold significant promise for enhancing the quality of life for older adults and individuals with disabilities, yet they also introduce substantial privacy challenges. Using the concept of "Privacy by Context" based on Nissenbaum's theory of contextual integrity, this paper proposes a privacy-preserving pipeline for omnidirectional cameras that aligns with EU legal frameworks. As part of the research, the ODIN dataset was created to facilitate the creation of machine learning models aimed at scene understanding with omnidirectional dioptric camera images..

Introduction

Active and Assisted Living (AAL) technologies have immense potential to enhance the lives of older adults and individuals with disabilities by supporting daily activities, health monitoring, and emergency assistance. However, integrating advanced imaging systems like omnidirectional cameras into personal spaces raises significant privacy concerns, potentially hindering user acceptance and adoption.

Contextual Integrity, a privacy framework proposed by Nissenbaum (2009), suggests that privacy is preserved when information flows adhere to the norms of specific social contexts, considering factors such as actors, transmission principles, and information types. Building on this, the concept of "Privacy by Context" emphasizes designing systems that respect these contextual norms, while "Privacy by Design" (Cavoukian, 2009) advocates incorporating privacy considerations into system architecture from the outset. Despite these frameworks, effectively implementing privacy-preserving mechanisms for omnidirectional imaging in AAL environments remains a challenge. This research aims to develop an end-to-end privacy-preserving pipeline for omnidirectional RGB cameras in AAL settings, compliant with EU legal regulations like the GDPR.

Research Questions

The goal of this research is to create an end-to-end privacy-preserving pipeline for omnidirectional RGB cameras in AAL settings that complies with EU legal regulations. Three research questions were formulated and explored as part of the research.

1. **RQ1**: Can contextual visual privacy be provided for individuals appearing in RGB images?
2. **RQ2**: Can privacy be maintained for individuals observed in zenithal-view dioptric omnidirectional camera images?
3. **RQ3**: Can a privacy by design pipeline be created for omnidirectional images adhering to legal regulations?

A taxonomy of visual Privacy Enhancing Technologies

The research involved developing a low-level taxonomy of Visual Privacy Enhancing Technologies aligning with the privacy by design elements at five different levels proposed in Mihailis and Colonna (2020): sensor, model, system, UI, and user. The proposed low-level taxonomy operates at five associated levels as well, namely Visual Obfuscation, Data Hiding, Secure Processing, Blind Vision, and Intervention methods. Further subdivisions of visual obfuscation technologies are also explored.

It was determined that in order to achieve visual privacy by design, a pipeline should provide privacy at all all five different levels of the schema. Particular emphasis was given to visual obfuscation methods, which operate at the UI level.

The pipeline proposed by Climent-Pérez and Florez-Revuelta (2021) was chosen as a foundation due to its comprehensive approach to privacy preservation in RGB images, involving segmentation masks, 3D pose estimates, and background processing. However, since this pipeline was designed for standard RGB images, it was adapted to accommodate the unique characteristics of top-view dioptric omnidirectional camera images. This involves the development of specialized models for semantic segmentation and 3D pose estimation that can handle the distortion and wide field of view inherent to omnidirectional images.

The creation of semantic segmentation and 3D pose estimation models for the omnidirectional pipeline required data that included associated pose and semantic segmentation labels for associated top-view omnidirectional images. This proved difficult as such data was scarce. The ODIN dataset was created for this purpose.

ODIN Dataset

The ODIN (OmniDirectional INdoor) dataset by Ravi et al. (2023) was developed to address the scarcity of data for training machine learning models. This dataset captures activities of daily living in real living and working settings with multiple synchronized modalities, including top-view omnidirectional and lateral-view RGB images, infrared and depth images captured from calibrated cameras. It captures a wide range of activities of daily living performed by 15 participants, providing rich annotations for scene understanding tasks such as pose estimation and semantic segmentation. ODIN also provides further labels and annotations making it suitable for other tasks as varied as activity recognition, 3D reconstruction and person tracking.

Legal Considerations and Challenges

The legal landscape, particularly the GDPR, presents challenges for Privacy Enhancing Technologies (PETs), as most are considered pseudonymisation due to their reversible nature. The research highlights the need for concrete guidelines on privacy by design under EU data protection laws. The EU AI Act is a step in the right direction. Mihaildis and Colonna (2020) provide the most comprehensive approach for integrating privacy into system design from inception.

Ongoing and Future Work

Ongoing work involves training lightweight models for pose estimation and semantic segmentation using the ODIN dataset and integrating them into the privacy by design pipeline. Collaborative efforts are also underway to develop activity recognition models using omnidirectional images.

Conclusion

This research advances the development of a privacy by design system for top-view omnidirectional cameras in AAL environments, addressing the critical need for contextual privacy preservation. By proposing a comprehensive taxonomy of visual privacy enhancing technologies and creating the ODIN dataset, we provide essential tools for developing and evaluating privacy-preserving models suitable for omnidirectional images.

This work not only contributes to technological innovation but also ensures compliance with legal frameworks, paving the way for the responsible deployment of AAL technologies. While challenges remain, particularly in optimizing models for real-time applications and expanding legal guidelines, our ongoing efforts aim to bridge these gaps. Ultimately, this research opens novel directions in the field of omnidirectional scene understanding with strong privacy guarantees.

Acknowledgments

This publication is based upon work from COST Action GoodBrother—Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living (CA19121), supported by COST (European Cooperation in Science and Technology). It is also part of the visuAAL project on Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living, funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 861091.

References

Cavoukian, A. (2009). Privacy by Design. Information and Privacy Commissioner of Ontario, Canada.

Climent-Pérez, P., & Florez-Revuelta, F. (2021). 'Protection of visual privacy in videos acquired with RGB cameras for active and assisted living applications', Multimedia Tools and Applications, vol. 80, no. 15, pp. 23649–23664.

Mihaildis, A., & Colonna, L. (2020). 'A Methodological Approach to Privacy by Design within the Context of Lifelogging Technologies', Rutgers Computer and Technology Law Journal, vol. 46, no. 1, pp. 1.

Nissenbaum, H. (2009). Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford University Press, Redwood City.

Ravi, S., Climent-Pérez, P., Morales, T., Huesca-Spairani, C., Hashemifard, K., & Florez-Revuelta, F. (2023). 'ODIN: An OmniDirectional INdoor Dataset Capturing Activities of Daily Living From Multiple Synchronized Modalities', Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6488–6497.

***Kooshan Hashemifard, and Francisco Florez-Revuelta (2024): Context Recognition for the Application of Visual Privacy**. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Context Recognition for the Application of Visual Privacy

Kooshan Hashemifard, and Francisco Florez-Revuelta
University of Alicante, Spain
k.hashemifard@ua.es, francisco.florez@ua.es

Abstract

The population of the world is aging. Studies show that in few decades there will be strong demographic changes which leads to an increase of older people population and triples the current number of over 85 years old people (Padilla-López, 2015). It is expected that more and more older people will be in need of professional healthcare, which will become a burden to families, healthcare systems and the society, so may rise various societal issues.

One of the promising ways to tackle healthcare challenges is Ambient Assisted Living (AAL). AAL systems are intended to be a constant part of the everyday life of older people and aim to improve people's life quality, well-being and safety, while preserving their autonomy (Climent-Pérez, 2020). By taking advantage of information and communication technologies, a variety of sensors can be integrated into home environments and lives of older people, from wearable health sensors (e.g. smart watches, glasses, heart rate sensors) to environmental sensors such as cameras, to detect risks and provide support. Visual components and video cameras are especially important in the context of AAL, since they are the most directed and natural way to record events and situations and also provide rich information (Padilla-López, 2015). Their importance and efficiency have massively increased

by the recent advances in intelligent systems and computer vision technologies so they can be utilized in automatic interpretation of the visual data, smart monitoring, and automatic human lifelogging. However, while cameras in public places have been tolerated, there are serious privacy issues and trust concerns for utilizing them in private spaces, since they enable the acquisition of a huge quantity of information that can be easily interpreted by unauthorized viewers. Therefore, the development of privacy aware video-based applications needs to consider ”privacy by design” since early stages. Furthermore, “privacy-by-context” (Padilla-López, 2015) approach is proposed which trades off privacy preservation and image understanding according to the context of the video. In such setups, users can adapt privacy to their preferences in different situations.

In order to achieve this goal, a reliable video context recognition system is a necessity. The focus of this work is to deploy the latest achievements in computer vision research in conjunction with studies on psychological aspects of privacy to develop a user-tailored adaption of visual AAL technologies fulfilling relevant needs and wishes, as well as taking into account existing concerns and fears. The context extracted from such system, should provide enough information to empower people to modify it to their preferences. Therefore, the first stage is to define privacy variables such as: observer, activity, appearance, location. A visualization setup will be applied to the original image according to the user preferences and the ongoing context in the video. There are various methods and filters which can come to aid regarding conceiling sensitive information. Using these filters in combination with context information can guarantee user privacy (Padilla-López, 2015). This work varies from dataset creation (Hashemifard, 2022) to segmentation and recognition systems and user studies regarding different variables. Also, some of the most important variables are tackled. For example, regarding the appearance, a nudity recognition system is proposed to detect unwanted body parts whenever exposed. Utilizing the latest deep learning methods in semantic segmentation and with advances in human body parts detection, an exposure map of body is obtained and some of the most important problems such as different illumination and skin tones are tackled (Hashemifard, 2024). By combining them with user studies and preferences, a level-based nudity recognition system is proposed, which can provide different visualizations for different situations in order to choose from (Maidhof, 2022). For activity recognition, daily activities and events in home environment are the most relevant and are taken into account. These activities can include normal behaviour such as drinking and sleeping, and also important events such as falling (Hashemifard, 2023), choking or other health problems. For all the variables the sensitivity of the system is adjusted according to the location and desired privacy presets. This task has been done using the advancements in deep learning techniques, prior knowledge about the objects involved in different activities and human poses in different activities. In the end, by merging the findings in each separate module and developing a comprehensive context recognition system, a privacy-by-context system for preserving visual privacy can be achieved.

Acknowledgments

This publication is based upon work from COST Action GoodBrother—Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living (CA19121), supported by COST (European Cooperation in Science and Technology). It is also part of the visuAAL project on Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living, funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 861091.

References

Climent-Pérez, P., Spinsante, S., Mihailidis, A., & Florez-Revuelta, F. (2020). A review on video-based active and assisted living technologies for automated lifelogging. *Expert Systems with Applications*, *139*, 112847.

Hashemifard, K., Climent-Perez, P., & Florez-Revuelta, F. (2024). Weakly supervised human skin segmentation using guidance attention mechanisms. *Multimedia Tools and Applications*, *83*(10), 31177-31194.

Hashemifard, K., & Florez-Revuelta, F. (2022, May). From garment to skin: the visuaal skin segmentation dataset. In *International conference on image analysis and processing* (pp. 59-70). Cham: Springer International Publishing.

Hashemifard, K., Florez-Revuelta, F., & Lacey, G. (2023). A Fallen Person Detector with a Privacy-Preserving Edge-AI Camera.

Maidhof, C., Hashemifard, K., Offermann, J., Ziefle, M., & Florez-Revuelta, F. (2022, June). Underneath your clothes: a social and technological perspective on nudity in the context of AAL technology. In *Proceedings of the 15th international conference on PErvasive technologies related to assistive environments* (pp. 439-445).

Padilla-López, J. R., Chaaraoui, A. A., Gu, F., & Flórez-Revuelta, F. (2015). Visual privacy by context: proposal and evaluation of a level-based visualisation scheme. *Sensors*, *15*(6), 12959-12982.

Padilla-López, J. R., Chaaraoui, A. A., & Flórez-Revuelta, F. (2015). Visual privacy protection methods: A survey. *Expert Systems with Applications*, *42*(9), 4177-4195.

Hassan Zaal, John Dinsmore, Owen Conlan, and Gerard Lacey (2024): Improving Home-Based Care Robots' Capabilities Using Natural Language. *In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Improving Home-Based Care Robots' Capabilities Using Natural Language

Hassan Zaal[1], John Dinsmore[1], Owen Conlan[1], Gerard Lacey[2]
[1] Trinity College Dublin, [2] Maynooth University
zaalh@tcd.ie , *Owen.Conlan@tcd.ie* , *dinsmorj@tcd.ie* , *Gerry.Lacey@mu.ie*

Abstract

In the EU, the old-age dependency ratio is expected to rise to 51.2% by 2070. Personal Assistant Robots (PARs) can assist individuals and extend their independent living. This work aims to enhance communication and interaction capabilities between the robot and human within the home environment, thereby supporting older adults to better ageing in place.

Introduction

In the EU, the old-age dependency ratio $\frac{\geq 65}{15\ to\ 64}$ is projected to grow from 29.6% in 2016 to 51.2% in 2070 [1]. In ageing societies, the demand for long-term care will increase alongside shortages in labor to meet this demand [2]. Recent studies indicate evidence that robotic interventions could support "ageing in place" [3]. PARs can assist individuals and extend their independent living. Natural language interaction with robots is challenging, especially when translating high-level abstract instructions to the robot's capabilities. Previous approaches vary from a set of predetermined instructions to explicit instructions such as Vision-and-Language Navigation (VLN) [4]. Recently, others have used a set of skills that the robot can perform and use Large Language Models (LLMs) to select from these skills [5]. LLMs can encode extensive semantic knowledge about the world. A significant weakness of LLMs is their lack of real-world experience, which makes it difficult to leverage them for decision-making within a given robot's capabilities. LLMs without knowing the environment and the robot's capabilities can produce irrelevant outcomes. The aim of this research is to improve the robot's ability to understand and execute user verbal commands, utilizing extracted information from the environment and the robot's capabilities i.e., what the robot can/can't do and what the robot can/can't sense. To achieve the aim, we utilized pre-trained LLMs to analyze the end-user instructions and convert them into structured low-level commands that the robot can understand and execute and evaluated the ability of the robot to adapt to the changes in the environment to perform the user instructions. Evaluation is performed via NVIDIA Isaac Sim. Outcomes advance understanding of the role of LLMs in facilitating the capabilities of robots to effectively navigate and interact with environments to support the activities of daily living of older adults.

Personal Assistant Robot

A review was conducted to understand the needs and perspectives of older adults living independently and the availability of robots to support these individuals. Outcomes from the review helped determine the gap between users' needs and what can be achieved (in terms of independent living), why some solutions are adopted while others are not, and which areas (needs and robotic support) are in need of improvement. Further to the above the viewpoints and requirements of other stakeholders (clinicians, formal and informal caregivers) were considered. Table I presents outcomes from the review on robots available to support various users' needs. It shows that although there are some specialized robots for specific tasks, there is a significant gap in multi-functional assistive robots that can provide broad care for users with diverse needs. This shows the need for more research and development on robotics to support ageing in place. It indicates that for those needing assistance in multiple areas, a combination of several robots would be required or developing robots that can perform multi-functionalities.

After discussions with an expert in the field of robotics for old adults' assistance, we decided to focus on the need for an old adult named Cara. Cara likes gardening, but she cannot carry on the gardening tools herself. She wants the robot to carry the tools for her. Additionally, she wants the robot to remind her to take her medication and to fetch them.

Table I. Users' Needs vs. Available Robotics Systems

Users' Needs	ASTRO	Jaco	Care-O-bot4	ReWalk	PARO	AIBO	AmazonAstro	Bomy
Dressing								
Personal Hygiene								
Feeding		X						
Toileting								
Transferring/ Walking	X			X				
Transportation/ Shopping								
Managing Medications			X					X
Housework								
Social					X	X	X	X
Indoor Object Transferring			X				X	

Embodied AI with LLMs

To enable robots to navigate and interact with the surrounding environment, they need to have a representation that captures the semantic and geometric aspects of the scene at multiple levels of abstraction (objects, rooms, etc). Semantic representation provides a high level of abstraction for the robot to understand and execute human instructions, while geometric representation enables robots to navigate safely and manipulate objects [6]. To enable robots to plan execution for human instructions of high-level abstraction they need to convert the high-level instruction to an ordered list of primitive tasks that the robots can perform. LLMs can provide high-level knowledge about how to perform and covert complex and temporally extended instructions into a plan (containing an ordered list of primitive tasks) for execution, that the robots can perform. However, this requires understanding what is in the environment and the robot's capabilities to get relevant responses. Conditioning LLMs with visual scenes can provide affordable and reasonable responses in the corresponding environment. This could speed up the adoption of robots that can easily interact with humans and the environment. Additionally, it could provide reasoning ability to the robot to inform the users about failure to perform the task and the possible reason. Figure 1 shows the proposed architecture for using LLMs to control the robot. In future work, we will test the proposed architecture using robots with different embodiments such as a robot with a manipulator and a robot without, and report how the plan changes based on the robot embodiments and the environments.

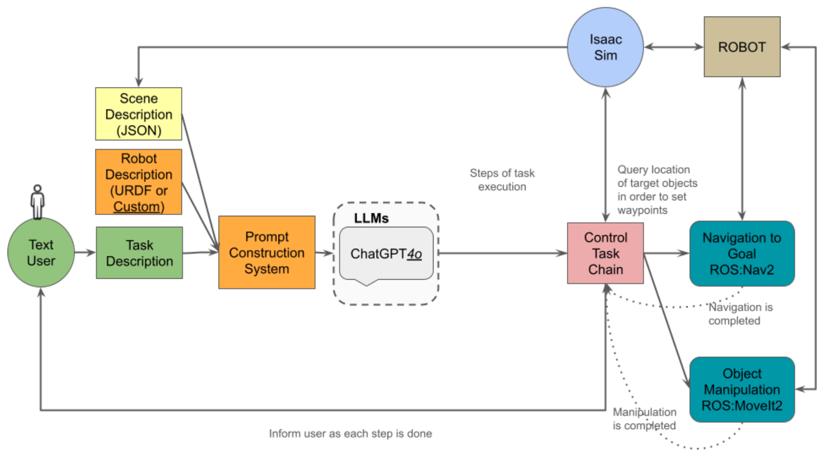


Figure 1. The Proposed Framework for Using LLMs to Control the Robot: The Prompt Construction System obtains its inputs from scene description, robot description, task description, and abstract user instruction. Then the constructed prompt is fed to the LLM to output a plan as a set of sub-tasks that the robot can perform based on its capabilities. When the sub-task requires navigation Nav2 can perform the navigation. While, when the sub-task requires manipulation MoveIt2 will perform it.

Conclusion

This work aims to enhance communication and interaction capabilities between the robot and human within the home environment, thereby supporting older adults to better ageing in place. This work is to parse abstract language instructions provided by the end-user and convert them into navigation commands within the constraints of the robot's capabilities by utilizing LLMs to provide a plan constrained by the robot's capabilities. Robots hold promise in enhancing the quality of life for older adults and caregivers by assisting with ADL, IADL, and medical needs.

Acknowledgments

This publication is part of the visuAAL project on Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 861091.

References

European Commission (2018), The 2018 Ageing Report, Economic and budgetary projections for the 28 EU Member States (2016-2070), Directorate-General for Economic and Financial Affairs.

New Job Opportunities in an Ageing Society : Paper Prepared for the 1st Meeting of the G20 Employment Working Group, 25-27 February 2019, Tokyo, Japan. 2019. Geneva: ILO.

Mois, G., & Beer, J. M. (2020). Robotics to support aging in place. In Living with robots (pp. 49-74). Academic Press.

Anderson, P., Wu, Q., Teney, D., Bruce, J., Johnson, M., Sünderhauf, N., ... & Van Den Hengel, A. (2018). Vision-and-language navigation: Interpreting visually-grounded navigation instructions in real environments. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3674-3683).

Brohan, A., Chebotar, Y., Finn, C., Hausman, K., Herzog, A., Ho, D., ... & Fu, C. K. (2023, March). Do as i can, not as i say: Grounding language in robotic affordances. In Conference on robot learning (pp. 287-318). PMLR.

Hughes, N., Chang, Y., Hu, S., Talak, R., Abdulhai, R., Strader, J., & Carlone, L. (2024). Foundations of spatial perception for robotics: Hierarchical representations and real-time systems. The International Journal of Robotics Research, 02783649241229725.

***Zhicheng He (2024): Bridging science and privacy: Health data governance in the age of digital health: The legal approaches contexts of the EU and China**. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Bridging science and privacy: Health data governance in the legal contexts of the EU and China

Zhicheng He
The Swedish Law and Informatics Research Institute, Stockholm University
zhicheng.he@juridicum.su.se

Abstract

This paper presents a comparative study of health data governance legal frameworks in the context of digital health technologies between the European Union (EU) and China, focusing on the assisted living sector. As healthcare systems undergo digital transformation, the widespread use of advanced ICT has led to extensive health data collection, raising privacy concerns. The paper examines the legal responses in the EU and China, identifying gaps and contradictions in their regulatory approaches, and provides recommendations on how existing laws should evolve to address these challenges effectively.

Health data processing in the age of digital health

The core content of this paper is a comparative study of the legal frameworks for health data protection in the context of digital health technologies between the European Union (EU) and China, with a focus on the assisted living sector.

Our healthcare systems are undergoing a wave of digital transformation characterized by the widespread application of information and communication technologies (ICT) in various health and care settings. Advanced digital technologies such as artificial intelligence, big data analytics, wearable devices, health monitoring sensors, mobile applications, and robotics are increasingly permeating the healthcare sector (World Health Organization, 2018). With the progress of ICT, healthcare services are extending beyond hospitals into private residences. For example, Ambient Assisted Living (AAL) technologies promise to enable elderly individuals to live more independently in their own homes, reducing their reliance on caregivers (Climent-Pérez et al., 2020).

The extensive use of ICT in healthcare has led to an unprecedented scale of health data collection and processing, with vast amounts of health data being stored in electronic health records and other platforms. Assistive technologies, often in the form of wearable devices and sensors, may pose privacy threats as they capture the user's health and well-being status by collecting large amounts of data, such as vital signs, daily activities, and even environmental data. The World Health Organization (WHO) has noted that the proliferation of wearable technologies, sensors, mobile health applications, social media, and genomic data collected through sequencing has significantly expanded the scope of health data collected (World Health Organization, 2021).

Health data processing in the age of digital health

In response to this situation, countries have enacted laws to strengthen data protection in the digital health transformation. For example, the EU has passed multiple data protection laws, including the General Data Protection Regulation (GDPR). Similarly, China enacted its first standalone data protection law, the Personal Information Protection Law, in 2021, which is expected to have a significant impact on technology and data processing. In an era of unprecedented health data processing, it is crucial to examine the impact of data protection and the extent to which existing laws can address new challenges.

Conclusions

The paper seeks to delineate the complex legal framework applicable to digital health technologies in the context of health data processing. It presents the legal regulation of health data in the EU and China, exploring similarities and differences between the two. The paper systematically presents the legal content according to the established hierarchy of legal sources, aiming to identify gaps and contradictions in the laws. While the research primarily focuses on traditional, positivist legal analysis, it also offers recommendations on how the law should act where appropriate.

Acknowledgments

This publication is part of the visuAAL project on Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 861091.

References

Climent-Pérez, P., Spinsante, S., Mihailidis, A., & Florez-Revuelta, F. (2020). A review on video-based active and assisted living technologies for automated lifelogging. *Expert Systems with Applications*, *139*, 112847. https://doi.org/10.1016/j.eswa.2019.112847

World Health Organization. (2018). *Classification of digital health interventions v1.0: A shared language to describle the uses of digital technology for health* (p. 1). World Health Organization. http://www.who.int/reproductivehealth/publications/mhealth/classification-digital-health-interventions/en/

World Health Organization. (2021). *Ethics and governance of artificial intelligence for health: WHO guidance*. World Health Organization. https://www.who.int/publications/i/item/9789240029200

***Maksymilian M, Kuźmicz (2024): AAL and the main approaches to balancing interests in European law**. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

AAL and the main approaches to balancing interests in European law

Maksymilian M. Kuźmicz
The Swedish Law and Informatics Research Institute, Stockholm University
Maksymilian.Kuzmicz@juridicum.su.se

Abstract

AAL systems offer valuable features like fall detection and medication reminders, yet they also raise privacy concerns, especially when utilizing visual data. Balancing competing interests emerges as a potential solution. However, the concept lacks precise definitions, leaving ambiguity about the balancing process. This contribution delves into the methodology of balancing in European law, starting with elucidating the principal dimensions of the balancing metaphor: proportionality and compromising.

Introduction

As the European population ages, the demand for supporting older adults in maintaining independence and quality of life grows exponentially. In his context, video-based Active and Assisted Living (AAL) technologies emerge as a beacon of hope. Numerous studies assert that AAL technologies will continue to improve their ability to assist individuals in need (Aleksic et al., 2022).

Such an optimistic prospect must be juxtaposed to a number of concerns (Ake-Kob et al., 2022). For example, the very sensors that provide critical assistance also pose a threat to the sanctity of personal space (Arning & Ziefle, 2015). That raises a question: what shall we sacrifice to receive the care we need? There are also other conflicts of interests, going beyond personal dilemmas. Can family or public authorities require older adults to use the most efficient ways of care despite personal objections? Furthermore, as AAL is subject to legal frameworks, to what extent do conflicts between norms and rights pose a problem?

Balancing is one of the plausible approaches to the conflicts of interests. That presents an issue of balancing as a conflict resolution tool, its forms, and effectiveness. In the legal context, the problem is more specific: what is balancing, and how could it be used to prevent or solve conflict? While the concept of balancing has been explored in legal scholarship (Cali, 2007; Van Der Sloot, 2016), it lacks a universally recognised procedure or a comprehensive inventory of legal instruments of balancing. Furthermore, previous discussions on balancing have predominantly occurred within the context of constitutional law and fundamental rights. Hence, the emergence of innovative technologies, typified by AAL, introduces a novel context to the already examined legal matter.

Balancing within European law

In European law, the concept of balancing has been used predominantly in the area of fundamental rights law. Invoking the concept of balancing does not entail the use of a uniform method of action. In most instances, the ECtHR conducts a seemingly quantitative balancing exercise that is both open-ended and ad hoc (Besson, 2017), entailing a concrete assessment of interests within the particular case (Gerards, 2009).

The conducted research identifies two primary approaches or dimensions of balancing in European law. The first, often referred to as proportionality, involves assigning comparative weights to the interests at stake to ascertain which should prevail (Greer, 2004; Van Der Sloot, 2016). The second approach aims to resolve conflicts of rights or interests by accommodating each involved interest partially. I propose labelling this method as compromising due to its resemblance to compromise.

Balancing as proportionality

Proportionality involves weighing competing interests to determine which should take precedence. This form of balancing primarily concerns the restriction of individual rights and their justification. It may also apply to clashes between stakeholders' interests (Leigh, 2012).

Proportionality involves evaluating a restriction of interest and the justification of the limitation. This method commences with the limitation under consideration and assesses its legality, necessity, suitability, and proportionality *sensu stricto* (PSS). The PSS is a core concept of balancing as proportionality. PSS entails ensuring that the least restrictive measure is applied and weighing the limitation of an interest caused by the least restrictive measure against the satisfaction of the other interest (Alexy, 2003; Greer, 2004).

Balancing as compromising

The alternative approach to balancing is compromising. Compromising seeks a potential solution for the problem at hand that would satisfy, at least partially, all the interests in question (Van Der Sloot, 2016). This approach commences with a particular problem at hand, for example, a conflict of interests or a technological phenomenon.

The process of compromising is governed by principles of 1) hierarchy of rights/interests, 2) consideration of all interests at stake, 3) equality of rights/interests, 4) avoidance of domination, and 5) specificity of balancing.

Conclusions

Proportionality and compromising represent two distinct dimensions of balancing, embodying disparate approaches, and functioning as disparate metaphors. However, they do not appear to be mutually exclusive. Rather, they may complement each other.

Proposing a method of balancing that reconciles proportionality and compromising would be a significant theoretical contribution with practical implications. Hence, it should be a subject of further research.

Acknowledgments

This publication is part of the visuAAL project on Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 861091.

References

Ake-Kob, A., Blazeviciene, A., Colonna, L., Čartolovni, A., Colantonio, S., Dantas, C., Fedosov, A., Florez-Revuelta, F., Fosch-Villaronga, E., He, Z., Klimczuk, A., Kuźmicz, M., Lukács, A., Lutz, C., Mekovec, R., Miguel, C., Mordini, E., Pajalic, Z., Pierscionek, B. K., … Tamò-Larrieux, A. (2022). State of the art on ethical, legal, and social issues linked to audio- and video-based AAL solutions. https://doi.org/10.5281/ZENODO.6793617

Aleksic, S., Atanasov, M., Calleja Agius, J., Camilleri, K., Čartolovni, A., Climent-Pérez, P., Colantonio, S., Cristina, S., Despotovic, V., Ekenel, H. K., Erakin, E., Florez-Revuelta, F., Germanese, D., Grech, N., Sigurðardóttir, S. G., Emirzeoğlu, M., Iliev, I., Jovanovic, M., Kampel, M., … Zgank, A. (2022). State of the Art of Audio- and Video-Based Solutions for AAL. https://doi.org/10.5281/ZENODO.6390709

Alexy, R. (2003). Constitutional Rights, Balancing, and Rationality. Ratio Juris, 16(2), Article 2. https://doi.org/10.1111/1467-9337.00228

Arning, K., & Ziefle, M. (2015). "Get that Camera Out of My House!" Conjoint Measurement of Preferences for Video-Based Healthcare Monitoring Systems in Private and Public Places. In A. Geissbühler, J. Demongeot, M. Mokhtari, B. Abdulrazak, & H. Aloulou (Eds.), Inclusive Smart Cities and e-Health (Vol. 9102, pp. 152–164). Springer International Publishing. https://doi.org/10.1007/978-3-319-19312-0_13

Besson, S. (2017). Human Rights in Relation: A Critical Reading of the ECtHR's Approach to Conflicts of Rights. In S. Smet & E. Brems (Eds.), When Human Rights Clash at the European Court of Human Rights (Vol. 1, pp. 23–37). Oxford University Press. https://doi.org/10.1093/oso/9780198795957.001.0001

Cali, B. (2007). Balancing Human Rights? Methodological Problems with Weights, Scales and Proportions. Human Rights Quarterly, 29(1), Article 1. https://doi.org/10.1353/hrq.2007.0002

Gerards, J. (2009). Judicial Deliberations in the European Court of Human Rights. In N. J. H. Huls, M. Adams, & J. Bomhoff (Eds.), The legitimacy of highest courts' rulings: Judicial deliberations and beyond. T.M.C. Asser Press ; Cambridge University Press [North American distributor].

Greer, S. (2004). "BALANCING" AND THE EUROPEAN COURT OF HUMAN RIGHTS: A CONTRIBUTION TO THE HABERMAS-ALEXY DEBATE. The Cambridge Law Journal, 63(2), Article 2. https://doi.org/10.1017/S0008197304006634

Leigh, I. (2012). Balancing Religious Autonomy and Other Human Rights under the European Convention. Oxford Journal of Law and Religion, 1(1), Article 1. https://doi.org/10.1093/ojlr/rwr015

Van Der Sloot, B. (2016). The Practical and Theoretical Problems with 'Balancing': Delfi, Coty and the Redundancy of the Human Rights Framework. Maastricht Journal of European and Comparative Law, 23(3), Article 3. https://doi.org/10.1177/1023263X1602300304

Irakli Pkhakadze (2024): Digital twins as a way to help ensure legal compliance of video-based AAL technologies*. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Digital Twins as a way to help ensure legal compliance of video-based AAL Technologies

Irakli Pkhakadze
The Swedish Law and Informatics Research Institute, Stockholm University
irakli.pkhakadze@juridicum.su.se

Abstract

The integration of Digital Twins (DTs) in Ambient Assisted Living (AAL) environments offers both opportunities and challenges, particularly in data management, privacy, and ethical considerations. AAL systems aim to enhance the quality of life for older adults through continuous monitoring, but they face significant issues such as data privacy, security, interoperability, and user compliance (Ganesan et al., 2019; Kim et al., 2013).

Data privacy and security are critical in AAL systems, which handle sensitive personal data prone to breaches (Morris et al., 2013). Ensuring secure data storage is challenging due to the diversity of devices. Data quality issues from sensor inaccuracies and the variability in data collection further complicate analysis (Reeder et al., 2013). Interoperability is a significant hurdle as AAL systems often integrate multiple devices, leading to compatibility challenges. Standardization is often lacking, making data aggregation difficult. User compliance is another issue, as elderly users may resist AAL technologies due to privacy concerns or complexity. Efficient data management is crucial but resource-intensive (Pirzada

et al., 2021). Ethical considerations also arise, particularly around autonomy and consent, requiring a balance between the benefits of data collection and individual rights (Manzeschke et al., 2015).

Addressing Challenges with Digital Twin Technology

DTs offer advanced solutions to these challenges. They can enhance data privacy and legal compliance in AAL environments by employing anonymization and pseudonymization techniques (Tao et al., 2019). Edge computing enables localized data processing, reducing the need for data transfer. Federated learning and privacy-preserving analytics, such as differential privacy and homomorphic encryption, allow secure, decentralized processing (Abadi et al., 2016). DTs can simulate environments, using synthetic data for testing and analysis without real-world data, thus improving data protection (Creswell et al., 2018).

DTs improve data quality and reliability by validating data from multiple sources, identifying anomalies before use. Acting as a central hub, DTs address interoperability and standardization issues, integrating data from various devices through standardized protocols (Grieves & Vickers, 2017). They enhance user compliance through personalized feedback and adaptive interfaces that demonstrate the benefits, encouraging adoption. DTs also manage large data volumes efficiently, leveraging cloud and edge processing for real-time analysis, while ethical data collection is ensured by incorporating consent management systems (Pirzada et al., 2021).

DTs adjust to different environments by creating dynamic models, ensuring consistent performance. They simulate scenarios to optimize system performance, improving reliability. Moreover, DTs overcome technological limitations by integrating advancements in sensor technology, ensuring systems are up-to-date and functioning optimally. By testing new technologies virtually, DTs identify potential issues and solutions before implementation, ensuring robust AAL systems.

Despite these advancements, ethical concerns remain, particularly regarding user independence and privacy. The extensive use of AI to monitor behavior could lead to a standardized lifestyle, potentially undermining autonomy (Manzeschke et al., 2015). Privacy concerns also arise from data collection within home environments, potentially leading to the medicalization of private spaces (Mortenson et al., 2015). Ensuring robust, adaptive regulatory frameworks is crucial to addressing these challenges, particularly in enhancing privacy protection and legal compliance.

Technical and Legal Framework

This research integrates technical analysis and doctrinal legal research to address these challenges. The technical analysis examines DTs' architecture and functionalities in AAL, assessing compliance with legal requirements. Collaboration with academic professionals provides practical insights, enhancing the study. Doctrinal research meticulously examines the GDPR, MDR, and AI Act, interpreting legislative texts to ensure compliance. The study also reviews existing societal implications of DTs in AAL. Ethical considerations, privacy concerns, and the potential social impact of deploying DTs are explored, focusing on balancing the benefits of these technologies with the need to protect individual rights. This comprehensive approach ensures well-rounded findings and recommendations for stakeholders involved in AAL technologies' development and regulation.

Conclusions

Proportionality and compromising represent two distinct dimensions of balancing, embodying disparate approaches, and functioning as disparate metaphors. However, they do not appear to be mutually exclusive. Rather, they may complement each other.

Proposing a method of balancing that reconciles proportionality and compromising would be a significant theoretical contribution with practical implications. Hence, it should be a subject of further research.

Acknowledgments

This publication is part of the visuAAL project on Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 861091.

References

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). "Deep learning with differential privacy," Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, vol. 308, pp. 308-318.

Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). "Generative adversarial networks: An overview," IEEE Signal Processing Magazine, vol. 35, no. 1, pp. 53-65.

Ganesan, B., Gowda, T., Al-Jumaily, A., Fong, K., Meena, S., & Tong, R. (2019). "Ambient assisted living technologies for older adults with cognitive and physical impairments: A review," European Review for Medical and Pharmacological Sciences, vol. 23, pp. 10470–10481.

Grieves, M., & Vickers, J. (2017). "Digital Twin: Mitigating unpredictable, undesirable emergent behavior in complex systems," in Transdisciplinary Perspectives on Complex Systems, Kahlen, J., Flumerfelt, S., & Alves, A. (eds.), Springer, Cham. https://doi.org/10.1007/978-3-319-38756-7_4.

Kim, M. J., Oh, M. W., Cho, M. E., Lee, H., & Kim, J. T. (2013). "A critical review of user studies on healthy smart homes," Indoor and Built Environment, vol. 22, no. 1, pp. 260–270.

Manzeschke, A. (2015). "MEESTAR: Ein Modell angewandter Ethik im Bereich assistiver Technologien," Technisierung des Alters - Beitrag zu einem guten Leben, vol. 7, pp. 263–283.

Morris, M. E., Adair, B., Miller, K., Ozanne, E., Hansen, R., Pearce, A. J., Santamaria, N., Viega, L., Long, M., & Said, C. M. (2013). "Smart-home technologies to assist older people to live well at home," Journal of Aging Science, vol. 1, no. 1, pp. 1–9.

Mortenson, W. B., Sixsmith, A., & Woolrych, R. (2015). "The power(s) of observation: theoretical perspectives on surveillance technologies and older people," Ageing and Society, vol. 35, no. 3, pp. 512–530.

Pirzada, P., Wilde, A., Doherty, G. H., & Harris-Birtill, D. (2021). "Ethics and acceptance of smart homes for older adults," Informatics for Health and Social Care. https://doi.org/10.1080/17538157.2021.1923500.

Reeder, B., Meyer, E., Lazar, A., Chaudhuri, S., Thompson, H. J., & Demiris, G. (2013). "Framing the evidence for health smart homes and home-based consumer health technologies as a public health intervention for independent aging: A systematic review," International Journal of Medical Informatics, vol. 82, no. 7, pp. 565–579.

Tao, F., Zhang, H., Liu, A., & Nee, A. (2019). "Digital Twin in Industry: State-of-the-Art," IEEE Transactions on Industrial Informatics, vol. 15, no. 4, pp. 2405-2415.

Tamara Mujirishvili, Anton Fedosov, Kooshan Hashemifard, Pau Climent-Pérez, and Francisco Florez-Revuelta (2024): Balancing Privacy and Safety: Stakeholder Perspectives on Video-Based Monitoring in Active and Assisted Living. *In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Balancing Privacy and Safety: Stakeholder Perspectives on Video-Based Monitoring in Active and Assisted Living

Tamara Mujirishvili[1], Anton Fedosov[2], Kooshan Hashemifard[1], Pau Climent-Pérez[1], Francisco Florez-Revuelta[1]

[1] University of Alicante (Spain), [2] University of Applied Sciences and Arts Northwestern Switzerland, Switzerland

{*tamar, k.hashemifard, pau.climent, francisco.florez}@ua.es, anton.fedosov@fhnw.ch*

Abstract

Active and Assisted Living (AAL) technologies aim to improve the quality of life for older adults. While video-based AAL solutions enhance healthcare management, they also present significant privacy risks. To address these, we developed a video monitoring system with privacy-preserving filters, deployed in an assistive technology center. A qualitative study involving older adults and care stakeholders revealed diverse perceptions and experiences with the technology, providing insights for further development. The findings highlight the privacy-safety trade-off in video technologies and how our system's privacy mechanisms alleviate concerns. Additionally, we identified varying stakeholder perspectives and potential development avenues for video-based monitoring in the AAL context.

Introduction

As the global population ages, there is a growing demand for technologies that support independent living and enhance the quality of life for older adults. Video-based monitoring systems (VMS) have emerged as a promising solution (Colantonio et al., 2018), providing comprehensive insights into the daily activities and health status of seniors. However, these systems raise significant privacy concerns due to the sensitive nature of visual data. To address these issues, the authors developed a VMS that incorporates various privacy-preserving filters, allowing for a balance between safety and privacy. The study investigates two primary research questions:

1. What are the perceptions and experiences of different stakeholders regarding the VMS, particularly concerning privacy?
2. What design opportunities exist for further development of the system?

Methodology

The research utilized a qualitative approach, user testing with prototype probes and in-depth interviews, gathering insights from three stakeholder groups: older adults, healthcare professionals, and technologists. User testing session were conducting with a total of 29 participants from the different stakeholder groups. This human-centered approach aimed to capture a comprehensive view of user experiences and expectations regarding the VMS.

Findings

The findings revealed diverse perspectives on the VMS among stakeholders. Older adults expressed concerns about being monitored and becoming "a number" rather than individuals, highlighting the need for systems that are humane and respect their dignity. Healthcare professionals recognized the potential benefits of the VMS in enhancing care but also acknowledged the importance of addressing privacy issues to foster trust among users. The study identified several design opportunities, including: Customization of Privacy Settings, User Education, Community Engagement. For the detailed results see the full paper (Mujirishvili et al., 2024).

Discussion

The research contributes to the understanding of privacy concerns in the context of video-based AAL technologies. It highlights the necessity of addressing these concerns to facilitate the adoption of such systems among older adults. By adopting a privacy-by-context approach (Chaaraoui et al., 2014), the study offers a framework for designing technologies that respect user privacy while providing essential care support. The findings underline the importance of balancing privacy and safety in the development of monitoring systems as also highlighted in previous studies (Schomakers et al., 2023). As technology continues to evolve, ongoing dialogue with stakeholders will be crucial in shaping solutions that enhance the quality of life for older adults without compromising their privacy.

Conclusion

In conclusion, the study emphasizes the need for privacy-preserving video monitoring systems in senior care. By understanding the diverse perspectives of stakeholders, developers can create more effective and acceptable technologies that support aging in place while respecting individual privacy rights. Future research should continue to explore user experiences and refine privacy mechanisms to enhance the trust and usability of AAL technologies.

Acknowledgments

- This publication is based upon work from COST Action GoodBrother - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living (CA19121), supported by COST (European Cooperation in Science and Technology)
- This publication is part of the visuAAL project on Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 861091.

References

Chaaraoui, Alexandros Andre, Padilla-López, José Ramón, Ferrández-Pastor, Francisco Javier, Nieto-Hidalgo, Mario, and Flórez-Revuelta, Francisco (2014): 'A vision-based system for intelligent monitoring: Human behaviour analysis and privacy by context', Sensors, vol. 14, no. 5, 2014, pp. 8895-8925.

Colantonio, Sara, Coppini, Giuseppe, Giorgi, Daniela, Morales, Maria-Aurora, and Pascali, Maria A. (2018): 'Computer vision for ambient assisted living: Monitoring systems for personalized healthcare and wellness that are robust in the real world and accepted by users, carers, and society', in Computer Vision for Assistive Healthcare, Elsevier, pp. 147-182.

Mujirishvili, Tamar, Fedosov, Anton, Hashemifard, Kooshan, Climent-Pérez, Pau, and Florez-Revuelta, Francisco (2024): '"I Don't Want to Become a Number": Examining Different Stakeholder Perspectives on a Video-Based Monitoring System for Senior Care with Inherent Privacy Protection (by Design)', Proceedings of the CHI Conference on Human Factors in Computing Systems, May 2024, pp. 1-19.

Schomakers, Eva-Maria and Ziefle, Martina (2023): 'Privacy vs. security: trade-offs in the acceptance of smart technologies for aging-in-place', International Journal of Human–Computer Interaction, vol. 39, no. 5, 2023, pp. 1043-1058.

Natalie An Qi Tham, Anne-Marie Brady, Martina Ziefle, and John Dinsmore (2024): Using behavioural insights to increase older adults' acceptance of camera-based active and assisted living technologies: An experimental medicine approach*. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Using behavioural insights to increase older adults' acceptance of camera-based active and assisted living technologies: An experimental medicine approach

Natalie An Qi Tham[1], Anne-Marie Brady[1], Martina Ziefle[2], John Dinsmore[1]
Trinity College Dublin[1], RWTH Aachen University[2]
ntham@tcd.ie

Abstract

Camera-based active and assisted living (AAL) technologies are a transformative solution to the challenges of population ageing but are hardly accepted by older adults. Few solutions have been offered to combat this issue of non-acceptance. The present research therefore employs behaviour change theory, couched in an experimental medicine approach, in an attempt to uncover strategies to increase older adults' acceptance of said technology. Results shed light on a fruitful target for interventions aimed at increasing said acceptance: future self-continuity – that is, individuals' sense of connectedness to their future self. Further insights are provided on novel means of targeting future self-continuity within acceptance-enhancing interventions.

Introduction

Many decisions involve making trade-offs between outcomes that arise in the present and those that attain in the future – these decisions are known are intertemporal decisions (Frederick et al., 2002). Intertemporal decisions are typically characterised by making small immediate sacrifices (e.g., foregoing dietary indulgences) for larger gains in the future (e.g., maintaining a healthy weight; Ersner-Hershfield et al., 2009; Rutchick et al., 2018). The decision to accept or reject camera-based active and assisted living (AAL) technologies is one such example. This is because the use of camera-based AAL technologies involves enduring upfront costs such as privacy violations (Mihailidis et al., 2008) and the stigmatising identity of being in receipt of assistance (Faucett et al., 2017). By contrast, the benefits of using the technology including improved health, safety, independence, wellbeing, and longevity, typically only manifest in the future (van den Broek et al., 2010). The decision to accept and use camera-based AAL technologies can thus be understood as an intertemporal one, requiring that older adults make costly immediate sacrifices in service of their long-run interests.

A remarkably frequent behavioural phenomenon is that when confronting intertemporal choice scenarios, people tend towards immediately gratification at the expense of larger future gains – a phenomenon known as temporal discounting (Fuchs, 1982; Lynch & Zauberman, 2006). In brief, people tend to discount the value of future gains, preferring instead to maximise immediate outcomes for themselves (Chapman & Elstein, 1995). From failing to save enough for retirement to neglecting healthy dietary practice, examples of temporal discounting abound in everyday life (Thaler & Benartzi, 2004).

An emerging line of research suggests that temporal discounting arises due to an inability to connect and empathise with one's future self. *Future self-continuity* can be understood as the degree to which individuals feel psychologically connected to their future selves, and is often elevated where the future self is: seen as similar to their current self, vividly envisioned, and positively appraised (Bixter et al., 2020; Hershfield et al., 2009; Sokol & Serper, 2019). According to a future self-continuity hypothesis, just as individuals have little reason to save resources for strangers to whom they do not feel connected, so too do they have little rational motive to reward their future self (Ersner-Hershfield et al., 2009; Hershfield, 2011; Hershfield et al., 2009). In line with this theorising, there is an extensive literature documenting the influence of future self-*dis*continuity on non-adaptive, impatient behaviour. For example, feeling dissimilar to the future self has shown associations with low retirement savings rates (Ersner-Hershfield et al., 2009) and low levels of physical activity (Rutchick et al., 2018), while individuals who report less (versus more) vivid imageries of their future selves tend to eat less healthily (Dalley, 2016) and save less for retirement (Ellen et al., 2012; Hershfield et al., 2011).

Discontinuity to the future self may reasonably account for older adults' disinclination towards camera-based AAL technologies. This is because if the above theorising holds that the use of said technology can be construed as an immediately costly action taken in service of future interests – that is, in service of the *future self's* interests – then older adults who feel disconnected to their future selves may be unwilling to take such costly action in present day.

Research Aim

Despite frequent and growing reports on older adults' poor acceptance of camera-based AAL technologies, few strategies have been devised to augment this acceptance. This programme of research therefore set out to explore, identify, and target within an intervention a hypothesised psychological barrier to older adults' acceptance of camera-based AAL technologies – future self-continuity – with a view to increasing the acceptance of said technology.

Methods

The experimental medicine approach, which advocates a theory-led, mechanisms focused approach to changing behaviour (Nielsen et al., 2018; Sheeran et al., 2017), was used as the research approach. The experimental medicine approach to behaviour change comprises four key steps: (i) to identify potentially malleable factors that can be targeted in an intervention to change behaviour; (ii) to validate the target by identifying measures of the target and examining whether changes in the target are associated with changes in the behavioural outcome of interest; (iii) to engage the target through experimental manipulation; and (iv) to conduct a full test of the proposed behaviour change model by investigating whether an intervention that contains those manipulations leads to change in the behavioural outcome of interest, and whether this change is mediated (i.e., explained by) changes in the target (Nielsen et al., 2018). Here, future self-continuity is the candidate target of focus within the programme of research.

Four key research activities were undertaken in line with the experimental medicine approach, with a view to increasing older adults' acceptance of camera-based AAL technologies. First, a scoping review was conducted to identify the barriers and facilitators to older adults' acceptance of camera-based AAL technologies that span important physical, environmental, social, and psychological domains. Second, a cross-sectional study was conducted to examine the association between an identified target (i.e., future self-continuity) and acceptance. Thereafter, qualitative work was undertaken to adapt and optimise a so-called "future-self intervention" designed to increase older adults' future self-continuity. Finally, an online randomised experimental study was examined the effect of the developed future-self intervention on acceptance.

Results

A scoping review identified 47 barriers and facilitators to older adults' acceptance of camera-based AAL technologies. Barriers included privacy concerns surrounding the technology, distaste for the technology's stigmatising quality, and skepticism about the technology's potential to displace cherished forms of human mediated care. Facilitators included the perceived usefulness of the technology for improving safety, independence, and overall wellbeing, older adults' perceived need for the technology, and assurance of data security. Of these, a particular barrier – i.e., the tendency to defer one's need for the technology into the future and onto other people – was the theoretical basis on which future self-continuity was identified as a putative candidate target for an acceptance-facilitating intervention.

A subsequent cross-sectional study (n = 183) revealed significant associations between future self-continuity and acceptance of camera-based AAL technologies among older adults. Specifically, older adults who reported experiencing their future selves in more vivid (odds ratio = 3.21, 95% CI [1.94, 5.31], p < .001) or positive (odds ratio = 3.09, 95% CI [1.76, 5.43], p < .001) terms were significantly more likely to say that they would install the technology in their homes today. This result justified the subsequent research activity, where semi-structured interviews with n = 7 older adults yielded a "future-self intervention" that was optimally usable, acceptable, relevant, and feasible for use with older adults. The intervention was evaluated in an online randomised controlled study (n = 181), which revealed a significant positive effect of the intervention on acceptance but only among those aged 64 years and below (B = .121, SE = .057, p = .037) – an effect mediated by felt positivity to the future self (B = .047, 95% CI [.005, .126]). No acceptance-facilitating effect of the intervention was found for those aged 65 and above (B = -.089, SE = .058, p = .122).

Discussion

In adopting an experimental medicine approach to behavior change, the present research isolated a novel psychological target for interventions aimed at increasing older adults' acceptance of camera-based AAL technologies – i.e., future self-continuity. Results stand to inform the development of acceptance-facilitating interventions, as they suggest that intervention strategies that increase older adults' felt connection to their future self – and in particular, those which increase the positivity with which future selves are appraised – should have beneficial effects on the acceptance of said technology.

Acknowledgments

This publication is part of the visuAAL project on Privacy-Aware and Acceptable Video-Based Technologies and Services for Active and Assisted Living. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 861091.

References

Bixter, M. T., McMichael, S. L., Bunker, C. J., Adelman, R. M., Okun, M. A., Grimm, K. J., Graudejus, O., & Kwan, V. S. Y. (2020). A test of a triadic conceptualization of future self-identification. PLOS ONE, 15(11), pp. e0242504.

Cardinaux, F., Bhowmik, D., Abhayaratne, C., & Hawley, M. S. (2011). Video based technology for ambient assisted living: A review of the literature. Journal of Ambient Intelligence and Smart Environments, 3(3), pp. 253–269.

Chapman, G. B., & Elstein, A. S. (1995). Valuing the future: temporal discounting of health and money. Medical Decision Making : An International Journal of the Society for Medical Decision Making, 15(4), pp. 373–386.

Dalley, S. E. (2016). Does the salience of possible selves mediate the impact of approach and avoidance temperaments on women's weight-loss dieting? Personality and Individual Differences, 88, pp. 267–271.

Ellen, P. S., Wiener, J. L., & Paula Fitzgerald, M. (2012). Encouraging People to save for their Future: Augmenting Current Efforts with Positive Visions of the Future. Journal of Public Policy & Marketing, 31(1), pp. 58–72.

Ersner-Hershfield, H., Garton, M. T., Ballard, K., Samanez-Larkin, G. R., & Knutson, B. (2009). Don't stop thinking about tomorrow: Individual differences in future self-continuity account for saving. Judgment and Decision Making, 4(4), pp. 280–286.

Faucett, H. A., Ringland, K. E., Cullen, A. L. L., & Hayes, G. R. (2017). (In)Visibility in Disability and Assistive Technology. ACM Transactions on Accessible Computing, 10(4), pp. 1–17.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. Journal of Economic Literature, 40(2), pp. 351–401.

Fuchs, V. (1982). Time Preference and Health: An Exploratory Study (pp. 93-120 BT-Economic Aspects of Health). National Bureau of Economic Research, Inc.

Hershfield, H. E. (2011). Future self-continuity: how conceptions of the future self transform intertemporal choice. Annals of the New York Academy of Sciences, 1235(1), pp. 30–43.

Hershfield, H. E., Goldstein, D. G., Sharpe, W. F., Fox, J., Yeykelis, L., Carstensen, L. L., & Bailenson, J. N. (2011). Increasing saving behavior through age-progressed renderings of the future self. Journal of Marketing Research, 48, pp. S23–S37.

Hershfield, H. E., Wimmer, G. E., & Knutson, B. (2009). Saving for the future self: Neural measures of future self-continuity predict temporal discounting. Social Cognitive and Affective Neuroscience, 4(1), pp. 85–92.

Lynch, J. G., & Zauberman, G. (2006). When do you Want It? Time, Decisions, and Public Policy. Journal of Public Policy & Marketing, 25(1), pp. 67–78.

Mihailidis, A., Cockburn, A., Longley, C., & Boger, J. (2008). The acceptability of home monitoring technology among community-dwelling older adults and baby boomers. Assistive Technology: The Official Journal of RESNA, 20(1), pp. 1–12.

Nielsen, L., Riddle, M., King, J. W., Aklin, W. M., Chen, W., Clark, D., Collier, E., Czajkowski, S., Esposito, L., Ferrer, R., Green, P., Hunter, C., Kehl, K., King, R., Onken, L., Simmons, J. M., Stoeckel, L., Stoney, C., Tully, L., & Weber, W. (2018). The NIH Science of Behavior Change Program: Transforming the science through a focus on mechanisms of change. Behaviour Research and Therapy, 101, pp. 3–11.

Rutchick, A. M., Slepian, M. L., Reyes, M. O., Pleskus, L. N., & Hershfield, H. E. (2018). Future self-continuity is associated with improved health and increases exercise behavior. Journal of Experimental Psychology: Applied, 24(1), pp. 72–80.

Sheeran, P., Klein, W. M. P., & Rothman, A. J. (2017). Health Behavior Change: Moving from Observation to Intervention. Annual Review of Psychology, 68, pp. 573–600.

Sokol, Y., & Serper, M. (2019). Development and Validation of a Future Self-Continuity Questionnaire: A Preliminary Report. Journal of Personality Assessment, 102(5), pp. 677–688.

Thaler, R. H., & Benartzi, S. (2004). Save more tomorrowTM: Using behavioral economics to increase employee saving. Journal of Political Economy, 112(1), pp. S164–S187.

van den Broek, G., Cavallo, F., & Wehrmann, C. (2010). AALIANCE Ambient Assisted Living Roadmap. IOS Press.

Maria J. Santofimia, Henry Llumiguano, Javier Dorado, Ángel Jesús Sáez, Vicente Díaz, and Álvaro Carrera (2024): MIRATAR: A Smart Mirror Solution for Remote Health Monitoring and Rehabilitation in Older Adults*. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

MIRATAR: A Smart Mirror Solution for Remote Health Monitoring and Rehabilitation in Older Adults

Maria J. Santofimia[1], Henry Llumiguano[1], Javier Dorado[1], Ángel Jesús Sáez[2], Vicente Díaz[3], Álvaro Carrera[4]
[1]Universidad de Castilla-La Mancha
[2]Centro de Investigación Biomédica en Red del Instituto de Salud Carlos III,
[3]Universidad Carlos III de Madrid, [4]Universidad Politécnica de Madrid
mariajose.santofimia@uclm.es, herny.llumiguano@uclm.es, Javier.dorado@uclm.es, pabizanda@sescam.jccm.es, vdgandas@polsoc.uc3m.es, carlosangel.iglesias@upm.es

Abstract

The MIRATAR project focuses on advancing the digital transformation of the care economy by leveraging innovative technologies such as a smart mirror and a virtual caregiver to address the frailty and multimorbidity challenges faced by older adults. By adopting a 4P approach (Predictive, Preventive, Personalized, and Participatory care), the project aims to empower individuals to manage their health proactively at home. With real-time data collection and AI-driven interventions, the system supports health literacy, behaviour change, and comprehensive care. This platform targets the challenges of promoting self-management, reducing institutionalisation, and improving quality of life for older adults.

Introduction

The demographic shift towards an ageing population presents unique challenges to modern healthcare systems, particularly in the management of frailty and multimorbidity. Frailty, characterized by diminished strength and endurance, increases vulnerability to adverse health outcomes, while multimorbidity, defined as the presence of two or more chronic conditions, complicates patient care due to complex treatment protocols and the risk of polypharmacy. Addressing these conditions requires a transformative approach in how care is delivered, moving from the current disease-centred model to a person-centred paradigm.

At the core of the project is the development of a smart mirror system, which leverages artificial intelligence (AI) and internet of things (IoT) technologies to create a personalised and interactive healthcare environment. This system includes a virtual caregiver, designed to support individuals in managing their frailty and multimorbidity through tailored interventions, health literacy initiatives, and behaviour change strategies. By placing the patient at the centre of the healthcare process, MIRATAR aligns with the broader objective of creating a digital, less institutionalised care economy.

The contributions of this work are outlined as follows: 1) Empowerment of individuals to self-manage frailty and multimorbidity risks through personalised self-management strategies, supported by multimodal and technology-based interventions; 2) Development of AI-based prediction models for frailty and multimorbidity risks, which will enable the delivery of personalised interventions, and the prediction of future health risks based on individualised data; 3) Integration of innovative ICT solutions into the care ecosystem, creating a cyber-physical platform that supports risk management and health monitoring; 4) Evaluation of the impact of these technologies on patient well-being, including the emotional and psychological aspects of care, with a specific focus on gender differences and public health policy recommendations; 6) Recommendations for public authorities, focusing on organisational and social changes needed to adopt and scale these technologies within healthcare systems.

The rise of digital health technologies, including wearable devices, virtual assistants, and AI-driven health platforms, has the potential to revolutionise chronic disease management and preventive care. While commercial products such as Fitbit, Apple Watch, and Google Fit have made strides in encouraging healthier lifestyles, their impact on frailty and multimorbidity populations remains underexplored. A systematic review has found that although wearable devices may increase motivation for physical activity, there is little evidence to suggest they significantly improve health outcomes in individuals with chronic conditions as stated in Burbach et al. (2019). Furthermore, AI and machine learning models have gained traction in healthcare, with applications in disease prediction, risk assessment, and decision support systems. However, these technologies face barriers to adoption, primarily due to the lack of integration into clinical workflows and issues of interoperability as stated in Jo et al. (2019). To overcome these limitations, MIRATAR focuses on its innovative smart mirror solution, which serves as an interactive platform for real-time health monitoring and personalised care interventions. The smart mirror integrates AI-driven predictive models and IoT technologies, allowing older adults to manage their frailty and multimorbidity at home. This system acts as a virtual caregiver, delivering tailored health advice and tracking key health metrics, while also fostering engagement through user-friendly, non-intrusive interactions.

Conclusions

The MIRATAR project offers an innovative solution for addressing the complex challenges of frailty and multimorbidity in older adults. Through its smart mirror system and virtual caregiver, the platform enables personalised, AI-driven interventions that promote self-management, health literacy, and proactive care.

Acknowledgments

This paper is founded by MCIN/ AEI /10.13039/501100011033 under Grant TALENT-BELIEF (PID2020-116417RB-C44) and the Project MIRATAR TED2021-132149B-C41 funded by MICIU/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR.

References

Burbach, L., Halbach, P., Plettenberg, N., Nakayama, J., Ziefle, M., & Valdez, A. C. (2019): 'Hey, Siri", "Ok, Google", "Alexa". Acceptance-relevant factors of virtual voice assistants', *IEEE International Professional Communication Conference*, vol. 2019, no. 101-111, pp. 101-111.

Jo, A., Coronel, B. D., Coakes, C. E., & Mainous III, A. G. (2019): 'Is there a benefit to patients using wearable devices such as Fitbit or health apps on mobiles? A systematic review', *The American Journal of Medicine*, vol. 132, no. 12, pp. 1394-1400.

Mustafa İzzet Muştu, and Hazım Kemal Ekenel (2024): Facial Attribute Based Text Guided Face Anonymization*. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Facial Attribute Based Text Guided Face Anonymization

Mustafa İzzet Muştu, Hazım Kemal Ekenel
Istanbul Technical University
{mustu18, ekenel}@itu.edu.tr

Abstract

The increasing prevalence of computer vision applications necessitates handling vast amounts of visual data, often containing personal information. While this technology offers significant benefits, it should not compromise privacy. Data privacy regulations emphasize the need for individual consent for processing personal data, hindering researchers' ability to collect high-quality datasets containing the faces of the individuals. This paper presents a deep learning-based face anonymization pipeline to overcome this challenge. Unlike most of the existing methods, our method leverages recent advancements in diffusion-based inpainting models, eliminating the need for training Generative Adversarial Networks. The pipeline employs a three-stage approach: face detection with RetinaNet, feature extraction with VGG-Face, and realistic face generation using the state-of-the-art BrushNet diffusion model. BrushNet utilizes the entire image, face masks, and text prompts specifying desired facial attributes like age, ethnicity, gender, and expression. This enables the generation of natural-looking images with unrecognizable individuals, facilitating the creation of privacy-compliant datasets for computer vision research.

Introduction

As computer vision applications become more common, they involve handling massive amounts of visual data, often containing individuals' information. There are many benefits of computer vision technology for the safe driving o autonomous vehicles, home security, and video calls. However, these should not be used at the cost of our privacy. Data privacy is a growing concern, with regulations like the General Data Protection Regulation (GDPR) in the European Union being passed to protect it. As Wolford (2020) stated, the GDPR affects all processing of personal data including images of individuals and requires the consent of the individual for processing their data. This makes it difficult for computer vision researchers to collect high-quality datasets that include people.

To overcome this concern, anonymization of people appearing in the acquired images becomes a must. While traditional anonymization methods like blurring, pixelization, or cropping faces are common ways to anonymize images, they distort the data so much that these images can not be used for any other purposes later. To prevent this problem, recently, realistic anonymization approaches have been proposed (Meden et al. (2017); Le et al. (2020); Li and Lin (2019); Liu et al. (2021); Maximov et al. (2020); Ma et al. (2021)). These methods utilize deep learning models to create realistic-looking faces to replace the original ones with them.

In this work, we present a deep learning-based three-stage face anonymization pipeline. Our pipeline starts with RetinaFace for face detection (Deng et al. (2019)). We employ VGG-Face[1] to extract facial attribute information (Parkhi et al. 2015)). We incorporate these facial attributes into text and perform anonymization with text guidance by using BrushNet (Ju et al. (2024)). In this way, we generate synthetic faces by conserving the age, gender, expression, and ethnicity information from the original image.

Related Work

In the current literature, there are two main types of image-generation methods. Most of the methods are derived from GAN which was first proposed by Goodfellow et al. (2014). The second main approach, the diffusion model, in image generation tasks comes from Ho et al. (2020). For the face anonymization task, the former is the most common method.

Anonymization. Hukkelås et al. (2019) detect faces using Dual Shot Face Detector (DSFD) and estimate face landmarks using a modified Mask R-CNN (Li et al. (2019); He et al. (2017)). They use a customized U-Net architecture in the GAN to generate images (Ronneberger et al. (2015)).

---

[1] We utilized the DeepFace repository for RetinaFace and VGG-Face (Serengil and Ozpinar (2021)).

Hukkelås and Lindseth (2023) propose an architecture that can replace faces as well as bodies with synthetic images. They detect and segment faces or whole bodies using three types of detectors. DSFD is used for face detection, Continuous Surface Embeddings (CSE), and Mask-RCNN are used for segmentation. Synthetic data is generated by a style-based GAN (Neverova et al. (2020)).

In the Latent Diffusion Face Anonymization (LDFA), Klemp et al. (2023) use the stable diffusion model from Rombach et al. (2022). The system works in two steps: first, it finds faces with RetinaFace and then replaces them with generated face images from stable diffusion.

Barattin et al. (2023) use a pre-trained GAN and proposed directly editing the latent space codes of the synthetic images to ensure a desired distance from the original identity while preserving facial attributes.

Method

We introduce a three-stage pipeline for the face anonymization task. We first find the faces with RetinaFace and create binary masks for the face regions. Then, we extract age, gender, ethnicity, and facial expression attributes using VGG-Face. These extracted features are converted to text for the last stage. At the last stage, we feed BrushNet with the original image, face masks, and text prompt (Ju et al. (2024)). The complete pipeline can be seen in Figure 1.

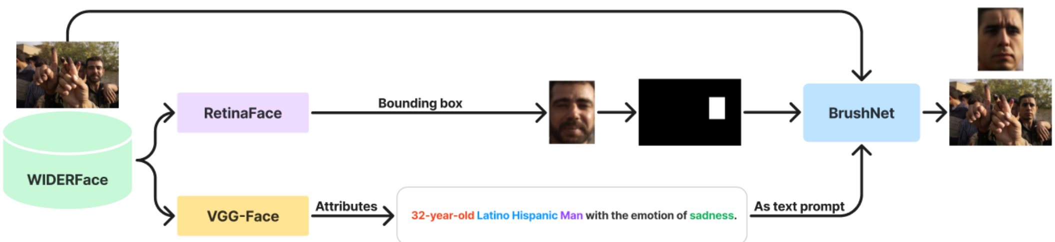


Figure 1. **Proposed pipeline.** Face detection is done by using RetinaFace. Facial attribute analysis is performed by VGG-Face and then these attributes are converted to text. Finally, faces are generated by BrushNet.

We use RetinaFace in the face detection stage. RetinaFace achieves state-of-the-art results for face detection on the WIDER Face dataset (Yang et al. (2016)). After we get face detections, we create binary masks for the detected facial areas.

We utilize VGG-Face to create text prompts that have facial attribute information. The VGG-Face model predicts the *age*, *ethnicity*, *gender*, and *expression* from the facial area. We then create texts from these predictions that are in a meaningful structure. One example text prompt is "*32-year-old Latino Hispanic Man with the emotion of sadness.*"

After we create masks and text prompts, we feed the BrushNet with them in addition to the original image. The model inpaint the original image in the mask area. We feed the model with each face mask in an image with the corresponding text prompt and generate a new face at each mask location.

Experiments

We run the pipeline on WIDER Face validation data and provide qualitative results. Example faces generated by our pipeline can be seen in Figure 2.

Figure 2. **Qualitative results.** Example face anonymization results. The first row is the originalimages, the second row is the results of LDFA and the last row is results of our pipeline.

From Figure 2, we see that our pipeline can generate faces that are different from the original identity. However, it preserves the facial attributes. Since our pipeline consists of three different models, our method depends on the capabilities of these models. This dependency causes some limitations to our pipeline. In one example scenario, when the resolution of the face region extracted from RetinaFace is less than 100 × 100 pixels, BrushNet struggles to generate fase images. In another scenario, BrushNet may generate faces that do not integrate into the image seamlessly. These drawbacks of the pipeline can be seen in Figure 3.

Figure 3. **Limitations.** The top row is the original images and the bottom row is generated images. The first column is related to the capabilities of the inpainting model and the second one is mostly related to the resolution of the facial region.

Conclusions and Future Work

In summary, with the growth of computer vision applications in our lives, computer vision researchers demand more image-based datasets than ever. Since the images collected in such datasets often include people's private information, researchers must ensure that data privacy regulations are employed when collecting images/videos. We introduce a new realistic face anonymization pipeline for this purpose. In our pipeline, we generate images that preserve the age, ethnicity, expression, and gender of the individuals by making use of facial attributes. Thus, by applying our pipeline to collected images for datasets, researchers can create natural-looking but unrecognizable people.

For future work, we plan to run quantitative experiments to analyze how much face recognition results change after image generation and how our pipeline affects facial attribute classification results. In addition, we plan to replace the rectangle-shaped masks with segmentation masks that are generated by a facial segmentation model in order to be able to integrate generated faces into the image more naturally.

Acknowledgments

This work was partially supported by the European Union's Horizon Europe research and innovation program under grant agreement No. 101135798 (My Personal AI Mediator for Virtual MEETtings BetWEEN People).

References

Barattin, S., C. Tzelepis, I. Patras, and N. Sebe (2023): 'Attribute-preserving Face Dataset Anonymization via Latent Code Optimization'. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8001–8010.

Deng, J., J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou (2019): 'Retinaface: Single-stage dense face localisation in the wild'. arXiv preprint arXiv:1905.00641. Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio (2014): 'Generative adversarial nets'. Advances in neural information processing systems, vol. 27.

He, K., G. Gkioxari, P. Dollar, and R. Girshick (2017): 'Mask R-CNN'. In: Proceedings of the IEEE international conference on computer vision. pp. 2961–2969. Hellmann, F., S. Mertes, M. Benouis, A. Hustinx, T.-C. Hsieh, C. Conati, P. Krawitz, and E. Andre (2023): 'Ganonymization: A gan-based face anonymization framework for preserving emotional expressions'. ACM Transactions on Multimedia Computing, Communications and applications.

Ho, J., A. Jain, and P. Abbeel (2020): 'Denoising diffusion probabilistic models'. Advances in neural information processing systems, vol. 33, pp. 6840–6851.

Hukkelas, H., R. Mester, and F. Lindseth (2019): 'DeepPrivacy: A Generative Adversarial Network for Face Anonymization'. In: Advances in Visual Computing. pp. 565–578, Springer International Publishing.

Hukkelas, H. and F. Lindseth (2023): ‘DeepPrivacy2: Towards Realistic Full-Body Anonymization’. In: 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). pp. 1329–1338.

Ju, X., X. Liu, X. Wang, Y. Bian, Y. Shan, and Q. Xu (2024): ‘BrushNet: A Plug-and-Play Image Inpainting Model with Decomposed Dual-Branch Diffusion’. Klemp, M., K. Rösch, R. Wagner, J. Quehl, and M. Lauer (2023): ‘LDFA: Latent Diffusion Face Anonymization for Self-Driving Applications’. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. pp. 3198–3204.

Le, M.-H., M. S. N. Khan, G. Tsaloli, N. Carlsson, and S. Buchegger (2020): ‘Anonfaces: Anonymizing faces adjusted to constraints on efficacy and security’. In: Proceedings of the 19th Workshop on Privacy in the Electronic Society. pp. 87–100.

Li, J., Y.Wang, C.Wang, Y. Tai, J. Qian, J. Yang, C.Wang, J. Li, and F. Huang (2019): ‘DSFD: Dual Shot Face Detector’. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Li, T. and L. Lin (2019): ‘Anonymousnet: Natural face de-identification with measurable privacy’. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. pp. 0–0.

Liu, Y., Q. Li, Z. Sun, and T. Tan (2021): ‘A 3 GAN: an attribute-aware attentive generative adversarial network for face aging’. IEEE Transactions on Information Forensics and Security, vol. 16, pp. 2776–2790.

Ma, T., D. Li,W.Wang, and J. Dong (2021): ‘CFA-Net: Controllable Face Anonymization Network with Identity Representation Manipulation’. arXiv e-prints, pp. arXiv–2105.

Maximov, M., I. Elezi, and L. Leal-Taixe (2020): ‘CIAGAN: Conditional Identity Anonymization Generative Adversarial Networks’. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

Meden, B., R. C. Mallı, S. Fabijan, H. K. Ekenel, V. Štruc, and P. Peer (2017): ‘Face deidentification with generative deep neural networks’. IET Signal Processing, vol. 11, no. 9, pp. 1046–1054.

Neverova, N., D. Novotny, M. Szafraniec, V. Khalidov, P. Labatut, and A. Vedaldi (2020): ‘Continuous surface embeddings’. Advances in Neural Information Processing Systems, vol. 33, pp. 17258–17270.

Parkhi, O., A. Vedaldi, and A. Zisserman (2015): ‘Deep face recognition’. In: BMVC 2015-Proceedings of the British Machine Vision Conference 2015.

Rombach, R., A. Blattmann, D. Lorenz, P. Esser, and B. Ommer (2022): ‘High-resolution image synthesis with latent diffusion models’. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684–10695.

Ronneberger, O., P. Fischer, and T. Brox (2015): ‘U-net: Convolutional networks for biomedical image segmentation’. In: Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. pp. 234–241.

Serengil, S. I. and A. Ozpinar (2021): ‘HyperExtended LightFace: A Facial Attribute Analysis Framework’. In: 2021 International Conference on Engineering and Emerging Technologies (ICEET). pp. 1–4.

Wolford, B. (2020): ‘What is GDPR, the EU’s new data protection law?’. https://gdpr.eu/what-is-gdpr. [Accessed 03-03-2024].

Yang, S., P. Luo, C. C. Loy, and X. Tang (2016): ‘WIDER FACE: A Face Detection Benchmark’. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Carina Dantas, Andrea Marinho, Miriam Cabrita, Miikka Keski-Säntti, Katja Nolvi, Oliver Díaz, Ana Sofia Carvalho, and Elísio Costa (2024): Liability regulations and respective challenges in the European Union regarding the use of Artificial Intelligence tools. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living

Liability regulations and respective challenges in the European Union regarding the use of Artificial Intelligence tools

Carina Dantas[1,4], Andrea Marinho[1], Miriam Cabrita[1], Miikka Keski-Säntti[2], Katja Nolvi[2], Oliver Díaz[3], Ana Sofia Carvalho[4], Elísio Costa[4]
SHINE 2Europe, Lda (Portugal)[1], Nordic Healthcare Group (Finland)[2], Universitat de Barcelona (Spain)[3], University of Porto (Portugal)[4]
{carinadantas, andreamartins, miriamcabrita}@shine2.eu, {miikka.keski-santti, katja.nolvi}@nhg.fi, oliver.diaz@ub.edu, aacarvalho@icbas.up.pt, emcosta@ff.up.pt

Abstract

Artificial Intelligence (AI) advancements present both innovative opportunities and potential risks, particularly concerning opaque decision-making. The European Union (EU) aims to address these challenges through proposed legal frameworks such as the AI Liability Directive and revisions to the Product Liability Directive, alongside the AI Act. These regulations respond to the complexities of assigning liability in AI-related harm cases, where traditional legal frameworks are difficult to apply due to AI's inherent characteristics, such as black-box and lack of causality. Ensuring conformity with fundamental rights and public interest, these frameworks seek to balance innovation with accountability. However, challenges persist in

determining liability when AI-driven systems cause harm, urging for the need to develop transparent AI models that document, explain, and validate decision-making processes. Robust oversight and ongoing education are essential for an ethical use of AI tools that ensure human agency and oversight. While these regulatory efforts signify progress, further steps are needed to establish a trustworthy AI ecosystem within the EU.

Artificial Intelligence (AI) is developing rapidly and especially data-driven learning approaches have been innovative in the progress of various technologies in several application fields. However, they also bring several potential risks, such as opaque decision-making (White Paper on AI, 2020). Even though there are safety regulations to reduce these risks, harm can occur (Fernández Llorca et al. 2023), and therefore, this transformation raises questions regarding liability laws in the European Union, leading to the need to explain the assignment of liability.

In general, there are three legal frameworks by which victims can be rewarded for product-induced harm (European Commission, 2019) and that are in general applicable also for AI-induced harm:

- **Fault-based liability**: in principle, the injured parties or claimants must prove that the defendant or wrongdoer caused the harm intentionally or negligently. This requires pointing out the applicable standard the accused should have fulfilled and proving that it was not satisfied.
- **Strict liability**: if a person can use a dangerous object or carry out a dangerous activity for their own objectives, then that person also bears the loss if the risk occurs (Karner, Koch, & Geistfeld, 2021). In this case, the victim does not have to prove the wrongdoing, but only that the risk emerging from the domain of the responsible party occurred.
- **Product liability**: where the injured party can formally request against the producer for a defect present at the time the product was put on the market. Victims must prove that the product was defective (regardless of fault) and the cause and effect linking the defect and the injury. (Vladeck, 2014).

AI systems are risky, namely because they are often opaque (High-level Advisory Body on AI, convened by the United Nations Secretary-General, 2023), unpredictable and lack causality, characteristics that may induce great difficulties when it comes to proving defect or fault and causal link with the harm in the different liability regimes (Fernández Llorca et al. 2023). Liability regimes should be adapted to mitigate the burden of proof on the injured parties when AI technologies are involved. It is still an open question for future investigation, on how to most usefully adjust the responsibility framework to converge the new challenges related to AI systems. Future work could include the development of an official model or procedure to guide possible lines of reasoning in liability cases involving AI systems (Fernández Llorca et al. 2023).

In this regard, the European Commission proposed the first-ever legal framework on AI, trying to address its risks (European Commission, 2021), with two regulations outlining the European approach to AI liability on 28 September 2022: a new AI Liability Directive (AILD) and a revision of the Product Liability Directive (PLD). Simultaneously with the recent AI Act, these new proposals are an advance in the direction of the EU´s efforts to make a unified approach to regulating the newest and risks of AI technologies in the EU (Mindy Nunez Duffourc & Sara Gerke, 2023).

This new AI framework establishes some limits to the EU Charter of Fundamental Rights (European Commission, 2012) on the freedom to conduct business (Article 16) and the freedom of art and science (Article 13) to guarantee conformity with fundamental reasons of public interest and the protection of other fundamental rights ('responsible innovation') when high-risk AI technology is developed and used. Those restraints are adequate and limited to the minimum necessary to prevent and reduce serious safety risks and likely violations of fundamental rights (European Commission, 2021).

The PLD and AILD are proposed Directives, which still need to be transposed into national laws (TFEU, Art.288), while the AI Act is a proposed regulation which will be directly applicable in all EU Member States. Once transposed into national laws, the two proposed Directives will operate in conjunction with EU Member States existing national (strict and fault-based) laws to govern liability for AI-caused injuries.

The AI and liability regulatory reform above-mentioned aim to respond to the difficulties of the injured person in satisfying the burden of proof on the defectiveness of the product or establishing the causal link between the defect and injuries suffered (Henrique Sousa Antunes, 2023).

AI blurs the lines between traditional product liability (for physical devices) and liability associated with services (software or data-driven services), making it hard to assign liability when an AI-driven device makes an error or causes harm, including identifying whether the responsibility lies with the manufacturer, the developer, the user of the device, or the AI algorithm itself.

The absence of transparency and explainability in AI systems not only hinders post-event accountability but also raises concerns about fairness, bias, and ethical implications. Biases inherent to training data can get amplified within black-box AI systems, potentially resulting in discriminatory decisions that escape scrutiny. Developing interpretable AI models that provide insights into how decisions are reached is crucial. This involves creating algorithms and models designed to elucidate their decision pathways, making the reasoning behind their outputs comprehensible to humans. Techniques like explainable AI (XAI) aim to demystify black-box AI systems by providing understandable justifications or explanations for their outputs. Methods such as feature importance analysis, model visualisation, or generating decision trees alongside AI predictions can offer insights into the factors influencing specific outcomes.

Moreover, establishing regulations and standards that compel developers and organisations to document, explain, and realistically validate AI models' decision-making processes could mitigate such risks, fostering trust, accountability, and ethical usage of AI, while enabling stakeholders to evaluate and understand AI-driven decisions more effectively.

Also, robust oversight mechanisms need to be established to ensure that AI tools are continuously monitored, validated, and updated to minimise errors and biases (European Commission, 2019). Additionally, ongoing education and training programs for professionals should emphasise the importance of critical thinking, ethical considerations, and the limitations of AI, enabling them to effectively collaborate with AI systems while retaining ultimate decision-making authority.

The new regulatory proposals are a step forward in these subjects at a European level, but challenges still remain (Claudia Nicastro, 2023) and thus additional efforts are needed towards a more trustworthy AI.

Acknowledgments

This research was partially performed under the RadioVal project, with the following main researchers: Oliver Díaz, Karim Lekadir, Philippe Lambin, Henry Woodruff, Ana Jiménez Pastor, Jose Munuera, Xavier Rafael, Gianna Tsakou, Stefanie Charalambous, Paris Laras, Carina Dantas, Miriam Cabrita, Harm op den Akker, Natália Machado, Andrea Martins, Luis Martí Bonmatí, Leonor Cerdá Alberich, Gloria Ribas Despuig, Ignacio Gómez-Rico Junquero, Rosa García Dosda, Fredrik Strand, Apostolia Tsirikoglou, Konstantinos Marias, Maciej Bobowicz, Katarzyna Gwoździewicz, Marlena Rygusik, Gordana Ivanac, Boris Brkljačić, Eugen Divjak, Katja Nolvi, Marketta Niemelä, Ira Haavisto, Miikka Keski-Säntti, Paola Clauser, Pascal A.T. Baltzer, Oguz Lafci, Daniel Mysler, Laura Cosaka, Africa Piñeiro, Lisandro Paganini, Federico Waisberg, Sara Tantawy, Pasant Aboelhoda, Eszter Balazs, Ancor Armas-Martin, Peter Gordebeke. RadioVal, International Clinical Validation of Radiomics Artificial Intelligence for Breast Cancer Treatment Planning, is funded by the European Union under Grant Agreement 101057699. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.

References

Amann, J., Blasimme, A., Vayena, E., Frey, D., Madai, V. (2020). *Explainability for artificial intelligence in healthcare: a multidisciplinary perspective.* https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7706019/

Antunes, H. S. (2023). *Non-contractual liability applicable to artificial intelligence: towards a corrective reading of the European intervention. In Católica Global School of Law Working Papers No.2/2023, page 2.* https://catolicalaw.fd.lisboa.ucp.pt/asset/2601/file

Duffourc, M.N. & Gerke, S. (2023). *The proposed EU Directives for AI liability leave worrying gaps likely to impact medical AI.* NPJ Digital Medicine, 6 Article number:77 published in partnership with Seoul National University Bundang Hospital. https://www.nature.com/articles/s41746-023-00823-w

European Commission (2019). *Expert group on liability and new technologies - new technologies formation, liability for artificial intelligence and other emerging digital technologies*. Publications Office of the European Union. https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf

European Commission (2020). *White Paper on Artificial Intelligence, A European approach to excellence and trust*, 65 final European Commission page 2. https://commission.europa.eu/system/files/2020-02/commission-white-paper-artificial-intelligence-feb2020_en.pdf

European Commission (2021). *Proposal for a Regulation of the European Parliament and the Council - Laying down harmonized rules on artificial Intelligence* (AI Act) and amending certain legislative acts, page 12. In an official EU website. https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF

European Commission (2022). *Proposal for a Directive of the European Parliament and of the Council on liability for defective products (PLD).* COM (2022) 495 final.

European Commission (2022). *Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence* (AI Liability Directive) (AILD). In an official EU website. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0496

European Commission (2022). *Questions & Answers: AI Liability Directive, in an official EU website.* https://ec.europa.eu/commission/presscorner/detail/en/qanda_22_5793

European Commission (2023). *New Product Liability Directive.* In briefing on EU Legislation in progress. https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739341/EPRS_BRI(2023)739341_EN.pdf

European Commission (2019). *Ethics Guidelines for Trustworthy Artificial Intelligence*. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

European Parliament (2023). *AI Act Briefing* https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf

European Parliament (2023b). Legislative Train Schedule. https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-regulation-on-artificial-intelligence

European Union (2012). *Charter of Fundamental Rights of the European Union*, 2012/C 326/02. https://www.refworld.org/docid/3ae6b3b70.html

Grant, L. & Pop,F. (2023). *An In-Depth Look at the EU´s AI Regulation and Liability Directive,* in EIPA - European Institute of Public Administration. https://www.eipa.eu/blog/an-in-depth-look-at-the-eus-ai-regulation-and-liability/

Hacker, P. (2023). *The European AI Liability Directives – Critique of a Half-Hearted approach and lessons for the future,* Working Paper, in European New School of Digital Studies, European University Viadrina Germany. https://arxiv.org/abs/2211.13960

Harwood, S. (2023). *EU Artificial Intelligence Liability Directive,* published online in Stephenson Harwood. https://www.shlegal.com/insights/eu-artificial-intelligence-liability-directive

Health Action International (2023). *Fact Sheet – The EU Medical Devices Regulation and the AI Act: a short comparison.* https://haiweb.org/wp-content/uploads/2023/03/MDR-AIAct_OnePager_FINAL.pdf

Hedderich, D.M., Weisstanner C., Van Cauter, S., Federau, C., Edjlali M., Radbruch, A., Gerke, S., Haller, S. (2023). *Artificial Intelligence tools in clinical neuroradiology: essential medico-legal aspects.* 65, 1091-1099. https://link.springer.com/article/10.1007/s00234-023-03152-7

High-level Advisory Body on AI, convened by the United Nations Secretary-General (December 2023) *Interim Report: Governing AI for Humanity*. https://www.un.org/sites/un2.un.org/files/ai_advisory_body_interim_report.pdf

Hoffmann, M. (2023). *The EU AI Act: A Primer.* https://cset.georgetown.edu/article/the-eu-ai-act-a-primer/

Karner. E., Koch, B.A., & Geistfeld, M.A., (2021) Comparative law study on civil liability for artificial intelligence. European Commission, Directorate-General for justice and Consumers. https://op.europa.eu/en/publication-detail/-/publication/8a32ccc3-0f83-11ec-9151-01aa75ed71a1/language-en

Llorca, D.F., Charisi, V., Hamon, R., Sánchez, I., Gómez, E. (2023). *Liability Regimes in the age of AI: a Use-Case Driven Analysis of the Burden of Proof,* in Journal of AI Research Vol.76 (2023) pages 613-644. https://dl.acm.org/doi/pdf/10.1613/jair.1.14565

Nicastro, C. (2023). *The EU Liability Directive and product liability Directive.* HAI – Health Action International, Background, page 2. https://haiweb.org/wp-content/uploads/2023/06/AI-Liability_Background.pdf

TFEU Art. 288, Consolidated version of the Treaty on the Functioning of the European Union 2012, OJ C326/47 (TFEU). https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:12012E/TXT:en:PDF

Vladeck, D.C., (2014). *Machines without principals: Liability rules and artificial intelligence.* Washington Law Review, 89(1), 117-150. https://digitalcommons.law.uw.edu/cgi/viewcontent.cgi?article=4800&context=wlr

Dinsmore J., Jacobs A., Söderström K., Harvey S., Teixeira F., Polak S., Coklar Okutkan I., Poll A., Ferreira Morais J., Byrne A., Hedenström I., Boschi T., Curran M., Donaghy D., Wffyels J., Verbeke N., Gonalez M., Brady AM., Boyle G., Diaz-Skeete Y., Tighe S., Doyle J. (2024): Advancing the use of digital health technologies for older adults self-managing at home with multimorbidity. *In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Advancing the use of digital health technologies for older adults self-managing at home with multimorbidity

Dinsmore John[1]., Jacobs An[2]., Söderström Kenth[3]., Harvey Séamus[7]., Teixeira Filipa[1]., Polak Sara[2]., Coklar Okutkan Isil[1]., Poll Anneleen[2]., Ferreira Morais Jéssica[2]., Byrne Aideen[4]., Hedenström Ida[3]., Boschi Tobia[8]., Curran Margaret[4]., Donaghy Deirdre[5]., Wffyels Julie[6]., Verbeke Naomi[6]., Gonzalez Manuel[3]., Brady Anne-Marie[1]., Boyle Gordon[7]., Diaz-Skeete Yohanca[7]., Tighe Sarah[7]., Doyle Julie[7].

[1] Trinity Centre for Practice and Healthcare Innovation (TCPHI), Trinity College Dublin, [2] imec-SMIT, Vrije Universiteit Brussel, [3] Region Vasterbotten, [4] Caredoc, [5] Home Instead Ireland, [6] Z-Plus, [7] NetwellCASLA Dundalk Institute of Technology (DkIT), [8] IBM Research Ireland

dinsmorj@tcd.ie, jacobs42@imec.be, kenth.soderstrom@regionvasterbotten.se, manuel.gonzalez@regionvasterbotten.se, margaret.curran@caredoc.ie, aideen.byrne@caredoc.ie, deirdre.donaghy@homeinstead.ie, julie.wyffels@z-plus.be, Shane.Gavin@dkit.ie, TEIXEIRF@tcd.ie, Jessica.Ferreira.Morais@vub.be, COKLAROI@tcd.ie, Seamus.Harvey@dkit.ie, sara.polak@imec.be, anneleen.poll@imec.be, ida.hedenstrom@regionvasterbotten.se, naomi.verbeke@z-plus.be *ABRADY4@tcd.ie, gordon.boyle@dkit.ie, Yohanca.Diaz@dkit.ie, Sarah.Tighe@dkit.ie, julie.doyle@dkit.ie*

Abstract

Multimorbidity has become an increasingly significant public health concern globally due to its increasing prevalence, particularly within older adult populations. For persons with multimorbidity, self-management of conditions at home is challenging. The SEURO project aims to evaluate via an effectiveness implementation hybrid trial (EIH) across 3 EU countries a novel digital health solution (PROACT) to enhance self-management practices and care for older adults living with multimorbidity.

Introduction

Multimorbidity, defined as the coexistence of two or more chronic conditions in an individual, has a global prevalence of 37.2%, affecting over half of adults over 60 (Chowdhury et al., 2023). Services for persons with multimorbidity (PwMs) are often poorly integrated, leading to inefficiencies, burdens, and potential safety risks, significantly affecting their quality of life (Ploeg et al., 2017; Eckerblad et al., 2015). There is a need to improve best practice around self-management and care for PwMs.

Funded across two European Commission Horizon 2020 projects (ProACT[2] and SEURO[3]), PROACT is a digital health platform co-designed with end-users to enhance self-management and care for older PwMs (65 years and over). The platform includes a tailored self-management application paired with a selection of devices to monitor health indicators such as blood pressure, blood glucose, and oxygen levels and well-being parameters like activity and sleep. This allows for personalised tracking, enabling PwMs to make informed decisions about their self-management and care through training and education tailored to their conditions.

Initially evaluated as part of the ProACT project (2016-2019) with n=120 PwMs across Ireland and Belgium (n=60 per region) in a longitudinal action research study (Dinsmore et al., 2021), outcomes show the majority of participants sustained engagement with the solution due to its usability and triage support, experiencing multiple benefits including improved self-management and well-being (Doyle et al., 2021; Sheng et al., 2023). The follow up SEURO project (2021-2025) will advance the platform and evaluate its effectiveness at scale across 3 EU countries.

2 https://proact2020.eu

3 https://seuro2020.eu

Methods

SEURO will employ a Type 1 Effectiveness Implementation Hybrid (EIH) trial design with n=720 PwMs across Ireland, Belgium and Sweden (n=240 per country)[4]. The EIH trial consists of a 3 arm (n=80 per arm per region) pragmatic Randomised Controlled Trial (pRCT) with embedded mixed-methods process evaluation (Curran et al., 2012). PwMs will be aged 65 and older, with two or more of the following conditions: diabetes; chronic respiratory disease; chronic heart failure; and/or chronic heart disease. Participants will be randomised (unblinded) to one of the three trial arms (**Arm 1:** Customised PROACT CareApp and devices + clinical triage support (observational only); **Arm 2:** Non-customised PROACT CareApp and devices; **Arm 3:** No technology/continue with their current self-management routine). Primary outcomes will focus on perceived quality of life and healthcare utilisation, whilst the secondary outcomes will explore PwMs' self-management of conditions across arms. The trials will not examine clinical effectiveness in terms of diagnosis or treatment etc.

Results and Discussion

Results will be reported upon the completion of the SEURO EIH trial in May 2025. Primary outcomes will assess the potential impact on PwMs and the health service, including cost-effectiveness. Secondary outcomes will explore how the platform influences patient engagement with self-management practices.

Conclusion

The PROACT platform marks a major advancement in digital health technologies designed to support older PwMs. It has the potential to significantly improve health outcomes and enhance the quality of life for this vulnerable population. The effectiveness of the platform will be evaluated through the SEURO EIH trial.

Acknowledgments

- ProACT: "The ProACT project has received funding from the EU's Horizon 2020 Research and Innovation Programme under grant agreement 689996".
- SEURO: "The SEURO project has received funding from the EU's Horizon 2020 Research and Innovation Programme under grant agreement 945449".

[4] ISRCTN34134007 Trial Registration: https://doi.org/10.1186/ISRCTN34134007

References

Chowdhury, S. R., Chandra Das, D., Sunna, T. C., Beyene, J., and Hossain, A. (2023): 'Global and regional prevalence of multimorbidity in the adult population in community settings: a systematic review and meta-analysis', EClinicalMedicine, vol. 57, February 2023, pp. 101860. doi: 10.1016/j.eclinm.2023.101860. PMID: 36864977; PMCID: PMC9971315.

Curran, G. M., Bauer, M., Mittman, B., Pyne, J. M., and Stetler, C.(2012): 'Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact', Medical Care, vol. 50, no. 3, March 2012, pp. 217-226. doi: 10.1097/MLR.0b013e3182408812. PMID: 22310560; PMCID: PMC3731143.

Dinsmore, J., Hannigan, C., Smith, S., Murphy, E., Kuiper, J., O'Byrne, E., Galvin, M., Jacobs, A., Sillevis Smitt, M., van Leeuwen, C., McAleer, P., Tompkins, L., Brady, A., McCarron, M., and Doyle, J. (2021): 'A digital health platform for integrated and proactive patient-centered multimorbidity self-management and care (ProACT): protocol for an action research proof-of-concept trial', JMIR Research Protocols, vol. 10, no. 12, 2021, pp. e22125. URL: https://www.researchprotocols.org/2021/12/e22125. DOI: 10.2196/22125.

Doyle, J., Murphy, E., Gavin, S., Pascale, A., Deparis, S., Tommasi, P., Smith, S., Hannigan, C., Sillevis Smitt, M., van Leeuwen, C., Lastra, J., Galvin, M., McAleer, P., Tompkins, L., Jacobs, A., Marques, M. M., Medina Maestro, J., Boyle, G., and Dinsmore, J. (2021): 'A digital platform to support self-management of multiple chronic conditions (ProACT): findings in relation to engagement during a one-year proof-of-concept trial', Journal of Medical Internet Research, vol. 23, no. 12, 2021, pp. e22672. URL: https://www.jmir.org/2021/12/e22672. DOI: 10.2196/22672.

Eckerblad, J., Theander, K., Ekdahl, A., Unosson, M., Wirehn, A. B., Milberg, A., Krevers, B., and Jaarsma, T. (2015): 'Symptom burden in community-dwelling older people with multimorbidity: a cross-sectional study', BMC Geriatrics, vol. 15, January 2015, pp. 1. doi: 10.1186/1471-2318-15-1. PMID: 25559550; PMCID: PMC4292813

Ploeg, J., Matthew-Maich, N., Fraser, K., Dufour, S., McAiney, C., Kaasalainen, S., Markle-Reid, M., Upshur, R., Cleghorn, L., and Emili, A. (2017): 'Managing multiple chronic conditions in the community: a Canadian qualitative study of the experiences of older adults, family caregivers and healthcare providers', BMC Geriatrics, vol. 17, no. 1, January 2017, pp. 40. doi: 10.1186/s12877-017-0431-6. PMID: 28143412; PMCID: PMC5282921.

Sheng, Y., Doyle, J., Bond, R., Jaiswal, R., Gavin, S., and Dinsmore, J. (2022): 'Home-based digital health technologies for older adults to self-manage multiple chronic conditions: a data-informed analysis of user engagement from a longitudinal trial', Digital Health, vol. 8, 2022, pp. 205520762211259. https://doi.org/10.1177/20552076221125957.

***Bruna Franceschini, Ângela Pinto, Vera Santos, and Maria Inês Pinto Santos (2024): Ethical checkpoint – a protocol for monitoring participants' engagement in the AAL4ALL project**. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Ethical checkpoint – a protocol for monitoring participants' engagement in the AAL4ALL project

Bruna Franceschini, Ângela Pinto, Vera Santos, Maria Inês Pinto Santos
Cáritas Coimbra
brunafranceschini@caritascoimbra.pt

Abstract

AAL technologies for care are an alternative to overloading professionals in the field, as well as promoting active and healthy ageing at home and in the community for as long as possible. However, there are several challenges surrounding their development in terms of ethical and legal suitability, particularly in relation to the security and privacy of end users. In addition, AAL technologies bring up the challenge of clarifying and informing these users who may not have contact with the system's control platforms. In other words, the very passivity of the system makes it difficult for people to understand how it works. This paper aims to demonstrate how the AAL4ALL - From Smart Home to Care (A4A) project approached monitoring users' understanding of how the system works and to report preliminary results obtained through questionnaires and interviews (interim phase) that took place during the implementation phase of the pilot study in Portugal, at Cáritas Coimbra. The pilot study involved installing the A4A solution in the homes of 15 older adults (65+) for 12 months. In Portugal, the AAL4All project is co-

funded by the European AAL Programme and by the Foundation for Science and Technology (FCT).

Introduction

The demographic change due to the ageing of the population, even though predictable, is still surprising in terms of “magnitude and speed”. By 2060, the population aged between 15-64 is expected to decline by nearly 42 million and older people are expected to account for 29,5% of the total population in the European Union (EU, 2014). This kind of projection raises several questions about the way we age and care for ourselves in the digital world.

Ambient Assisted Living technologies show up as a set of technological solutions that help older people in their daily lives, promote their independence and safety, and contribute to enabling people to live autonomously for longer in their own homes. These technologies integrate sensors and intelligent devices capable of being portable and environmental, collecting data in the older adult’s living space to analyse physical and comfort parameters and thus support monitored and personalised care. The assets of these technological developments relate to the quality of life of older people and carers, since AAL environments expect to ease their workload and burden (Calvaresi et al, 2017).

Within this scope, the AAL4ALL- From Smart Home to Care Home (A4A) project aims to enable older persons to stay independent for longer in their own homes by incorporating an ADLAI Platform (for detection of abnormal activities of daily living (ADLs)), room sensor devices (RSD), motion sensors and front door sensors, for data validation and correlation with the RSD data-driven information, and the Anyware Smartphone application for informal and formal care providers. The A4A solution does not include cameras, video or voice recording. The ultimate goal of the project is to democratise the social and healthcare systems by providing those involved with the tools and resources they need to optimise their approach to care.

In the framework of this paper, the idea is to share the approach towards a methodological checkpoint regarding the subjective impressions participants held in terms of clarity about the project mostly related to privacy, safety and data sharing.

Methodological overview

The A4A project encompasses a longitudinal, correlational study, in which the researchers will collect and analyse quantitative and qualitative data from the selected sample of participants during the field trial period (12 months). Due to the observational design of the study, there will be no control group. The pilot phase will take place in four different countries (Portugal, Romania, Switzerland and Denmark), and the intervention will focus on a participatory and multicentric design, as decisions underlying the field trials process are the results of co-creation and involvement of different partners (business, research and end-users and end-users organisations), and it will be an iterative cycle throughout the project. Beyond that, a lab test and a demo phase were conducted (before the actual field trial implementation) to ensure technological readiness and ethics by design (EC, 2021).

The data collection will be carried out at 70 test homes (70 older adults) in the field trials across the 4 end-user countries. The involvement of the older adults in the trials is in dyads – one older adult plus one caregiver (formal or informal).

To be part of the study the person must be able and willing to provide written consent. The older people able to join the study are (inclusion criteria): the one who demonstrates interest and acceptance of the solution; with an engaged secondary end-user who accepts to be involved; the living alone (or who lives with others but spends most of the day alone); 65+ (however, the participant can be younger if proven need of support is detected, i.e. rehabilitation).

A selection of assessment tools was selected and developed to investigate which are the impacts of the solutions in the piloted social and care settings. Older people should be inquired about their sociodemographic background, quality of life, ADLS deterioration, on the safety, acceptance and functionality of the system (baseline, mid-term and end-line).

The core of this paper reflects on the initiative to include a checkpoint of comprehension and clarity regarding the terms first presented in the Informed Consent as a way to investigate (from a qualitative point of view) their perceptions.

Assessing participants' impressions at the mid-term level

Empirical experience with older adults using passive AAL systems shows some level of difficulty for the primary end-user to comprehend the system and its functionalities (or to retain the information for longer periods). AAL systems as time goes by – and the person does not deal with the control dashboard directly – tend to become a stranger at home. It is even borderline when family members ask for clarifications on what is the system about or why it is there and the primary end-user does not know how to respond anymore.

A4A is a project that included a whole setting of sensors in a person's house, it also included sensors in private areas such as bedrooms and toilets, which complexifies the testing phase. The importance of the information on projects of such nature is even higher, not only in the beginning (for the consent phase) but also throughout the implementation.

It is noteworthy that the informed consent signature already followed a thorough, structured and user-centred protocol to be as friendly, inclusive and clear as possible (CEN, 2023). Beyond that, a very important step of the project was to decide, together with the older adults, which data they would like to share in the system (all participants decided to test the whole system, but they were aware they could turn on/off specific sensors and decide who would have access to the control dashboard – Anyware Software).

The tentative methodological approach the A4A project brought was to include qualitative inquiries in mid-term analysis to perceive how participants were feeling about the solution and if they were still aware of the terms detailed explained and agreed upon during the recruitment phase (for the Informed Consent signature).

Preliminary results

For the purpose of this paper, only the questions developed to address the themes of safety, privacy and data protection will be analysed.

The tentative methodological approach the A4A project brought was to include qualitative inquiries in mid-term analysis to perceive how participants were feeling about the solution and if they were still aware of the terms detailed explained and agreed upon during the recruitment phase (for the Informed Consent signature).

For the mid-term assessment, 9 evaluation statements/questions (multiple choices and open answers) were included in the interviews regarding participants' feelings and perceptions of the system usage for that timeframe.

The statements/questions and results were (n=14):

1. The A4A solution respects my autonomy and privacy.	57, 1% Totally Agree; 28,6% Agree; 14,3% Neither agree nor disagree
2. I feel safer when I use A4A solution and its equipment.	7,1% Totally Agree; 50% Agree; 35,7% Neither agree nor disagree; 7,1% Totally disagree
3. How safe did A4A make you feel at home?	35,7% higher feeling of safety; 64,3% no changes in the feeling of safety
4. A4A has improved my feeling of safety and peace.	50% Agree; 42,9% Neither agree nor disagree; 7,1% Disagree
5. How often did you remember A4A solution was installed at your home?	7,1% Always; 35,7% Frequently; 14,3% Sometimes; 42,9% Ocassionally
6. Are you aware you are the one in charge and who authorizes the data sharing?	100% Yes
7. Explain answer 6. (open answer)	7 people referred they are **aware** of the data sharing; 2 people referred

	they signed the **Informed Consent**; 1 person referred that when the project was presented it was explained; 1 person referred knowing the equipment is installed and that the collected data can be accessed by the researchers; 3 people did not reply.
8. How important do you consider to be in charge of data sharing?	35,7% Very important; 57,1% Important; 7,1% Neutral
9. Which were the most meaningful changes since you joined the project? (open answer)	5 people referred no meaningful changes; 3 people referred they feel safer; 2 people referred they feel they are helping the project; 1 person referred feeling less lonely (due to the contact with the research team); 1 person referred it is an asset for frail people; 1 person referred if the equipment were fully functional he/she would feel safer; 1 person did not reply.

Discussion

The results are only preliminary and are not intended to be an abstract generalization capable of reflecting all tests of AAL sensors for older people. On the contrary, the small sample, selected for convenience, and the qualitative data are intended to reveal the profile of this specific sample and their perception of the system throughout the intervention phase, so that activities could be administered meanwhile to provide clarification and information, if necessary. The methodological approach connects to the good practices for the ethical engagement of participants in digital innovation projects.

According to the results obtained, the participants showed some positive attachment to security and privacy due to the system under test (statements 1, 2, 4), while the majority also reported not often remembering that the system was installed in their home (question 5) - which is precisely what is intended with systems of this type.

However, it is worth noting that in a question of the same nature - about security - the majority of participants answered that they perceived no difference in their feeling of security (question 3). Combined with the final open-ended question (question 9) on the changes felt, the majority of participants reported feeling no difference.

Regarding the selection of data sharing (question 6), it was interesting to see that the entire sample is aware that, although they do not have direct contact with the sensor control application, they are the main ones responsible for deciding how to share the data collected by the system, and can configure and customize it as they wish. The open question was interesting in demonstrating that most of the participants were aware of the information and even remembered where/when they had access to it (informed consent and project presentation).

Conclusion

AAL systems carry with them a paradox that complicates their testing phase in research projects. On one hand, their passive nature should be seen as an asset in terms of comfort and practicality, when it comes to clarifying user understanding this becomes a complicating factor. This is because people end up not interacting with the technology and therefore don't understand its functionalities and associated benefits. Depending on the target group, for example, older people, they may even forget what it's for.

Given this practical challenge, an effort is needed to include active and participatory methods in study protocols to integrate primary users into the testing process in a free and informed manner, while at the same time not promoting discrimination or stigma.

In addition, the testing of sensors in a private environment raises several ethical-legal issues that also need to be addressed in a user-friendly way for the peace of mind and safety of participants, family members and support networks in general.

Therefore, the development of specific materials for the project and its objectives (especially Informed Consent), and a team that is aware of the need to clearly explain the project using appropriate language are essential elements for participation.

In the end, what the paper addresses is the use of the interim phase as an ethical checkpoint to verify that the intended messages have been received by the interlocutors, i.e. that the participants are aware of sensitive issues relevant to a given project so that the research team will still have time and opportunity to act within the intervention phase and the data will serve to qualitatively analyse the system from a person-centred perspective.

References

Calvareci, D.; Cesarini, D.; Sernani, P.; Marinoni, M.; Dragoni, A.; Sturm, A. (2017). Exploring the ambient assisted living domain: a systematic review. Journal of Ambient Intelligence and Humanized Computing 8(2). Available at: https://www.researchgate.net/publication/301901787_Exploring_the_ambient_assisted_living_domain_A_systematic_review

CEN. (2023). Digital health innovations – Good practice guide for obtaining consente for the use of personal health information for research and innovations. CWA 17933. Available at: https://www.cencenelec.eu/media/CEN-CENELEC/News/Workshops/2023/2023-03-15%20-%20Digital%20Health/draftcwa.pdf

European Commission. (2014). Population ageing in Europe. Available at: https://op.europa.eu/en/publication-detail/-/publication/1e7549b4-2333-413b-972c-f9f1bc70d4cf

European Commission. (2021). Ethics by design and Ethics of Use Approaches for Artificial Intelligence. Version 1.0. Available at: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf

Aysegul Ilgaz (2024): Balancing Autonomy and Protection: Ethical and Privacy Issues in Active and Assisted Living Technologies for Older People*. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Balancing Autonomy and Protection: Ethical and Privacy Issues in Active and Assisted Living Technologies for Older People

Aysegul Ilgaz
Akdeniz University Faculty of Nursing, Department of Public Health Nursing
aozer2012@gmail.com

Abstract

Elderly people's life can be significantly enhanced by active and assistive living (AALs) technologies. With the goal of enhancing older people's quality of life and lessening the load on caregivers, AALs technologies encompass any gadget that supports the psychological and physiological changes that come with age. There is a growing discourse over the potential of technological methods to facilitate assistance with everyday tasks. One common argument for AALs technologies is that they help people become more autonomous. A person's autonomy isn't always reduced just because they utilize an assistive gadget. Including senior citizens in decision-making on a regular basis may guarantee that AALs are viewed as an important instrument to support senior citizens' independence and autonomy. And also, the rights to personal data protection (privacy), as well as the rights and inclusion of older people, are major issues. Security and privacy issues are brought up by the fact that AALs technologies frequently gather sensitive personal data in

order to operate. It is vital to guarantee that these gadgets follow stringent privacy regulations and utilize strong security protocols to safeguard user information.

It is assumed that all AALs technologies have both positive and negative effects on the circumstances of how environments function, which may have a vital impact on people's health and quality of life. These technologies are touted for their alleged benefits. Nevertheless, privacy concerns and ethical standards are disregarded in its utilization. It is necessary to consider the ethical principles such as autonomy, non-maleficence, beneficence, fairness, interdependence, and privacy. Devices that support personal autonomy include new AALs technologies that can be used to track routine daily patterns or capture physiological data. In order to balance autonomy and protection, this study explores the ethical and privacy concerns related to the use of AALs technologies in depth, underline these issues in detail from an ethical perspective, and offer potential recommendations.

References

Martin, S., Bengtsson, J.E., Dröes, RM. (2010). Assistive Technologies and Issues Relating to Privacy, Ethics and Security. In: Mulvenna, M., Nugent, C. (eds) Supporting People with Dementia Using Pervasive Health Technologies. Advanced Information and Knowledge Processing. Springer, London, pp. 63-76.

Morte, Ricardo., Toboso, Mario., Aparicio, Manuel., Ausín, Txetxu., Monasterio, Aníbal., López, Daniel. (2020). Personal autonomy in elderly and disabled: How assistive technologies impact on it". Aging between Participation and Simulation: Ethical Dimensions of Socially Assistive Technologies in Elderly Care, edited by Joschka Haltaufderheide, Johanna Hovemann and Jochen Vollmann, Berlin, Boston: De Gruyter, pp. 185-198.

Schicktanz, S., Schweda, M. (2021). Aging 4.0? Rethinking the ethical framing of technology-assisted eldercare. HPLS, vol. 43, no. 93, pp. 1-19.

Wangmo, T., Lipps, M., Kressig, R. W., & Ienca, M. (2019). Ethical concerns with the use of intelligent assistive technology: findings from a qualitative study with professional stakeholders. *BMC Medical Ethics,* vol. 20, no. 1, pp. 98.

Rodrigo Perez-Vega, Niloofar Borguei-Razavi, Malcolm Fisk, and Kate Wigley (2024): The Independent Living Project: Challenges and Opportunities in Implementation and the Evaluation of AAL technologies*. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

The Independent Living Project: Challenges and Opportunities in Implementation and the Evaluation of AAL technologies

Rodrigo Perez-Vega[1], Niloofar Borguei-Razavi[1], Malcolm Fisk[2], Kate Wigley[3]
[1] Henley Business School, [2] University of Central Lancashire, [3] Reading Borough Council
r.perezvega@henley.ac.uk, niloofar.borgheirazavi@reading.ac.uk, mfisk@uclan.ac.uk, kate.wigley@reading.gov.uk

Abstract

The Independent Living Project aims to test the use of assistive living technologies in Reading, United Kingdom. The project aims to install sensors in people homes, and to assess the impact that the technology can have on the wellbeing of service users, the decisions that the adult social care team make, and the circle of support providing informal care to service users. This article outlines the plan to evaluate the impact of these technologies in this project and some of the early challenges faced by the project team.

Introduction

Older adults increasingly require assisted living technology to enhance their quality of life, maintain independence, and ensure safety in their daily activities (Sweeting et al., 2024). As the population ages, the integration of smart home technologies and other assistive devices has become instrumental in addressing the unique challenges faced by this demographic. The Independent Living Project is a 2 year project that aims to assess the impact that it has in multiple stakeholders, including service users, circle of support, and people working in adult social care within the Reading area.

Evaluation approach

The evaluation of Reading Borough Council's project will adopt the "Impact and Outcomes Analysis" approach, a method that is extensively utilized in fields like public policy, social sciences, and the non-profit sector (LBRO, 2010). This approach focuses in assessing the effectiveness of interventions, not just on a surface level, but in a manner that provides a holistic understanding of short-term and long-term results. In the first phase, the identification of specific impacts and outcomes that the Care Technology Solutions project is aiming for was conducted. In the second phase, a systematic linkage mapping the various elements such as activities, inputs, partners, and outputs to their respective outcomes and impacts. This ensures that every component of the project can be directly related to its intended effect, both immediate and prolonged. Phase three involves the identification of specific indicators that can measure these elements. A thorough assessment will be made to differentiate between potential indicators and the indicators that are currently in use, facilitating an integration of new metrics with existing ones. As part of this process, some of the learnings included that even though some of these metrics were being measured, the cycles of measurement did not align with the scope of the project, and therefore new data collection mechanisms needed to be implemented. A mixed of qualitative and quantitative measures are used. To develop the qualitative measures, the research team developed interview guides based on the outcomes being measured, and taking into account interview guides from other sites trialing AAL technology. To measure the service user satisfaction an adapted version of QUEST 2.0 has been used (Aledda et al., 2024) together with an adapted version of the WHO Quality of Life survey (WHOQOL) (World Health Organization, 2013). To assess the impact on workload on practitioners, the NASA TLX scale was used (Hoonakker et al., 2011). The project also aimed to determined cost savings and cost avoidance opportunities. Lastly, in the fourth phase, all the data accumulated is synthesized to derive meaningful insights.

Challenges in the project

Several challenges were identified as part of the project implementation. Firstly, there was refusal from service users and the circle of support to install the technology. Limited understanding of the benefits and privacy concerns are some of the main reasons why people do not accept trialing the technology. As part of finding ways to address these limitations, the evaluating partner involved in this project conducted a focus group with the target audience to identify some barriers and enablers of adoption. As a consequence of this work, a series of interventions are currently being explored for implementation.

Conclusions

The Independent Living Project aims to test the impact of assistive living technologies among different stakeholders, including service users, family members and circle of support, and practitioners involved in social care. This short paper summarized the approach adopted by the evaluation team to assess the impact of the technology, and outlined some of the early challenges that the project has faced.

Acknowledgments

This publication is based upon work from COST Action CA19121 – GoodBrother, Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living (https://goodbrother.eu/), supported by COST (European Cooperation in Science and Technology). In addition, the Independent Living Project received support from the Adult Social Care Digital Transformation Fund.

References

Aledda, S., Galeoto, G., Fabbrini, G., Lucibello, L., Tofani, M., Conte, A., & Berardi, A. (2024). A systematic review of the psychometric properties of Quebec user evaluation of satisfaction with assistive technology (QUEST). Disability and Rehabilitation: Assistive Technology, 19(4), 1228-1235.

Hoonakker, P., Carayon, P., Gurses, A. P., Brown, R., Khunlertkit, A., McGuire, K., & Walker, J. M. (2011). Measuring workload of ICU nurses with a questionnaire survey: the NASA Task Load Index (TLX). IIE transactions on healthcare systems engineering, 1(2), 131-143.

LBRO (2010). Impacts and outcomes toolkit [report]. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/262540/10-1406-oi-toolkit-summary.pdf

Sweeting, A., Warncken, K. A., & Patel, M. (2024). The role of assistive technology in enabling older adults to achieve independent living: past and future. Journal of medical internet research, 26, e58846.

World Health Organization (2013). WHOQOL User Manual [Report]. Available at https://www.who.int/publications/i/item/WHO-HIS-HSI-Rev.2012-3.

***Martin Kampel and Irene Ballester (2024): AI as a socio-technical system for care**. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

AI as a socio-technical system for care

Martin Kampel, Irene Ballester
Computer Vision Lab, Vienna University of Technology
martin.kampel@tuwien.ac.at, irene.ballester@tuwien.ac.at

Abstract

Artificial Intelligence (AI) as a socio-technical system holds promise for enhancing caregiving. The "ToiletHelp" system exemplifies this by using AI to assist users during toilet use, improving safety and efficiency. However, integrating AI in care settings raises significant ethical challenges, including risks of depersonalization, bias, and the potential dehumanization of care. A robust governance model ensures fairness, transparency, and accountability. The successful implementation of AI in care requires technological innovation and careful consideration of ethical, social, and practical implications to serve best all stakeholders involved.

AI system

Artificial Intelligence (AI) as a socio-technical system offers transformative potentials in care, particularly by facilitating the daily work of caregivers, enhancing the safety of those in need, and contributing significantly to the health and well-being of users. Technologies, integrated within care environments, aim to increase the productivity of care activities and reduce the long-term costs associated with rapid demographic changes.

One such technology is “ToiletHelp” [1], which uses a rule-based engine and convolutional neural network to detect difficulties in toilet’s use for people with dementia. ToiletHelp utilizes depth maps to analyze the scene and determine if the user requires assistance. Depending on the user's current action, the system predicts the subsequent step and awaits the user's completion. Should the user encounter difficulties, ToiletHelp activates both video and verbal instructions to assist the user in proceeding to the next action [2].

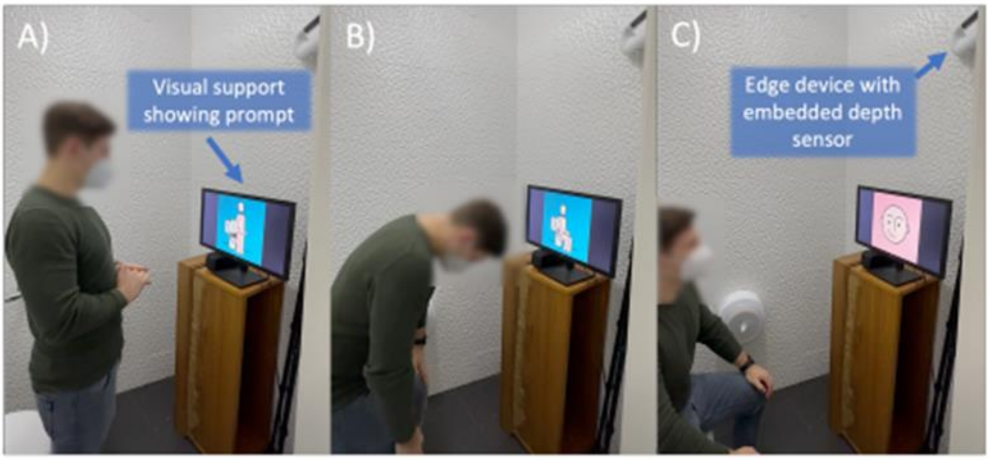


Figure 1 ToiletHelp in a care center [1]. User interaction sequence with ToiletHelp.a) Recognition of the user next to the toilet, prompting them to sit. b) User follows the prompt. c) Confirmation message reassures the user

Figure 1 illustrates the ToiletHelp system installed in a restroom. The device, a white edge device equipped with an integrated depth sensor, is mounted at an elevated position on the wall or ceiling to ensure a comprehensive view of the room, covering both the toilet bowl and the washbasin. An interaction device, consisting of a monitor coupled with a speaker, is strategically placed to be easily visible to the user. ToiletHelp processes information through four principal components: the depth sensor, the scene recognition module, the action recognition module, and the interaction module.

Governance of AI in Care

However, the integration of AI in care settings comes with its own set of challenges and ethical considerations. Risks such as the depersonalization of care through algorithm-based standardization, discrimination against minority groups due to biased algorithms, the dehumanization of care relationships through automation, and coercive control of users through surveillance technologies need careful consideration. These ethical risks require a robust governance model that ensures AI applications in healthcare are fair, trustworthy, and accountable.

The governance of AI in health care should include data governance panels to oversee data collection and usage, ensuring that AI models are designed to uphold procedural and distributive justice [3]. It is essential to educate both caregivers and patients about the functionalities and limitations of AI, ensuring fully informed consent is obtained for the use of such technologies. Furthermore, AI applications must be regulated at various stages—approval, introduction, and deployment—to ensure accountability and transparency in decision-making processes.

AI systems in care are part of complex networks that include material infrastructures such as sensors and the care infrastructure, as well as human activities and organizational practices. Different stakeholders—developers, caregivers, managers, and residents—bring diverse perspectives to the use of technology in care, influencing its development and implementation.

Care work itself is a complex network that necessitates an interdisciplinary approach, combining insights from care, technology, law, and social sciences to tackle the legal, cultural, and practical challenges associated with AI. Questions about who defines risky behaviors and what criteria lead to someone being labeled fall-prone are crucial. These discussions extend to the transparency and fairness of AI systems, particularly how behaviors are classified and the implications for those in care.

As primary users of these AI systems, older adults must be actively considered in the development processes. How older individuals perceive, understand, and interact with AI systems is fundamental to the design of these technologies. Ensuring that AI systems are understandable and explainable across different levels of technological competence among older adults is vital for their acceptance and efficacy.

The practical challenges of implementing technologies in care settings demand consideration not only of their technical aspects but also of their social implications. Legal assessments are needed to determine the classification of data used, while acknowledging cultural and regional differences in technology acceptance is crucial to ensure a comprehensive and inclusive approach to AI in care.

In conclusion, AI as a socio-technical system in care settings represents a promising frontier with the potential to enhance care quality and efficiency. However, it also demands careful consideration of ethical, social, and practical aspects to ensure that technology serves the best interests of all stakeholders involved.

References

I. Ballester, M. Gall, T. Münzer and M. Kampel, "Vision-Based Toilet Assistant for People with Dementia in Real-Life Situations," 2024 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), Biarritz, France, 2024, pp. 141-147, doi: 10.1109/PerComWorkshops59983.2024.10503526.

Lumetzberger J., Ginzinger F., Kampel M., “Sensor-based toilet instructions for people with dementia”, In: Proceedings of the AHFE 2021 Virtual Conference on Human Factors and Systems Interaction, July 25–29, 2021, USA.

Vera Gallistl et al.. 2024. Addressing the Black Box of AI – A Model and Research Agenda on the Co-constitution of Aging and Artificial Intelligence. The Gerontologist, 2024

Nipun Sandamal, Stefania Cristina, and Kenneth P. Camilleri (2024): Robust Iris Centre Localisation for Assistive Eye-Gaze Tracking*. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Robust Iris Centre Localisation for Assistive Eye-Gaze Tracking

Nipun Sandamal, Stefania Cristina, Kenneth P. Camilleri
Department of Systems and Control Engineering, Faculty of Engineering, University of Malta, Msida, Malta
Contact Author: nipun.ranasekara@um.edu.mt

Abstract

In this research work, we address the problem of robust iris centre localization in unconstrained conditions as a core component of our eye-gaze tracking platform. We investigate the application of U-Net variants for segmentation-based and regression-based approaches to improve our iris centre localisation, which was previously based on Bayes' classification. The achieved results are comparable to or better than the state-of-the-art, offering a drastic improvement over those achieved by the Bayes' classifier, and without sacrificing the real-time performance of our eye-gaze tracking platform.

Introduction

Video-based eye-gaze tracking technology offers an assistive solution to individuals with limited mobility by enabling them to interact with a device, such as a computer, through their eye and head movements alone. Fuelled by an interest in tracking the eye-gaze under unconstrained real-life conditions, years of research work have led us to develop a real-time passive eye-gaze tracking platform that estimates the eye-gaze by means of a geometrical model, which integrates information about the image position of the iris centre and the user's head pose (Cristina and Camilleri, 2016). In our work, iris centre localisation was carried out by first segmenting the iris region using a Bayes' classifier and subsequently finding the centroid of the resulting binary blob (Cristina and Camilleri, 2014). This approach worked robustly when the iris region was sufficiently illuminated and visible, but was challenged by real-life factors such as heavy eyelid occlusion and shadow, which often led the Bayes' classifier to incorrectly segment the iris region and consequently displaced the iris centre estimation.

Recent advancements in Convolutional Neural Networks (CNNs) led to such models being increasingly applied to the problem of iris segmentation over the years (Arsalan et al., 2017; Wang et al., 2020; Chen et al., 2022), with increasing complexity and computational cost. For instance, Arsalan et al. (2017) proposed a two-stage approach, which performs image enhancement and noise removal operations in the first stage, and then feeds the pre-processed image to a VGG-face model for pixel-level classification. More recently Chen et al. (2022) proposed a Dual Attention Densely Connected Network that incorporates two attention modules and skip connections.

Another interesting approach formulates iris/pupil centre localisation as a heatmap regression problem, where a heatmap is generated for any given image and the iris centre coordinates are found as the 'hottest' point in the heatmap (Xia et al., 2019; Choi et al., 2019). The model of Xia et al. (2019), for instance, is based on a VGG-16 architecture, but their performance was found to be susceptible to eye region occlusions, reflections, and shadow. Choi et al. (2019), on the other hand, made use of a U-Net model with an added skip connection between the encoder and decoder.

In light of the challenges for eye-gaze tracking posed by real-life conditions, the research work presented in this paper focuses on investigating segmentation-based and regression-based deep learning models for iris centre estimation. Our aim is to improve the accuracy of iris centre localisation, as a core component of our eye-gaze tracking platform, keeping the inference time to a minimum to retain the platform's real-time capabilities and simplifying the processing pipeline.

Method

Following preliminary experiments with several models, we have identified the U-Net family of models to be providing the most promising results for both segmentation and heatmap regression. Hence, the research work that we present in this section is based upon variants of the U-Net model.

Iris Centre Localisation via Segmentation

The U2-Net model features a two-level nested 'U' structure comprising encoder, decoder and bridge components, with symmetric encoder and decoder blocks (Qin et al., 2020). Unlike other segmentation models, the U2-Net model does not rely on a pre-trained backbone to abstract local and global features. This reduces the computational cost, making this model more suitable for real-time applications such as ours. In training the U2-Net model, we have tailored the loss function to combine region-based and boundary-aware loss functions (Terven et al., 2023), as follows:

$$\text{Total Loss} = \text{Region-Based Loss} + \text{Boundary-Aware Loss} \quad (1)$$

The Dice Loss function was used as the region-based loss function, which measures the similarity of the ground truth and predicted segmentation masks, with the aim of training the model to produce spatially coherent segmentation results. This is combined with the Boundary Loss function to emphasise accurate segmentation of the iris boundary. Following segmentation, the iris centre was localised by finding the centroid of the resulting binary blob, as in our approach in Cristina and Camilleri (2014).

Iris Centre Localisation via Heatmap Regression

A U-Net model (Ronneberger et al., 2015) was trained to generate a heatmap for a cropped eye region image, in which the iris centre was then localised as the 'hottest' point in the heatmap. The vanilla U-Net model was modified slightly by introducing CoorConv operations in its initial convolutional layers of its encoder block, with the aim of encouraging the model to learn spatial relationships between pixels. During training, a Mean Squared Loss function was used. Furthermore, the ground truth heatmaps were generated by centering a Gaussian kernel on the ground truth iris coordinates, as explained by Xia et al. (2019).

Results and Discussion

The U2-Net segmentation model was trained and tested on the MICHE 2 dataset (De Marsico et al., 2017), this being a challenging dataset featuring various eye rotations under strong shadow and reflections, and comprising ground truth segmentation masks. Table I tabulates the achieved segmentation results as *Ours (Seg)*, and compares them to the results of several state-of-the-art methods. These state-of-the-art methods have been tested on the MICHE 1 dataset (De Marsico et al., 2015), which is however encompassed and extended in the MICHE 2. Hence, for this reason the MICHE 2 is considered to be the most challenging of the two datasets. Furthermore, since the MICHE 2 dataset does not include the iris centre coordinates, these were found via centroid detection on the ground truth and predicted segmentation masks. This permitted us to compute the iris centre localisation error of the segmented masks, as tabulated in Table II by *Ours (Seg) + Centroid*, to be able to draw comparisons with the results of the U-Net.

Table I: Comparison of iris segmentation results.

Iris Segmentation				
Method	**E1**	**Precision**	**Recall**	**F1-Score**
MICHE 1 DATASET				
Arsalan et al. (2017)	0.0035	-	-	-
Wang et al. (2020)	0.82	-	-	-
Chen et al. (2022)	-	-	-	0.9319
MICHE 2 DATASET				
Ours (Seg)	0.00627	0.9499	0.9533	0.9526

Table II: Comparison of iris centre localisation results.

Iris Centre Localisation			
BioID DATASET	$d \leq 0.25$	$d \leq 0.10$	$d \leq 0.05$
Bayes' + Centroid	85.71%	57.72%	35.27%
Xia et al. (2019)	100%	99.90%	94.40%
Cai et al. (2019)	-	-	92.80%
Choi et al. (2019)	-	-	93.30%
Lee et al. (2020)	100%	98.95%	96.71%
Ours (Seg) + Centroid	99.12%	93.29%	52.18%
Ours (Reg)	100%	99.41%	96.33%
GI4E DATASET			
Bayes' + Centroid	89.91%	35.96%	15.35%
Xia et al. (2019)	100%	100%	99.10%
Cai et al. (2019)	-	-	99.50%
Choi et al. (2019)	-	-	99.60%
Lee et al. (2020)	100%	99.84%	99.84%
Ours (Seg) + Centroid	97.36%	97.36%	96.92%
Ours (Reg)	100%	99.62%	99.21%
I2Head DATASET			
Bayes' + Centroid	92.02%	59.92%	26.34%
Ours (Seg) + Centroid	99.44%	99.44%	98.51%
Ours (Reg)	100%	99.62%	98.71%
MPIIGaze DATASET			
Bayes' + Centroid	93.75%	59.38%	37.5%
Ours (Seg) + Centroid	100%	99.21%	97.65%
Ours (Reg)	100%	100%	98.23%

The U-Net model for heatmap regression was, on the other hand, trained and tested on the BioID (Jesorsky et al., 2001), GI4E (Ponz et al., 2012), I2Head (Martinikorena et al., 2018) and MPIIGaze (Zhang et al., 2015) datasets, since these included the required iris centre coordinates as ground truth information. These datasets yielded a total of 10,850 images, which were partitioned into training, validation, and testing sets with proportions of 75%, 10%, and 15%, respectively. Table II tabulates the results achieved by the trained U-Net model as *Ours (Reg)*, along with those reported by several state-of-the-art methods which regress the iris centre coordinates directly at the output. In the literature, most evaluations are carried out on the BioID and GI4E datasets, to which we have also included the I2Head and MPIIGaze to test the generalisability of our model. Our earlier approach employing Bayes' classification followed by centroid detection (Cristina and Camilleri, 2014) was also tested for comparison. The Maximum Normalised Error, denoted by d, is used as the evaluation metric, where the maximum error between left and right eye predictions is normalised by the distance between the eye centres (Cai et al., 2019). When $d \leq 0.05$, this indicates that the error is within the pupil diameter. If d ≤ 0.10, the error is within the iris diameter, whereas if d ≤ 0.25, the error is within the distance between the eye corners.

The segmentation results in Table I indicate that the trained U2-Net model performs well with respect to the state-of-the-art, albeit on an extended and more challenging dataset. This is corroborated by the qualitative results in Figure 1, showing challenging examples from the MICHE 2 dataset comprising contrasting variations of light and reflections.

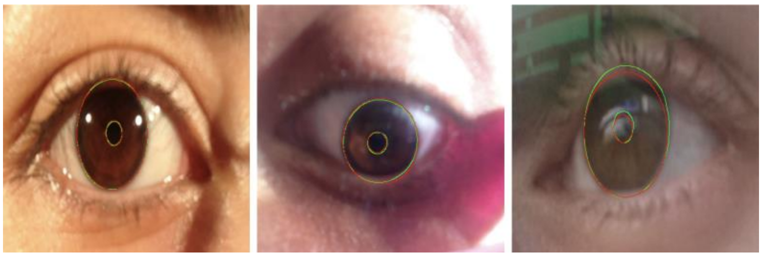


Figure 1: Segmentation showing ground truth (green) and predicted (red) iris boundaries.

The regression results in Table II indicate that the approach based on Bayes' classification performed very poorly. This was expected since the datasets under consideration contain iris occlusions and illumination variations that proved to be too challenging to the Bayes' classifier. The U-Net model trained for heatmap regression, on the other hand, can be seen to perform at par or better than the state-of-the-art. Its architecture is also sufficiently lightweight permitting it to process a cropped eye region image in an average of 20ms on an Intel Core i7-10750H machine with an NVIDIA GeForce GTX 1660 Ti GPU. It may also be noticed that the approach of performing centroid detection on the U2-Net segmentation results (*Ours (Seg) + Centroid*) also performed relatively well when considering that the U2-Net model was only trained on the MICHE 2 dataset. The error in this case is possibly stemming from the fact that the centroid does not necessarily represent the iris centre well given the convexity of the eyeball surface.

Conclusion

In this work, we have investigated the application of U-Net variants for segmentation-based and regression-based approaches to iris centre localisation. The results indicate that the heatmap regression U-Net, albeit having a relatively simpler processing pipeline, achieves comparable or better results than the state-of-the-art without sacrificing real-time performance. Future research work will be looking at integrating the heatmap regression U-Net into our real-time eye-gaze tracking platform, and testing this out within an assistive context.

Acknowledgments

This work is part of the LuminEye project, financed by the Malta Council for Science & Technology, for and on behalf of the Foundation for Science and Technology, through the FUSION: R&I Research Excellence Programme 2022.

References

Arsalan, M., H. G. Hong, R. A. Naqvi, M. B. Lee, M. C. Kim, D. S. Kim, C. S. Kim, and K. R. Park (2017): 'Deep learning-based iris segmentation for iris recognition in visible light environment'. Symmetry, vol. 9, no. 11, pp. 263.

Cai, H., B. Liu, Z. Ju, S. Thill, T. Belpaeme, B. Vanderborght, and H. Liu (2019): 'Accurate eye center localization via hierarchical adaptive convolution'.

Chen, Y., H. Gan, Z. Zeng, and H. Chen (2022): 'DADCNet: Dual attention densely connected network for more accurate real iris region segmentation'. International Journal of Intelligent Systems, vol. 37, no. 1, pp. 829–858.

Choi, J. H., K. Il Lee, Y. C. Kim, and B. Cheol Song (2019): 'Accurate Eye Pupil Localization Using Heterogeneous CNN Models'. In: 2019 IEEE International Conference on Image Processing (ICIP). pp. 2179–2183.

Cristina, S. and K. Camilleri (2014): 'Cursor Control by Point-of-Regard Estimation for a Computer With Integrated Webcam'. In: Proceedings of the 8th International Conference on Advanced Engineering Computing and Applications in Sciences (ADVCOMP). pp. 126–131.

Cristina, S. and K. P. Camilleri (2016): 'Model-based head pose-free gaze estimation for assistive communication'. Computer Vision and Image Understanding, vol. 149, pp. 157–170.

De Marsico, M., M. Nappi, and H. Proença (2017): 'Results from MICHE II – Mobile Iris Challenge Evaluation II'. Pattern Recognition Letters, vol. 91, pp. 3–10.

De Marsico, M., M. Nappi, D. Riccio, and H. Wechsler (2015): 'Mobile Iris Challenge Evaluation (MICHE)-I, biometric iris dataset and protocols'. Pattern Recognition Letters, vol. 57, pp. 17–23.

Jesorsky, O., K. J. Kirchberg, and R. W. Frischholz (2001): 'Robust face detection using the hausdorff distance'. In: Audio-and Video-Based Biometric Person Authentication: Third International Conference, AVBPA 2001 Halmstad, Sweden. pp. 90–95.

Lee, K. I., J. H. Jeon, and B. C. Song (2020): 'Deep learning-based pupil center detection for fast and accurate eye tracking system'. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIX 16. pp. 36–52.

Martinikorena, I., R. Cabeza, A. Villanueva, and S. Porta (2018): 'Introducing I2Head Database'. In: Proceedings of the 7th Workshop on Pervasive Eye Tracking and Mobile Eye-Based Interaction. pp. 1–7.

Ponz, V., A. Villanueva, and R. Cabeza (2012): 'Dataset for the evaluation of eye detector for gaze estimation'. In: Proceedings of the 2012 ACM Conference on Ubiquitous Computing. New York, NY, USA, p. 681–684, Association for Computing Machinery.

Qin, X., Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand (2020): 'U2-Net: Going deeper with nested U-structure for salient object detection'. Pattern recognition, vol. 106, pp. 107404.

Ronneberger, O., P. Fischer, and T. Brox (2015): ‘U-Net: Convolutional Networks for Biomedical Image Segmentation’. Terven, J., D. M. Cordova-Esparza, A. Ramirez-Pedraza, and E. A. Chavez-Urbiola (2023): ‘Loss Functions and Metrics in Deep Learning’.

Wang, C., Y. Wang, B. Xu, Y. He, Z. Dong, and Z. Sun (2020): ‘A lightweight multi-label segmentation network for mobile iris biometrics’. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1006–1010.

Xia, Y., H. Yu, and F.-Y. Wang (2019): ‘Accurate and robust eye center localization via fully convolutional networks’. IEEE/CAA Journal of Automatica Sinica, vol. 6, no. 5, pp. 1127–1138.

Zhang, X., Y. Sugano, M. Fritz, and A. Bulling (2015): ‘Appearance-based gaze estimation in the wild’. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 4511–4520.

***Vladimir Despotovic, Peter Pocta, and Andrej Zgank (2024): Self-supervised learning approach for heart sound classification**. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

Self-supervised learning approach for heart sound classification

Vladimir Despotovic[1], Peter Pocta[2], Andrej Zgank[3]
[1] Luxembourg Institute of Health (Luxembourg), [2] University of Zilina (Slovakia),
[3] University of Maribor (Slovenia)
Contact Author: vladimir.despotovic@lih.lu

Abstract

This study contributes to the ongoing efforts in leveraging computational approaches for remote monitoring of cardiovascular diseases based on heart sounds, paving the way for more efficient and accurate clinical decision support systems. We propose a model for heart sound classification pretrained in a self-supervised learning fashion using a collection of heart sound datasets recorded in clinical and non-clinical conditions. Self-supervised learning in the domain of phonocardiogram recordings was a key step that enabled the model to capture richer and more robust representations by leveraging in-domain knowledge, thus, improving generalization when the model was exposed to previously unseen data. Our model outperforms the baseline BYOL-A model pretrained on general audio by 4% in terms of balanced accuracy, and 5% measured by F1 score.

Introduction

Cardiovascular diseases (CVD) currently stand as the primary cause of mortality worldwide, responsible for roughly one-third of global deaths (Chen et al. (2021)).

Cardiac auscultation serves as a first-line, non-invasive method for screening CVDs, allowing detection of structural heart abnormalities from heart sounds. However, highly proficient and trained cardiologists are required to reach the desired level of diagnostic skills with cardiac auscultation (Mangione and Nieman (1997)), prompting the utilization of computer-assisted systems for early diagnosis and treatment of CVD patients (Reyna et al. (2022)).

Early computer-assisted systems for classification of human heart sounds were based on classical machine learning approaches, such as Hidden Markov Models (Romero-Vivas and White (2002)), Support Vector Machines (SVM) (Jiang et al. (2007)), or self-organizing maps (Dokur and Ölmez (2008)). More recently, deep learning has been also playing a key role, either based on the Convolutional Neural Networks (CNN) (Dominguez-Morales et al. (2017); Oh et al. (2020)), the Recurrent Neural Networks (RNN) (Latif et al. (2018)), or the combination of CNN and RNN (Deng et al. (2020); Fakhry and Gallardo-Antolín (2024)).

The development of automatic classification of heart sounds was often driven by benchmark challenges together with the open access datasets released for this purpose, such as the Pascal (Gomes et al. (2013)), the PhysioNet/Computing in Cardiology Challenge 2016 (Clifford et al. (2016)), the Interspeech Computational Paralinguistics Challenge (ComParE) 2018 (Schuller et al. (2018)), with a sub-challenge on heart sound classification, and the 2022 George B. Moody PhysioNet Heart Sound Challenge (Reyna et al. (2022)).

In this paper, we pre-trained the model for heart sound classification in a self-supervised fashion using a set of six publicly available heart sound datasets, allowing the model to acquire sufficient domain-specific knowledge. We evaluated a performance of the proposed model on the CinC16 dataset used in the PhysioNet 2016 challenge, achieving state-of-the-art results.

Heart sound datasets

Self-supervised learning requires a vast amount of data to train the model. We have identified and collected all the publicly available heart sound datasets up-to-date, with the characteristics of the datasets highlighted in Table I. Six datasets (Pascal, HSCT11, CDHS, Ephnogram, Open Heart, CirCor22) containing 15058 PCG recordings in total were used for pretraining the model in the self-supervised fashion, whereas the CinC16 dataset was used for an external model evaluation. The CinC16 contains heart sound recordings collected from both healthy and pathological subjects labeled into normal and and abnormal categories. As the official test set was never publicly released, we use the validation set to evaluate the model performance, which contains approximately 10% of the CinC16 dataset.

Employing diverse datasets requires harmonizing the recordings. All the PCG recordings were resampled to a sampling rate of 2 kHz. Given their length, the Ephnogram recordings were divided into 30 seconds segments, whereas the Open

Heart and Pascal recordings shorter than 6 seconds were concatenated with their copies to meet the feature extraction requirements.

Table I: Characteristics of the collected heart sound datasets.

	Participants	**Recordings**	**File length [s]**	**Duration [h]**	**Dataset use**
Pascal	NA	832	0.8 – 27.9	1.6	Pretraining
HSCT11	206	412	17.9 – 71.1	5.2	Pretraining
CDHS	76	3,875	15.0 – 34.1	18.1	Pretraining
Ephnogram	24	69	1,800	30.6	Pretraining
Open Heart	NA	1,000	1.1 – 4.0	0.7	Pretraining
CirCor22	1,568	5,272	4.8 – 80.4	33.5	Pretraining
CinC16	1,072	3,240	5.3 – 122.0	20.2	Evaluation

Heart sound classification

We have incorporated three transfer learning approaches for a heart sound classification, i.e., a feature extraction, fine-tuning, and training from scratch, with the aim to evaluate how the classification performance is affected by the domain used for model training – general sounds vs heart sounds. The Bootstrap Your Own Latent for Audio (BYOL-A) self-supervised learning model pretrained on the AudioSet dataset, containing over 2 million YouTube audio recordings, was used as a base model. The input data for BYOL-A were log-mel spectrograms, and the output was a feature vector with 3072 elements (Niizumi et al. (2023)).

When it comes to the feature extraction, the base BYOL-A model was employed directly, with only a classification head trained on the PCG recordings coming from the CinC16 dataset. In case of the fine-tuning, all the layers of the base BYOL-A model were unfrozen, and it was retrained with the PCG recordings coming from the six datasets (Pascal, HSCT11, CDHS, Ephnogram, Open Heart, CirCor22). In the training from scratch, the model with the same architecture as the base BYOL-A model, but with randomly initialized weights, was trained using the same six datasets as in the fine-tuning case. The same classification head as deployed in the feature extraction case, composed of a single hidden layer with the ReLU activation function, dropout layer, and an output layer, was used. The hyperparameters of the classification head were optimized via grid search with 5×5 cross-validation using the training CinC16 dataset, within the ranges shown in Table II. The model was trained with the AdamW optimizer, cosine annealing learning rate scheduler and binary cross-entropy loss function.

The evaluation was done with the held-out CinC16 validation dataset using the model with the optimal hyperparameters. Balanced accuracy, sensitivity (recall), specificity, F1 score and Area Under the Receiver Operating Characteristic curve (AUROC) were used as performance metrics, reported with 95% confidence intervals determined via bootstrapping with 1000 bootstrap samples.

Table II: Hyperparameter tuning.

Hyperparameter	Range
Mini-batch size	16, 32, 64, 128
Number of hidden layer nodes	64, 128, 256
Learning rate	0.00001, 0.0001, 0.001
Number of epochs	10, 20, 30
Dropout rate	0.3, 0.4, 0.5

Results and discussion

The automated heart sound classification is an important task in cardiovascular diagnostics that could be used to assist healthcare professionals in timely diagnosis, risk stratification, and monitoring of cardiovascular conditions. In this study, we explored various transfer learning strategies, including the feature extraction, fine-tuning and training the model from scratch in a self-supervised manner, in order to accurately classify heart sounds into clinically relevant categories of normal and abnormal heart sounds.

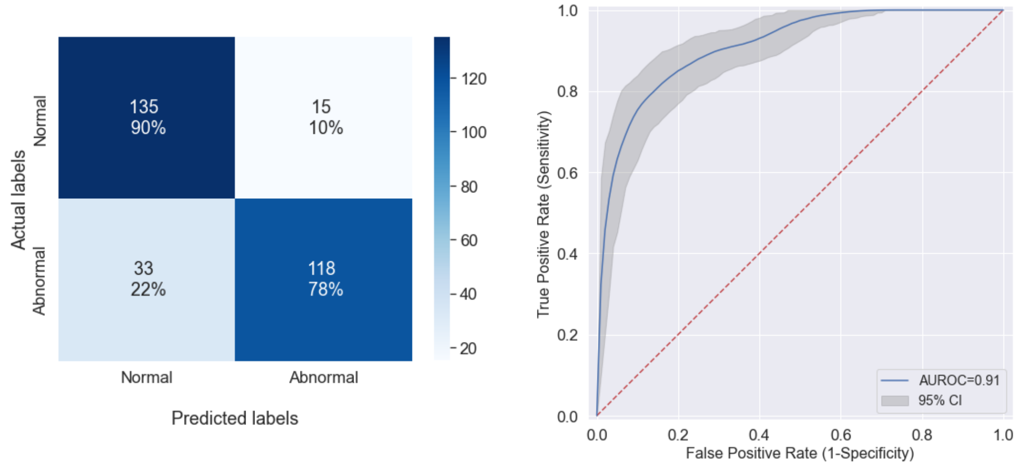
Given that the training set was highly imbalanced, with over 80% of normal PCG recordings, cost-sensitive training with the weighted binary cross-entropy cost function was deployed. Class weights were computed as the inverse class frequency normalized to value between zero and one. By incorporating class weights into the loss function, the model is penalized more for misclassifying minority class instances compared to majority class instances. On the other hand, the validation set was rather well balanced.

Performance of the heart sound classifier using the feature extraction, fine-tuning, and training from scratch is shown in Table III. Even in a zero-shot learning scenario represented by the feature extraction, where the model is pretrained on the audio data coming from entirely different domain, i.e., domain of general audio (e.g., human/animal sounds, common environmental sounds, musical instruments, etc.), it is able to learn reasonable heart sound representations, proving the fact that low level audio features are indeed shared across different audio domains. When the model is exposed to the PCG recordings during the training stage, either by fine-tuning the base BYOL-A model pretrained on the AudioSet, or training the model from scratch on the six PCG datasets, it is able to leverage the in-domain knowledge, boosting the model performance with the best balanced accuracy (0.84), F1 score (0.83) and AUROC (0.91).

Table III: Performance of the heart sound classification.

	Bal. accuracy	Sensitivity	Specificity	AUROC	Precision	F1 score
Feat. extract.	0.80 (0.75-0.84)	0.73 (0.67-0.81)	0.86 (0.80-0.91)	0.90 (0.86-0.93)	0.84 (0.78-0.90)	0.78 (0.73-0.84)
Fine-tuning	**0.84 (0.80-0.88)**	**0.78 (0.71-0.84)**	0.90 (0.85-0.94)	**0.91 (0.88-0.94)**	**0.88 (0.83-0.94)**	**0.83 (0.78-0.87)**
Train. scratch	0.83 (0.78-0.87)	0.75 (0.68-0.82)	**0.91 (0.86-0.95)**	**0.91 (0.88-0.94)**	**0.89 (0.83-0.94)**	0.81 (0.76-0.86)

All the models were better at predicting the negative class ("Normal"), but failed more often to identify the instances of the positive class ("Abnormal"), resulting in the higher specificity than the sensitivity. This behavior can be also observed in Figure 1a. Since detecting abnormal heart sounds is clinically more relevant, the fine-tuned model with the best balance between the sensitivity and specificity should be preferred. Solid model performance is confirmed by the ROC curve of the best performing model in Figure 1b, with a good trade-off between the sensitivity and specificity at different threshold levels and narrow confidence intervals, indicating that the cost-sensitive training has solved the problem of the imbalanced class distribution.



(a) (b)

Figure 1: a) Confusion matrix; b) ROC curve of the best performing model (fine-tuning with mini-batch size 32, 128 hidden layer nodes, learning rate 0.001, 30 epochs, dropout rate 0.4).

Conclusions

Our results demonstrate the potential of self-supervised learning in extracting informative features from PCG recordings and leveraging them for accurate heart sound classification. Through rigorous experimentation and evaluation, we have shown that pretraining or fine-tuning the model in domain of heart sounds with multiple PCG datasets recorded in different conditions (clinical and non-clinical) allows the model to capture richer and more robust representations by leveraging the diverse information, thus, providing improved generalization when the models are tested on previously unseen data.

Acknowledgments

This publication is based upon work from COST Action CA19121 – GoodBrother, Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living (https://goodbrother.eu/), supported by COST (European Cooperation in Science and Technology) (https://www.cost.eu/).

References

Chen, W., Q. Sun, X. Chen, G. Xie, H. Wu, and C. Xu (2021): 'Deep learning methods for heart sounds classification: a systematic review'. Entropy, vol. 23, no. 6, pp. 667., and H. Liu (2019): 'Accurate eye center localization via hierarchical adaptive convolution'.

Clifford, G. D., C. Liu, B. Moody, D. Springer, I. Silva, Q. Li, and R. G. Mark (2016): 'Classification of normal/abnormal heart sound recordings: The PhysioNet/Computing in Cardiology Challenge 2016'. In: 2016 Computing in Cardiology Conference (CinC). pp. 609–612.

Deng, M., T. Meng, J. Cao, S. Wang, J. Zhang, and H. Fan (2020): 'Heart sound classification based on improved MFCC features and convolutional recurrent neural networks'. Neural Networks, vol. 130, pp. 22–32.

Dokur, Z. and T. Ölmez (2008): 'Heart sound classification using wavelet transform and incremental self-organizing map'. Digital Signal Processing, vol. 18, no. 6, pp. 951–959.

Dominguez-Morales, J. P., A. F. Jimenez-Fernandez, M. J. Dominguez-Morales, and G. Jimenez-Moreno (2017): 'Deep neural networks for the recognition and classification of heart murmurs using neuromorphic auditory sensors'. IEEE transactions on biomedical circuits and systems, vol. 12, no. 1, pp. 24–34.

Fakhry, M. and A. Gallardo-Antolín (2024): 'Elastic net regularization and gabor dictionary for classification of heart sound signals using deep learning'. Engineering Applications of Artificial Intelligence, vol. 127, pp. 107406.

Gomes, E. F., P. J. Bentley, E. Pereira, M. T. Coimbra, and Y. Deng (2013): 'Classifying Heart Sounds - Approaches to the PASCAL Challenge'. In: International Conference on Health Informatics.

Jiang, Z., S. Choi, and H. Wang (2007): 'A new approach on heart murmurs classification with SVM technique'. In: 2007 International Symposium on Information Technology Convergence (ISITC 2007). pp. 240–244.

Latif, S., M. Usman, R. Rana, and J. Qadir (2018): 'Phonocardiographic sensing using deep learning for abnormal heartbeat detection'. IEEE Sensors Journal, vol. 18, no. 22, pp. 9393–9400.

Mangione, S. and L. Z. Nieman (1997): 'Cardiac Auscultatory Skills of Internal Medicine and Family Practice Trainees: A Comparison of Diagnostic Proficiency'. JAMA, vol. 278, no. 9, pp. 717–722.

Niizumi, D., D. Takeuchi, Y. Ohishi, N. Harada, and K. Kashino (2023): 'BYOL for Audio: Exploring Pre-trained General-purpose Audio Representations'. IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 31, pp. 137—-151.

Oh, S. L., V. Jahmunah, C. P. Ooi, R.-S. Tan, E. J. Ciaccio, T. Yamakawa, M. Tanabe, M. Kobayashi, and U. R. Acharya (2020): 'Classification of heart sound signals using a novel deep WaveNet model'. Computer Methods and Programs in Biomedicine, vol. 196, pp. 105604.

Reyna, M. A., Y. Kiarashi, A. Elola, J. Oliveira, F. Renna, A. Gu, E. A. P. Alday, N. Sadr, A. Sharma, S. Mattos, et al. (2022): 'Heart murmur detection from phonocardiogram recordings: The George B. Moody Physionet challenge 2022'. In: 2022 Computing in Cardiology (CinC), Vol. 498. pp. 1–4.

Romero-Vivas, E. and P. White (2002): ‘Heart sound segmentation by hidden Markov models’. In: IEE Seminar Digest: Medical Applications of Signal Processing, Vol. 110. p. 19.

Schuller, B. W., S. Steidl, A. Batliner, P. B. Marschik, H. Baumeister, F. Dong, S. Hantke, F. B. Pokorny, E.-M. Rathner, K. D. Bartl-Pokorny, et al. (2018): ‘The Interspeech 2018 computational paralinguistics challenge: Atypical & self-assessed affect, crying & heart beats’. In: INTERSPEECH 2018.

***Malcolm Fisk (2024): People, technologies and trust – Scripts and theories**. In: Proceedings of the Joint visuAAL-GoodBrother Conference on trustworthy video- and audio-based assistive technologies – COST Action CA19121 - Network on Privacy-Aware Audio- and Video-Based Applications for Active and Assisted Living*

People, technologies and trust – Scripts and theories

Malcolm Fisk
School of Social Work, Care and Community, University of Central Lancashire
mfisk@uclan.ac.uk

Abstract

The role of technologies in the context of care and health is changing rapidly. Challenges relate to (a) the way that people access and use technologies; and (b) the way that care and health service providers employ or deploy technologies for those who may have support needs. This paper takes steps towards identifying theoretical perspectives that help us with the development of a framework for socio-technology theory to underpin our approach to telehealth. Each perspective, albeit different, has relevance at the intersection of technologies, people and services. Six theorists are pointed to, and their work is briefly noted. Six parameters, that will be detailed further in a book, are then listed.

The Search for Theories

Many theories are concerned with people's usage of and relationships with technologies, but very few of them relate to care and health. This author has searched, therefore, for relevant theories at the intersection of (a) technologies; (b) people; and (c) care and health services. Further outcomes of this work will be published (through Palgrave Macmillan) in 2025.

It is argued that, given the pace and nature of technological developments in the health and care sector, a theoretical framework is necessary to anchor our thinking. Such an anchorage will help to ensure the correctness and appropriateness of our understandings and will position us in relation to the users of such technologies.

A Selection of Theorists

The focus is on six theorists (some other 'thinkers' are mentioned), each of whom help point to matters that might be part of a socio-technical theoretical framework for 'telehealth' (i.e. technologies and services accessed by people in relation to their care and health needs). A useful chronological marker in the search is the Second World War - with its demonstration of technology's awful power.

First is Martin Heidegger (German philosopher and Nazi) who asked the question about how technologies 'enframe' us (does things to us). He was concerned with the instrumental role of technologies and their use in the exercise of power. He signalled our need to consider these in the context of governments and politics (Heidegger, 1977).

We can note, around this time, the work of philosophers (mostly Jewish) in the Frankfurt School who moved to the United States when Hitler came to power. Their work carries the umbrella title of 'critical theory' - drawing on the work of Marx to critique people's oppression in the context of the often-authoritarian impact of capitalism. The philosophers, it can be noted, were offended by the brash and vulgar culture of commercialism that was evident in the United States and saw a challenge arising from the expanding bureaucratisation of society (Bottomore, 1984).

In the ensuing period we can point to the contribution of our second and third theorists, Michel Foucault (French philosopher) and Stanley Joel Reiser (American academic). The context for them was one of postwar optimism. Hence some of their attention being given to societal change (including health reforms). Foucault was focused on clinical health (especially clinical settings). He gave us the notion of the 'medical gaze' as testimony to the increasingly intrusive nature of medicine, the treatments invoked, and the power exercised over people's bodies (Foucault, 2003).

Reiser was more rooted in the kind of healthcare systems that we now recognise. He was prescient in seeing technology as both a 'savior and a threat'. In the 1970s, he heralded the computer as an 'eventual replacement for the physician … and (as) an instrument which would soon eliminate the human element in data acquisition, processing and interpretation'. Importantly, he called for policy frameworks (for 'technological medicine') that would 'effectively govern its empire of machines' (Reiser, 1978).

Towards the millennium, we can note the work of Madeleine Akrich on 'scripts'. She saw these as integral to the thinking of technology designers and embodying different belief systems, biases, and practices - and potentially contributing to technological determinism (Akrich, 1994).

Our fourth and fifth theorists are Joan Tronto (American academic) and Deborah Lupton (Australian academic). Tronto has not addressed technologies, but she has provided a backdrop, with her 'moral values' of attentiveness and responsiveness, that can legitimise technology use (Tronto, 2013). Lupton, for the broader context of 'digital health' (rather than telehealth) explored different theoretical approaches – from Foucault to Marx. In so doing she drew our attention to issues of power. But she also pointed to how power is now being increasingly exercised not by physical force but through the gathering, analysis and usage of personal data (Lupton, 2018). Finally, the 'sixth' among our theorists, viz Alexander Peine and Louis Neven (Dutch academics) who have promoted the notion of socio-gerontechnology. They, in part, argued similarly to Lupton (and drawing on Akrich's scripts) by pointing to the 'very deliberate acts' of designers. In focusing on older people, they have highlighted the importance of many issues around design (for technologies and services) in the context of people who are vulnerable (Peine et al, 2021).

Six Parameters

The exploration has drawn on the six theorists above, and careful cognisance has been taken of other contemporary agendas. This has enabled six parameters to be set out that will underpin the socio-technical framework for telehealth to be detailed in the coming book. These are as follows (though liable to some adjustment):

- Human rights – regarding people's equal rights when using telehealth services regardless of personal and systemic factors.
- Power and responsibility – regarding the control of practices to ensure professional and organisational power is appropriately managed.
- Data and measurement – regarding understandings of the relevance and use of networks and how people's personal data are impacted.
- Care and relationships – regarding the meanings and importance of care and relationships in the context of technology or service use.
- Design (technology and services) – regarding care and health technology service designs and how these relate to role and function.
- Environmental sustainability – regarding the sustainability imperative, in common (for all technologies and services) to minimise adverse impact.

Each carries importance. Consideration of them will, it is considered, support responsibility in the future design of technologies and related services.

References

Akrich M (1994) 'The De-scription of Technical Objects' in Bijker W and Law J (Eds) 'Shaping Technology/Building Society', MIT Press, Cambridge pp205-224.

Bottomore T (1984) 'The Frankfurt School' Ellis Horwood Ltd and Tavistock Publications, London.
Foucault M (2003) 'The Birth of the Clinic', Routledge Classics, Abingdon. First published in 1963.
Heidegger M (1977) 'The Question Concerning Technology', Garland Publishing, New York.
Lupton D (2018) 'Digital Health: Critical and Cross-Disciplinary Perspectives', Routledge, London.
Peine A, Marshall BL, Martin W and Neven L (2021) 'Socio-Gerontechnology: Key Themes, Future Agendas' in Peine A, Marshall BL, Martin W and Neven L (Eds) 'Socio-Gerontology: Interdisciplinary Critical Studies of Ageing and Technology', Routledge, Amsterdam. pp1-23.
Reiser SJ (1978) 'Medicine and the Reign of Technology', Cambridge University Press.
Tronto J (2013) 'Caring Democracy: Markets, Equality and Justice', New York University Press.