Proceedings of the 4th International Conference on Computing Innovation and Applied Physics
DOI: 10.54254/2753-8818/79/2025.19943

Advancing Emotion Recognition in Wearable Devices: Core Technologies, Challenges and Practical Solutions

Yuwei Ni[1,a,*]

[1]*Department of Computer Science and Technology, Xiamen University Malaysia, Sepang, Selangor, 43900, Malaysia*
a. CST2109192@xmu.edu.my
**corresponding author*

Abstract: This paper reviews the advancements and challenges in emotion recognition systems, focusing on their applications in wearable devices. It examines core technologies—multimodal data fusion, deep learning models, and personalized frameworks—in addressing critical issues such as data quality, computational efficiency, and privacy protection. By highlighting innovative solutions, including edge computing, lightweight architectures, and advanced privacy-preserving techniques, the paper emphasizes their role in enhancing the robustness, scalability, and security. By integrating real-world applications with state-of-the-art methodologies, this review underscores the transformative potential of wearable emotion recognition technologies for enhancing well-being and daily life. This review also examines current challenges in emotion recognition and provides potential solutions to address these issues.

Keywords: Wearable Devices, Emotion Recognition, Multimodal Data Fusion, Deep Learning , Personalized Models

1. Introduction

Emotions are a fundamental driving force behind human behavior and decision-making [1]. Traditionally, identifying emotional states has depended on trained professionals, which limits accessibility and often overlooks the emotional well-being of individuals in suboptimal health conditions. This limitation highlights the growing need for rapid, cost-effective, and continuous emotion monitoring solutions. The integration of wearable technology and artificial intelligence (AI) has transformed emotion recognition systems, enabling real-time collection of emotion-related data through multimodal sensors that capture heart rate, skin conductance, facial expressions, and more. Especially in psychological health and stress monitoring, early intervention and personalized care can significantly promote better mental health and overall well-being.

However, the complexity and variability of human emotions present significant challenges. Raw sensor data is prone to low signal-to-noise ratios, contextual dependencies, and interpersonal variability, complicating accurate inference. Moreover, external factors such as environmental noise and user activity further hinder the generalizability of emotion recognition models across diverse scenarios [2].

This study examines the progression of emotion recognition technologies, tracing their evolution from early physiological signal analysis to advanced AI-driven systems. It focuses on addressing key

DOI: 10.54254/2753-8818/79/2025.19943

challenges through multimodal data fusion, lightweight AI models for edge computing, and personalized frameworks, providing a comprehensive review of existing methodologies and their practical applications.

2. Key Technologies for Enhancing Wearable Emotion Recognition

As a pivotal research direction in artificial intelligence, emotion recognition poses significant challenges due to the multidimensional, dynamic, and individualistic nature of emotions [2]. Emotions integrate factors such as physiology, psychology, and behavior, creating a complex system where reliable feature extraction is highly challenging. This complexity is compounded by the dynamic nature of emotions, which can fluctuate dramatically over short periods and are heavily influenced by environmental conditions. For instance, sensor data often suffers from noise caused by factors like lighting variations, which can lead to the misrecognition of facial expressions, and user activities, where expressions during motion may be misinterpreted [3]. Furthermore, individual differences in emotional expression—shaped by physiological traits, cultural backgrounds, and contextual factors—limit the ability of general models to perform consistently across diverse users and scenarios [2]. Finally, emotion recognition heavily depends on high-quality labeled datasets, but their collection and standardization are constrained by high costs and subjective variability [4].

To address these challenges, researchers have turned to multimodal fusion, deep learning, and personalized modeling. Multimodal data fusion integrates complementary information from diverse sources, providing a comprehensive foundation for analyzing complex emotional states [5]. Deep learning models build upon this foundation, offering powerful tools to uncover intricate patterns and relationships within the fused data [6]. Techniques such as transfer learning further alleviate the reliance on large, labeled datasets by reusing knowledge from pretrained models, while personalized models dynamically adapt system behavior to individual users and contexts, making them particularly effective in applications like wearable devices or emotion-aware health monitoring [7]. Together, these frameworks enhance the adaptability and generalizability of emotion recognition systems, creating a synergistic effect: multimodal fusion enriches data diversity, deep learning strengthens pattern extraction, and personalized modeling optimizes user-specific performance.

2.1. Multimodal Data Fusion

Multimodal data fusion is a cornerstone technique in emotion recognition, integrating diverse modalities such as speech, facial expressions, and physiological signals to enhance accuracy and robustness [5]. By addressing challenges like noise, individual variability, and incomplete data, this approach enables a more comprehensive understanding of emotional states. Modality refers to a specific data source, such as physiological signals (e.g., heart rate and skin conductance), speech (e.g., pitch and tone), or visual cues (e.g., facial expressions). Each modality contributes distinct and complementary information: physiological signals reveal internal states, while speech and visual cues capture external behavioral expressions. Integrating these sources allows systems to provide a holistic and nuanced emotional assessment.

To integrate these diverse modalities, researchers have developed various fusion strategies tailored to specific challenges. Feature-level fusion combines low-level features from each modality into a unified representation, exploiting inter-modal correlations for richer feature extraction. However, this approach demands precise alignment between modalities, increasing computational complexity [8]. Decision-level fusion, in contrast, processes each modality independently and aggregates predictions through methods like weighted voting or ensemble learning. Although more flexible in handling asynchronous or missing modalities, it risks losing valuable inter-modal interactions [9]. Bridging these approaches, hybrid fusion employs a two-stage process: modality-specific features are first

DOI: 10.54254/2753-8818/79/2025.19943

aligned, then aggregated through decision-making steps. This method strikes a balance between preserving inter-modal relationships and maintaining modularity, making it a versatile solution [10].

Dynamic fusion, driven by attention mechanisms, enhances adaptability by adjusting modality weights dynamically based on contextual cues. For instance, in noisy environments where the reliability of speech data may be compromised, attention mechanisms can prioritize physiological signals, ensuring robustness in varied scenarios. By continually adapting to the quality and relevance of input modalities, dynamic fusion offers a flexible and context-aware solution, particularly for real-time emotion recognition tasks [5].

2.2. Deep Learning Models

Deep learning models play a central role in emotion recognition, offering powerful tools to process complex data and uncover patterns indicative of emotional states. These models can be broadly categorized into three types: general-purpose models, specialized models, and generative models.

Convolutional Neural Networks (CNNs) are among the most prominent general-purpose models in emotion recognition. Initially developed for image classification, CNNs excel at extracting spatial features from visual data. In emotion recognition, CNNs are widely applied to analyze static inputs, such as single-frame facial expressions [11]. These models detect subtle spatial changes, such as muscle movements in static images, enabling accurate classification of basic emotions like happiness, anger, or sadness. Beyond static inputs, CNNs are also adapted for dynamic data by processing time-series information through sequential frame analysis or converting time-series signals (e.g., EEG or heart rate variability) into 2D spatial representations. For example, CNN-based approaches have been employed to analyze changes in facial expressions over time, providing insights into emotional transitions in video sequences [12]. Similarly, in physiological signal processing, 2D representations like spectrograms allow CNNs to extract meaningful spatial patterns from temporal data. This distinction between static and dynamic input highlights CNNs' adaptability across a range of scenarios. Dynamic input analysis expands CNNs' utility to tasks requiring temporal context, such as monitoring stress levels or analyzing emotional shifts during conversations [13].

Specialized models are designed to address unique challenges posed by specific data characteristics. Autoencoders (AEs), for instance, are unsupervised models that learn compact feature representations by compressing and reconstructing input data. In emotion recognition, AEs effectively denoise physiological signals, such as EEG, extracting features that retain critical information while minimizing noise impact [14]. Graph Neural Networks (GNNs) provide another example, offering the capability to model spatial and temporal dependencies in graph-structured data, such as EEG channel networks. These models are particularly valuable in neuroscience studies, as they capture complex spatial relationships that significantly enhance emotion classification tasks [15]. Capsule Networks, by preserving spatial hierarchies, offer an alternative to CNNs and are particularly suitable for capturing micro-expressions, which often involve subtle spatial variations critical for accurate emotion detection. Unlike CNNs, Capsule Networks maintain part-whole relationships, enabling more robust analysis of spatial hierarchies and transformations [16].

Generative models have emerged as powerful tools for addressing data scarcity, a common challenge in emotion recognition. Generative Adversarial Networks (GANs), for example, generate synthetic data samples, such as facial expressions or physiological signals, to augment small datasets and improve model generalization. By generating diverse and realistic examples, GANs effectively address data imbalance and enhance training robustness. Variational Autoencoders (VAEs) complement GANs by offering probabilistic data generation and latent space exploration, further enhancing model performance in scenarios with limited training data [17].

In summary, CNNs are highly efficient for static image analysis and are widely used for facial expression recognition, but they are less suited for detecting micro-expressions or capturing complex

DOI: 10.54254/2753-8818/79/2025.19943

spatial dependencies in sequential data. Capsule Networks, while superior in handling micro-expressions, face challenges in computational efficiency and training stability. GNNs, on the other hand, excel in handling graph-structured data like EEG signals, but are limited by their computational complexity. Therefore, the choice of model depends on the specific requirements of the emotion recognition task and the available computational resources.

Certain models may overlap multiple categories, demonstrating the versatility of deep learning frameworks in addressing complex challenges in emotion recognition.

2.3. Personalized Models and Transfer Learning

Personalized models and transfer learning provide mechanisms to adapt systems to specific contexts and optimize their performance across diverse domains. By focusing on adaptability, efficiency, and scalability, these approaches have become essential in addressing the demands of modern emotion recognition frameworks.

Personalized models tailor system behavior to user-specific characteristics by capturing unique physiological and behavioral patterns. These systems fine-tune base models using small amounts of individual data, enabling adjustments to reflect personal traits, such as heart rate variability or EEG signal dynamics. Over time, incremental learning mechanisms dynamically incorporate new data by updating model parameters without overwriting previously acquired knowledge, ensuring both stability and adaptability [7]. Personalized models have demonstrated particular effectiveness in scenarios requiring precise and real-time user feedback. In psychological health monitoring, personalized models analyze deviations in baseline physiological signals, such as heart rate or skin conductance, to provide early warnings of stress or emotional imbalance [18]. Similarly, in intelligent learning environments, these models adapt teaching strategies to a learner's emotional state, ensuring optimal engagement and improving overall learning outcomes [19].

Transfer learning complements personalized models by leveraging knowledge from pretrained models to enhance performance and efficiency in new tasks or domains. By reusing high-level representations learned from large-scale datasets, transfer learning minimizes reliance on extensive labeled data in target domains. Feature-based maps transfer generic representations to specific applications, such as using visual features from image datasets to improve facial emotion recognition tasks [20]. Fine-tuning further optimizes pretrained models with task-specific data, addressing unique environmental requirements. Additionally, domain adaptation techniques align feature distributions between source and target domains, enabling models to generalize effectively across varied contexts, such as transitioning from lab-controlled data to real-world scenarios [21]. These strategies play a critical role in multimodal emotion recognition, enabling seamless integration of diverse inputs like speech, facial expressions, and physiological signals for comprehensive emotional analysis.

The combination of personalized models' fine-grained individual adaptation and transfer learning's generalizability across domains creates robust frameworks for emotion recognition. Together, they ensure efficient scaling to diverse applications while balancing the need for personalized insights with broad applicability.

3. Addressing Key Challenges in Wearable Emotion Recognition

Emotion recognition on wearable devices encounters unique challenges stemming from their mobile and user-centered nature. Issues such as noisy sensor data, resource limitations, and privacy concerns complicate the implementation of stable systems. Overcoming these barriers requires the integration of targeted advancements, enabling enhanced performance in areas like data quality, computational efficiency, and user data security.

DOI: 10.54254/2753-8818/79/2025.19943

3.1. Data Quality

The reliability and robustness of emotion recognition systems on wearable devices hinge on the quality of the multimodal data they process. Signals from heart rate, skin conductance, speech, and facial expressions must be accurately captured and integrated to infer emotional states effectively. Ensuring data quality is not just a technical requirement—it is the foundation upon which the accuracy and generalizability of emotion recognition systems depend.

One significant challenge is noise in sensor data, often resulting from motion artifacts, environmental interference, or inconsistent sensor calibration. For instance, physical activities frequently distort heart rate variability measurements, while vocal signals are particularly vulnerable to background noise in uncontrolled environments. Advanced filtering techniques, such as Kalman filters and wavelet transforms, dynamically adapt to changing signal conditions. Kalman filters smooth time-series physiological data, improving signal clarity by over 20% in real-world motion scenarios, while wavelet transforms excel in isolating frequency-specific noise components, particularly in GSR signals [22, 23].

Another pervasive issue is missing or incomplete data. Multimodal systems are highly susceptible to gaps when sensors fail, data streams become asynchronous, or specific modalities are unavailable due to user-related factors. Techniques like tensor completion have been shown to improve classification accuracy by up to 15% in scenarios with substantial signal loss, particularly in GSR and EEG datasets [23]. Lightweight generative models, such as VAEs (Variational Autoencoders), also play a vital role by synthesizing plausible replacements for visual or auditory inputs, addressing modality gaps while ensuring data consistency. These methods collectively enhance the robustness of multimodal data pipelines, ensuring reliable downstream processing [17].

Multimodal fusion networks further enhance the utility of noisy and incomplete data by integrating signals from diverse modalities. These networks employ joint representation learning and attention mechanisms to align and combine heterogeneous inputs seamlessly. For example, Multimodal Transformers utilize self-attention to capture complex dependencies across speech, visual cues, and physiological signals. Cross-modal attention mechanisms dynamically prioritize reliable modalities, such as focusing on visual cues when audio data is affected by noise, ensuring robust emotion detection even in challenging scenarios [24].

By addressing noise, missing data, and multimodal alignment, these techniques collectively improve the quality of data processed by wearable emotion recognition systems. This foundational enhancement not only ensures accurate and robust emotion detection but also enables downstream tasks, such as real-time inference and user-specific adaptation, to operate reliably and effectively.

3.2. Computational Efficiency

Wearable devices are inherently constrained by limited computational resources, energy efficiency, and physical size, making it challenging to deploy high-complexity models for real-time emotion recognition.

One primary challenge lies in the resource-intensive nature of deep learning models. Traditional architectures such as large-scale CNNs and Transformers require significant memory and processing power, often exceeding the capabilities of wearable devices. To address this, lightweight architectures like MobileNet and TinyML have been developed. These models, employing depth-wise separable convolutions, reduce parameters and computations by more than half, making them suitable for resource-constrained environments [25]. Knowledge distillation further enhances efficiency by transferring critical insights from a large teacher model to a smaller student model. Using knowledge distillation, a joint architecture demonstrated a significant reduction in computational cost—over 10

DOI: 10.54254/2753-8818/79/2025.19943

times—and model size—over 8 times—while maintaining negligible accuracy loss, highlighting its potential for resource-constrained devices [26].

Energy consumption poses another critical challenge in wearable systems, where real-time data processing and model inference can rapidly deplete limited battery reserves. Techniques like model pruning reduce computational complexity by eliminating redundant parameters and layers, while quantization, such as 8-bit integer operations, replaces high-precision floating-point calculations to enhance energy efficiency. For instance, an emotion recognition system designed for smartwatches employed hybrid quantization techniques, combining 8-bit integer quantization for core computations and binary weights for less critical layers, reducing overall energy consumption by over 40% without compromising detection accuracy. This selective approach balances computational demand with energy efficiency, making it particularly effective in resource-constrained wearable devices [27].

Edge computing plays a pivotal role in overcoming computational limitations by offloading resource-intensive tasks to nearby edge nodes or servers. This approach reduces the computational load on wearable devices, enabling real-time emotion recognition without sacrificing performance. Additionally, edge-optimized architectures like embedded RNNs, designed for low-power environments, enhance the practicality of wearable systems. For example, edge-aware CNNs leverage optimized convolutional operations and sparse attention mechanisms to reduce latency while maintaining robustness in handling multimodal inputs. To further adapt to wearable-specific requirements, these models often integrate hardware-aware optimizations, such as support for low-power chipsets and memory-efficient processing. Such designs ensure high reliability in applications like health monitoring, where low latency and accuracy are critical [28].

By integrating these strategies, wearable emotion recognition systems achieve a balance between computational efficiency and functional robustness. The combination of lightweight architectures, model optimization techniques, and edge computing frameworks ensures these systems can operate effectively within the stringent resource constraints of wearable devices.

3.3. Privacy and Security

Privacy and security are paramount concerns for wearable devices, as these systems handle highly sensitive data, including physiological signals and behavioral patterns. Advanced privacy-preserving techniques and security measures have significantly bolstered their reliability, ensured strong protection while enabled seamless functionality in diverse applications.

A key advancement in this domain is federated learning, which trains models directly on user devices, eliminating the need to transmit raw data to central servers. For example, Google's Federated Averaging algorithm enables collaborative learning without exposing individual data [29]. This approach not only safeguards privacy but also aligns with the decentralized nature of wearable devices. Some optimizations, such as quantized updates and sparse gradient transmission, have further enhanced federated learning's efficiency. Quantized updates reduce the size of transmitted gradients by representing numerical values with fewer bits, thereby decreasing communication overhead without significantly impacting accuracy. Sparse gradient transmission focuses on transmitting only the most significant updates, effectively minimizing the volume of data exchanged while maintaining model performance. These techniques enable faster convergence and better scalability in resource-constrained environments [30].

Differential privacy offers an intensive framework for securely sharing data in research and development contexts. By introducing controlled noise into datasets, differential privacy masks user-specific details while retaining aggregate data utility. Privacy budgets play a critical role in this process by specifying allowable privacy leakage during data utilization, enabling a tunable balance between privacy protection and data accuracy [31].

DOI: 10.54254/2753-8818/79/2025.19943

Hardware-based solutions such as Trusted Execution Environments (TEEs) complement these software-based methods by creating secure zones for sensitive computations. Lightweight cryptographic algorithms like elliptic-curve cryptography further enhance these efforts by minimizing overhead performance, ensuring real-time data encryption aligns with wearable devices' computational constraints [32].

4. Conclusion

This paper systematically examines the technological advancements and challenges in emotion recognition for wearable devices, highlighting the essential contributions of multimodal data fusion, deep learning models, and personalized frameworks. Through a detailed analysis, we explored how these foundational technologies address key issues, such as noisy sensor data, computational limitations, and privacy concerns, while presenting innovative solutions like edge computing, lightweight model optimization, and privacy-preserving methods.

As wearable emotion recognition systems continue to advance, their potential extends far beyond the realm of personal health monitoring. In healthcare, these systems can revolutionize mental health diagnostics and interventions, enabling continuous, real-time emotional health tracking and early identification of psychological disorders. In the workplace, emotion recognition systems can enhance productivity by fostering emotionally intelligent environments that adapt to employees' mental states, potentially reducing stress and improving job satisfaction. Furthermore, in daily life, the integration of emotion-aware technologies into wearable devices can significantly contribute to personal well-being, offering tailored insights that promote emotional regulation and overall happiness.

Looking ahead, privacy protection remains a critical concern. The need for more robust privacy-preserving models, especially in distributed networks, will continue to shape the evolution of emotion recognition technologies. Balancing high performance with data privacy is a societal necessity, as users demand greater control over their sensitive emotional data. In parallel, future research should focus on improving the scalability and adaptability of these systems, making them more efficient and accessible in diverse real-world applications.

References

[1] *A. S. Fox, R. C. Lapate, A. J. Shackman, and R. J. Davidson, The nature of emotion: Fundamental Questions. Oxford University Press, 2018.*

[2] *R. W. Picard, Affective computing. MIT Press, 2000.*

[3] *Z. Wang, S. Wang, and Q. Ji, "Capturing complex spatio-temporal relations among facial muscles for facial expression recognition," 2, Jun. 2013, doi: 10.1109/cvpr.2013.439.*

[4] *S. Koelstra et al., "DEAP: A database for Emotion Analysis; Using Physiological Signals," IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 18–31, Jun. 2011, doi: 10.1109/t-affc.2011.15.*

[5] *S. Poria, E. Cambria, R. Bajpai, and A. Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion," Information Fusion, vol. 37, pp. 98-125, 2017.*

[6] *Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436-444, 2015.*

[7] *M. Rescigno, M. Spezialetti, and S. Rossi, "Personalized models for facial emotion recognition through transfer learning," Multimedia Tools and Applications, vol. 79, no. 47–48, pp. 35811–35828, Aug. 2020, doi: 10.1007/s11042-020-09405-4.*

[8] *N. Vouitsis, S. Zhao, and M. Chen, "Data-Efficient Multimodal Fusion on a Single GPU," IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 12, pp. 5678–5690, Dec. 2023.*

[9] *S. Roheda, H. Krim, Z.-Q. Luo, and T. Wu, "Decision Level Fusion: An Event Driven Approach," in Proceedings of the 2018 IEEE International Conference on Image Processing (ICIP), Athens, Greece, Sep. 2018, pp. 2715-2719.*

[10] *F.-Z. Nakach, A. Idri, and E. Goceri, "A comprehensive investigation of multimodal deep learning fusion strategies for breast cancer classification," Artificial Intelligence Review, vol. 57, no. 12, Oct. 2024, doi: 10.1007/s10462-024-10984-z.*

[11] *N. Mehendale, "Facial emotion recognition using convolutional neural networks (FERC)," SN Applied Sciences, vol. 2, no. 3, Feb. 2020, doi: 10.1007/s42452-020-2234-1.*

DOI: 10.54254/2753-8818/79/2025.19943

[12] K. K. Talluri, M.-A. Fiedler, and A. Al-Hamadi, "Deep 3D Convolutional Neural Network for Facial Micro-Expression Analysis from Video Images," Applied Sciences, vol. 12, no. 21, p. 11078, Nov. 2022, doi: 10.3390/app122111078.

[13] L. Liakopoulos, N. Stagakis, E. I. Zacharaki, and K. Moustakas, "CNN-based stress and emotion recognition in ambulatory settings," in Proceedings of the 2021 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), Messina, Italy, Jun. 2021, pp. 1-6.

[14] Q. Li, Y. Liu, Y. Shang, Q. Zhang, and F. Yan, "Deep sparse autoencoder and recursive neural network for EEG emotion recognition," Entropy, vol. 24, no. 9, p. 1187, Aug. 2022, doi: 10.3390/e24091187.

[15] P. Zhong, D. Wang, and C. Miao, "EEG-Based emotion recognition using regularized graph neural networks," IEEE Transactions on Affective Computing, vol. 13, no. 3, pp. 1290–1301, May 2020, doi: 10.1109/taffc.2020.2994159.

[16] X. Shu, J. Li, L. Shi, and S. Huang, "RES-CapsNet: an improved capsule network for micro-expression recognition," Multimedia Systems, vol. 29, no. 3, pp. 1593–1601, Mar. 2023, doi: 10.1007/s00530-023-01068-z.

[17] Y. Luo, L.-Z. Zhu, Z.-Y. Wan, and B.-L. Lu, "Data augmentation for enhancing EEG-based emotion recognition with deep generative models," Journal of Neural Engineering, vol. 17, no. 5, p. 056021, Sep. 2020, doi: 10.1088/1741-2552/abb580.

[18] A. Tazarv, S. Labbaf, S. M. Reich, N. Dutt, A. M. Rahmani, and M. Levorato, "Personalized Stress Monitoring using Wearable Sensors in Everyday Settings," 2021 43rd Annual International Conference of the IEEE Engineering in Medicine &Amp; Biology Society (EMBC), pp. 7332–7335, Nov. 2021, doi: 10.1109/embc46164.2021.9630224.

[19] B. Grawemeyer, M. Mavrikis, W. Holmes, S. Gutiérrez-Santos, M. Wiedmann, and N. Rummel, "Affective learning: improving engagement and enhancing learning with affect-aware feedback," User Modeling and User-Adapted Interaction, vol. 27, no. 1, pp. 119–158, Feb. 2017, doi: 10.1007/s11257-017-9188-z.

[20] A. Sultana, S. K. Dey, and Md. A. Rahman, "Facial emotion recognition based on deep transfer learning approach," Multimedia Tools and Applications, vol. 82, no. 28, pp. 44175–44189, May 2023, doi: 10.1007/s11042-023-15570-z.

[21] T. Rajapakshe, R. Rana, and S. Khalifa, "Domain Adapting Speech Emotion Recognition Models to Real-World Scenarios with Deep Reinforcement Learning," arXiv preprint arXiv:2207.12248, 2022.

[22] T. Kautz and B. Eskofier, "A Robust Kalman Framework with Resampling and Optimal Smoothing," Sensors, vol. 15, no. 3, pp. 4975–4995, Feb. 2015, doi: 10.3390/s150304975.

[23] J. Yuan, Y. Yuan, F. Liu, Y. Pang, and J. Lin, "An improved noise reduction algorithm based on wavelet transformation for MEMS gyroscope," Frontiers of Optoelectronics, vol. 8, no. 4, pp. 413–418, Feb. 2015, doi: 10.1007/s12200-015-0474-2.

[24] R. G. Praveen and J. Alam, "Recursive Joint Cross-Modal Attention for Multimodal Fusion in Dimensional Emotion Recognition," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2024, pp. 4803–4813.

[25] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv preprint arXiv:1704.04861, 2017.

[26] Z.-R. Wang and J. Du, "Joint Architecture and Knowledge Distillation in Convolutional Neural Network for Offline Handwritten Chinese Text Recognition," arXiv preprint arXiv:1912.07806, Dec. 2019. [Online]. Available: https://arxiv.org/abs/1912.07806

[27] K. Choksi, H. Chen, K. Joshi, S. Jade, S. Nirjon, and S. Lin, "SensEmo: Enabling Affective Learning through Real-time Emotion Recognition with Smartwatches," arXiv preprint arXiv:2407.09911, Jul. 2024. [Online]. Available: https://arxiv.org/abs/2407.09911

[28] V. Pandelea, E. Ragusa, T. Apicella, P. Gastaldo, and E. Cambria, "Emotion recognition on edge Devices: training and deployment," Sensors, vol. 21, no. 13, p. 4496, Jun. 2021, doi: 10.3390/s21134496.

[29] Google, "Federated Learning: Collaborative Machine Learning without Centralized Training," retrieved from: https://research.google/blog/federated-learning-collaborative-machine-learning-without-centralized-training-data/

[30] A. F. Aji and K. Heafield, "Sparse Communication for Distributed Gradient Descent," in Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP), Copenhagen, Denmark, 2017, pp. 440–445.

[31] M. Kilpala, T. Kärkkäinen, and T. Hämäläinen, "Differential Privacy: an umbrella review," in Springer eBooks, 2022, pp. 167–183. doi: 10.1007/978-3-031-15030-2_8.

[32] C. Shepherd and K. Markantonakis, Trusted Execution environments. Springer, 2024.