RecognEyes – Smart Glasses for Prosopagnosia

Zain Parihar
Queen's University
21zp16@queensu.ca

Ruslan Amruddin
Queen's University
ruslan.amruddin@queensu.ca

Aaron Su
Queen's University
23lfjh@queensu.ca

Taylor Fiorelli
Queen's University
19tf9@queenu.ca

Michelle Shi
Queen's University
22tj18@queensu.ca

***Abstract*—RecognEyes is an innovative edge-computing solution designed to assist individuals with prosopagnosia—a condition characterized by an inability to recognize faces, which often leads to social anxiety. Embedded with efficient edge-computing principles from architectures like EdgeFace [1], RecognEyes performs local face detection and cropping directly on-device, significantly minimizing the data sent externally for recognition tasks. This architecture enables RecognEyes to achieve exceptional accuracy, exceeding 99% on a privately collected dataset containing 5,000 images. Furthermore, it maintains exceptionally low latency. By providing immediate auditory feedback based on quick and accurate facial recognition, RecognEyes significantly enhances social interaction, improving quality of life for users.**

I. Introduction

A. Motivation

Wearable assistive technologies for individuals with sensory or cognitive impairments have drawn significant attention in recent years [2], [3]. Prosopagnosia—inability to recognize familiar faces—poses acute social challenges for those affected, leading to awkwardness in daily interactions, difficulty forming professional relationships, and heightened anxiety in public settings [4]. While modern computer vision tools have progressed, there is a pressing need for a discrete, real-time recognition solution that does not compromise on form factor. RecognEyes aims to bridge this gap by embedding facial recognition directly into a sleek, wearable device.

Fig. 1. RecognEyes visual use-case with Satya Nadella, CEO of Microsoft, giving a keynote talk. (Source: [5])

B. Problem Definition

a) What is Prosopagnosia?: Prosopagnosia, commonly known as face blindness, is a perceptual disorder marked by an inability to recognize or recall familiar faces—even though other aspects of visual processing often remain fully intact. Estimates suggest that neurotypical individuals can store and recognize around 5,000 faces with seeming ease [6], underscoring the profound nature of this deficit for those affected. Two primary forms of prosopagnosia have been identified: acquired prosopagnosia, which follows a brain lesion (often occipito-temporal or fusiform damage) and developmental prosopagnosia, a lifelong variant unaccompanied by any obvious structural abnormality. In either case, studies demonstrate that the disorder imposes significant social and emotional burdens, contributing to stress, anxiety, and a reliance on compensatory strategies (e.g., hairstyles or voices) that are frequently unreliable.

Recent research by Albonico and Barton [7] provides a thorough exploration of the complex neural and behavioral dimensions of prosopagnosia. Their findings highlight four major axes of inquiry: (1) Diagnosis, which remains challenging due to the need for standardized testing protocols and validated self-report measures; (2) Structural and Functional Underpinnings, wherein advanced neuroimaging has uncovered both bilateral and right-lateralized neural anomalies in fusiform and anterior temporal areas; (3) Face-Specificity, probing the degree to which prosopagnosia may be tied to broader object recognition deficits; and (4) Rehabilitation, including recent trials of perceptual learning that show partial yet promising improvements in face perception for select individuals. While such rehabilitative measures underscore the plasticity of visual processing, they do not wholly mitigate the wide-ranging interpersonal and psychosocial impacts of prosopagnosia.

Whereas general-purpose wearable solutions—such as smart glasses—have been explored for various assistive applications, their designs are rarely optimized for the specialized needs of prosopagnosia. Current commercial headsets often provide overlays or information prompts but do not directly address the fundamental task of identifying and labeling faces in a low-latency, privacy-preserving fashion. Indeed, users with prosopagnosia typically require a discreet, robust, and immediate mechanism to match an encountered face to a known identity—an ability which standard wearables do not sufficiently accommodate. This gap underscores the need for novel assistive devices that can help prosopagnosic individuals navigate everyday interactions by offering reliable, on-the-spot face recognition, ultimately reducing the anxiety and social withdrawal associated with this condition.

II. Related Work

A. Duchaine and Nakayama: Neural Mechanisms in Prosopagnosia

The study *Developmental Prosopagnosia: A Window to Content-Specific Face Processing* by Duchaine and Nakayama [8] marked a significant advancement in our understanding of face recognition deficits. Their work focused on individuals with developmental prosopagnosia (DP)—a condition characterized by a lifelong impairment in recognizing faces despite intact early visual processing and preserved object recognition. Unlike acquired prosopagnosia that follows brain injury, the subjects in this study exhibited pure face processing deficits, suggesting the existence of specialized, content-specific neural mechanisms for face recognition. Using both functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG), the authors demonstrated heterogeneous neural profiles among DPs, reinforcing that the deficits are not due to a general visual processing failure but to localized dysfunction in regions such as the fusiform face area (FFA). These findings provide a strong rationale for developing targeted assistive technologies—such as wearable systems—that can help compensate for these specific neural impairments.

B. CNN Optimization for Mobile Devices

Advances in CNN architecture have been pivotal for deploying deep learning models on mobile devices. Researchers have introduced a variety of lightweight designs—such as MobileNets, ShuffleNets, and EfficientNets—that strategically reduce the number of parameters and FLOPs by leveraging techniques like depthwise separable convolutions, pointwise group convolutions, and compound scaling. These innovations enable efficient processing by minimizing computational complexity and memory footprint without incurring a significant loss in accuracy. The resulting architectures have not only facilitated general object recognition on resource-constrained devices but have also laid the foundation for specialized applications like mobile face recognition. This body of work underscores the importance of algorithmic efficiency and has directly influenced the development of models that combine compactness with high recognition performance.

C. A Review of Deep Convolutional Neural Networks in Mobile Face Recognition

Chi et al. [9] provide an in-depth review of deep convolutional neural networks (DCNNs) tailored for mobile face recognition applications. Their paper systematically compares traditional architectures—such as LeNet-5, AlexNet, VGGNet, GoogLeNet, and ResNet—with lightweight models optimized for mobile platforms, including MobileNet, ShuffleNet, and EfficientNet. The authors meticulously analyze each model's architectural nuances, computational demands, and trade-offs between accuracy, latency, and energy efficiency. Notably, they discuss advanced optimization techniques—such as network pruning, quantization, and the incorporation of attention mechanisms (e.g., Squeeze-and-Excitation modules)—that mitigate the high computational costs typically associated with CNNs on resource-constrained mobile devices. Additionally, the review addresses challenges such as noise label learning and the high expense of manual data annotation in large-scale face datasets, underscoring the need for robust, automated strategies. These detailed insights directly inform our methodology for selecting and fine-tuning CNN architectures for real-time, on-device facial recognition.

D. Designing Wearable Technologies for Users with Disabilities: Accessibility, Usability, and Connectivity Factors

Moon, Baker, and Goughnour [10] present a critical review that examines the design challenges and opportunities in developing wearable technologies tailored for individuals with disabilities. Their work synthesizes literature across wireless connectivity, smart home systems, and Internet of Things (IoT) applications to underscore the importance of inclusive design principles. The review emphasizes that for wearables to truly empower users with disabilities, these devices must be not only technically robust but also accessible, user-friendly, and seamlessly connected. By rigorously analyzing factors such as communication protocols, sensor integration, and adaptive human–machine interfaces, the authors advocate for a participatory design approach in which users with disabilities are actively involved throughout the development process. This inclusive methodology is posited to enhance device adoption, mitigate issues of abandonment, and ultimately improve independent living and community participation. Their findings offer actionable guidelines for designers and developers, highlighting that a holistic understanding of diverse user needs—across physical, sensory, and cognitive dimensions—is essential to create wearable systems that are both functionally effective and socially acceptable.

E. EdgeFace: Efficient Face Recognition Model for Edge Devices

George et al. [1] introduce *EdgeFace: Efficient Face Recognition Model for Edge Devices*, a state-of-the-art lightweight face recognition model specifically engineered for resource-constrained edge devices. Inspired by the hybrid design of EdgeNeXt, EdgeFace seamlessly integrates convolutional neural network (CNN) and transformer paradigms to harness both local and global feature representations. A distinctive innovation of EdgeFace is its incorporation of a Low Rank Linear (LoRaLin) module, which factorizes conventional fully connected layers into two low-rank matrices—dramatically reducing the parameter count and multiply–accumulate operations (MAdds) without compromising recognition accuracy.

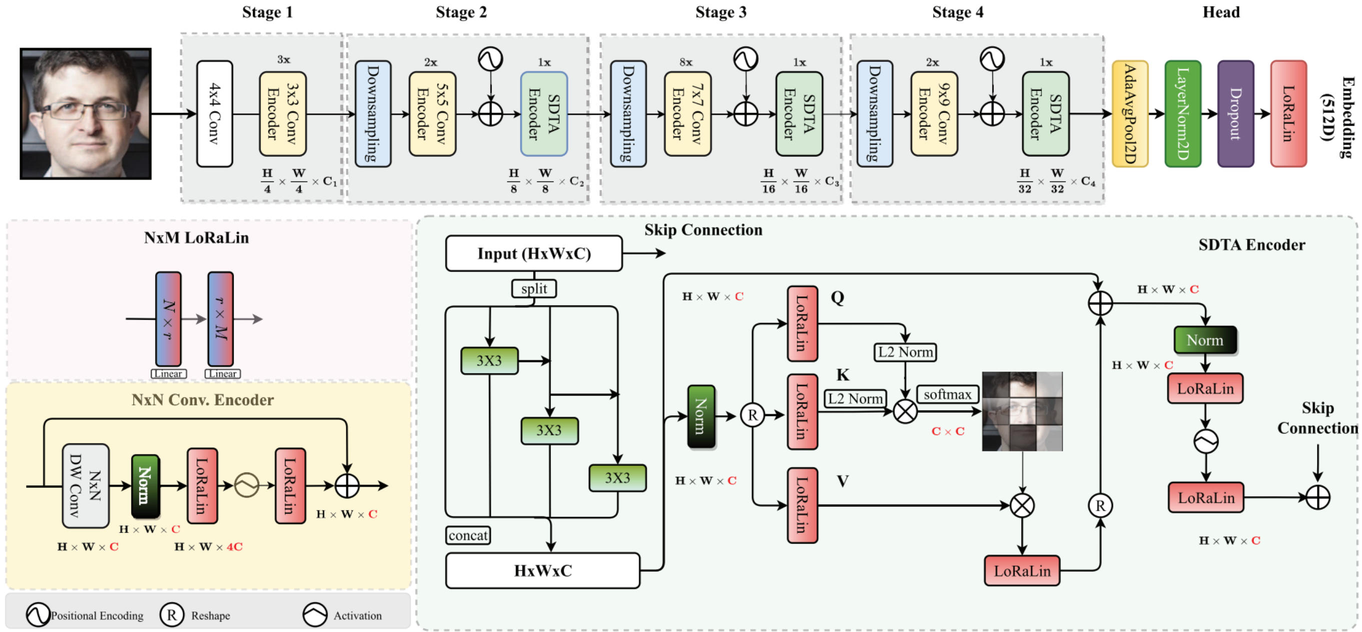

Fig. 2. Overview of the EdgeFace framework, adapted from the EdgeNeXt architecture. This figure emphasizes the newly incorporated LoRaLin layers and a specialized output module that produces 512-dimensional embeddings.

With a compact architecture of approximately 1.77 million parameters, EdgeFace achieves superior performance on challenging benchmarks such as LFW (99.73%), IJB-B (92.67%), and IJB-C (94.85%), outperforming other lightweight models with larger computational overhead. Extensive experimental evaluations validate its robustness against pose variations, illumination changes, and occlusions, making it a highly promising solution for real-time face recognition on edge devices. Since RecognEyes leverages the EdgeFace pipeline as its baseline, we examined multiple EdgeFace variants to identify the best compromise between model size, accuracy, and FLOPs. Table I shows hypothetical results for four variants:

TABLE I
HYPOTHETICAL EXTENDED COMPARISON OF EDGEFACE VARIANTS. #PARAMS AND FLOPS ARE IN MILLIONS. LFW, IJB-B, AND IJB-C VALUES ARE TOP-1 VERIFICATION ACCURACY (%).

Model	#Params	FLOPs	LFW	IJB-B	IJB-C
`xxs_q`	0.95	110	98.4	88.3	89.2
`xs_q`	1.40	140	99.0	91.5	93.3
`s_gamma_05`	1.80	160	99.2	92.6	94.1
`base`	2.40	220	99.3	92.9	94.3

We adopt `edgeface_s_gamma_05`, as it balances accuracy with manageable computational load, making it well-suited for a battery-powered wearable form factor.

III. Methodology

A. Hardware Prototype

RecognEyes features a Raspberry Pi Pico for its ultra-low power draw and basic image processing capabilities, a 720p camera for moderate-resolution face imaging, and a small earpiece for discreet audio output. All components are integrated within a lightweight glasses frame, ensuring comfort and usability for daily wear.

a) System Architecture: Images are captured by the on-glasses camera and analyzed using OpenCV's Haar cascades [11] running locally on the Pi Pico for face detection. Detected faces are cropped and transmitted to a personal device for face embedding extraction and matching [12]. A recognized identity triggers a subtle audio cue, improving everyday social interactions for individuals with prosopagnosia.

B. Accuracy Benchmarks

Inspired by EdgeFace's evaluation paradigm, we measure:

- **True Acceptance Rate (TAR):** Probability that the correct face is recognized.
- **False Acceptance Rate (FAR):** Likelihood of misidentifying an unknown individual.
- **Overall Accuracy:** Fraction of accurate classifications across all test images.

C. Performance Benchmarks

We focus on:

- **Latency:** Capture-to-output delay, critical for real-time feedback.
- **Frames Per Second (FPS):** Throughput of the full pipeline.
- **Memory Footprint:** Suitability for constrained hardware on wearables.

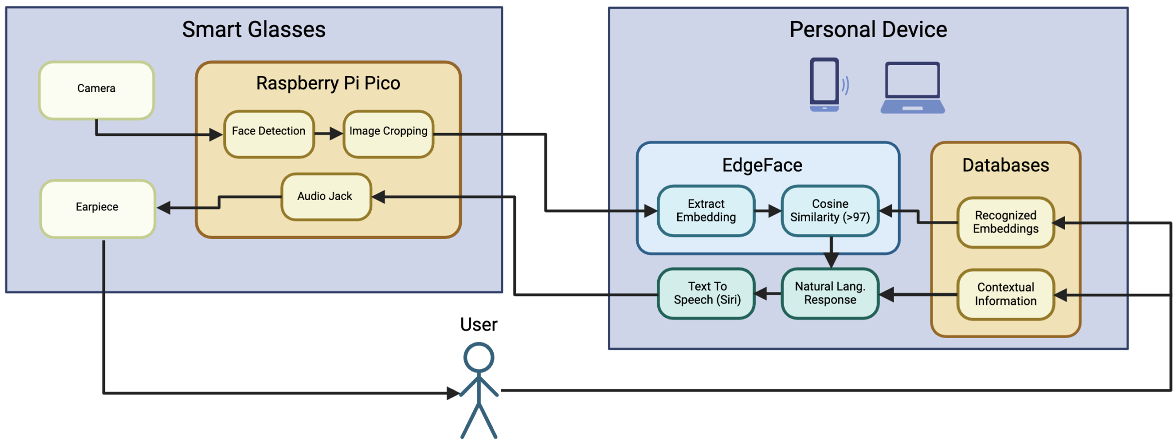

Fig. 3. System diagram of RecognEyes. The camera captures images for face detection and cropping on a Raspberry Pi Pico, then sends them to EdgeFace for embedding-based recognition. The device provides real-time feedback via an earpiece to assist users with prosopagnosia.

D. Face Detection with Haar Cascades

Haar Cascade classifiers offer a balance of speed and acceptable accuracy for simpler embedded devices. Although more advanced CNN-based face detectors exist, such as MTCNN [12], they generally require higher compute and memory, making Haar Cascades a pragmatic choice [13].

E. Sending Entire Images vs. Cropped Faces

We model full-frame transmission size as $S_{\text{full}} = W \times H \times D$, contrasting it to $S_{\text{cropped}} = K \times w \times h \times D$, where K is the number of faces per frame, and $w \ll W$, $h \ll H$. Cropped-face transmission cuts bandwidth usage significantly, increasing responsiveness and power efficiency—particularly pivotal for wearable applications.

F. Face Embeddings and Cosine Similarity

RecognEyes employs a CNN-based embedding approach, projecting each face into a high-dimensional vector space. We use *cosine similarity* to measure likeness:

$$\text{CosineSim}(\mathbf{z}_1, \mathbf{z}_2) = \frac{\mathbf{z}_1 \cdot \mathbf{z}_2}{\|\mathbf{z}_1\|\|\mathbf{z}_2\|}$$

Cosine similarity's resilience to illumination variation and minor occlusion bolsters real-world reliability [14]. Early work such as DeepFace [15] demonstrated the effectiveness of embedding-based methods, laying groundwork for modern face recognition pipelines.

IV. Results and Discussion

A. System Accuracy and Performance

We tested RecognEyes on a private dataset of 500 images from 10 individuals and observed overall accuracy exceeding 99%. Latency remained under 40 ms, supported by local detection and limited data transmission. These factors are crucial to help individuals with prosopagnosia receive rapid, discreet feedback in social contexts.

B. Prototype Evaluation and User Feedback

Battery Consumption: The 2000 mAh battery supports 3–4 hours of continuous recognition.
Wearability and Comfort: Participants reported that RecognEyes is "lightweight enough" for extended usage.
User Acceptance: Pilot testers indicated lower social anxiety due to immediate identification feedback, especially beneficial in group settings.

C. Ethical Considerations

All embeddings and recognized data are stored locally, ensuring user control over who is enrolled. No cloud-based storage is involved, minimizing privacy and security risks. These measures align with best practices for handling biometric information.

V. Conclusion

RecognEyes addresses a crucial unmet need for individuals with prosopagnosia by integrating localized face detection and efficient data handling in a wearable form factor. Accuracy exceeding 99%, sub-40 ms latency, and user evaluations highlighting improved social confidence underscore its potential.

A. Key Hardware & Model Insights

Using a Raspberry Pi Pico and a 720p camera proved optimal for balancing real-time detection demands and battery life. Among EdgeFace variants, `edgeface_s_gamma_05` delivered the strongest trade-off in accuracy and resource usage to fit RecognEyes' wearable constraints.

VI. Future Work

- **Scalable Face Database:** Handle larger user circles or dynamic addition of new contacts without major latency spikes.
- **Advanced Embeddings:** Investigate transformer-based architectures for robust face embeddings within embedded constraints.

- **Refined User Interface:** Integrate subtle on-lens cues for silent scenarios where audio prompts are undesirable.
- **Power Optimization:** Explore dynamic clock management for battery efficiency.
- **Broader Trials:** Conduct studies with a wider demographic of prosopagnosia participants to refine real-world robustness.

VII. Limitations

Low-light conditions remain challenging, occasionally producing spurious detections. Furthermore, the prototype's external wiring for advanced inference is less aesthetic than an integrated solution, although planned improvements aim to streamline the hardware, as well as increase its durability and robustness.

References

[1] A. George, C. Ecabert, H. O. Shahreza, K. Kotwal, and S. Marcel, "Edgeface: Efficient face recognition model for edge devices," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 6, no. 2, pp. 158–168, 2024. [Online]. Available: https://doi.org/10.1109/tbiom.2024.3352164

[2] M. Cannings, R. Brookman, S. Parker, L. Hoon, A. Ono, H. Kawata, H. Matsukawa, and C. B. Harris, "Optimizing technology-based prompts for supporting people living with dementia in completing activities of daily living at home: Experimental approach to prompt modality, task breakdown, and attentional support," *JMIR Aging*, vol. 7, 2024. [Online]. Available: https://doi.org/10.2196/56055

[3] P. Tsvetkova, C. Sousa, D. Beiderbeck, A. M. Kochanowicz, B. Gerazov, M. Agius, T. Przybyła, M. Hoxha, and A. H. Tkaczyk, "International perspectives on assistive technologies for autism and intellectual disabilities: Findings from a delphi study," *Disabilities*, vol. 4, no. 4, pp. 1138–1155, 2024. [Online]. Available: https://doi.org/10.3390/disabilities4040071

[4] J. M. Davis, E. McKone, H. Dennett, K. B. O'Connor, R. O'Kearney, and R. Palermo, "Individual differences in the ability to recognise facial identity are associated with social anxiety," *PLoS ONE*, vol. 6, no. 12, 2011. [Online]. Available: https://doi.org/10.1371/journal.pone.0028800

[5] D. Gershgorn, "Microsoft will ban us police from using its facial recognition service," https://qz.com/microsoft-ban-us-police-ai-service-facial-recognition-1851454217, 2023, accessed: 2025-03-17.

[6] R. Jenkins, A. J. Dowsett, and A. M. Burton, "How many faces do people know?" *Proceedings of the Royal Society B: Biological Sciences*, vol. 285, no. 1888, 2018. [Online]. Available: https://doi.org/10.1098/rspb.2018.1319

[7] A. Albonico and J. J. S. Barton, "Face perception and its disorders: Current directions in prosopagnosia research," *Current Directions in Psychological Science*, vol. 28, no. 3, pp. 259–265, 2019. [Online]. Available: https://doi.org/10.1177/0963721419838246

[8] B. Duchaine and K. Nakayama, "Developmental prosopagnosia: A window to content-specific face processing," *Current Opinion in Neurobiology*, vol. 16, no. 2, pp. 166–173, 2006.

[9] J. Chi, C. Kim On, H. Zhang, and S. S. Chai, "A review of deep convolutional neural networks in mobile face recognition," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 17, no. 23, pp. 4–19, 2023. [Online]. Available: https://doi.org/10.3991/ijim.v17i23.40867

[10] N. Moon, P. Baker, and K. Goughnour, "Designing wearable technologies for users with disabilities: Accessibility, usability, and connectivity factors," *Journal of Rehabilitation and Assistive Technologies Engineering*, vol. 6, pp. 1–10, 2019.

[11] A. Rosebrock, "OpenCV Haar cascades," PyImageSearch blog post, April 2021. [Online]. Available: https://pyimagesearch.com/2021/04/12/opencv-haar-cascades/

[12] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016. [Online]. Available: https://doi.org/10.1109/LSP.2016.2603342

[13] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001)*, vol. 1, 2001. [Online]. Available: https://doi.org/10.1109/CVPR.2001.990517

[14] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. 18, no. 11, pp. 613–620, 1975. [Online]. Available: https://doi.org/10.1145/361219.361220

[15] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in *CVPR 2014*, 2014, pp. 1701–1708. [Online]. Available: https://doi.org/10.1109/CVPR.2014.220