Towards Face De-identification for Wearable Cameras

Bunyarit Puangthamawathanakun
bunyarit.puangthamawathanakun2@mail.dcu.ie
Dublin City University
Dublin, Ireland

Chonlameth Arpnikanondt
chonlameth@sit.kmutt.ac.th
School of Information Technology, King Mongkut's University of Technology Thonburi
Bangkok, Thailand

Worarat Krathu
worarat.kra@sit.kmutt.ac.th
School of Information Technology, King Mongkut's University of Technology Thonburi
Bangkok, Thailand

Graham Healy
Cathal Gurrin
graham.healy@dcu.ie
cathal.gurrin@dcu.ie
Dublin City University
Dublin, Ireland

ABSTRACT

Wearable cameras provide valuable new sources of data for health and wellness monitoring, however, such visual data brings privacy concerns. This paper proposes a prototype egocentric face de-identification system for wearable camera images by swapping the original faces with synthetic faces. The motivation of this paper is to: (1) de-identify faces in egocentric images and (2) preserve the existence of each identity in images where the source identity is altered. The system incorporates our proposed method, which promises a privacy-aware and cost-effective approach. We evaluated the system on the Ego4D audio-visual PoV diarization training set by analysing six activities where faces are visible in wearable camera data. The results show promising de-identification on the source faces while most existences remain.

CCS CONCEPTS

• **Security and privacy** → **Privacy protections**.

KEYWORDS

Lifelog, Face De-identification, Privacy

ACM Reference Format:
Bunyarit Puangthamawathanakun, Chonlameth Arpnikanondt, Worarat Krathu, Graham Healy, and Cathal Gurrin. 2023. Towards Face De-identification for Wearable Cameras. In *20th International Conference on Content-based Multimedia Indexing (CBMI 2023), September 20–22, 2023, Orléans, France.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3617233.3617276

1 INTRODUCTION

Wearable sensing technologies have become commonplace in recent years. These wearable sensors range from wrist-worn fitness trackers to wearable PoV cameras. Fitness trackers quantify the physical biometrics of the individual in terms of numbers (e.g. heart rate, footsteps, etc.). PoV cameras on the other hand are typically worn on a lanyard around the neck of the wearer and gather video or periodic images that capture what the wearer sees. While fitness trackers and similar devices 'look inwards' and record data about the physical activities of the individual, PoV wearable cameras 'look outwards' and gather data which includes images or videos of other people and environments around the individual. Wearing such PoV cameras has been referred to in the literature as Lifelogging [10]. This brings concerns around privacy and data governance of people captured by the wearable cameras [17]. Notwithstanding such privacy concerns, many studies have shown that capturing the 'lived experience' of an individual using PoV cameras provides a valuable information source that can be used for personal health & wellness analytics [1], memory support [11] or epidemiological (public health) studies [18, 20].

In order to realise the value of lifelogs and PoV cameras in data analytics for well-being insights for the individual and larger-scale epidemiological studies of the 'lived-experience' of populations, one needs to 'solve' the issue of privacy so that individuals feel comfortable to share their PoV image data with healthcare professionals and public-health specialists. One obvious action here is to simply remove all faces from the image data before it is shared or seen by people. It is our opinion that this is a blunt instrument and the value of the image for the wearer and the aesthetics of the image for the healthcare professionals would be reduced if faces are removed. Additionally, many memory or wellness applications could benefit if the faces in images are allowed to remain intact. Hence we propose a new type of lifelog data analysis in which faces are made unidentifiable, yet the image remains aesthetically pleasing. The contribution of this paper is a system design and surrogate face strategy for PoV privacy support to save computational cost and retain links of the de-identified version while diminishing face recognition capability.

2 RELATED WORK

2.1 Face De-identification

The prevalence of digital cameras with associated data analysis and sharing platforms has resulted in a need for the de-identification of visual images. Faces within digital images or videos can be manipulated by various methods: adding noise, blackout, pixelation,

CBMI 2023, September 20–22, 2023, Orléans, France

ACM ISBN 979-8-4007-0912-8/23/09.
https://doi.org/10.1145/3617233.3617276

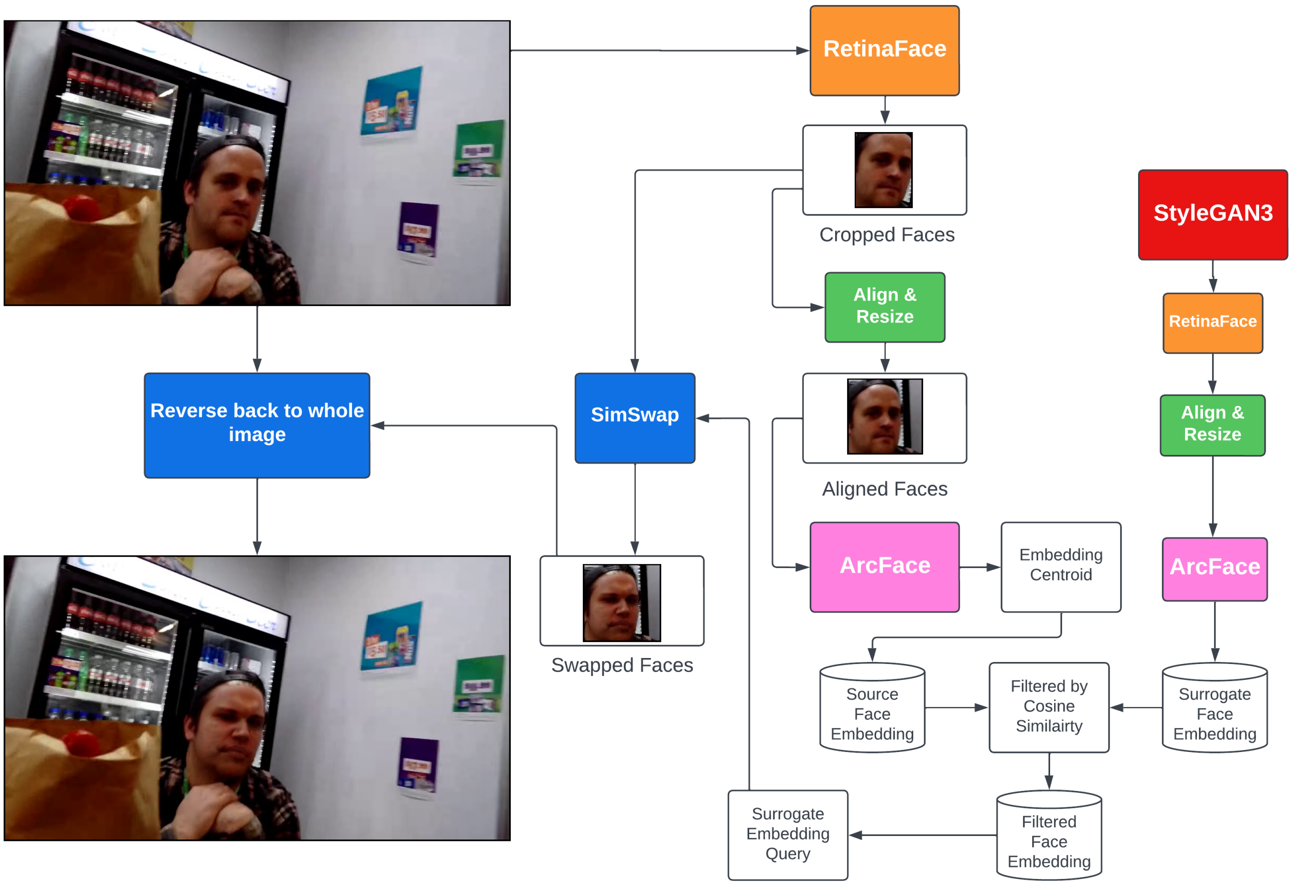

Figure 1: Overview of lifelog face de-identification system (green bounding boxes in the original image are the face detection result where red boxes indicate missing detection).

or blurring. While effective for anonymising, the results from such methods destroy the original facial information such as facial expression, gender, and race. The K-same algorithm introduced the k-anonymity property which aims to prevent face recognition from working efficiently while maintaining certain facial attributes [15]. However, the output of such methods often shows blurry content and insufficient levels of face de-identification. Many state-of-the-art approaches rely on deep-learning-based models with a number of techniques used such as inpainting and face swapping [13, 14], which better preserve facial attributes.

Face swapping [16] refers to a process to transfer an identity of a target face onto the source face while the facial attributes of the source face remain. It has wide applications in the film industry, privacy protection, etc. Regardless of attribute preservation, generalisation is also very challenging in order to be able to swap any arbitrary faces [16]. My Face My Choice (MFMC) [21] is taking advantage of face swapping in privacy enhancement by utilising synthetic faces for swapping a source image with a specific surrogate (target) face selection strategy to enhance privacy anonymisation when sharing or rendering photos on a social platform. As a result, faces in sequential images of a person might probably be replaced with different synthetic identities. Besides, images uploaded to the platform tend to have a higher quality than PoV images in terms of the positions of faces, occlusions, and blurriness. In contrast, PoV cameras automatically capture faces in the wild. Our system adopts MFMC in replacing such faces with synthetic faces while ensuring that the faces of the same person are replaced with the same synthetic face.

2.2 Egocentric Datasets

Egocentric vision has drawn great attention from researchers to study various activities from the view of the first person. There are multiple tasks studied in this field, for example, location-based segmentation [5], hand-object interaction [6], and hand gesture [19]. Often datasets are collected and published after faces are anonymised, usually using blurring, as is the case with the lifelog datasets from the LSC - Lifelog Search Challenge [9]. Such datasets are of limited use because the original facial attributes are destroyed.

Ego4D, a massive-scale egocentric dataset collected in 74 worldwide locations and containing 3,670 hours of daily-life activity video captured using PoV cameras. The dataset provides an audio-visual (AV) diarisation sub-dataset (1,945 minutes of PoV video data) that exposes bystander faces in multiple contexts, such as eating in a group, having a meeting, and social activities [7]. Each video is a single perspective of a participant that has been segmented into multiple clips. Ego4D split them into training, validation, and test set. In the training set of the sub-dataset, there are a total of 153 videos. Some videos consist of multiple sequential and non-sequential clips, while others consist of only one clip. Therefore, these clips represent short activities recorded by the PoV wearers. Table 1 and Table 2 illustrate statistics of the Ego4D AV training set and its activities' insights. We choose to use this dataset, as opposed to any other existing PoV dataset because this sub-dataset contains PoV video content that contains faces in an un-blurred manner.

3 LIFELOG FACE DE-IDENTIFICATION FRAMEWORK

3.1 Overview of LFD

Fig. 1 depicts the overview of the proposed system, which consists of six components: face detection (orange), face alignment (green), face recognition model (pink), synthetic face generator (red), and face swapping model (blue), and our proposed method, embedding centroid. Initially, the generation of synthetic faces is a critical step prior to commencing the de-identification process. A synthetic face generator is employed to produce images containing synthetic faces. Subsequently, a face detection model is used to detect and crop faces from every image/frame. After that, the cropped faces must be aligned and resized to compute a face embedding. This process is applied to both wearable camera images and synthetic images to retrieve embeddings from faces that serve as source and surrogate embeddings, respectively. When a surrogate embedding and a source face are chosen, the swapping model will replace the source face with the selected surrogate embedding and return the output face back to its original location. Each component in the system will be described in later sections in detail.

3.2 Face Detection and Face Embedding

For face detection and recognition, we employ RetinaFace [3] and ArcFace [4] implemented in InsightFace [8]. First, all faces are from each selected frame. After cropping and aligning all the faces, then we group them by their original identities at the video-clip level. Next, we extract face embedding using ArcFace [4], which employs a novel loss function giving higher compactness of intra-class and diversity for inter-class than a softmax loss function.

3.3 Embedding Centroid and Mapping Database

In order to retain the existence of each identity in every moment where the identity might appear, we believe that a mapping database and a strategy to choose a surrogate face are required. MFMC [21] uses a source face embedding to determine a surrogate face from a filtered surrogate face set based on the source embedding where randomisation and a proper threshold are applied to ensure that embeddings in the surrogate face set are not recognisable as the source face. When each source face of the same identity's faces results in a different filtered set, multiple surrogate identities will probably be chosen for a single identity. As a result, the faces of a person will be replaced with multiple surrogate identities and the existence throughout the entire clip are unlikely to be recognised as the same person.

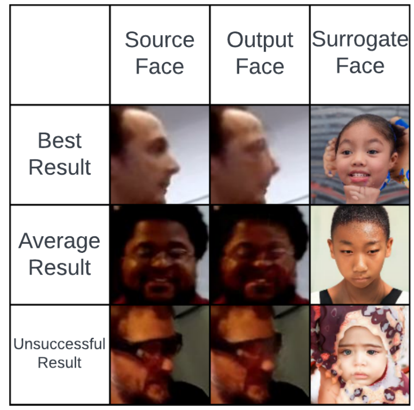

Figure 2: Example swapped faces from the proposed method

In this paper, we extend this work and propose a method to choose only one surrogate face for a single identity, which aims to facilitate de-identified individual tracking over time. First, we calculate an embedding centroid for each identity before choosing a surrogate face. The equation is as:

$$C(S) = \frac{1}{n}\sum_{i=1}^{n} E(S_i) \quad (1)$$

where $S = S_1, S_2, S_3, ..., S_n$ represents the source faces of the same identity in a single clip, C(S) is the S's embedding centroid, and E(.) is the face embedding.

Next, the cosine similarity between a surrogate face and a source face should be lower than a certain threshold to ensure that the surrogate face cannot be recognised as the source face. After calculating the embedding centroid and selecting a surrogate face, they will be saved in the mapping database for future use. This approach can optimize computational efficiency by avoiding the need to choose a surrogate face from a large surrogate face pool for each face for an identity, instead comparing the embedding centroid with existing ones. Ideally, the embedding centroid of the same individual should be consistent across all potential positions where their faces appear if the majority of the front-face images of each identity are captured by the PoV cameras. Our method will average the faces in different locations and positions in the same clip where a person might show different views which aims to assist face recognition in finding the existing source identity in the mapping database and selecting a proper surrogate face. Unfortunately, the Ego4D dataset lacks annotation of identities across the entire collection of video clips, so we will calculate an "ideal centroid" by utilizing all faces that belong to the same individual within a given clip only.

Table 1: Overview of Ego4D AV training set

# Videos	# Clips	Length (Minutes)	# Frames	# Unique Activities
153	389	1945	3,500,671	6

Table 2: Overview of activities

Activity	# Videos	# Clips	# Length(Minutes)	# Frames	# Frame with face(s)
Grocery Shopping Indoors	74	153	765	1,377,000	25,790
Play Cards/Games	42	151	755	1,358,789	40,639
Talking	38	92	460	827,743	24,367
TA session	6	17	85	152,969	5,062
Cooking	6	13	65	116,961	1,448
Outdoor social (including campfire)	4	10	50	89,990	877

3.4 Synthetic Face Pool

In this paper, we employ StyleGAN [12] to generate 24,000 synthetic faces for our surrogate face pool. It does not provide parameters to control the facial attributes over the synthetic faces, however, the bias of synthetic faces in terms of gender and age has little impact on the cosine similarity between face embeddings after face swapping [21]. When employing StyleGAN, some synthetic images may not contain detectable faces or exhibit multiple faces, which can result in improper face embeddings. So, we apply face detection and alignment in Sec. 3.2 to remove the synthetic images that do not contain only one face to ease the process. After this stage, not every generated faces can be used. A synthetic face is usable only if the cosine similarity to the source face is below a certain threshold to ensure that the face recognition model cannot reliably recognise the source face.

3.5 Swapping Model

SimSwap [2] has the best performance among three other models in MFMC [21]. For this reason, SimSwap was chosen to replace the source image with the surrogate identity. Theoretically, the ideal face-swapping model will replace only identity where other facial attributes remain. As a result, the cosine similarity between the surrogate face and the swapped face should be really high, whereas the cosine similarity with the source face is low. When surrogate face embedding is chosen, the associate aligned cropped face of the source identity will be entered into SimSwap in order to swap faces. The result will be an aligned swapped face which requires to be post-processed, which SimSwap has already provided, back to the location where the source face is.

4 EXPERIMENT

4.1 Data Preprocessing

Many lifelog datasets are limited by device constraints, hence most lifeloggers opt to record still images for every n (typically 30-60) seconds instead. To mimic a lifelog dataset, the Ego4D AV training set is employed in this experiment by extracting one frame for each second in each clip, which has a fixed 30 frame-per-second, 5 minutes duration. In PoV lifelog datasets, face tracking is very challenging because wearable cameras may be moving arbitrarily in different locations and positions. Therefore, in this work, we utilise Ego4D face tracking annotations to guide tracking a person in each clip. With face tracking and bounding box annotation of Ego4D, we merge them with bounding boxes from Insightface to retrieve the entire face tracking for each identity per video clip.

4.2 Surrogate Face Selection Strategy

The surrogate face selection strategy plays an important role in face swapping as it determines which surrogate face will be swapped onto the source face. If the source and surrogate face are highly similar, they can be recognised as the same person. Besides, if the surrogate face varies across different faces of the same identity despite being appropriately chosen to be swapped with the source face, it is improbable that the identity will be traced using the surrogate face.

Therefore, we explore three different surrogate face selection strategies. Two are from MFMC: random and furthest selection which mimic that surrogate face varies across each identity's faces. Both strategies employ the same way to make a surrogate face set created as the original paper used [21]. The third one is our proposed method, an embedding centroid, that maintains the same surrogate face for different faces of the same identity. As mentioned in Sec. 3.3, an ideal centroid will be computed for each identity for every single clip. The surrogate face will be chosen only if the cosine similarity to a threshold is lower enough to ensure unreliably recognition.

4.3 Evaluation

Since our motivation is to alleviate privacy concerns about the original identity while preserving the existence of a now unidentified individual through the entire video clip, we will evaluate our proposed method, ideal centroid, with the other two methods from My Face My Choice [21].

On one hand, we evaluate different surrogate face selection strategies by measuring the cosine similarity of each identity's face embeddings in the same clip with the corresponding output (swapped) face embedding. This will show the cosine similarity distribution in different techniques between source and output face, source and surrogate face, and surrogate and output face. The lower the cosine similarity between the source and output face, the better, as it decreases the probability of the output face being perceived as the source face. In contrast, face replacement aims to transfer an identity into the source face which is supposed to show high cosine similarity between the output face and the surrogate face. Different thresholds may be picked according to the desired confidence level on face recognition. However, the distribution of the cosine similarity between the surrogate and the source face reflects the strategy of surrogate face selection. If the output (swapped) face has a higher cosine similarity to the surrogate face than the source face, it indicates that the strategy favours a surrogate face that less resembles the source face.

On the other hand, to ensure the output faces from the same identity will be recognised as the same synthetic identity, we measure the cosine similarity of inter-class and intra-class embedding between the source face and the other three different techniques' results. We also compare the fine-tuned thresholds after the source faces are swapped by each method. Different thresholds will result in different accuracy if such a threshold is chosen to determine whether two unidentified faces are from the same identity.

4.4 Result and Analysis

Overall, all of the surrogate face selection strategies enhance the privacy of the source face. However, there are interesting results. In Fig. 3 and Fig. 4, our proposed method and MFMC random selection show the cosine similarity between the source and the swapped faces which looks more likely to have the same identity than the swapped image and the surrogate face. This can be implied that the surrogate identity cannot be properly transferred onto the source face regardless of the surrogate face selection strategy employed. Ideally, it should be the opposite because the swapped faces are supposed to resemble the surrogate faces more than the source face. In contrast to both mentioned strategies, Fig. 5, shows that MFMC furthest selection gives the lowest cosine similarity between the source and surrogate faces as the strategy's aim while the source faces are less recognisable.

Fig. 6 and Fig. 7 show the cosine similarity among inter-class and intra-class, respectively. For each surrogate face selection strategy, the cosine similarity of swapped faces is close to 1 for both classes while the cosine similarity of source face faces is distributed around -0.2 to 0.8 for the intra-class and -0.2 to 0.4 for the inter-class source faces. Typically, a threshold used for ArcFace is recommended at about 0.6 or higher for better confidence, which is inappropriate for both output faces and source faces. The reason is that the mentioned threshold could get almost perfect accuracy on inter-class and intra-class is fairly low for the source faces, whereas the output faces show the opposite. Hence, the threshold should be re-adjusted for the output faces. Table 3 shows the accuracy on output faces of each method for inter-class and intra-class regarding different thresholds. Using an output face embedding as a representative of a new identity of the source identity can be effective, but it requires a new threshold despite the surrogate face selection strategy.

Fig. 2 contains the example results of the proposed method, ideal centroid. We ranked the result based on the difference between the cosine similarity between the surrogate and the output (swapped) face and the cosine similarity between the source and the output face. The best result indicates that the output face resembles the surrogate face more than the source face. Conversely, the unsuccessful result demonstrates that the output face resembles the source face more than the surrogate face. In the last row, only glasses are altered which makes de-identification fail. This shows swapping faces cannot be done effectively when the surrogate identity is transferred to the source face while occlusion (e.g. PoV camera glasses) that cover the majority of the face is present.

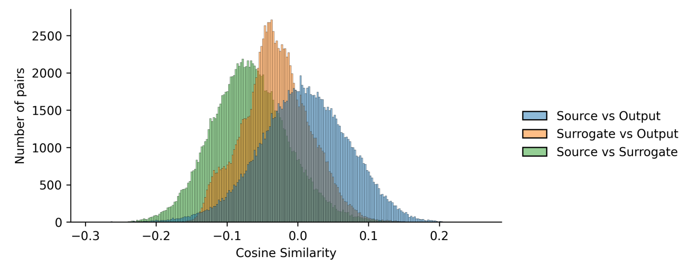

Figure 3: Embedding cosine similarity when using our proposed method

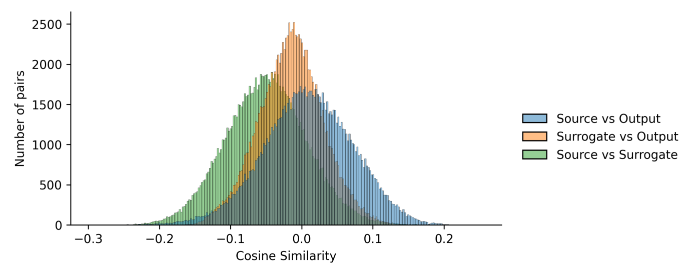

Figure 4: Embedding cosine similarity when using MFMC with random surrogate face selected

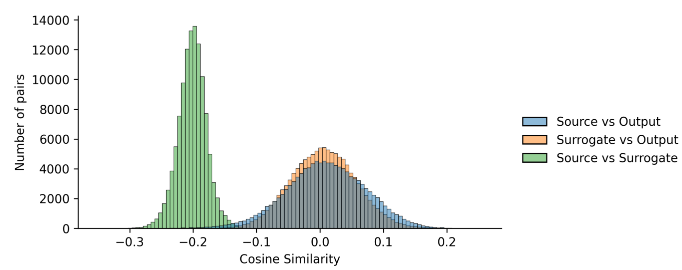

Figure 5: Embedding cosine similarity when using MFMC with the furthest surrogate face selected

4.5 Conclusion

In this paper, we have introduced the concept of face replacement in PoV wearable camera data. We have shown a prototype system and method that utilises state-of-the-art components and achieves a result that is at least worthy of additional investigation.

Our system has shown promise to support the de-identication of faces in the egocentric dataset with our system where the source is translated into the synthetic faces generated by StyleGAN. By using

Table 3: Accuracy of identity recognition on output faces for each method regarding fine-tuned threshold

Threshold	MFMC Random		MFMC Furthest		Proposed	
	Interclass	Intraclass	Interclass	Intraclass	Interclass	Intraclass
0.8800	0.7960	0.722	0.7978	0.718	0.8006	0.720
0.8850	0.7787	0.749	0.7807	0.746	0.7837	0.748
0.8900	**0.7597**	**0.775**	**0.7619**	**0.772**	**0.7653**	**0.775**
0.8950	0.7388	0.800	0.7413	0.798	0.7450	0.801
0.9000	0.7159	0.825	0.7186	0.822	0.7227	0.826

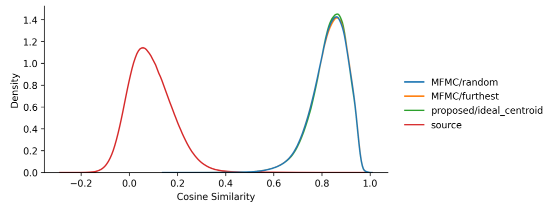

Figure 6: Cosine similarity between every pair of inter-class faces in the same clip

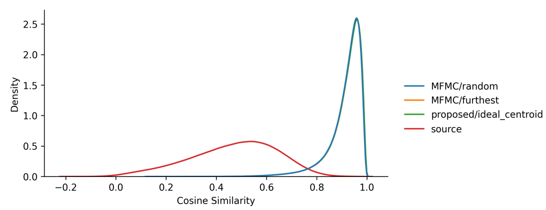

Figure 7: Cosine similarity between every pair of intra-class faces in the same clip

our proposed method, computational resources can be reduced and the majority of the existence of the same identity's faces can be maintained after being de-identified.

It is our conjecture that this approach shows promise and is a starting point in the increasingly important topic of PoV camera de-identification. We have referenced some early research that has been using PoV cameras for healthcare and epidemiological studies. As the use of PoV cameras to capture the lived experience becomes more popular, such a de-identification system will be necessary. While our initial approach is very preliminary, we believe that it shows promise and we will continue to develop new approaches and work with practitioner partners to solve real-world healthcare challenges in a privacy-aware manner.

ACKNOWLEDGMENTS

This research was conducted with the financial support of Science Foundation Ireland under Grant Agreement Nos. 18/CRT/6223 (SFI Centre for Research Training in Artificial Intelligence) and SFI/13/RC/2106_P2 (Adapt Centre).

REFERENCES

[1] S Cartledge, M Rogerson, T K R Singh, V Huynh Huu, D Phung, C Gurrin, C Neil, K Ball, and Maddison. 2019. Seeing is believing: the feasibility and acceptability of using wearable cameras to enhance self-management of heart failure. *European Heart Journal* 40 (10 2019). https://doi.org/10.1093/eurheartj/ehz747.0094

[2] Renwang Chen, Xuanhong Chen, Bingbing Ni, and Yanhao Ge. 2020. SimSwap: An Efficient Framework For High Fidelity Face Swapping. In *Proceedings of the 28th ACM International Conference on Multimedia*. 2003–2011. https://doi.org/10.1145/3394171.3413630 arXiv:2106.06340 [cs].

[3] Jiankang Deng, Jia Guo, Evangelos Ververas, Irene Kotsia, and Stefanos Zafeiriou. 2020. Retinaface: Single-shot multi-level face localisation in the wild. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 5203–5212.

[4] Jiankang Deng, Jia Guo, Jing Yang, Niannan Xue, Irene Kotsia, and Stefanos Zafeiriou. 2022. ArcFace: Additive Angular Margin Loss for Deep Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 10 (Oct 2022), 5962–5979. https://doi.org/10.1109/TPAMI.2021.3087709

[5] Antonino Furnari, Sebastiano Battiato, and Giovanni Maria Farinella. 2018. Personal-Location-Based Temporal Segmentation of Egocentric Video for Lifelogging Applications. *Journal of Visual Communication and Image Representation* 52 (2018), 1–12. https://doi.org/10.1016/j.jvcir.2018.01.019

[6] Guillermo Garcia-Hernando, Shanxin Yuan, Seungryul Baek, and Tae-Kyun Kim. 2018. First-Person Hand Action Benchmark With RGB-D Videos and 3D Hand Pose Annotations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[7] Kristen Grauman, Andrew Westbury, Eugene Byrne, Zachary Chavis, Antonino Furnari, Rohit Girdhar, Jackson Hamburger, Hao Jiang, Miao Liu, Xingyu Liu, Miguel Martin, Tushar Nagarajan, Ilija Radosavovic, Santhosh Kumar Ramakrishnan, Fiona Ryan, Jayant Sharma, Michael Wray, Mengmeng Xu, Eric Zhongcong Xu, Chen Zhao, Siddhant Bansal, Dhruv Batra, Vincent Cartillier, Sean Crane, Tien Do, Morrie Doulaty, Akshay Erapalli, Christoph Feichtenhofer, Adriano Fragomeni, Qichen Fu, Abrham Gebreselasie, Cristina Gonzalez, James Hillis, Xuhua Huang, Yifei Huang, Wenqi Jia, Weslie Khoo, Jachym Kolar, Satwik Kottur, Anurag Kumar, Federico Landini, Chao Li, Yanghao Li, Zhenqiang Li, Karttikeya Mangalam, Raghava Modhugu, Jonathan Munro, Tullie Murrell, Takumi Nishiyasu, Will Price, Paola Ruiz Puentes, Merey Ramazanova, Leda Sari, Kiran Somasundaram, Audrey Southerland, Yusuke Sugano, Ruijie Tao, Minh Vo, Yuchen Wang, Xindi Wu, Takuma Yagi, Ziwei Zhao, Yunyi Zhu, Pablo Arbelaez, David Crandall, Dima Damen, Giovanni Maria Farinella, Christian Fuegen, Bernard Ghanem, Vamsi Krishna Ithapu, C. V. Jawahar, Hanbyul Joo, Kris Kitani, Haizhou Li, Richard Newcombe, Aude Oliva, Hyun Soo Park, James M. Rehg, Yoichi Sato, Jianbo Shi, Mike Zheng Shou, Antonio Torralba, Lorenzo Torresani, Mingfei Yan, and Jitendra Malik. 2022. Ego4D: Around the World in 3,000 Hours of Egocentric Video. *arXiv:2110.07058 [cs]* (Mar 2022). http://arxiv.org/abs/2110.07058 arXiv:2110.07058.

[8] Jia Guo, Jiankang Deng, Xiang An, Jack Yu, and Baris Gecer. 2023. InsightFace: 2D and 3D Face Analysis Project. https://github.com/deepinsight/insightface

[9] Cathal Gurrin, Klaus Schoeffmann, Hideo Joho, Bernd Munzer, Rami Albatal, Frank Hopfgartner, Liting Zhou, and Duc-Tien Dang-Nguyen. 2019. A Test Collection for Interactive Lifelog Retrieval. In *MultiMedia Modeling*, Ioannis Kompatsiaris, Benoit Huet, Vasileios Mezaris, Cathal Gurrin, Wen-Huang Cheng, and Stefanos Vrochidis (Eds.). Springer International Publishing, Cham, 312–324.

[10] Cathal Gurrin, Alan F. Smeaton, and Aiden R. Doherty. 2014. LifeLogging: Personal Big Data. *Found. Trends Inf. Retr.* 8, 1 (jun 2014), 1–125. https://doi.org/10.1561/1500000033

[11] Steve Hodges, Lyndsay Williams, Emma Berry, Shahram Izadi, James Srinivasan, Alex Butler, Gavin Smyth, Narinder Kapur, and Ken Wood. 2006. SenseCam: A Retrospective Memory Aid. In *UbiComp 2006: Ubiquitous Computing*, Paul Dourish and Adrian Friday (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 177–193.

[12] Tero Karras, Miika Aittala, Samuli Laine, Erik Härkönen, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2021. Alias-free generative adversarial networks. *Advances in Neural Information Processing Systems* 34 (2021), 852–863.

[13] Yongxiang Li, Qianwen Lu, Qingchuan Tao, Xingbo Zhao, and Yanmei Yu. 2021. SF-GAN: Face De-Identification Method Without Losing Facial Attribute Information. *IEEE Signal Processing Letters* 28 (2021), 1345–1349. https://doi.org/10.

1109/LSP.2021.3067517

[14] Jiacheng Lin, Yang Li, and Guanci Yang. 2021. FPGAN: Face de-identification method with generative adversarial networks for social robots. *Neural Networks* 133 (Jan 2021), 132–147. https://doi.org/10.1016/j.neunet.2020.09.001

[15] E.M. Newton, L. Sweeney, and B. Malin. 2005. Preserving privacy by de-identifying face images. *IEEE Transactions on Knowledge and Data Engineering* 17, 2 (Feb 2005), 232–243. https://doi.org/10.1109/TKDE.2005.32

[16] Thanh Thi Nguyen, Quoc Viet Hung Nguyen, Dung Tien Nguyen, Duc Thanh Nguyen, Thien Huynh-The, Saeid Nahavandi, Thanh Tam Nguyen, Quoc-Viet Pham, and Cuong M Nguyen. 2022. Deep learning for deepfakes creation and detection: A survey. *Computer Vision and Image Understanding* 223 (2022), 103525.

[17] Kieron O'Hara, Mischa Tuffield, and Nigel Shadbolt. 2009. Lifelogging: Privacy and Empowerment with Memories for Life. *Identity in the Information Society* 1 (12 2009), 155–172. https://doi.org/10.1007/s12394-009-0008-4

[18] Louise Signal, James Stanley, Moira Smith, M. Barr, Tim Chambers, J. Zhou, Aaron Duane, Cathal Gurrin, Alan Smeaton, Christina Mckerchar, Amber Pearson, Janet Hoek, Gabrielle Jenkin, and C. Mhurchu. 2017. Children's everyday exposure to food marketing: An objective analysis using wearable cameras. *International Journal of Behavioral Nutrition and Physical Activity* 14 (10 2017). https://doi.org/10.1186/s12966-017-0570-3

[19] Yifan Zhang, Congqi Cao, Jian Cheng, and Hanqing Lu. 2018. Egogesture: a new dataset and benchmark for egocentric hand gesture recognition. *IEEE Transactions on Multimedia* 20, 5 (2018), 1038–1050.

[20] Qianling Zhou, Di Wang, Cliona Mhurchu, Cathal Gurrin, Jiang Zhou, Yu Cheng, and Haijun Wang. 2019. The use of wearable cameras in assessing children's dietary intake and behaviours in China. *Appetite* 139 (04 2019). https://doi.org/10.1016/j.appet.2019.03.032

[21] Umur A. Çiftçi, Gokturk Yuksek, and İlke Demir. 2023. My Face My Choice: Privacy Enhancing Deepfakes for Social Media Anonymization. 1369–1379. https://openaccess.thecvf.com/content/WACV2023/html/Ciftci_My_Face_My_Choice_Privacy_Enhancing_Deepfakes_for_Social_Media_WACV_2023_paper.html