Fatigue Detection and Face Recognition Using Deep Learning: A Survey

www.surveyx.cn

Abstract

This survey paper explores the integration of deep learning and computer vision technologies in fatigue detection and face recognition systems, emphasizing their transformative impact on accuracy and operational efficiency. The adoption of advanced neural network architectures, such as Convolutional Neural Networks (CNNs), has significantly enhanced the ability to extract intricate facial features, crucial for both fatigue detection and face recognition tasks. The paper highlights the importance of data augmentation techniques in improving the robustness and generalization capabilities of these systems, as well as the role of lightweight CNN architectures in facilitating real-time applications. Furthermore, the survey underscores the critical need for addressing biases in face recognition algorithms to ensure equitable performance across diverse demographic groups. It also discusses the challenges posed by adversarial attacks and the necessity for robust detection methods to safeguard biometric systems. The integration of multimodal and biometric approaches is identified as pivotal for enhancing recognition accuracy and robustness. The survey concludes by emphasizing the importance of continued research and development to address existing challenges, explore new opportunities, and drive innovation in fatigue detection and face recognition technologies, ultimately enhancing safety, security, and user interaction across various domains.

1 Introduction

1.1 Importance of Fatigue Detection

Fatigue detection is crucial in safety-critical environments, particularly in transportation, where driver drowsiness significantly contributes to vehicular accidents. In 2014, drowsiness was linked to 846 fatalities, underscoring the urgent need for effective fatigue detection systems to mitigate accidents due to inattention [1]. The prevalence of fatigue-related accidents necessitates the development of accurate detection mechanisms [2]. Drowsiness impairs reaction times, posing substantial risks on the road, thus highlighting the essential role of fatigue detection in enhancing road safety [3].

The relevance of fatigue detection extends beyond transportation to various safety-critical settings, such as industrial environments, where identifying fatigue can prevent hazardous situations and enhance user interaction [4]. Fatigue adversely affects both physical performance and cognitive functions, contributing to traffic accidents, medical errors, and decreased workplace productivity [5]. Consequently, detecting fatigue through physiological indicators, particularly eye-related responses, is vital for road safety and is commonly utilized in fatigue detection systems [6].

Preventing over 1.3 million annual fatalities from road accidents due to drowsy driving necessitates robust fatigue detection systems, which are essential for reducing accidents linked to fatigue and distraction [7]. This need emphasizes the critical implications of fatigue detection in safety-critical environments [8]. Furthermore, fatigue detection impacts user engagement and productivity; systems that accurately assess attentiveness through visual cues can enhance productivity and well-being in computer-based tasks, crucial for mental health and burnout prevention. The integration of thermal

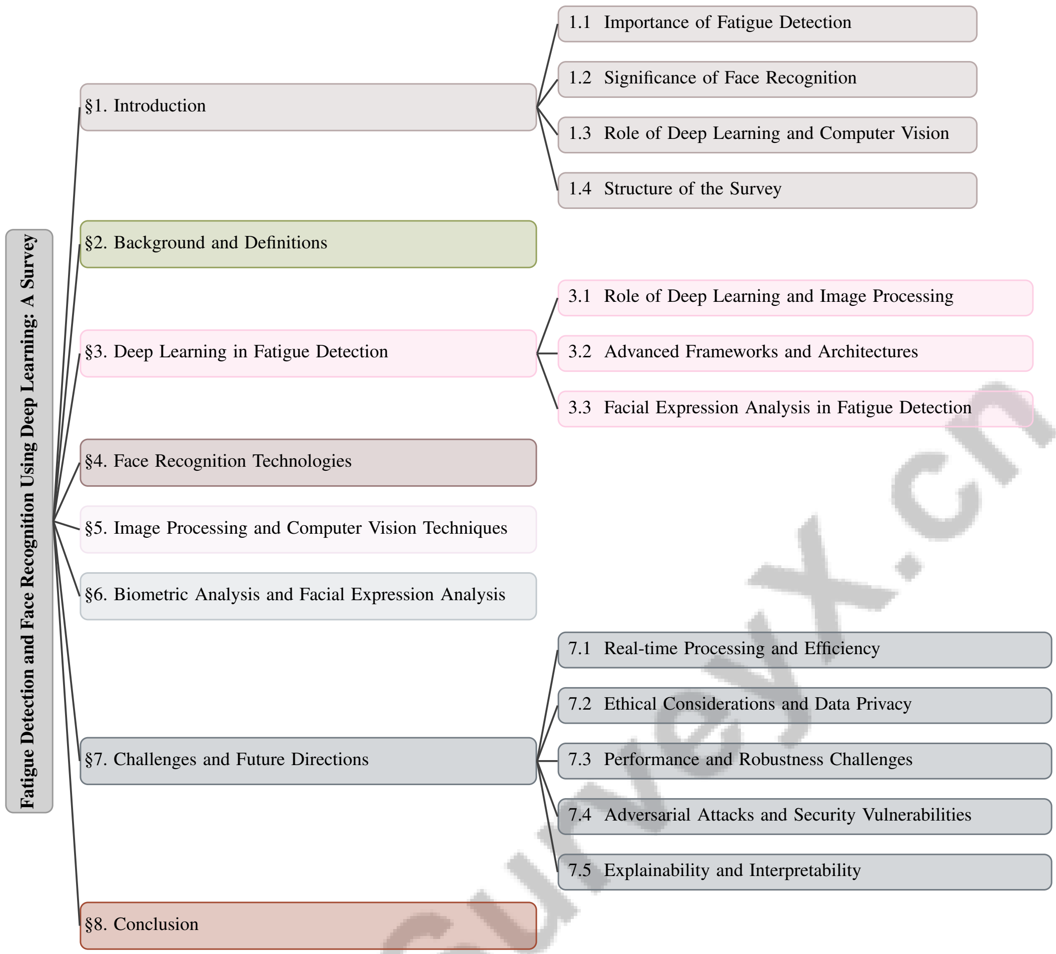


Figure 1: chapter structure

imaging in fatigue detection systems opens new research avenues in biometrics and non-verbal communication analysis [9].

1.2 Significance of Face Recognition

Face recognition has emerged as a pivotal technology in security and authentication, surpassing traditional biometric systems like fingerprint and iris recognition due to its non-intrusive nature and capability to function at a distance [10]. This feature is particularly beneficial in surveillance and access control, where thermal infrared imaging enhances robustness against variations in illumination, addressing challenges such as changes in light and disguises. In security contexts, face recognition is essential for identifying individuals and preventing unauthorized access, thereby bolstering security protocols in areas like border control and banking [11]. Its effectiveness in open-set scenarios, such as those encountered with body-worn cameras, further emphasizes its importance in dynamic environments [12]. The necessity for real-time inference and privacy considerations is paramount in these applications [12], as the ability to distinguish individuals and respond according to their privileges is vital for smart security systems [13].

In multimedia applications, face recognition addresses challenges like variations in pose, illumination, and expression [14]. Its significance extends to emotion analysis and biometric security, where precise recognition is critical [15]. Despite vulnerabilities, such as aging, ongoing research aims to enhance the resilience of face recognition technologies, particularly in age-invariant recognition, where intra-class variations due to aging pose significant challenges. The adoption of deep learning methodologies has markedly improved the accuracy and efficiency of face recognition systems

AI-generated, for reference only.

[16], facilitating seamless interactions across diverse applications, from e-commerce to personalized content delivery [17].

However, face recognition systems encounter challenges in accurately identifying individuals wearing masks, raising concerns about the validity of in-house datasets and operational viability [18]. As the demand for effective identification systems grows, face recognition remains at the forefront of biometric research, evolving to meet the challenges posed by diverse environments [19].

1.3 Role of Deep Learning and Computer Vision

Deep learning and computer vision have profoundly transformed fatigue detection and face recognition technologies by enabling sophisticated visual data analysis. The integration of deep learning models, particularly Convolutional Neural Networks (CNNs), has improved the ability to identify intricate patterns within facial features, essential for both fatigue detection and face recognition tasks [1]. These advanced models facilitate real-time classification of driver drowsiness states, such as normal, yawning, and drowsy, illustrating the transformative impact of deep learning on fatigue detection systems [20].

In fatigue detection, deep learning has enabled the development of comprehensive systems that combine image processing with biometric data, including thermal imaging and pose estimation, to enhance detection accuracy. By leveraging deep neural networks, these systems can efficiently evaluate drowsiness states from video frames, significantly improving safety measures in transportation [8]. The application of deep learning in gaze tracking exemplifies its impact, as image processing algorithms estimate gaze direction, enhancing human-computer interaction [9].

For face recognition, deep learning has revolutionized the field by creating models capable of recognizing facial expressions across varying image resolutions, thus enhancing recognition accuracy in diverse conditions [21]. The adaptation of lightweight CNN architectures for soft-biometrics prediction underscores deep learning's transformative role in biometric analysis, providing robust solutions even in challenging scenarios, such as facial mask usage [19]. Moreover, employing visual psychophysics in face recognition algorithms addresses the need for explainability, enabling controlled stimulus manipulation to study model responses [22].

Advancements in deep learning and computer vision have significantly enhanced the accuracy and efficiency of technologies for detecting facial fatigue and recognizing faces. Innovations have led to the development of sophisticated frameworks utilizing non-local attention mechanisms and CNNs, effectively analyzing facial dynamics and detecting subtle fatigue signs in real-world scenarios. The integration of extensive audiovisual datasets, such as the Daily-Life Fatigue Dataset (DLFD), has improved the ability to accurately assess fatigue levels. These breakthroughs not only enhance current biometric analysis and facial expression recognition systems but also pave the way for future innovations, addressing challenges like low-resolution images, occlusions, and varying facial expressions in diverse environments [20, 23, 24, 16, 25]. These technologies continue to evolve, offering new possibilities for improving user interaction, safety, and security across various applications.

1.4 Structure of the Survey

This survey is meticulously structured to provide a comprehensive overview of the current state of fatigue detection and face recognition technologies, emphasizing the transformative role of deep learning and computer vision. The introduction highlights the critical role of advanced computer vision technologies in enhancing safety, security, and user interaction, particularly in the context of the COVID-19 pandemic, which necessitated the development of innovative solutions for challenges such as face recognition with masked individuals, automated analysis of visual data for public health measures, and the enhancement of existing biometric authentication systems [13, 26, 18, 9, 27]. It establishes the critical need for effective fatigue detection and face recognition systems across various applications.

Section 2 delves into the background and definitions, providing clear explanations of key concepts such as fatigue detection, face recognition, deep learning, image processing, computer vision, biometric analysis, and facial expression analysis. This foundational section is essential for understanding the subsequent discussions and analyses presented in the survey.

AI-generated, for reference only.

Section 3 focuses on the application of deep learning in fatigue detection, exploring the integration of deep learning with image processing techniques, reviewing advanced frameworks and architectures, and examining the role of facial expression analysis in enhancing fatigue detection accuracy.

Section 4 offers an in-depth examination of face recognition technologies, emphasizing deep learning-based approaches. The discussion encompasses various methodologies, including large-scale evaluations of face recognition systems, analysis of CNNs, and assessment of datasets and metrics related to face identification accuracy. It also addresses the influence of external factors, such as demographic biases and image quality, on recognition performance, highlighting the implications of training data inclusivity and the distinctiveness of biometric features across populations [28, 11, 29, 30].

The survey then transitions to Section 5, which discusses image processing and computer vision techniques crucial for enhancing both fatigue detection and face recognition. This section covers feature extraction techniques, image enhancement, preprocessing methods, and data augmentation and generative techniques.

Section 6 examines the application of biometric analysis and facial expression analysis within advanced technologies, highlighting the evolution of face verification methodologies, the integration of explainable AI principles in biometric systems, and the impact of various factors on facial uniqueness. It also addresses the challenges posed by face masks on recognition performance and discusses the benefits of fusing face and periocular biometrics to improve accuracy in recognition systems [11, 31, 32, 33, 34]. This section reviews multimodal and biometric approaches for enhanced detection and recognition.

In Section 7, the survey identifies current challenges and potential future directions in the field, encompassing critical aspects of face recognition and biometric systems. This includes challenges related to real-time processing and operational efficiency, ethical implications surrounding data privacy and societal trust in AI technologies, the need for robust performance against adversarial attacks, and the significance of explainability and interpretability in enhancing detection and recognition systems. It emphasizes understanding biometric uniqueness, the impact of training data inclusion on identification accuracy, and the necessity for effective countermeasures against morphing and adversarial threats, ensuring alignment with ethical standards and societal norms [35, 30, 31, 11, 36].

The concluding Section 8 provides a comprehensive summary of key findings, emphasizing the transformative impact of advancements in deep learning and computer vision across various applications, including driverless technology, medical imaging, and handwritten digit recognition. It underscores the necessity for ongoing research and development to address the complexities and challenges inherent in these rapidly evolving fields, particularly in enhancing model accuracy and automating processes for improved efficiency and reliability in real-world applications [37, 22, 29].The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Definitions and Key Concepts

Fatigue detection involves assessing an individual's alertness, crucial in contexts like driving, where lapses in attention can have severe consequences. This process monitors physical and cognitive indicators, such as eye states and body posture, to evaluate alertness [38]. In driver monitoring systems, fatigue detection is essential for classifying drowsiness states through facial expressions, utilizing computer vision techniques to identify signs of fatigue and distraction [20].

Face recognition is a biometric technique identifying individuals based on facial features, offering a non-intrusive method effective from a distance [8]. Despite challenges like variations in illumination and pose, face recognition remains vital for security and authentication applications [19]. The task is further complicated by the need to match faces across different spectral domains, such as thermal and visible imagery [39], and the introduction of face masks, which necessitates adaptations in recognition algorithms [18].

Facial expression recognition (FER) involves classifying basic expressions—such as happiness, sadness, anger, and surprise—based on facial cues. This classification is influenced by identity-specific attributes, pose, illumination, and expression intensity [21]. FER enhances human-computer interaction systems by predicting emotional states from facial images, thereby impacting usability and performance.

Biometric recognition encompasses classifying individuals based on physiological and behavioral traits, including soft biometrics, which involve non-intrusive features like age and gender [40]. Integrating biometric recognition with well-being monitoring presents challenges due to the variability in biometric data across individuals and conditions [41]. These concepts form the basis for understanding the interconnectedness of fatigue detection, face recognition, and related biometric technologies, emphasizing the need for robust and adaptive systems across various applications.

2.2 Interrelation of Concepts

The interrelation between fatigue detection and face recognition involves converging technologies and methodologies that enhance both processes. Infrared imaging significantly contributes to face recognition by addressing challenges such as illumination changes and facial disguises, thereby improving system robustness in diverse lighting conditions [42]. This capability is particularly relevant to fatigue detection, where accurate facial feature monitoring under varying conditions is critical for assessing drowsiness.

Driver physiological parameters, facial features, and vehicle dynamics are integral to fatigue detection systems [43]. These elements are interconnected, collectively contributing to a comprehensive understanding of a driver's state and facilitating precise fatigue detection, thereby enhancing transportation safety.

Emotion recognition and face recognition are interconnected, with advancements in face recognition networks enhancing emotion recognition accuracy [44]. This relationship is crucial for fatigue detection, as emotional states can significantly influence cognitive and physical performance. The quality of facial images also impacts recognition performance, making high-quality imagery essential for both security applications and fatigue monitoring [26].

Challenges associated with synthetic data and the objectivity of facial Action Units (AUs) complicate accurate facial expression interpretation, which is vital for both fatigue detection and face recognition [45]. This underscores the need for sophisticated methods to effectively leverage these aspects for improved system performance.

The limitations of knowledge-blind machine learning approaches highlight the necessity for deeper reasoning in artificial intelligence, particularly regarding the integration of multiple modalities for fatigue and face recognition [46]. Engagement detection, involving visual cues and physiological signals, exemplifies the interconnectedness of these concepts, crucial for understanding user interaction across various applications [47].

Soft biometrics enhance person recognition systems by providing additional context in unconstrained scenarios, although this area remains underexplored [48]. The variability in individual experiences and environmental factors poses challenges for reliable fatigue detection, necessitating user-friendly monitoring technologies that effectively integrate into real-world applications [5].

Lastly, the challenges posed by GAN-generated deep morph videos emphasize the need for robust benchmarks to enhance detection accuracy in face recognition systems [40]. These interconnected concepts underscore the importance of a holistic approach in developing advanced technologies for fatigue detection and face recognition, ensuring their efficacy and reliability across diverse applications.

3 Deep Learning in Fatigue Detection

The integration of deep learning techniques has revolutionized fatigue detection, particularly in safety-critical fields like transportation and healthcare. This section explores deep learning's pivotal role, especially through image processing, in enhancing the precision and reliability of fatigue detection methods. Figure 2 illustrates the hierarchical structure of deep learning applications in fatigue detection, highlighting the roles of image processing, advanced frameworks, and facial expression analysis. It showcases the integration of convolutional neural networks (CNNs), multimodal approaches, and traditional techniques, emphasizing their contributions to enhancing detection accuracy

AI-generated, for reference only.

and system robustness. The following subsection highlights deep learning's application in recognizing fatigue-related facial features and states.

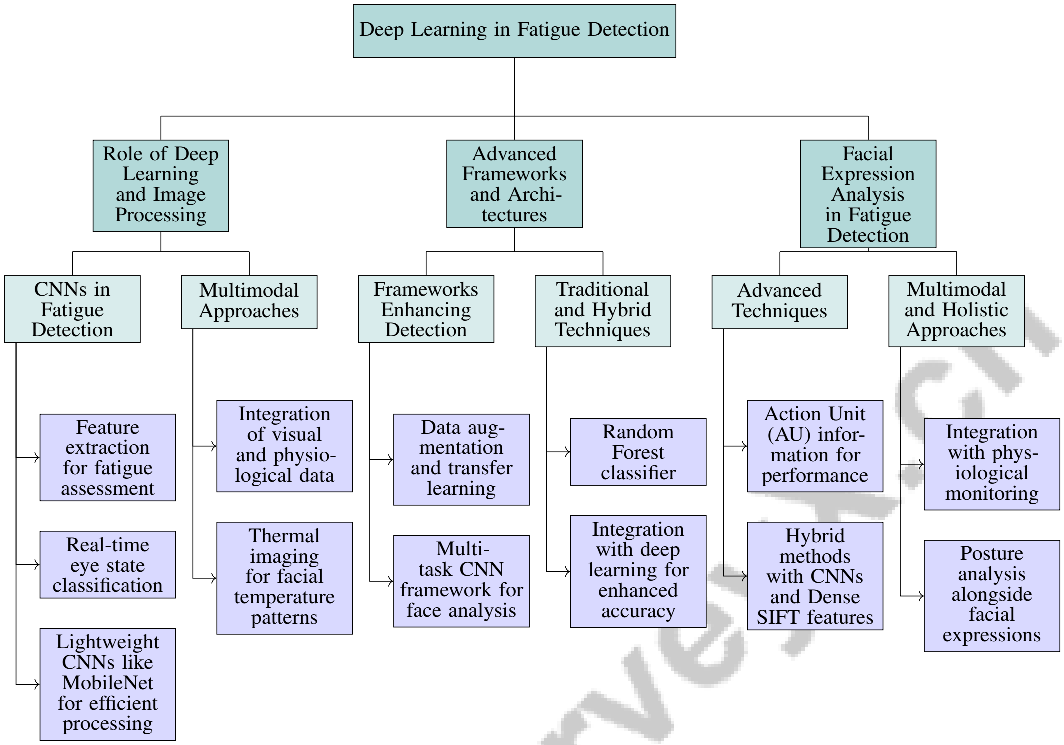


Figure 2: This figure illustrates the hierarchical structure of deep learning applications in fatigue detection, highlighting the roles of image processing, advanced frameworks, and facial expression analysis. It showcases the integration of CNNs, multimodal approaches, and traditional techniques, emphasizing their contributions to enhancing detection accuracy and system robustness.

3.1 Role of Deep Learning and Image Processing

The combination of deep learning and image processing has transformed fatigue detection, enabling advanced real-time analysis of fatigue-indicative facial features. Convolutional Neural Networks (CNNs) are instrumental in extracting features for fatigue assessment, distinguishing between facial states such as eye openness, essential for detecting drowsiness [7]. CNN models optimized for real-time eye state classification exemplify the synergy between deep learning and image processing in enhancing detection accuracy.

Lightweight CNNs like MobileNet offer efficient real-time processing while maintaining accuracy, demonstrating the practicality of such models in real-world scenarios [20]. Advanced frameworks like the T-SCAF model employ shared backbones for feature extraction, promoting multi-task learning that improves fatigue monitoring [8].

Machine learning combined with facial expression analysis effectively detects driver drowsiness, with CNNs processing critical facial cues for real-time physiological monitoring [4]. Multimodal approaches like VisioPhysioENet, integrating visual and physiological data, exemplify the robustness of data fusion in fatigue detection systems [47].

Deep learning also extends to thermal imaging, as seen in the Thermal Facial Fatigue Detection (TFFD) method, which uses facial temperature patterns for classification, highlighting deep learning's adaptability across modalities [38]. A proposed taxonomy categorizes methods into appearance-based, model-based, learning-based, and hand-crafted techniques, offering a structured view of fatigue detection and face recognition [19].

The integration of deep learning and image processing has led to significant advancements in monitoring physiological parameters and facial features, emphasizing sophisticated feature extraction and data fusion in achieving comprehensive fatigue detection. These approaches enhance the efficacy of fatigue monitoring systems, particularly in intelligent vehicle systems where driver fatigue is a major accident contributor [6, 20].

Figure 3 illustrates the role of deep learning in fatigue detection, highlighting key CNN models, multimodal approaches, and advancements in taxonomy. The CNN models section includes eye state detection, MobileNet efficiency, and the T-SCAF framework. Multimodal approaches encompass VisioPhysioENet and thermal imaging techniques. Lastly, the taxonomy and categorization section covers face recognition taxonomy and advancements in fatigue detection.

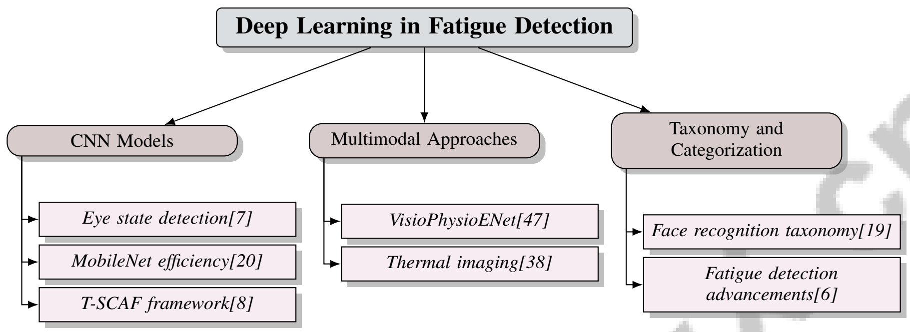


Figure 3: This figure illustrates the role of deep learning in fatigue detection, highlighting key CNN models, multimodal approaches, and taxonomy advancements. The CNN models section includes eye state detection, MobileNet efficiency, and the T-SCAF framework. Multimodal approaches encompass VisioPhysioENet and thermal imaging techniques. Lastly, the taxonomy and categorization section covers face recognition taxonomy and advancements in fatigue detection.

3.2 Advanced Frameworks and Architectures

Advancements in frameworks and architectures have been crucial in enhancing the precision and robustness of fatigue detection systems. Techniques such as data augmentation and transfer learning, particularly with models like YOLOv8, have improved accuracy and real-time processing in image classification and object detection. CNNs achieve near-human accuracy in complex tasks, while methods like Data Fine-tuning optimize performance without altering core parameters [49, 29, 50].

The multi-task CNN framework consolidates face analysis tasks—detection, alignment, pose estimation, and recognition—into a single system, enhancing fatigue detection through comprehensive facial feature analysis [15]. The Intelligent Fatigue Detection and Automatic Vehicle Control System (IFD-AVCS) integrates image processing with physiological monitoring, using heart rate sensors to detect early fatigue signs in train drivers, preventing collisions by controlling the train [20, 51].

The m-CNN method employs dynamic-weighting in its multi-task learning framework, optimizing fatigue detection while managing resources across related tasks [24, 20, 8]. The MS-FRCNN architecture enhances facial region detection through a multi-scale approach, using non-local attention mechanisms to capture local and long-range features, crucial for identifying mild fatigue in challenging environments [23, 52, 38, 6, 24].

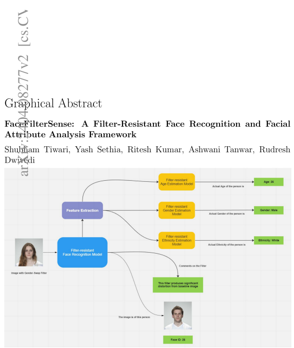
Traditional machine learning techniques, like the Random Forest classifier, achieve high accuracy in detecting driver fatigue, demonstrating the effectiveness of conventional methods in addressing fatigue-related incidents [20, 4, 43]. Integrating traditional methods with deep learning, as in the NLMDA-Net, enhances fatigue detection accuracy through advanced facial alignment networks and MobileNet-based CNNs for recognizing distracted driving behaviors [24, 20].

The DeepID2 feature extraction technique uses deep convolutional networks to enhance face recognition performance by minimizing intra-personal variations and maximizing inter-personal differences, achieving a 99.15

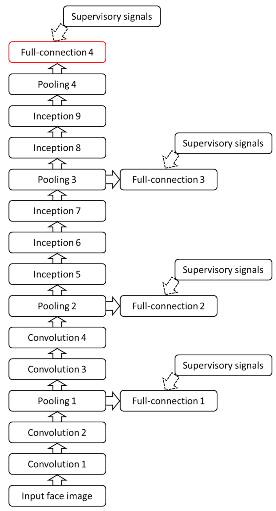
The MultiFace method accelerates face recognition training by using an ensemble of low-dimensional features, improving detection accuracy through enhanced intra-class compactness and inter-class separability [52, 53, 54, 49, 55].

The Single Morphing Attack Detection (S-MAD) method identifies morphed images by analyzing visual artifacts without an authentic reference image, highlighting the importance of image quality in fatigue detection systems [26, 6, 38].

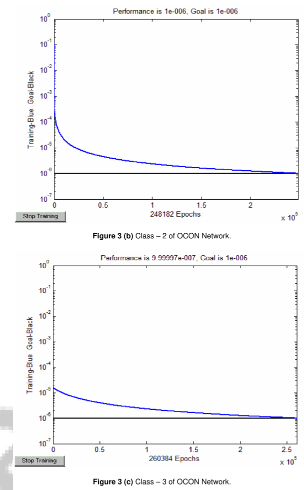
These frameworks and architectures showcase diverse methodologies in fatigue detection, emphasizing their strengths and limitations. Continuous model development enhances fatigue detection systems' efficacy, with deep learning playing a crucial role in hierarchical feature extraction and analysis [22].



(a) Graphical Abstract[56]



(b) The image depicts a flowchart of a neural network architecture, specifically focusing on the use of supervision signals.[57]



(c) Performance is 1e-006, Goal is 1e-006[58]

Figure 4: Examples of Advanced Frameworks and Architectures

As shown in Figure 4, deep learning applications in fatigue detection have facilitated the development of sophisticated frameworks and architectures. The graphical abstract from the "FaceFilterSense" framework illustrates a novel approach to enhancing face recognition and facial attribute analysis through filter-resistant models. Complementing this, the flowchart of a neural network architecture highlights the intricate design of a deep learning model that utilizes supervision signals for performance optimization. Lastly, performance graphs of an OCON network underscore the precision and goal-oriented nature of these advanced architectures, collectively showcasing innovative strategies in deep learning that drive advancements in fatigue detection through the integration of filter-resistant models, supervision signals, and performance optimization.

3.3 Facial Expression Analysis in Fatigue Detection

Facial expression analysis enhances fatigue detection systems by accurately identifying subtle drowsiness indicators, such as eye movements and facial cues [59]. Traditional methods often struggle with these nuances; however, advanced techniques have been developed to overcome these challenges. The AU-CVT framework, for instance, utilizes Action Unit (AU) information to improve facial expression recognition (FER) performance by jointly training on auxiliary datasets with AU or pseudo AU labels, thus enhancing the detection of fatigue-related expressions [45].

Monitoring eye closedness is central to many fatigue detection methods, significantly contributing to accuracy by minimizing false negatives and positives [60]. This focus is complemented by hybrid approaches that combine CNNs with Dense SIFT features, demonstrating marked improvements in facial expression recognition accuracy [61].

The disentanglement of identity, pose, and expression features, as seen in the IPD-FER method, allows for effective separation of these components, leading to improved fatigue detection by focusing on expression changes correlating with drowsiness [62]. The Optical Strain Suppression (OSS) method enhances fatigue detection accuracy by neutralizing facial expressions through the replacement of high strain pixels with those from a neutral expression reference frame, thereby reducing noise in expression analysis [63].

Integrating facial expression analysis with physiological monitoring, such as cardiovascular activity via remote photoplethysmography (rPPG), offers a multimodal approach to fatigue detection, enhancing engagement detection through comprehensive facial feature extraction [47]. Additionally, methods that analyze posture alongside facial expressions provide a more holistic assessment of fatigue, moving beyond reliance on eye states alone [6].

Leveraging independent representations from multiple facial regions facilitates the synthesis of high-quality visible images, aiding in better recognition and verification, further contributing to the robustness of fatigue detection systems [39]. The integration of advanced facial expression analysis techniques with deep learning architectures ensures that fatigue detection systems remain robust and reliable, enhancing safety and performance in environments where fatigue poses a significant risk.

4 Face Recognition Technologies

4.1 Deep Learning-Based Face Recognition

Deep learning has revolutionized face recognition by significantly enhancing recognition accuracy and robustness across diverse conditions through sophisticated neural network architectures. Convolutional Neural Networks (CNNs) are pivotal in this transformation, enabling the automatic extraction of intricate facial features from large datasets. Lightweight models like LightCNN-29 demonstrate high performance with reduced computational demands, maintaining competitive accuracy while offering faster processing compared to complex models such as ResNet-50 [64].

Advancements like the Neighborly De-convolutional Neural Network (NbNet) enhance recognition accuracy by reconstructing face images from deep templates [65]. The Deep 3DMM Regression (D3MMR) method uses deep CNNs to regress 3D morphable model parameters, addressing facial variations robustly [66]. Few-shot learning techniques improve recognition accuracy in scenarios with limited datasets, crucial for practical applications [49]. The MaskTheFace study highlights enhanced recognition for masked individuals, increasing true positive rates significantly in post-pandemic contexts [18].

Innovative approaches like Spatial Pyramid Pooling for Face Recognition (SPP-FR) enable robust recognition across varied conditions by pooling local patches from facial images [67]. CorrRISE generates saliency maps to enhance interpretability in recognition systems' decision-making processes [68]. Attention-based mechanisms, such as the attention-aware wavelet-based detection method, improve system accuracy by focusing on critical features [69].

These methodologies underscore deep learning's transformative impact on face recognition, ensuring efficacy and reliability across diverse applications. The shift from traditional methods to deep learning approaches leveraging large-scale datasets has reshaped the face recognition landscape, improving accuracy and efficiency, particularly in managing extensive datasets and coping with unconstrained conditions [70, 71].

4.2 Datasets and Evaluation Metrics

The evaluation of face recognition systems relies on diverse datasets and robust metrics to ensure comprehensive assessments across real-world scenarios. The Labeled Faces in the Wild (LFW) dataset, with 13,233 images of 5,749 subjects, serves as a benchmark for face verification accuracy [72]. The VGGFace2 dataset, comprising over 3.31 million images, is critical for training and evaluating CNN architectures like MobileFaceNet and ResNet50 [49].

Specialized datasets like BWCFace, with 178,000 facial images, provide insights into recognition systems in dynamic environments [12]. The 300W-LP and Lock3DFace databases aid in evaluating performance under variations in facial structure and pose [76]. Datasets like the IIITD Plastic Surgery Face and Disguised Faces in the Wild highlight the importance of handling facial alterations due to surgery and disguises [77]. Benchmarks like LFW, YTF, and IJB-A offer comprehensive evaluations across diverse conditions [66].

Benchmark	Size	Domain	Task Format	Metric
LFW[72]	13,233	Face Recognition	Face Verification	Verification Accuracy
FRL-Bench[49]	1,793,119	Facial Analysis	Face Recognition	NME, ICC
BWC[12]	178,000	Face Recognition	Face Identification	Rank-1 Accuracy, Inference Time
FFS[56]	2,040	Face Recognition	Facial Attribute Analysis	Accuracy, F1 Score
UMDFaces[73]	3,735,476	Face Recognition	Face Verification	ROC Curve
DLFD[24]	20,000	Facial Fatigue Detection	Video-based Fatigue Detection	Accuracy, F1-score
ID-Overlap[74]	150,000	Face Recognition	Image Pair Classification	Accuracy, F1-score
MAD[75]	2,500	Facial Recognition	Morphing Attack Detection	D-EER, BPCER-10

Table 1: Table ef presents a comprehensive overview of representative benchmarks in face recognition and analysis. It includes datasets spanning various domains such as face verification, identification, and morphing attack detection, detailing their sizes, task formats, and evaluation metrics. The table serves as a critical resource for understanding the diversity and application of datasets in face recognition research.

Common evaluation metrics include accuracy, precision, recall, and F1-score, providing clear assessments of model performance. The Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC) offer nuanced insights into model performance across different thresholds [39]. Experiments with benchmark datasets like VGG-Face and Multi-PIE underscore their significance in evaluating models such as NbNet [65].

As illustrated in Figure 5, the strategic use of diverse datasets and evaluation metrics is vital for advancing face recognition research. This figure highlights the hierarchical structure of datasets and evaluation metrics used in face recognition systems, showcasing key datasets, common evaluation metrics, and advanced techniques such as multi-modal fusion and infrared imaging. Integrating physiological features with advanced multi-modal fusion techniques enhances recognition accuracy, as demonstrated by methods leveraging visual and physiological signals for tasks like face recognition, gender, race, and age detection. Infrared imaging has emerged as a promising avenue to mitigate challenges like illumination and pose variations, highlighting the potential for ongoing advancements in face recognition technologies [42, 52, 78, 47, 79]. Table 1 provides a detailed overview of the representative benchmarks utilized in face recognition systems, highlighting the diversity of datasets and the evaluation metrics employed to assess model performance.

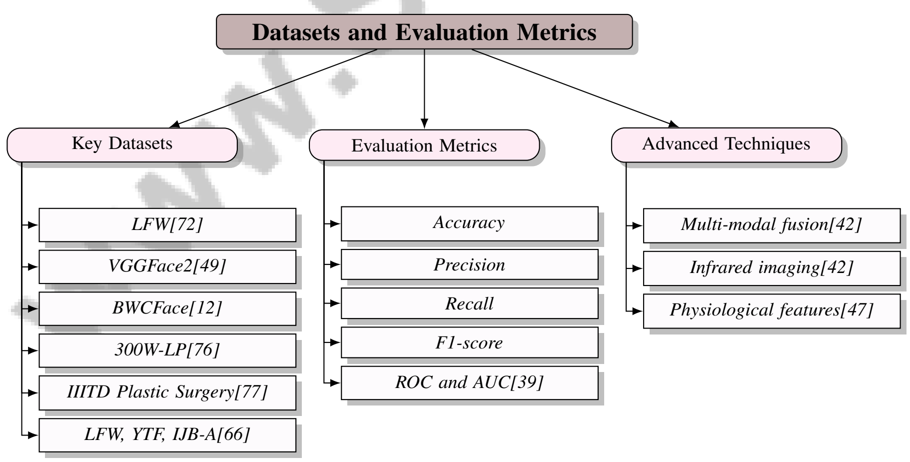


Figure 5: This figure illustrates the hierarchical structure of datasets and evaluation metrics used in face recognition systems, highlighting key datasets, common evaluation metrics, and advanced techniques such as multi-modal fusion and infrared imaging.

4.3 Impact of External Factors on Recognition Accuracy

Face recognition systems are significantly affected by external factors such as lighting, pose, and occlusion, which present challenges that can undermine accuracy. Variations in lighting can alter facial features, complicating recognition algorithms' ability to maintain accuracy. Techniques like spatial pyramid pooling have shown robustness against image variations, suitable for real-world applications with unpredictable lighting [67].

Pose variations introduce complexity, as deviations from frontal views can lead to discrepancies between captured images and stored templates, affecting recognition performance [64]. Ensuring robustness to such variations is crucial for reliable performance in unconstrained environments.

Occlusion, from accessories like glasses or masks, complicates recognition. Masks, in particular, impair detection and recognition accuracy, exacerbated by their widespread use [64]. Addressing these challenges requires innovative approaches to enhance system robustness across diverse conditions.

Image quality also impacts recognition accuracy. Lower-quality images degrade performance, though robustness varies across systems [72]. Metrics like Normalized Mean Error (NME) and Intraclass Correlation Coefficient (ICC) assess accuracy and correlation of facial representations, emphasizing the need for high-quality data [49].

The figure Figure 6 illustrates the impact of external factors on face recognition accuracy, categorizing the challenges into lighting variations, pose and occlusion, and image quality, each affecting the system's performance in unique ways. External factors' influence on recognition accuracy highlights the need for ongoing methodological improvements, particularly given findings on demographic biases and the influence of individual inclusion in training datasets. Studies show higher accuracy for individuals present in training data, with biases favoring certain racial groups [28, 11, 30]. Developing robust algorithms and frameworks to handle variations in lighting, pose, and occlusion is crucial for ensuring reliable and effective face recognition technologies in real-world applications.

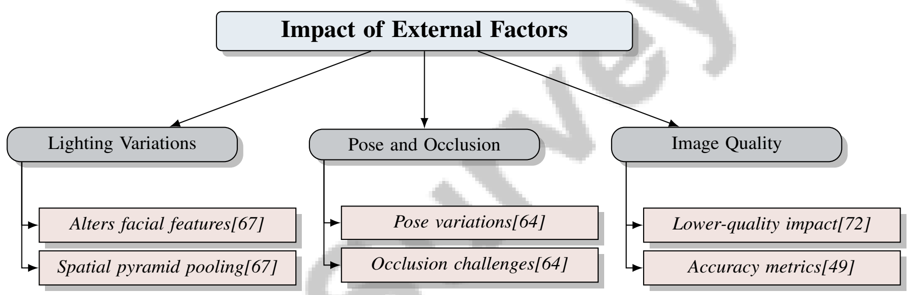


Figure 6: This figure illustrates the impact of external factors on face recognition accuracy, categorizing the challenges into lighting variations, pose and occlusion, and image quality, each affecting the system's performance in unique ways.

5 Image Processing and Computer Vision Techniques

Category	Feature	Method
Feature Extraction Techniques	Augmentation and Variability	CF[53], DALIFR[80]
	Data Fusion and Integration	VPEN[47], WD-LBP[34]
	Algorithm Selection and Adaptation	HCA[81], DNDIM[76], MTF[18]
	Optical and Strain Techniques	OSS[63]
Image Enhancement and Preprocessing	Correlation Analysis	FPM[82]
Data Augmentation and Generative Techniques	Generative Techniques	SAT[83], SFRB[84], FRA[85], FF[54]

Table 2: This table provides a comprehensive summary of various techniques and methods employed in image processing, specifically focusing on feature extraction, image enhancement, and data augmentation. It categorizes these techniques into feature extraction, image enhancement and preprocessing, and generative techniques, highlighting the specific methods and references associated with each category. The table serves as a reference for understanding the diversity and application of these techniques in enhancing fatigue detection and face recognition systems.

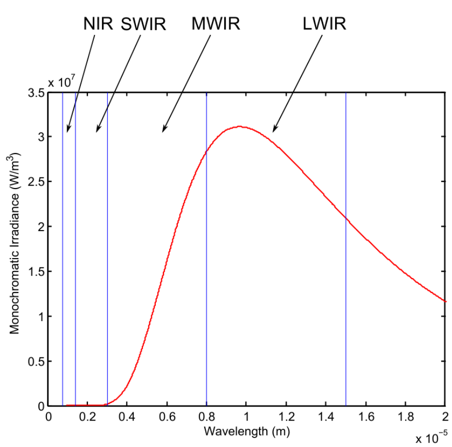
The optimization of fatigue detection and face recognition systems is heavily dependent on sophisticated feature extraction techniques, which enhance accuracy and reliability by distilling meaningful information from visual data. Table 2 presents a detailed classification of methods used in image processing and computer vision, emphasizing their roles in feature extraction, image enhancement, and data augmentation for improving these systems. Additionally, Table 4 presents a comparative analysis of various image processing methods, detailing their core techniques, purposes, and unique features pertinent to enhancing fatigue detection and face recognition systems.

5.1 Feature Extraction Techniques

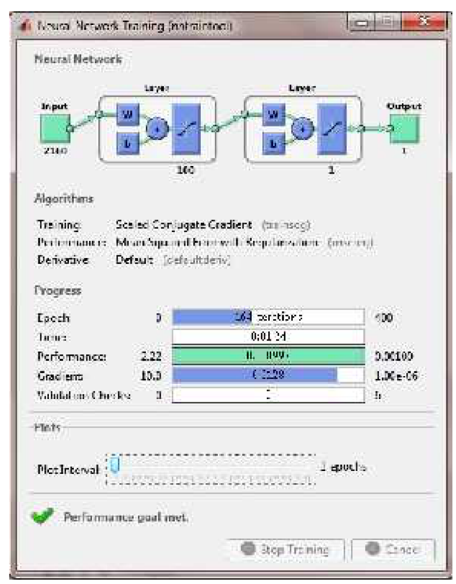
Feature extraction is pivotal for enhancing the accuracy and efficiency of fatigue detection and face recognition systems. The Hybrid Convolution Algorithm (HCA) exemplifies advanced methodologies by dynamically selecting between Fast Fourier Transform (FFT) and Winograd algorithms based on convolution parameters, thereby minimizing latency and enhancing processing efficiency [81]. Facial landmark detection and tracking are crucial for methods like Optical Strain Suppression (OSS), which generates strain maps for each video frame, improving feature extraction for recognition [63]. Accurate extraction of facial landmarks and eye metrics is essential for improved engagement detection [47].

Data augmentation significantly enhances face recognition systems' robustness. The ComplexFace approach utilizes innovative data augmentation and fusion techniques to achieve superior accuracy from limited data [53]. Techniques like synthetic augmentation improve system resilience against pose and lighting variations [80]. The Wavelet Decomposed Local Binary Pattern (WD-LBP) method combines wavelet decomposition with local binary patterns for robust feature extraction from face and periocular images [34].

For fatigue detection, extracting facial landmark coordinates is vital, especially in systems analyzing real-time eye states using CNNs. Normalizing facial depth images to a frontal pose and neutral expression, followed by robust feature extraction via deep CNNs, underscores the importance of preprocessing for accurate recognition [76]. Techniques adapting to challenges like mask usage, through landmark detection algorithms, further improve recognition accuracy [18].



(a) The image shows a graph of monochromatic irradiance (W/m^2)$asafunctionofwavelength(m)fordifferentspectralregions.$[42]



(b) Neural Network Training Progress[86]

Figure 7: Examples of Feature Extraction Techniques

As shown in Figure 7, feature extraction techniques are crucial in analyzing and interpreting visual data. The first image illustrates how spectral analysis aids in identifying features within different spectral regions, essential for applications like infrared face recognition. The second image highlights machine learning's role in feature extraction, reflecting diverse methodologies from spectral analysis to advanced neural network techniques [42, 86].

5.2 Image Enhancement and Preprocessing

Image enhancement and preprocessing are critical for preparing facial data, directly influencing fatigue detection and face recognition systems' effectiveness. Techniques such as facial landmarking and pose correction enhance recognition accuracy but must be evaluated alongside input data quality, affected by demographic biases and variations [26, 28, 87, 88].

The Face Prediction Model (FPM) captures temporal and spatial correlations in facial images, refining inputs for advanced neural network architectures and improving recognition accuracy [82]. Preprocessing typically involves normalizing and aligning facial images to reduce discrepancies from pose, lighting, and expression variations, vital for training deep learning models like CNNs [49, 88]. Techniques such as histogram equalization and contrast adjustment enhance image clarity.

Noise reduction is crucial, using methods like Gaussian and median filtering to eliminate artifacts obscuring facial details. These methods enhance face image quality, particularly for face image quality assessment (FIQA) methods aimed at improving recognition performance [28, 88]. Data augmentation during preprocessing significantly enhances recognition systems' robustness, addressing biased datasets by introducing variability in pose and lighting conditions through semi-synthetic image generation [80, 50, 49, 85, 29].

Image enhancement and preprocessing are foundational to fatigue detection and face recognition systems' success, ensuring optimal data preparation. These techniques address challenges associated with fluctuating image quality and diverse conditions, enhancing biometric technologies' reliability and effectiveness. Emphasizing biometric quality metrics, such as those in the FaceQnet framework, ensures accurate assessment and utilization of facial images for recognition purposes [89, 31].

5.3 Data Augmentation and Generative Techniques

Method Name	Techniques Utilized	Challenges Addressed	Future Directions
FRA[85]	Generative Techniques	Limited Data Availability	Advanced Augmentation Techniques
SAT[83]	Policy Gradient Method	Limited Data Availability	Improving Simulator Realism
FF[54]	Class Proxy Merging	Conflicting Identities	Refining Merging Process
SFRB[84]	Synthetic Dataset Generation	Algorithmic Bias Measurement	Annotation Process Enhancements

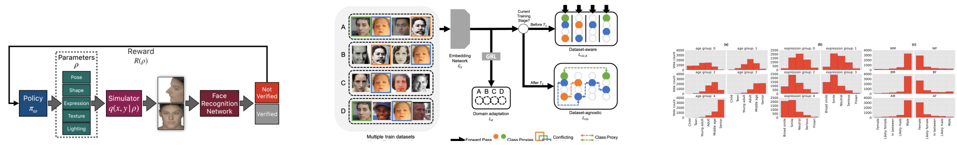
Table 3: Overview of methodologies in data augmentation and generative techniques for face recognition. The table summarizes various methods, highlighting the techniques utilized, challenges addressed, and future directions for each approach. These insights contribute to advancing the robustness and effectiveness of facial analysis systems.

Data augmentation and generative techniques are vital for enhancing fatigue detection and face recognition systems' robustness and generalization capabilities. These methods address limited data availability through Data Fine-tuning, enhancing training datasets with controlled perturbations, increasing input diversity and variability. Large-scale unsupervised pre-training on diverse, uncurated datasets consistently yields accuracy improvements across multiple facial analysis tasks [49, 50].

A systematic analysis of augmentation techniques, including basic manipulations and advanced generative methods, has proven effective in enhancing face recognition from limited data [85]. Generative techniques, particularly using Generative Adversarial Networks (GANs), create realistic synthetic data that closely mimics real-world images, increasing training data volume and introducing crucial variations for improving model robustness. However, the vulnerability of systems to GAN-generated deep morph videos highlights the need for robust benchmarks and advanced detection methods [40].

Future research should focus on advanced detection techniques incorporating interpretable deep learning models and robust methodologies, such as wavelet-based approaches, to enhance morphing attack identification. Expanding datasets to include diverse morphing techniques, particularly those exploiting facial landmarks, will improve detection systems' effectiveness [69, 90, 36]. Integrating data augmentation and generative techniques into fatigue detection and face recognition systems is essential for improving reliability and effectiveness, addressing challenges such as limited datasets and variations in pose and lighting conditions. Combining generative methods with basic augmentation techniques significantly boosts performance, while advanced machine learning approaches effectively detect fatigue by analyzing facial features and dynamics [20, 38, 80, 24, 85].

AI-generated, for reference only.



(a) DeepFace: A Deep Learning Approach for Face Verification[83]

(b) Domain Adaptation for Multi-Label Classification with Conflicting Identity Labels[54]

(c) Age and Expression Group Distribution of Votes in a Dataset[84]

Figure 8: Examples of Data Augmentation and Generative Techniques

As shown in Figure 8, the exploration of image processing and computer vision techniques, particularly in data augmentation and generative techniques, is exemplified through three distinct studies. The first example presents a flowchart elucidating the deep face verification process via a deep learning framework. The second example illustrates a framework managing datasets with conflicting identity labels through domain adaptation. Lastly, the third example is depicted through bar charts visualizing vote distributions across various age and expression groups, highlighting innovative applications of data augmentation and generative techniques in enhancing computer vision systems' efficacy and adaptability [83, 54, 84]. Additionally, Table 3 provides a comprehensive summary of recent methodologies in data augmentation and generative techniques, illustrating their applications, addressed challenges, and prospective advancements in the context of face recognition.

Feature	**Hybrid Convolution Algorithm (HCA)**	**Optical Strain Suppression (OSS)**	**ComplexFace**
Core Technique	Fft And Winograd	Strain Maps	Data Augmentation
Purpose	Minimize Latency	Improve Recognition	Achieve Superior Accuracy
Unique Feature	Dynamic Selection	Facial Landmark Tracking	Limited Data Fusion

Table 4: This table provides a comparative analysis of three advanced image processing techniques: Hybrid Convolution Algorithm (HCA), Optical Strain Suppression (OSS), and ComplexFace. It highlights their core techniques, primary purposes, and unique features, emphasizing their roles in minimizing latency, improving recognition, and achieving superior accuracy in fatigue detection and face recognition systems.

6 Biometric Analysis and Facial Expression Analysis

Facial expression analysis is integral to biometric systems, significantly enhancing detection efficacy. This section delves into the integration of biometric features, illustrating how combining various cues enhances detection accuracy and robustness. By exploring the synergy between distinct biometric indicators like facial expressions and physiological signals, we gain insights into advancements in detection methodologies. Key methodologies and advancements highlighting the importance of these integrations for comprehensive user state assessments are discussed.

6.1 Integration of Biometric Features for Enhanced Detection

Integrating biometric features significantly enhances the accuracy and robustness of fatigue detection and face recognition systems. By incorporating diverse biometric cues—such as facial expressions, eye movements, and physiological signals—these systems achieve comprehensive user state assessments. For example, thermal signatures associated with fatigue can enhance detection accuracy without requiring wearable sensors [38], demonstrating the effectiveness of non-invasive monitoring methods.

Advanced methodologies, such as fusing thermal and visual images, extract synergistic information that enhances feature extraction and recognition under varying conditions [79]. This multimodal approach enriches facial feature representation, improving classification accuracy and robustness against environmental variations. The effectiveness of this integration is evident in face anti-spoofing models, which emphasize pixel-wise supervision and generalization techniques to bolster detection accuracy [69].

Tailoring thresholds to individual facial features is crucial for enhancing the robustness of face recognition systems [91]. This adaptability ensures high accuracy in dynamic environments. Methods that control identity-specific variations, such as disentangling identity-specific features, further enhance system performance [21].

In 3D face shape estimation, proposed methods demonstrate superior accuracy compared to existing techniques, enhancing detection system performance [66]. Integrating advanced biometric features ensures systems are accurate and resilient to challenges posed by diverse and dynamic environments.

Strategic integration of biometric features is crucial for developing reliable and secure detection systems across applications like border control and banking security, addressing societal concerns regarding the trustworthiness of artificial intelligence. This integration must consider factors influencing biometric modality uniqueness, such as image resolution, feature representation, and demographic variables, as highlighted by studies on face recognition algorithms. By employing explainable AI principles, developers can create transparent systems that foster public confidence in biometric technologies [11, 31]. These methodologies enable the creation of systems that effectively adapt to real-world scenarios, providing robust biometric analysis.

6.2 Multimodal and Biometric Approaches

Integrating multimodal and biometric approaches in face recognition and fatigue detection systems is crucial for enhancing recognition accuracy and robustness. These approaches leverage multiple data sources—such as visual, thermal, and physiological signals—to provide comprehensive analyses of biometric features. Fusing these modalities allows systems to effectively distinguish between relevant and irrelevant features, improving the efficiency of machine learning tasks like face recognition [92].

A significant advantage of multimodal approaches is their adaptability to diverse environmental conditions, essential for maintaining high recognition accuracy. Controlled image transformations, as demonstrated in visual psychophysics, generate item-response curves that reveal algorithm performance across different perturbation levels [93], emphasizing the need for robust recognition algorithms.

The effectiveness of Single Morphing Attack Detection (S-MAD) in multimodal contexts is noteworthy. By leveraging high-level features learned from images, S-MAD effectively detects morphing artifacts and contextual information, enhancing the detection of adversarial manipulations [94]. This capability is crucial for ensuring the integrity and security of face recognition systems amidst evolving threats.

Future research should refine detection algorithms and explore additional countermeasures to enhance the robustness of deep learning models against a broader range of adversarial attacks [95]. Introducing a taxonomy of adversarial attacks and defenses provides a framework for categorizing existing research, emphasizing the need for comprehensive strategies to address these challenges [35].

Moreover, evaluating biometric systems, including face recognition, through case studies highlights challenges in providing understandable and justifiable explanations for system decisions [31]. This underscores the necessity for explainable AI to ensure that multimodal and biometric approaches are effective, transparent, and accountable.

The integration of multimodal and biometric approaches enhances recognition systems by leveraging diverse data types, such as facial features, soft biometrics (gender, age, ethnicity), and multi-view images. This fusion improves classification accuracy—evidenced by increases of up to 22.2

7 Challenges and Future Directions

7.1 Real-time Processing and Efficiency

Real-time processing in detection and recognition systems is challenged by high computational demands and the need for extensive annotated datasets. The scarcity of labeled data for supervised learning limits the effectiveness of face recognition methods, as generating large datasets requires specialized expertise [22]. Innovative approaches, such as unsupervised pre-training on uncurated datasets, have improved accuracy in facial tasks, addressing some real-time processing challenges [49]. Efficient resource utilization is crucial, particularly in dynamic environments, as demonstrated by the Multi-Region Thermal-to-Visible Synthesis approach, which optimizes real-time processing without custom thermal detectors [39]. However, reliance on accurate real-time data can be limiting, especially in complex urban contexts [79].

Real-time fatigue detection using EEG-based sensors shows promise but remains challenging on mobile devices [5]. Adversarial training has improved scalability and flexibility, yet limitations persist with extreme facial expressions and occlusions not well represented in training data [21]. Image quality significantly impacts face recognition performance, necessitating further research to enhance robustness against variations [72]. Real-time detection methods must also distinguish between morphed and genuine images, as existing approaches often struggle with the subtlety of morphing, leading to potential false acceptances in security systems [69].

Addressing real-time processing challenges requires a multifaceted approach combining innovative methodologies with empirical evaluations. Overcoming these obstacles can enhance the accuracy and reliability of detection and recognition systems across various applications [8].

7.2 Ethical Considerations and Data Privacy

The deployment of face recognition and fatigue detection technologies raises significant ethical and data privacy concerns due to their pervasive use and potential impact on individual rights. The integration of monitoring cameras for driver behavior poses privacy challenges, as surveillance may occur without explicit consent, raising ethical dilemmas regarding the balance between safety and privacy [20]. Reliance on social media data for training face recognition models presents ethical issues, as such data may not be representative, potentially leading to biases and privacy violations [13].

Bias in face recognition models, particularly against certain demographics, remains a critical concern, resulting in misclassification and unequal performance across demographic groups. Continuous monitoring and algorithm adjustments are necessary to address these disparities. Furthermore, the absence of liveness detection in some systems poses risks of spoofing attacks, complicating the ethical landscape by undermining the security of biometric systems [96]. Challenges related to occlusion, such as glasses or masks, underscore additional ethical considerations regarding user privacy and data collection, affecting the accuracy and reliability of detection systems [49]. Security and privacy vulnerabilities in systems relying on deep templates raise ethical implications concerning unauthorized access and data misuse [65]. Robust security measures and regulatory frameworks are essential to protect personal data and ensure responsible deployment.

The lack of interpretability in deep learning models, often operating as 'black boxes', poses challenges for user trust and understanding of decision-making processes [68]. This lack of transparency can hinder accountability and exacerbate ethical concerns, particularly in critical applications. Addressing ethical considerations and data privacy requires a comprehensive approach prioritizing transparency, accountability, and fairness. By ensuring datasets are ethically sourced, inclusive, and adhering to explainable AI principles, stakeholders can safeguard individual privacy rights and foster public trust [30, 46, 31, 18, 11].

7.3 Performance and Robustness Challenges

Face recognition and fatigue detection systems face challenges from diverse environmental conditions and intrinsic dataset variability, significantly impacting performance and robustness. High-quality data acquisition is crucial, yet systems often struggle in low-light conditions or when facial features are obscured, necessitating solutions that maintain reliability in such contexts [3]. Performance variability across algorithms and demographic factors, such as race, complicates assessments of face recognition systems and introduces potential biases that must be addressed for consistent performance [97].

Adaptive thresholding techniques have shown promise in enhancing accuracy by addressing threshold selection issues, providing robust solutions for dynamic environments [91]. However, reliance on the quality of base classifiers remains a challenge, as poor performance can undermine overall system effectiveness [98]. Future research in fatigue detection could explore improving algorithm robustness across diverse environments and integrating additional physiological signals to enhance accuracy [6]. Visual changes, such as facial tattoos and paintings, negatively affect performance, highlighting the need for robust algorithms capable of managing these variations [99]. Additionally, performance can degrade with expressions from extremely low-resolution images, underscoring the necessity for methods to handle such variations [12]. Metrics for assessing model effectiveness in identifying faces and real-time performance on mobile devices are crucial for ensuring reliability in practical applications [12].

Current studies often rely on simulations rather than real-life scenarios, which may not fully capture the complexities of real-world fatigue experiences [5]. Future research should focus on retraining models like FaceNet with diverse datasets to reduce bias, implementing liveness detection mechanisms, and automating user classification for digital onboarding [96]. The NbNet method demonstrates higher accuracy and robustness in reconstructing face images, addressing performance challenges under diverse conditions [65]. However, challenges with diverse mask types and real-world conditions indicate ongoing performance and robustness challenges [18].

The robustness and discriminative power of approaches estimating 3D face shapes from single images highlight potential advantages over traditional methods [66]. Yet, current research often lacks comprehensive solutions for addressing specific challenges, such as occlusion and environmental variations in traditional methods [64]. Emphasizing model architecture is vital for robustness against image quality variations [72]. Future research should develop solutions exploiting new imaging technologies and enhance algorithm adaptability for real-time applications in embedded systems [19].

To effectively tackle performance and robustness challenges, a comprehensive strategy is essential, incorporating innovative methodologies such as dataset fusion and data fine-tuning while emphasizing empirical evaluation techniques to assess model behavior under real-world conditions [87, 11, 50, 54]. Overcoming these obstacles can lead to high accuracy and reliability across diverse applications and environmental conditions.

7.4 Adversarial Attacks and Security Vulnerabilities

Adversarial attacks pose significant challenges to the security and reliability of face recognition systems, exploiting vulnerabilities to induce misclassification and compromise system integrity. These attacks involve subtle modifications to input data, resulting in dramatic reductions in recognition accuracy [95]. Developing robust defenses against such attacks remains a pressing issue, necessitating effective detection methods that maintain performance while ensuring security [35].

Morphing attacks, which manipulate images to blend features from multiple identities, represent a severe threat to biometric verification processes, potentially deceiving systems into granting unauthorized access [94]. Detecting these morphing attacks is critical for safeguarding biometric systems, as they exploit weaknesses in current recognition models.

The vulnerability of face recognition systems to deep-morphed videos generated using advanced Generative Adversarial Networks (GANs) underscores the need for new detection methodologies capable of addressing these technological advancements [40]. Existing systems often lack the robustness needed to accurately detect and mitigate adversarial manipulations, emphasizing the necessity for ongoing research and development [83].

Moreover, the absence of standardized datasets for training and evaluating drowsiness detection systems contributes to performance vulnerabilities, as models may not generalize well across diverse conditions [4]. This lack of standardization can lead to inconsistencies in performance and exacerbate risks associated with adversarial attacks.

The robustness of face recognition systems against circumvention techniques, such as occlusion, remains an area with unanswered questions, highlighting the need for improved algorithms to handle such challenges [64]. Future research should enhance the robustness of face recognition algorithms against environmental variations and integrate additional modalities, such as gesture recognition, to provide comprehensive and secure human-machine interactions [78]. Addressing these challenges is critical for developing resilient detection systems capable of withstanding evolving adversarial threats and ensuring the security of biometric technologies.

7.5 Explainability and Interpretability

The pursuit of explainability and interpretability in detection and recognition systems is essential for their ethical and effective deployment. Explainability is crucial in AI-based face recognition systems, as it elucidates decision-making processes while fostering user trust and enhancing transparency. As face recognition technology integrates into various applications, understanding how these systems operate becomes increasingly important, especially given societal concerns surrounding privacy and ethical use. Recent advancements, such as model-agnostic explanation methods like CorrRISE, highlight the importance of visual saliency maps in providing insights into factors influencing recognition outcomes. By addressing explainability challenges, these frameworks can help ensure AI systems align with societal norms, contributing to responsible deployment in real-world scenarios [30, 31, 68, 26, 100]. Current methods often focus on face verification, but future research should incorporate saliency information into representations to enhance performance through localization of important facial features.

Developing intelligent face recognition systems capable of adapting to evolving threats and improving reliability is another critical focus area. Exploring 3D face recognition systems presents a promising avenue for enhancing robustness and interpretability [64]. Furthermore, integrating unsupervised and semi-supervised learning techniques could reduce reliance on annotated data, improving adaptability and explainability [22].

In fatigue detection, future research should prioritize real-life driving scenarios and ergonomic device design for user-friendliness. Exploring combined measures for fatigue detection can also contribute to more comprehensive and interpretable systems, enhancing their applicability in practical settings [5].

Emphasizing explainability and interpretability is crucial for developing powerful technologies that align with societal norms and foster trust in artificial intelligence (AI). As these systems become integrated into operational environments, addressing the challenges of creating algorithms that provide clear explanations of their decision-making processes is essential. This focus on explainable AI is particularly important in biometric technologies, such as face recognition, where understanding the uniqueness of faces and the factors influencing biometric identification—such as image resolution, feature representation, and demographic variables—can significantly impact effectiveness and reliability [11, 31]. Prioritizing these attributes can enhance the ethical deployment of these systems, ultimately improving their acceptance and utility across various applications.

8 Conclusion

This survey highlights the profound impact of deep learning and computer vision on the evolution of fatigue detection and face recognition technologies. The deployment of advanced neural network frameworks, such as Convolutional Neural Networks (CNNs) and Deep Neural Networks (DNNs), has substantially improved the precision and functionality of these systems. Notably, the use of lightweight CNN models for age and gender prediction from ocular images demonstrates their efficacy in mobile applications, enhancing real-time user experiences.

The importance of data augmentation techniques is also underscored, as they significantly contribute to the robustness and adaptability of face recognition systems. By employing diverse augmentation strategies, these systems exhibit improved performance, addressing the challenges of data limitations and model robustness. Additionally, the integration of behavioral analysis with machine learning in drowsiness detection highlights the imperative for expansive, standardized datasets to propel further advancements in this area.

The introduction of the TetraLoss method marks a significant stride in strengthening face recognition systems against morphing threats, enhancing security measures beyond existing capabilities. This advancement is crucial for the practical deployment of secure biometric systems.

Addressing algorithmic biases in face recognition is critical, with the survey stressing the need for individualized bias assessments and a nuanced understanding of item difficulty to ensure fair performance across various demographic groups. Future research directions include refining disentanglement techniques and exploring broader applications beyond facial expression recognition, thereby expanding the scope and applicability of these technologies.

References

[1] Bhargava Reddy, Ye-Hoon Kim, Sojung Yun, Chanwon Seo, and Junik Jang. Real-time driver drowsiness detection for embedded system using model compression of deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 121–128, 2017.

[2] Jianchao Lu, Xi Zheng, Tianyi Zhang, Michael Sheng, Chen Wang, Jiong Jin, Shui Yu, and Wanlei Zhou. Can steering wheel detect your driving fatigue?, 2020.

[3] Rateb Jabbar, Mohammed Shinoy, Mohamed Kharbeche, Khalifa Al-Khalifa, Moez Krichen, and Kamel Barkaoui. Driver drowsiness detection model using convolutional neural networks techniques for android application. In *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, pages 237–242. IEEE, 2020.

[4] Mkhuseli Ngxande, Jules-Raymond Tapamo, and Michael Burke. Driver drowsiness detection using behavioral measures and machine learning techniques: A review of state-of-art techniques. *2017 pattern recognition Association of South Africa and Robotics and mechatronics (PRASA-RobMech)*, pages 156–161, 2017.

[5] Prabin Sharma, Joanna C. Justus, Megha Thapa, and Govinda R. Poudel. Sensors and systems for monitoring mental fatigue: A systematic review, 2023.

[6] Ashish Verma, Ankush Goyal, and Davinderjit Kaur. Fatigue detection, 2019.

[7] Israt Jahan, KM Aslam Uddin, Saydul Akbar Murad, M Saef Ullah Miah, Tanvir Zaman Khan, Mehedi Masud, Sultan Aljahdali, and Anupam Kumar Bairagi. 4d: a real-time driver drowsiness detector using deep learning. *Electronics*, 12(1):235, 2023.

[8] Shulei Qu, Zhenguo Gao, Xiaowei Chen, Na Li, Yakai Wang, and Xiaoxiao Wu. Multi-task learning for fatigue detection and face recognition of drivers via tree-style space-channel attention fusion network, 2024.

[9] Mohamed Fawzi Abdelshafie Abuhussein, Ashraf Darwish, and Aboul Ella Hassanien. Exploring thermography technology: A comprehensive facial dataset for face detection, recognition, and emotion, 2024.

[10] Yang Li and Sangwhan Cha. Implementation of robust face recognition system using live video feed based on cnn, 2018.

[11] Michal Balazia, S L Happy, Francois Bremond, and Antitza Dantcheva. How unique is a face: An investigative study, 2022.

[12] Ali Almadan and Ajita Rattani. Towards on-device face recognition in body-worn cameras, 2021.

[13] Trung Nguyen, Barth Lakshmanan, and Weihua Sheng. A smart security system with face recognition, 2018.

[14] Changxing Ding and Dacheng Tao. Robust face recognition via multimodal deep face representation, 2015.

[15] Rajeev Ranjan, Swami Sankaranarayanan, Carlos D. Castillo, and Rama Chellappa. An all-in-one convolutional neural network for face analysis, 2016.

[16] Sukhada Chokkadi, Sannidhan M S, Sudeepa K B, and Abhir Bhandary. A study on various state of the art of the art face recognition system using deep learning techniques, 2019.

[17] Pangoth Santhosh Kumar and Garika Akshay. Deep learning-based spatio temporal facial feature visual speech recognition, 2023.

[18] Aqeel Anwar and Arijit Raychowdhury. Masked face recognition for secure authentication, 2020.

[19] Alireza Sepas-Moghaddam, Fernando Pereira, and Paulo Lobato Correia. Face recognition: A novel multi-level taxonomy based survey, 2019.

[20] Dongjiang Wu. Improving automatic detection of driver fatigue and distraction using machine learning, 2024.

[21] Xiaofeng Liu. Disentanglement for discriminative visual recognition, 2020.

[22] Muhammad Imran Razzak, Saeeda Naz, and Ahmad Zaib. Deep learning for medical image processing: Overview, challenges and the future. *Classification in BioApps: Automation of decision making*, pages 323–350, 2017.

[23] Janelle Domantay. How facial features convey attention in stationary environments, 2021.

[24] Zeyu Chen, Xinhang Zhang, Juan Li, Jingxuan Ni, Gang Chen, Shaohua Wang, Fangfang Fan, Changfeng Charles Wang, and Xiaotao Li. Machine vision detection to daily facial fatigue with a nonlocal 3d attention network, 2021.

[25] Tiago de Freitas Pereira, Dominic Schmidli, Yu Linghu, Xinyi Zhang, Sébastien Marcel, and Manuel Günther. Eight years of face recognition research: Reproducibility, achievements and open issues, 2022.

[26] Torsten Schlett, Christian Rathgeb, Olaf Henniger, Javier Galbally, Julian Fierrez, and Christoph Busch. Face image quality assessment: A literature survey, 2021.

[27] Fevziye Irem Eyiokur, Alperen Kantarcı, Mustafa Ekrem Erakın, Naser Damer, Ferda Ofli, Muhammad Imran, Janez Križaj, Albert Ali Salah, Alexander Waibel, Vitomir Štruc, and Hazım Kemal Ekenel. A survey on computer vision based human analysis in the covid-19 era, 2022.

[28] Žiga Babnik and Vitomir Štruc. Assessing bias in face image quality assessment, 2022.

[29] Rezoana Bente Arif, Md. Abu Bakr Siddique, Mohammad Mahmudur Rahman Khan, and Mahjabin Rahman Oishe. Study and observation of the variations of accuracies for handwritten digits recognition with various hidden layers and epochs using convolutional neural network, 2018.

[30] Chris Dulhanty and Alexander Wong. Investigating the impact of inclusion in face recognition training data on individual face identification, 2020.

[31] P. Jonathon Phillips and Mark Przybocki. Four principles of explainable ai as applied to biometrics and facial forensic algorithms, 2020.

[32] Ožbej Golob. Analysis of face detection, face landmarking, and face recognition performance with masked face images, 2022.

[33] V. Karthikeyan, Manjupriya, C. K. Chithra, and M. Divya. Face verification using kernel principle component analysis, 2013.

[34] N. S. Lakshmiprabha. Fusing face and periocular biometrics using canonical correlation analysis, 2016.

[35] Fatemeh Vakhshiteh, Ahmad Nickabadi, and Raghavendra Ramachandra. Adversarial attacks against face recognition: A comprehensive study, 2021.

[36] Eduarda Caldeira, Pedro C. Neto, Tiago Gonçalves, Naser Damer, Ana F. Sequeira, and Jaime S. Cardoso. Unveiling the two-faced truth: Disentangling morphed identities for face morphing detection, 2023.

[37] Ming Li and ChenHao Guo. Automatic annotation of visual deep neural networks, 2021.

[38] Detecting exercise-induced fatig.

[39] Benjamin S. Riggan, Nathaniel J. Short, and Shuowen Hu. Thermal to visible synthesis of face images using multiple regions, 2018.

[40] Pavel Korshunov and Sébastien Marcel. Vulnerability of face recognition to deep morphing, 2019.

[41] João Ribeiro Pinto. Seamless multimodal biometrics for continuous personalised wellbeing monitoring, 2023.

[42] Reza Shoja Ghiass, Ognjen Arandjelovic, Hakim Bendada, and Xavier Maldague. Infrared face recognition: a comprehensive review of methodologies and databases, 2014.

[43] Sheng-Yang Shi, Wen-Zhong Tang, and Yan-Yang Wang. A review on fatigue driving detection. In *ITM Web of Conferences*, volume 12, page 01019. EDP Sciences, 2017.

[44] Boris Knyazev, Roman Shvetsov, Natalia Efremova, and Artem Kuharenko. Convolutional neural networks pretrained on large face recognition datasets for emotion classification from video, 2017.

[45] Shuyi Mao, Xinpeng Li, Junyao Chen, and Xiaojiang Peng. Au-supervised convolutional vision transformers for synthetic facial expression recognition, 2022.

[46] Moshe BenBassat. Wikipedia for smart machines and double deep machine learning, 2018.

[47] Alakhsimar Singh, Nischay Verma, Kanav Goyal, Amritpal Singh, Puneet Kumar, and Xiaobai Li. Visiophysioenet: Multimodal engagement detection using visual and physiological signals, 2025.

[48] Ester Gonzalez-Sosa, Julian Fierrez, Ruben Vera-Rodriguez, and Fernando Alonso-Fernandez. Facial soft biometrics for recognition in the wild: Recent works, annotation, and cots evaluation, 2022.

[49] Adrian Bulat, Shiyang Cheng, Jing Yang, Andrew Garbett, Enrique Sanchez, and Georgios Tzimiropoulos. Pre-training strategies and datasets for facial representation learning, 2022.

[50] Saheb Chhabra, Puspita Majumdar, Mayank Vatsa, and Richa Singh. Data fine-tuning, 2018.

[51] Monali Gulhane and P. S. Mohod. Intelligent fatigue detection and automatic vehicle control system, 2014.

[52] Wei Li and Zhigang Zhu. Learning and fusing multimodal features from and for multi-task facial computing, 2016.

[53] Guiying Zhang, Yuxin Cui, Yong Zhao, and Jianjun Hu. Complexface: a multi-representation approach for image classification with small dataset, 2020.

[54] Chiyoung Song and Dongjae Lee. Facefusion: Exploiting full spectrum of multiple datasets, 2023.

[55] Jing Xu, Tszhang Guo, Yong Xu, Zenglin Xu, and Kun Bai. Multiface: A generic training mechanism for boosting face recognition performance, 2021.

[56] Shubham Tiwari, Yash Sethia, Ritesh Kumar, Ashwani Tanwar, and Rudresh Dwivedi. Face-filtersense: A filter-resistant face recognition and facial attribute analysis framework, 2024.

[57] Yi Sun, Ding Liang, Xiaogang Wang, and Xiaoou Tang. Deepid3: Face recognition with very deep neural networks, 2015.

[58] M. K. Bhowmik, Debotosh Bhattacharjee, M. Nasipuri, D. K. Basu, and M. Kundu. A parallel framework for multilayer perceptron for human face recognition, 2010.

[59] Qi Shen, Shengjie Zhao, Rongqing Zhang, and Bin Zhang. Robust two-stream multi-feature network for driver drowsiness detection, 2020.

[60] Ken Alparslan, Yigit Alparslan, and Matthew Burlick. Towards evaluating driver fatigue with robust deep learning models, 2021.

[61] Mundher Al-Shabi, Wooi Ping Cheah, and Tee Connie. Facial expression recognition using a hybrid cnn-sift aggregator, 2017.

[62] Jing Jiang and Weihong Deng. Disentangling identity and pose for facial expression recognition, 2022.

[63] Ghada Zamzmi, Gabriel Ruiz, Matthew Shreve, Dmitry Goldgof, Rangachar Kasturi, and Sudeep Sarkar. A method to suppress facial expression in posed and spontaneous videos, 2018.

[64] Andrew Jason Shepley. Face recognition in unconstrained conditions: A systematic review, 2019.

[65] Guangcan Mai, Kai Cao, Pong C. Yuen, and Anil K. Jain. On the reconstruction of face images from deep face templates, 2018.

[66] Anh Tuan Tran, Tal Hassner, Iacopo Masi, and Gerard Medioni. Regressing robust and discriminative 3d morphable models with a very deep neural network, 2016.

[67] Fumin Shen, Chunhua Shen, and Heng Tao Shen. Face image classification by pooling raw features, 2014.

[68] Yuhang Lu, Zewei Xu, and Touradj Ebrahimi. Towards a comprehensive visual saliency explanation framework for ai-based face recognition systems, 2024.

[69] Poorya Aghdaie, Baaria Chaudhary, Sobhan Soleymani, Jeremy Dawson, and Nasser M. Nasrabadi. Attention aware wavelet-based detection of morphed face images, 2021.

[70] Daniel Sáez Trigueros, Li Meng, and Margaret Hartnett. Face recognition: From traditional to deep learning methods, 2018.

[71] Zitong Yu, Yunxiao Qin, Xiaobai Li, Chenxu Zhao, Zhen Lei, and Guoying Zhao. Deep learning for face anti-spoofing: A survey, 2022.

[72] Klemen Grm, Vitomir Štruc, Anais Artiges, Matthieu Caron, and Hazim Kemal Ekenel. Strengths and weaknesses of deep learning models for face recognition against image degradations, 2017.

[73] Ankan Bansal, Carlos Castillo, Rajeev Ranjan, and Rama Chellappa. The do's and don'ts for cnn-based face verification, 2017.

[74] Haiyu Wu, Sicong Tian, Jacob Gutierrez, Aman Bhatta, Kağan Öztürk, and Kevin W. Bowyer. Identity overlap between face recognition train/test data: Causing optimistic bias in accuracy measurement, 2024.

[75] Ulrich Scherhag, Christian Rathgeb, Johannes Merkle, and Christoph Busch. Deep face representations for differential morphing attack detection, 2020.

[76] Ziqing Feng and Qijun Zhao. Robust face recognition with deeply normalized depth images, 2018.

[77] Saksham Suri, Anush Sankaran, Mayank Vatsa, and Richa Singh. On matching faces with alterations due to plastic surgery and disguise, 2018.

[78] Johannes Reschke and Armin Sehr. Face recognition with machine learning in opencv$_{fusion of the results with the localization data of an acoustic camera for speaker identification}$, 2017.

[79] Mrinal Kanti Bhowmik, Debotosh Bhattacharjee, Dipak Kumar Basu, and Mita Nasipuri. Polar fusion technique analysis for evaluating the performances of image fusion of thermal and visual images for human face recognition, 2011.

[80] Daniel Crispell, Octavian Biris, Nate Crosswhite, Jeffrey Byrne, and Joseph L. Mundy. Dataset augmentation for pose and lighting invariant face recognition, 2017.

[81] Chuanhao Zhuge, Xinheng Liu, Xiaofan Zhang, Sudeep Gummadi, Jinjun Xiong, and Deming Chen. Face recognition with hybrid efficient convolution algorithms on fpgas, 2018.

[82] Poonam Yadav. Face prediction model for an automatic age-invariant face recognition system, 2015.

[83] Nataniel Ruiz, Adam Kortylewski, Weichao Qiu, Cihang Xie, Sarah Adel Bargal, Alan Yuille, and Stan Sclaroff. Simulated adversarial testing of face recognition models, 2022.

[84] Hao Liang, Pietro Perona, and Guha Balakrishnan. Benchmarking algorithmic bias in face recognition: An experimental approach using synthetic faces and human evaluation, 2023.

[85] Aleksei Zhuchkov. Analyzing the effectiveness of image augmentations for face recognition from limited data, 2021.

[86] Meftah Ur Rahman. A comparative study on face recognition techniques and neural network, 2012.

[87] Erdi Sarıtaş and Hazım Kemal Ekenel. Analyzing the effect of combined degradations on face recognition, 2024.

[88] Sandipan Banerjee, Joel Brogan, Janez Krizaj, Aparna Bharati, Brandon RichardWebster, Vitomir Struc, Patrick Flynn, and Walter Scheirer. To frontalize or not to frontalize: Do we really need elaborate pre-processing to improve face recognition?, 2018.

[89] Javier Hernandez-Ortega, Javier Galbally, Julian Fierrez, and Laurent Beslay. Biometric quality: Review and application to face recognition with faceqnet, 2021.

[90] Mathias Ibsen, Lázaro J. González-Soler, Christian Rathgeb, and Christoph Busch. Tetraloss: Improving the robustness of face recognition against morphing attacks, 2024.

[91] Hsin-Rung Chou, Jia-Hong Lee, Yi-Ming Chan, and Chu-Song Chen. Data-specific adaptive threshold for face recognition and authentication, 2018.

[92] Mohammad Pezeshki, Sajjad Gholami, and Ahmad Nickabadi. Distinction between features extracted using deep belief networks, 2014.

[93] Brandon RichardWebster, So Yon Kwon, Christopher Clarizio, Samuel E. Anthony, and Walter J. Scheirer. Visual psychophysics for making face recognition algorithms more explainable, 2018.

[94] Joana Pimenta, Iurii Medvedev, and Nuno Gonçalves. Impact of image context for single deep learning face morphing attack detection, 2023.

[95] Gaurav Goswami, Nalini Ratha, Akshay Agarwal, Richa Singh, and Mayank Vatsa. Unravelling robustness of deep learning based face recognition against adversarial attacks, 2018.

[96] Ben Wycliff Mugalu, Rodrick Calvin Wamala, Jonathan Serugunda, and Andrew Katumba. Face recognition as a method of authentication in a web-based system, 2021.

[97] Jacqueline G. Cavazos, P. Jonathon Phillips, Carlos D. Castillo, and Alice J. O'Toole. Accuracy comparison across face recognition algorithms: Where are we on measuring race bias?, 2020.

[98] Ahmad H. A. Eid. Combined classifiers for invariant face recognition, 2016.

[99] Mathias Ibsen, Christian Rathgeb, Thomas Fink, Pawel Drozdowski, and Christoph Busch. Impact of facial tattoos and paintings on face recognition systems, 2021.

[100] Yuhang Lu, Zewei Xu, and Touradj Ebrahimi. Towards visual saliency explanations of face verification, 2023.

Disclaimer:

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.