Facial Landmark Detection and Related Technologies in Computer Vision: A Survey

www.surveyx.cn

Abstract

Facial landmark detection is a cornerstone in computer vision, enabling applications ranging from face alignment and 3D modeling to virtual makeup and security. This survey paper provides a comprehensive review of the evolution and current state of facial landmark detection technologies. It explores methodologies from traditional techniques to advanced deep learning frameworks, emphasizing the integration of convolutional neural networks (CNNs) and their role in enhancing precision across challenging conditions such as occlusions and pose variations. The survey highlights innovations in 3D facial landmarks, face morphing, and pose estimation, underscoring their significance in improving facial analysis accuracy. Furthermore, it addresses the application of these technologies in multimedia and daily life, particularly in enhancing user experiences through virtual makeup and personalized interfaces. The paper identifies key challenges, including the need for robust models that can generalize across diverse environments and the importance of comprehensive datasets for training. Future directions suggest focusing on optimizing real-time performance and integrating advanced techniques to improve robustness and generalization. The insights gained aim to guide future research and development, ensuring these technologies continue to evolve and meet modern application demands, while also considering ethical implications in their deployment. Overall, this survey underscores the pivotal role of facial landmark detection in advancing computer vision applications and highlights the ongoing efforts to address persistent challenges in the field.

1 Introduction

1.1 Importance of Facial Landmark Detection

Facial landmark detection is crucial in computer vision, enabling accurate localization of facial features for applications such as animation, recognition, and expression analysis [1]. Its importance is amplified in challenging conditions, where factors like lighting and occlusion can hinder accuracy [2]. Despite advancements, issues persist, particularly with severe occlusion and large head poses [3].

The adoption of convolutional neural networks (CNNs) represents a significant advancement, providing robust solutions for landmark detection [4]. These approaches are vital for applications requiring precise landmark detection, such as facial expression analysis and head pose estimation, which enhance overall system performance [4]. However, semantic ambiguity from inconsistent annotations remains a challenge, negatively impacting detection performance [5].

Facial landmark detection is also essential for aligning facial features across varying poses, critical for reliable face recognition and manipulation [6]. Its relevance extends beyond human faces to animal affective computing, where automated systems are developed for species like cats [7].

This technology underpins advancements in computer vision by enabling precise tracking of facial features essential for applications such as augmented reality and emotion analysis. The field

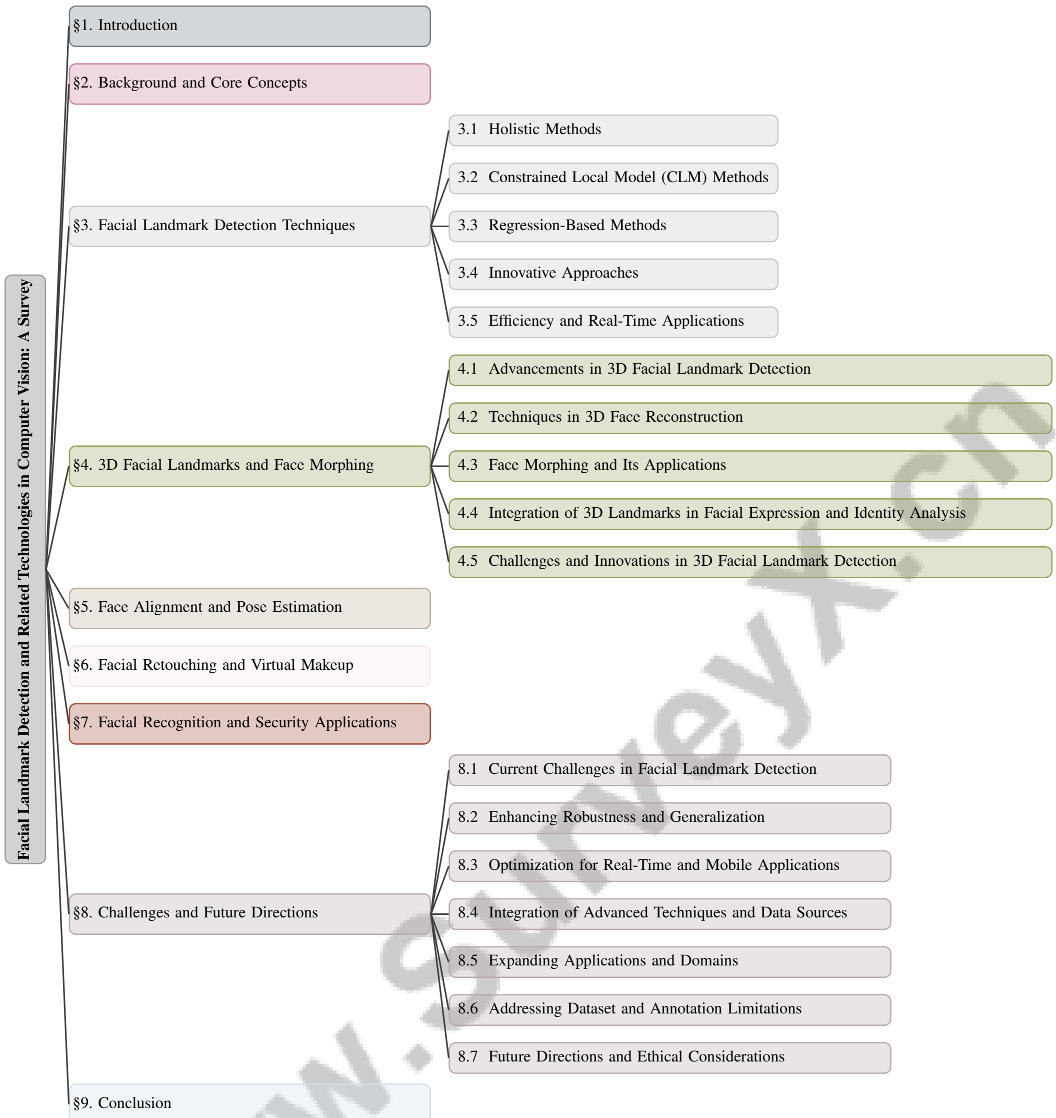


Figure 1: chapter structure

has evolved through diverse algorithms, including holistic, Constrained Local Model (CLM), and regression-based methods, each adept at capturing facial deformations caused by expressions and head movements. Recent innovations, including knowledge distillation and novel CNN architectures, have further enhanced model accuracy and efficiency, particularly in real-world scenarios characterized by varying expressions and orientations. Addressing challenges like resource limitations and the need for robust performance across diverse populations, facial landmark detection continues to drive innovation across multiple domains [8, 9, 10, 11, 4].

1.2 Applications in Multimedia and Daily Life

Facial landmark detection has diverse applications across multimedia and daily life, impacting sectors such as entertainment, healthcare, and social media. In multimedia, it is fundamental for creating realistic animations and special effects in films and video games, where precise localization of facial features is essential for lifelike character expressions. The rise of augmented reality (AR) has further broadened its utility, exemplified by applications like FaceAtlasAR, which accurately localizes facial acupuncture points for educational and self-treatment purposes [12].

In everyday contexts, facial landmark detection enhances social media experiences through features like virtual makeup and facial retouching, allowing real-time alterations of user appearances. It also aids in privacy protection by enabling automatic blurring or masking of faces in shared photos and videos, thus safeguarding identities in social media and legal settings [13].

Moreover, this technology facilitates personalized user interfaces in smart devices by recognizing user identities and adapting functionalities accordingly. In healthcare, it enhances non-invasive diagnostic tools and therapeutic aids by analyzing facial expressions and detecting subtle variations in features, improving emotion recognition and patient assessments [13, 14, 15, 16, 17]. The integration of facial landmark detection into various aspects of multimedia and daily life showcases its versatility and transformative potential in enhancing user interactions and experiences.

1.3 Scope and Relevance of the Survey

This survey offers a comprehensive analysis of facial landmark detection technologies, tracing their historical evolution and current integration across computer vision domains. It encompasses a wide range of methodologies, from traditional techniques to advanced deep learning frameworks, such as the joint AU detection and face alignment model, which improves accuracy by leveraging task correlations [18]. A key focus is the challenge of failure detection in landmark systems, which is essential for the reliability of applications relying on these technologies [19].

The survey emphasizes robustness by examining detection model performance under challenging conditions, such as motion blur and occlusions, which significantly affect accuracy [20]. It highlights the need for lightweight and efficient models suitable for mobile and edge device deployment, with benchmarks assessing model size, parameters, and inference time [21]. This focus is critical for applications in resource-constrained environments, balancing computational efficiency with high accuracy [22].

Additionally, the survey explores 3D surface imaging technologies, including stereophotogrammetry and laser-based scanning, with applications in craniofacial research [23]. It discusses the integration of facial landmark detection in real-time facial expression recognition systems, highlighting its role in enhancing system performance and responsiveness [24].

Face alignment techniques are categorized into generative and discriminative methods, identifying existing knowledge gaps [25]. The review covers recent advancements in neural network-based algorithms for facial landmark detection from 2018 to 2021, alongside older algorithms for context [26].

The survey also addresses the lack of dense facial landmarks in existing datasets, which are vital for detailed facial analysis and applications [27]. It considers benchmarks for evaluating the effectiveness of fusing human examiner decisions in digital face manipulation detection, improving upon existing automated systems [28]. Furthermore, it discusses various face models used in face alignment, including 2D and 3D models, while excluding specific applications outside of face alignment methods [29].

This survey aims to provide a thorough review of the current state and future directions of facial landmark detection technologies, underscoring their pivotal role in advancing computer vision applications and addressing enduring challenges such as occlusions, pose variations, and demographic biases. The insights derived are intended to guide future research and development efforts, ensuring that these technologies evolve to meet the demands of modern applications [30].

1.4 Structure of the Survey

The survey is systematically organized to thoroughly explore facial landmark detection and its associated technologies within computer vision. The introductory section highlights the importance of facial landmark detection, its applications, and the survey's scope and relevance. Following this, the paper delves into the background and core concepts, providing insights into the historical development and fundamental methodologies of facial landmark detection.

The survey transitions into a detailed examination of various facial landmark detection techniques, including holistic methods, Constrained Local Model (CLM) methods, and regression-based approaches, assessing their advantages and limitations. This section is complemented by discussions on innovative approaches and the efficiency of these techniques in real-time applications.

Subsequent sections focus on advanced topics such as 3D facial landmarks, face morphing, face alignment, pose estimation, and facial retouching and virtual makeup. Each section explores respective technologies, methodologies, and implications for enhancing digital media and user experiences.

This analysis explores the critical role of facial recognition in security applications, emphasizing how integrating facial landmark detection techniques enhances recognition accuracy by effectively capturing key facial features and accommodating variations in head movements, expressions, and environmental conditions [8, 19, 9, 11, 26]. The survey concludes with a discussion on challenges and future directions in the field, offering insights into current limitations and potential research trajectories.

This structured approach provides an in-depth exploration of facial landmark detection technologies, offering readers a comprehensive understanding of various algorithms—from holistic methods to deep learning-based techniques—while addressing their strengths, weaknesses, and performance across diverse conditions. By categorizing these methods and analyzing their applications in fields such as facial recognition and forensic analysis, the text guides readers through the complexities and recent advancements in facial landmark detection, ultimately highlighting future research directions for improved accuracy and reliability in real-world scenarios [9, 31, 4, 1].The following sections are organized as shown in Figure 1.

2 Background and Core Concepts

2.1 Historical Development of Facial Landmark Detection

Facial landmark detection has advanced significantly alongside computer vision, initially focusing on face alignment within limited pose variations but struggling with angles beyond 90° [32]. The mid-2010s marked a turning point, with improved alignment performance on benchmarks like the 300W dataset [33], driving research into enhancing localization precision through deep learning, although 2D alignment challenges persisted [34]. Specific applications, such as neonatal landmark detection, highlighted the inadequacy of models trained on adult data, underscoring the need for specialized datasets for diverse populations [35]. This historical trajectory reflects ongoing efforts to address challenges such as pose variations and demographic diversity, employing holistic, CLM, and regression-based strategies. Deep learning innovations have improved detection accuracy and efficiency, yet issues like annotation noise and robust failure detection remain critical [9, 8, 19, 36]. These developments have led to sophisticated models leveraging deep learning to enhance accuracy and robustness across varied contexts.

2.2 Core Concepts in Facial Landmark Detection

Facial landmark detection is crucial for precise facial feature localization in applications like recognition, augmented reality, and emotion analysis. Traditional methods, such as ASMs and AAMs, focus on holistic and local variations but struggle with generalization across diverse appearances, prompting innovative solutions [1]. Recent advancements incorporate CNNs, effectively managing complex scenarios with significant pose variations and occlusions [4]. Current methodologies include regression and heatmap approaches, with their integration enhancing accuracy and reliability. A unified approach combining face detection and landmark localization is critical for addressing challenges of diverse appearances [6], optimizing detection and localization tasks. Specialized datasets, like those with 48 cat facial landmarks, contribute to diverse training data [7]. Innovative methods, such as using 3DMM parameters in joint detection and localization, enhance accuracy through shared learning, iteratively refining predictions against occlusions [9, 3, 37]. These advancements highlight the evolution of methodologies addressing pose, occlusion, and demographic diversity, ensuring accurate and reliable detection in varied conditions.

2.3 Integration with Computer Vision Applications

Facial landmark detection is integral to computer vision applications, enhancing system accuracy in recognition, emotion detection, and augmented reality. Its role in face alignment is critical, with

AI-generated, for reference only.

models like ERCLM addressing challenges of varying expressions and poses [38]. Accurate detection under occlusions and large pose variations remains challenging [39], with datasets like one comprising 20,600 images serving as robust test-beds [40]. Incorporating 3D structures improves localization, addressing limitations of 2D approaches [41]. Addressing variability in appearances due to dynamic features and occlusions is crucial for advancing recognition applications [25]. Techniques like boundary-aware alignment enhance robustness across poses and occlusions, reducing inconsistencies [42]. Integration with emotion recognition systems benefits from multiple modalities, including 3D data, enhancing detection accuracy [16]. Despite advancements, real-time performance is challenged by computational demands, significant for applications needing immediate feedback [43]. Additionally, vulnerabilities to morphing attacks necessitate robust techniques distinguishing genuine from manipulated data [44].

3 Facial Landmark Detection Techniques

Category	Feature	Method
Holistic Methods	Projection Techniques Model Flexibility	3DSTN[41] BRF[45]
Constrained Local Model (CLM) Methods	Hypothesis Evaluation Sequential Regression Part-Based Approaches	ERCLM[38] BCR[46] CPA[47]
Regression-Based Methods	Sequential Processing Enhancements Task Integration and Alignment 3D Reconstruction and Modeling	ACE-Net[48], MSM[49] BAFR[50], TCR[51] JFAR[52], PIFA[53]
Innovative Approaches	Pose and Landmark Enhancement Feature and Representation Optimization Model Accuracy and Robustness	BAFA[42], i2p[54], RHT[55] DSAT[56] CE-CLM[57]
Efficiency and Real-Time Applications	Data and Annotation Techniques Resource Optimization Accuracy Enhancement	AvS[2], SA[5], PWC+Disc[4] MGR[58], MFN[59], MHM[6] TS3[60], SAN[61], RCRF[3], RePFormer[62]

Table 1: This table provides a comprehensive overview of various facial landmark detection techniques categorized into holistic methods, constrained local model methods, regression-based methods, innovative approaches, and efficiency-focused techniques. Each category is detailed with specific features and methods, highlighting their contributions to enhancing system accuracy, robustness, and applicability across diverse scenarios. The table serves as a reference for understanding the advancements and methodologies in facial landmark detection.

Facial landmark detection encompasses diverse methodologies aimed at enhancing system accuracy and robustness. This section explores key approaches, beginning with holistic methods that leverage comprehensive facial geometry to address challenges such as pose variations and occlusions. As illustrated in **??**, the hierarchical categorization of facial landmark detection techniques highlights these key methodologies. The figure details various categories, including holistic methods, constrained local model methods, regression-based methods, innovative approaches, and efficiency-focused techniques. Each category is further elaborated with specific methods and models that contribute to improved accuracy, robustness, and applicability across different scenarios. Additionally, Table 4 offers a detailed comparison of the key facial landmark detection techniques discussed in this section, emphasizing their categorization and specific methods employed to improve accuracy and robustness. Table 1 provides a detailed summary of these techniques.

3.1 Holistic Methods

Holistic methods, focusing on entire facial geometry, effectively address complexities associated with varying structures and conditions, excelling in scenarios with significant pose variations and occlusions [33]. Neural network architectures have advanced these methods, with the Backbone Replaceable Fine-tuning Framework (BRF) enhancing video performance [45], and Memory-efficient Global Refinement (MGR) reducing memory requirements through quantization [58]. The 3D Spatial Transformer Network (3DSTN) models camera projection for accurate alignment, especially in non-frontal views [41], while deep learning approaches using heatmaps demonstrate superiority over traditional models [29]. Dynamic Neural Networks inspire the DSAT method to enhance detection through efficient learning [56], and the Ensemble Landmark Detector (ELD) improves accuracy through a magnifying ensemble method [7]. Holistic methods integrate advanced models capturing global facial appearance, enhancing robustness and precision in applications like facial analysis and emotion recognition [8, 63, 64, 9, 4].

3.2 Constrained Local Model (CLM) Methods

Method Name	Methodology Framework	Robustness Techniques	Application Domains
ERCLM[38]	Hypothesize-and-evaluate	Occlusion Reasoning	Face Alignment
CPA[47]	Piece-wise Planar	Spatial Relationships Alignment	Face Recognition
CE-CLM[57]	Convolutional Experts Network	Mixture OF Experts	Landmark Detection
BCR[46]	Branching Cascaded Regression	Branching Cascaded Regression	Face Alignment

Table 2: Comparison of Constrained Local Model (CLM) methods, detailing their methodological frameworks, robustness techniques, and application domains. The table highlights the diverse approaches employed by each method to enhance accuracy and robustness in facial analysis tasks.

Constrained Local Models (CLM) utilize local appearance variations around feature points to enhance accuracy [1]. Table 2 presents a comparative analysis of various Constrained Local Model (CLM) methods, illustrating their distinct methodologies, robustness techniques, and application domains within the field of facial feature detection and recognition. The Enhanced Regression Constrained Local Model (ERCLM) uses a hypothesize-and-evaluate framework for joint inference of landmark positions and occlusion labels [38]. Techniques like Constrained Part-based Alignment (CPA) improve robustness against variations by aligning facial parts while considering spatial relationships [47]. The Convolutional Experts Constrained Local Model (CE-CLM) employs a Convolutional Experts Network for enhanced localization accuracy [57]. Branching in cascaded regression sequentially solves simpler regression problems, improving handling of varying head poses [46]. CLM methods advance detection by integrating local appearance models with robust frameworks adaptable to facial variations, enhancing precision in applications like facial recognition and forensic analysis [64, 36, 9, 19, 31].

3.3 Regression-Based Methods

Method Name	Method Integration	Generalization Techniques	Robustness Enhancements
JFAR[52]	Joint Method	Cascaded Regressor Framework	Iterative Refinement Process
TCR[51]	-	Transductive Cascaded Regression	-
MSM[49]	Face Alignment, 3D Reconstruction	Transductive Cascaded Regression	Spatial Transformer Networks
ACE-Net[48]	3D Morphable Model	-	Two-module System
BAFR[50]	Alignment Learning Recognition	-	Spatial Transformer Network
PIFA[53]	3D Deformable Model	Cascaded Coupled-regressor	3D Modeling Integration

Table 3: Overview of regression-based methods for facial landmark detection, highlighting their integration approaches, generalization techniques, and robustness enhancements. The table includes methods such as JFAR, TCR, MSM, ACE-Net, BAFR, and PIFA, each contributing uniquely to the field through various methodologies and innovations.

Table 3 presents a comprehensive comparison of various regression-based methods used in facial landmark detection, focusing on their integration, generalization, and robustness strategies. Regression-based methods predict landmark positions by mapping image features to facial points [1], often involving iterative refinement for enhanced accuracy and robustness. The Joint Face Alignment and 3D Face Reconstruction (JFAR) method integrates face alignment with 3D reconstruction for improved localization [52]. The Transductive Cascaded Regression (TCR) method facilitates annotation transfer across datasets, enhancing generalization [51]. Multi-stage models employ spatial transformer networks and stacked hourglass networks for enhanced robustness [49], while ACE-Net leverages fine-level anchors for superior accuracy [48]. The Balanced Alignment and Face Recognition (BAFR) model combines alignment learning with recognition tasks [50], and the Pose-Invariant 3D Face Alignment (PIFA) method enhances accuracy with a 3D deformable model [53]. These methods efficiently map image features to landmark positions, addressing challenges like pose variations and diverse datasets, enhancing accuracy across applications [9, 10, 64, 36].

3.4 Innovative Approaches

Recent advancements have introduced innovative methodologies enhancing detection accuracy, robustness, and efficiency. The img2pose method integrates pose estimation and face detection for real-time processing [54]. DSAT partitions samples into subsets for specialized feature learning [56], while Semantic Alignment optimizes a latent variable for improved alignment [5]. The Reference

Heatmap Transformer (RHT) models facial shape constraints for enhanced accuracy [55], and RePFormer improves localization robustness through landmark query refinement [62]. Boundary-aware approaches use boundary heatmaps for improved alignment tasks [42], and the Mixture of Expert Layer in CEN improves predictions through expert output combinations [57]. The Joint Multi-view Hourglass Model (MHM) employs a coarse-to-fine approach for robust estimation [6], and the Aggregation via Separation (AvS) method advances traditional methods through disentangled representation learning [2]. These innovations signify a transformative shift in the field, enhancing accuracy, efficiency, and adaptability across applications like augmented reality and emotion analysis [8, 64, 36, 9, 26].

3.5 Efficiency and Real-Time Applications

Efficiency and real-time applicability are critical for deploying systems on devices with limited resources. The RePFormer method balances efficiency and accuracy by modeling dependencies between landmarks [62]. MGR reduces memory usage while maintaining accuracy, ideal for constrained environments [58]. TS3 employs a feedback loop for more accurate predictions [60], and the Style Aggregated Network (SAN) enhances prediction accuracy through a dual-input strategy [61]. The Multi-task Facial Network (MFN) demonstrates high accuracy suitable for mobile devices [59]. The Robust Cascade Regression Framework (RCRF) improves detection in challenging images [3], and AvS enhances accuracy through data augmentation [2]. Semantic Alignment method reduces annotation noise, crucial for real-time applications [5], while CNN-based approaches involve pixel-wise classification for enhanced real-time applicability [4]. The Joint Multi-view Hourglass Model (MHM) enhances real-time efficiency and accuracy [6], illustrating the potential of technologies to transform real-time applications by improving efficiency and accuracy, enabling deployment in resource-constrained environments without compromising performance.

Feature	Holistic Methods	Constrained Local Model (CLM) Methods	Regression-Based Methods
Accuracy Enhancement	Heatmap-based Models	Local Appearance Models	Iterative Refinement
Robustness Strategy	Pose And Occlusion	Spatial Relationships	Dataset Generalization
Application Domain	Facial Analysis	Facial Recognition	3D Face Reconstruction

Table 4: This table provides a comparative analysis of facial landmark detection methods, categorized into holistic methods, constrained local model (CLM) methods, and regression-based methods. It highlights the key features of each category, including accuracy enhancement techniques, robustness strategies, and application domains, offering a comprehensive overview of their distinct approaches and contributions to the field.

4 3D Facial Landmarks and Face Morphing

4.1 Advancements in 3D Facial Landmark Detection

Recent advancements in 3D facial landmark detection have markedly enhanced the precision and robustness of facial analysis systems, effectively addressing challenges in predicting 3D facial structures. The introduction of the nonlinear 3D Morphable Model (N3DMM) significantly improves representation power over traditional linear models, enabling effective reconstruction of facial features from unconstrained images [65]. This advancement highlights the potential of nonlinear methods in capturing complex facial geometries.

The application of neural radiance fields (NeRFs) for direct prediction of 3D face landmarks exemplifies the shift towards advanced neural representations, improving accuracy across diverse scenarios and overcoming the limitations of previous 2D landmark-based methods [66]. Additionally, RingNet demonstrates the ability to estimate 3D face shapes from single images without 3D supervision, outperforming methods reliant on 3D training data [67].

Integrating geometric angle-based and texture-based features enhances model generalization in real-time applications, as shown in facial expression recognition systems [24]. The Backbone Replaceable Fine-tuning Framework (BRF) further advances detection by incorporating temporal coherence, significantly stabilizing predictions in dynamic video contexts [45]. This adaptability underscores the evolution of 3D facial landmark detection methods in dynamic scenarios.

The development of realistic 3D morphable models (3DMM) enables intuitive facial attribute editing, crucial for robust facial recognition across various poses and expressions, achieving state-of-the-art performance in face alignment and 3D reconstruction [52]. Moreover, generating synthetic training samples through style translation addresses challenges in maintaining facial geometry integrity during landmark detection [2].

These advancements represent a collaborative effort to enhance accuracy and efficiency in 3D facial landmark detection. By integrating holistic methods, Constrained Local Models (CLMs), and regression-based techniques, researchers tackle challenges such as facial deformations due to head movements and expressions. Innovative strategies, including spatial transformer networks for optimal face normalization and knowledge distillation for lightweight model creation, pave the way for robust applications in augmented reality, facial recognition, and emotion analysis. This comprehensive refinement not only improves performance in controlled environments but also enhances adaptability in real-world conditions, leading to more sophisticated applications in computer vision [9, 8, 1, 64].

4.2 Techniques in 3D Face Reconstruction

3D face reconstruction is a pivotal technique in computer vision, enabling the creation of detailed and realistic facial models from 2D images using landmark data. The Weakly-supervised Multi-face 3D Reconstruction (WM3DR) framework stands out by employing a shared ResNet-50 backbone as the encoder and a 3D Morphable Model (3DMM) as the decoder, effectively reconstructing facial shape and texture [68]. This approach underscores the potential of weak supervision and shared architectures to enhance reconstruction fidelity.

The integration of multi-task learning frameworks facilitates the reconstruction of facial features using shared parameters, thereby enhancing efficiency and accuracy in 3D face modeling [69]. Automated systems capturing facial geometry and appearance from multi-camera photographs, followed by automatic landmarking and deformation transfer, have proven effective in creating animated meshes [70].

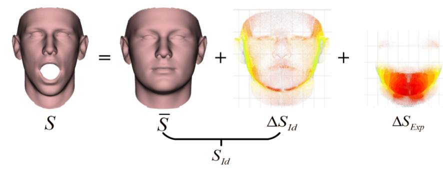
FLNeRF, a multi-scale, coarse-to-fine model, predicts 3D face landmarks directly from neural radiance fields, utilizing 3D convolutional networks for facial feature sampling, illustrating the shift towards advanced neural representations [66]. The encoder-decoder architecture proposed by [71] combines parameter estimation with detailed reconstruction processes, effectively estimating shape and albedo parameters.

RingNet processes multiple images in parallel, learning a consistent mapping from 2D facial features to 3D face shape parameters, thus enhancing robustness and consistency in reconstructions [67]. The nonlinear 3D Morphable Model (N3DMM) approach, involving an encoder-decoder architecture, significantly enhances representation power, facilitating accurate reconstructions [65]. The part-based 3D Morphable Model (P3DMM) segments the face into independent parts, applying principal component analysis (PCA) for localized editing and reconstruction, crucial for detailed facial modifications [72].

These methodologies illustrate a variety of innovative approaches to 3D face reconstruction, leveraging landmark data to improve precision, realism, and versatility across numerous applications. For instance, one method predicts ten times more landmarks than traditional techniques, covering the entire head and enabling detailed performance capture. Another approach integrates joint face alignment with 3D reconstruction, allowing simultaneous generation of pose-and-expression-normalized 3D faces, significantly enhancing recognition capabilities. The synergy between 3D Morphable Models and 3D landmarks facilitates comprehensive predictions of facial geometry, pushing the boundaries of 3D face modeling while addressing challenges such as expressions, head poses, and occlusions [73, 74, 9, 52].

As shown in Figure 2, the integration of facial landmarks and morphing techniques significantly advances our understanding and manipulation of facial expressions. The first image, "Facial Expressions: A Visual Representation of Different Emotions," showcases a sequence of 3D-rendered faces, each depicting a distinct emotion. This visual progression from neutrality to expressive states, such as a wide smile, underscores the nuanced representation achievable through 3D modeling. The uniform blue rendering and soft lighting accentuate facial contours, clearly depicting emotional transitions. The second image, "Facial Expression Recognition Using a 3D Human Head Model," delves into technical aspects, featuring a 3D model segmented into identity-specific and expression-

(a) Facial Expressions: A Visual Representation of Different Emotions[75]



(b) Facial Expression Recognition Using a 3D Human Head Model[76]

Figure 2: Examples of Techniques in 3D Face Reconstruction

specific components, enhancing precision in recognizing and analyzing facial expressions. Together, these examples underscore the sophistication and potential of 3D facial landmarking and morphing techniques in reconstructing and interpreting human facial expressions [75, 76].

4.3 Face Morphing and Its Applications

Face morphing, a transformative technique in computer vision, blends two or more facial images to create a new, intermediate face that retains features from the original inputs. This process is pivotal in various applications, from entertainment to security. A notable advancement is the generation of 3D morphing point clouds, which aligns and averages the coordinates and colors of two input point clouds using Bayesian Coherent Point Drift (BCPD) for registration [77]. This method enhances realism and accuracy, making morphed faces suitable for applications requiring detailed 3D representations.

The automated pipeline developed by [70] exemplifies face morphing's utility in creating facial meshes that retain unique surface details, essential for background characters in virtual environments, effectively conveying emotions. The ability to edit facial features locally without affecting unrelated areas, as proposed by [72], significantly enhances user experience in applications like digital makeup and personalized avatar creation, allowing precise modifications.

Face morphing also plays a critical role in security applications, particularly regarding morphing attacks, which exploit the process to create deceptive facial images that can mislead biometric systems. This raises significant privacy concerns, as morphed images can facilitate misleading identities in digital media, undermining identity verification processes and enabling unauthorized access to sensitive information. The manipulation affects personal privacy and poses security threats, particularly in identity documents, where a single morphed image may be associated with multiple identities, complicating detection of fraudulent activities [13, 78, 79].

Face morphing is a highly adaptable technology that poses significant security challenges, particularly in electronic identity verification, while also offering diverse applications across fields such as biometric security, facial anonymization, and image quality enhancement. Its potential for misuse necessitates robust detection methods, as evidenced by ongoing research aimed at improving morphing attack detection systems and exploring innovative uses in journalism and privacy protection [13, 78, 80, 81, 82]. Its ability to alter facial appearances while maintaining realistic features makes it invaluable in entertainment, digital media, and security, where innovation and precision are paramount.

4.4 Integration of 3D Landmarks in Facial Expression and Identity Analysis

The integration of 3D facial landmarks into the analysis of facial expressions and identity has revolutionized the accuracy and depth of facial analysis in computer vision. This integration allows for a nuanced understanding of facial dynamics, capturing subtle variations in expressions and identity often lost in 2D analysis. 3D landmarks provide a robust framework for analyzing facial expressions by accounting for geometric and structural changes during different expressions [24].

A key advantage of 3D landmark integration is the ability to accurately capture and analyze facial expressions across a broad range of poses and lighting conditions. This capability is crucial in applications like emotion recognition, where precise localization of facial landmarks significantly enhances expression analysis accuracy [66]. The use of neural radiance fields (NeRFs) in predicting

3D face landmarks exemplifies the shift towards advanced neural representations for improved accuracy in facial analysis across diverse scenarios [66].

In identity analysis, 3D landmarks provide a comprehensive representation of facial geometry, enabling more accurate and reliable identification processes. Integrating 3D facial data enhances the robustness of identity verification systems by adding spatial information that complements traditional 2D methods [65]. This integration is crucial for applications requiring high precision, such as biometric authentication and security systems.

Moreover, the ability to reconstruct 3D facial shapes from single images without requiring 3D supervision, as demonstrated by methods like RingNet, highlights the potential of 3D landmarks in identity analysis [67]. These advancements facilitate the creation of detailed 3D facial models for applications like virtual reality and personalized avatar creation.

The integration of 3D landmarks into facial expression and identity analysis marks a significant advancement in computer vision, enhancing the accuracy and comprehensiveness of understanding facial dynamics. This approach leverages dense landmark representations, capturing intricate facial features and movements that sparse landmarks often overlook, such as cheek contours and the outer eye region. Utilizing advanced techniques like differentiable rendering and morphable models, researchers achieve state-of-the-art results in 3D face reconstruction and performance capture. Additionally, incorporating 3D landmarks facilitates improved temporal stability and accuracy in landmark detection, providing a robust framework for analyzing expressions and identities across diverse scenarios [9, 73, 64]. This integration continues to drive innovation, enabling the development of sophisticated applications that leverage the full potential of 3D facial data.

4.5 Challenges and Innovations in 3D Facial Landmark Detection

Despite significant strides in 3D facial landmark detection, challenges persist, driving ongoing innovation. A primary challenge is the reliance on linear assumptions and the need for extensive controlled datasets, which are costly and difficult to obtain, limiting existing methods' adaptability in diverse environments [83]. This highlights the necessity for models capable of operating effectively in less controlled settings.

Handling occlusions and large pose variations remains critical. Techniques like FoxNet leverage global semantic information and multi-scale features to manage occlusions and accommodate varying face numbers [84]. The synergy between 3D Morphable Models (3DMM) and 3D landmarks improves accuracy in predicting facial geometry, better handling occlusions and enhancing performance in large-pose scenarios [74].

The computational demands for producing high-resolution outputs and real-time applications, especially with diverse facial expressions, pose additional obstacles [85]. Landmark-free approaches like ExpNet provide robust facial expression estimation under varying scales and occlusions, addressing some limitations of traditional landmark-based methods [86].

Shadow removal is crucial for improving the robustness of facial landmark detection, particularly in severe shadow conditions [87]. This underscores the importance of preprocessing techniques in enhancing detection accuracy.

Innovative methodologies like weakly-supervised multi-face 3D reconstruction frameworks reduce computational time and simplify deployment while accurately recovering relative positions of multiple faces in a unified camera model [68]. These advancements facilitate more efficient and accessible 3D facial analysis.

Real-time applicability remains a focal point of innovation, with methods achieving significant improvements in processing speed, enhancing utility in dynamic environments [88]. Future research could explore machine learning techniques to identify nonlinear mappings and incorporate fine-scale details like wrinkles into models [72]. Additionally, RingNet exemplifies advancements in learning without 3D supervision, offering robustness across diverse imaging conditions and comprehensive reconstruction of the full head and neck [67].

Despite ongoing challenges, recent advancements in algorithm design and imaging technology are significantly improving the accuracy, efficiency, and versatility of 3D facial landmark detection systems. Innovations such as joint training of spatial transformer networks for optimal face normalization,

AI-generated, for reference only.

enhancements in output predictions to infer landmarks in canonical 3D space, and the development of semantic correction networks to address dataset inconsistencies drive these improvements. Furthermore, advancements in 3D surface imaging systems enable clinicians to select purpose-specific technologies that enhance craniofacial research and applications, broadening the scope of 3D facial analysis across various fields [9, 23, 64].

5 Face Alignment and Pose Estimation

Face alignment and pose estimation are pivotal in computer vision, essential for applications in security systems and augmented reality. Their efficacy is closely linked to the precision of these processes, necessitating a thorough exploration of the methodologies employed. This section delves into various face alignment techniques, focusing on advancements that enhance the precision and robustness of facial feature normalization, thereby laying the foundation for understanding their practical applications.

5.1 Methods for Face Alignment

Face alignment is crucial for normalizing facial features, facilitating accurate analysis across varying conditions, including diverse poses, occlusions, and expressions. Numerous techniques have been developed to bolster the robustness and precision of face alignment. Dual-dimensional networks that integrate 2D and 3D features leverage multidimensional data, enhancing alignment robustness [89]. The img2pose method efficiently regresses the 6DoF pose for all faces in an image without prior face detection, optimizing real-time applications [54]. Boundary-aware techniques, utilizing heatmaps to represent facial structures, refine landmark position regression, crucial for precise alignment [42].

Incorporating 3D face models has proven effective, with techniques optimizing lighting distribution to ensure consistent alignment across different lighting conditions [90]. Iterative methods that alternate between updating 2D landmarks and refining 3D shapes enhance alignment [52]. The Convolutional Experts Constrained Local Model (CE-CLM) combines Convolutional Experts Network (CEN) strengths with a constrained optimization framework, improving landmark localization, particularly for irregular shapes [57]. The FaceAtlasAR model improves the accuracy of anatomical landmark identification, benefiting applications like acupuncture [12]. Additionally, the Multi-task Facial Network (MFN) optimizes alignment for multiple faces while addressing computational complexity [59].

Benchmarks assessing the impact of face alignment on image quality in facial analysis tasks underscore the necessity of precise alignment for effective analysis [91]. The diverse methodologies in face alignment reflect a multifaceted effort to enhance facial feature normalization, ranging from traditional landmark-based approaches to advanced deep learning techniques, addressing challenges like pose variation and occlusions. This ongoing research aims to improve performance metrics in tasks such as recognition and expression detection while ensuring high-quality outputs across various datasets and applications [29, 91, 33, 25, 92].

5.2 Pose Estimation Techniques

Pose estimation is critical for accurately determining head orientation, essential for face recognition and augmented reality applications. Various techniques have been developed to enhance pose estimation's precision and robustness. The integration of the Perspective-n-Point (PnP) algorithm within the YOLOMT framework exemplifies advancements, facilitating accurate head pose estimation through detected landmarks [93]. This highlights the significance of combining landmark detection with geometric algorithms for precise pose estimation.

The Face Shape-Guided Deep Feature Alignment (FS-GDFA) framework enhances face recognition performance under misalignment conditions by integrating pixel and feature alignment processes guided by face shape priors [94]. Deep learning techniques, particularly convolutional neural networks (CNNs), have significantly improved pose estimation accuracy and efficiency. Recent advancements in 3D face pose estimation show that CNN-based models can effectively regress 6DoF poses without preliminary face detection, simplifying the estimation process. Comparative analyses of pre-trained deep learning models for facial landmark localization demonstrate their effectiveness

AI-generated, for reference only.

in challenging scenarios, such as occluded faces, enhancing pose estimation precision and enabling real-time processing [95, 54].

The integration of 3D models into pose estimation further improves accuracy by considering spatial relationships and geometric constraints not captured in 2D approaches. Ongoing advancements in pose estimation techniques are driving innovation in facial analysis, ensuring precise head orientation determination across various applications. Recent developments include a real-time multi-task learning system that integrates face detection, landmark detection, and head pose estimation, effectively addressing challenges posed by extreme head postures using the YOLOv8 framework. This system demonstrates improved performance on benchmark datasets, even under complex conditions like facial occlusion. Comprehensive surveys of facial landmark detection algorithms emphasize the importance of accurate fiducial point localization and the need for future research to combine different methodologies for enhanced real-world performance [93, 37, 9, 1].

5.3 Challenges in Face Alignment and Pose Estimation

Face alignment and pose estimation face several challenges that hinder accuracy and efficiency. Managing occlusions and pose variations significantly impacts landmark predictions. Techniques incorporating boundary information enhance robustness against occlusions and large pose variations, improving alignment accuracy [96]. Current methods often struggle with maintaining spatial structure during coordinate prediction, leading to poor localization performance in complex scenarios [97]. This issue is exacerbated by the inability to handle landmark occlusions and appearance changes at extreme angles, where traditional approaches falter [98]. Pose-invariant methods leveraging deeper feature extraction without handcrafted inputs offer promising solutions by improving landmark detection accuracy across varying poses [99].

The dependency on input signal quality is a notable limitation of methods like Teacher-Student Asynchronous Learning (TSAL), where low-quality pseudo-labels can adversely affect performance [100]. Additionally, landmark labeling ambiguities lead to inconsistent error distributions, hindering face alignment performance [101]. In pose estimation, achieving high accuracy in landmark detection is challenging due to the complexities of facial data. Techniques such as Quantum-Assisted Support Vector Regression (QA-SVR) aim to improve landmark localization precision [102]. The computational cost associated with fusion blocks in methods like LiteHRNet poses challenges for real-time applications, necessitating the development of more efficient architectures to reduce processing time [103].

Existing methods often require re-optimization when face detectors are updated, complicating the alignment process [54]. Furthermore, current benchmarks may not account for all real-world variations affecting face image quality, influencing the generalizability of findings [91]. The fixed model architecture and shared computational path in existing landmark detectors hinder performance in complex scenarios [56]. The longer training times required compared to CNN-based models introduce additional parameters, posing potential drawbacks for some applications [104].

Addressing these challenges necessitates continuous development of innovative methodologies and technologies that effectively manage issues related to occlusions, pose variations, labeling ambiguities, and computational inefficiencies. These advancements are critical for enhancing the precision and versatility of facial analysis systems, particularly in fields such as entertainment, virtual media, and cosmetic surgery, by integrating machine learning techniques to assess and enhance facial attractiveness based on established beauty canons, while also emphasizing the critical role of face alignment in optimizing image quality for applications like face recognition and emotion detection [17, 91].

6 Facial Retouching and Virtual Makeup

Facial retouching and virtual makeup have become central to digital aesthetics, driven by technological progress and the demand for enhanced visual experiences. This section explores the methodologies and innovations underpinning these applications, beginning with real-time tracking techniques that facilitate seamless virtual makeup application. Cutting-edge computer vision technologies are enhancing user interaction, ensuring the fidelity and realism of digital makeup applications.

6.1 Real-Time Tracking for Virtual Makeup

Real-time tracking in virtual makeup applications utilizes advanced computer vision technologies to create dynamic user experiences. A key approach combines SegNet for precise landmark detection with the Kanade-Lucas-Tomasi (KLT) tracker for rapid point tracking, effectively addressing point loss issues [105]. This integration is crucial for accurate facial feature tracking, which is essential for realistic virtual makeup application.

The Facial Attribute Transformer (FAT) enhances this field by facilitating makeup transfer through color and spatial transformations, effectively modifying facial attributes like color and shape to improve realism [106]. Additionally, texture descriptors and deep face representations have been employed to detect retouched images, with evaluations under varying compression scenarios highlighting the significance of image quality in virtual makeup applications [107]. Facial retouching detectors provide retouching labels that guide image restoration, emphasizing the importance of accurate detection in maintaining authenticity and quality during virtual makeup application [108]. These technologies collectively enhance user experiences by ensuring efficiency and effectiveness, addressing challenges through deep learning-based facial landmark detection and innovative makeup transfer techniques [105, 106].

6.2 Facial Aesthetic Quality Enhancement

Digital retouching for enhancing facial aesthetic quality is vital in contemporary computer vision applications, particularly within virtual makeup and digital media. A systematic approach to analyzing and enhancing the aesthetic quality of 3D facial models allows for precise modifications that improve overall appearance [17]. Advanced techniques ensure that digital retouching yields natural and visually appealing results.

One innovative method involves style transfer, which enhances morphed images by applying artistic styles that are challenging for existing morphing attack detection systems to identify [80]. This technique not only improves aesthetic quality but also introduces complexity that enriches the visual appeal of morphed faces. A critical challenge in digital retouching is the sensitivity of texture descriptors to pixel variations from image compression, which can lead to misleading detection scores [107]. Ensuring robustness against such variations is essential for preserving aesthetic quality. Furthermore, retouching labels guide the restoration process, correcting imperfections while maintaining authenticity and aesthetic appeal [108]. By leveraging these labels, digital retouching can achieve targeted enhancements, leading to superior aesthetic outcomes.

The ongoing advancement in digital retouching techniques is crucial for applications in virtual makeup and the beauty industry, addressing the demand for realistic facial representations while incorporating established beauty canons such as symmetry and proportions [13, 106, 17]. These innovations ensure that retouched images are visually appealing and realistic, meeting the need for high-quality digital content.

6.3 Facial Attribute Transformer (FAT) for Makeup Applications

The Facial Attribute Transformer (FAT) is a pivotal innovation in virtual makeup applications, providing a sophisticated mechanism for transforming facial attributes while preserving individual identity [106]. This technology effectively maintains the authenticity of facial features during virtual makeup application, ensuring that unique characteristics are retained even as aesthetic modifications are applied.

FAT operates by estimating and transferring facial attributes, enabling seamless integration of makeup effects such as color adjustments and texture modifications. This capability is essential for developing realistic and personalized virtual makeup experiences, incorporating advanced techniques that ensure high-fidelity color transfer and geometric transformations [106, 17]. The integration of color fidelity and spatial transformations within the FAT framework significantly enhances its effectiveness in makeup transfer applications, allowing precise morphing of makeup attributes while accommodating variations in facial geometry and lighting conditions [13, 106, 109, 17, 90]. By enabling detailed adjustments to facial attributes, FAT produces outcomes that are visually appealing and highly personalized, making it an invaluable tool for digital makeup artists and developers seeking to create immersive virtual makeup solutions.

The FAT plays a crucial role in enhancing makeup transfer applications by accurately modeling the relationships between reference and source faces, facilitating high-fidelity transformations that preserve identity while addressing challenges in the digital cosmetics field [13, 106, 17, 30, 56]. This robust framework for attribute transformation enables the creation of high-quality virtual makeup experiences that cater to diverse user preferences.

6.4 Restoration of Retouched Facial Images

Restoration of retouched facial images is vital for maintaining authenticity in digital media applications. This process involves reversing alterations made during digital retouching to recover the original appearance of facial images. An effective approach utilizes facial retouching detectors to obtain labels that guide the restoration process, ensuring that restored images retain their natural features [108].

Texture descriptors are significant for detecting retouched images; however, challenges arise due to their sensitivity to pixel variations from compression [107]. Addressing these challenges is essential for achieving high-quality restoration outcomes, ensuring that restored images are free from artifacts introduced during retouching and compression. Advancements in machine learning and computer vision have led to sophisticated models capable of identifying and reversing specific retouching effects. These models leverage deep learning frameworks to analyze and reconstruct original facial features, enhancing the accuracy of biometric applications such as facial recognition and virtual makeup [13, 88, 30]. By utilizing techniques like generative adversarial networks (GANs) and 3D morphable models, restoration methods can improve the realism and quality of reconstructed images.

The restoration of retouched facial images is essential for the credibility of digital content. By employing cutting-edge detection and reconstruction techniques, this process ensures that facial images maintain authenticity and integrity, enhancing their applicability in digital media production, security systems, and identity verification protocols. This approach addresses challenges posed by image compression and manipulation detection, ensuring reliable performance across various scenarios [13, 107, 1, 75, 82].

6.5 Integration of Detection and Motion Retargeting

Integrating detection and motion retargeting into facial retouching applications enhances the realism and accuracy of virtual makeup experiences. This integration is particularly relevant for applications requiring real-time processing and precise shape representation [105]. Advanced detection techniques enable accurate tracking of facial features, facilitating dynamic adjustments that align with user movements.

Motion retargeting ensures that virtual makeup maintains a consistent and natural appearance, adapting to the user's facial expressions. This technique employs algorithms to capture and transfer facial attributes, enabling seamless integration of makeup effects that respond to changes in facial geometry and expressions [106, 110, 59, 85]. This capability is vital for authentic virtual makeup applications, allowing cosmetic effects to continuously adapt to real-time facial dynamics.

The integration of detection and motion retargeting marks a significant advancement in virtual cosmetics, enhancing the authenticity of retouched images while addressing challenges related to identity verification and misinformation on social media platforms. By employing a unified framework that combines facial retouching detection, image restoration, and motion retargeting, these applications accurately preserve facial identity while managing color fidelity and geometric transformations [106, 75, 59, 108]. Leveraging these technologies allows developers to create sophisticated applications that offer high-quality, realistic virtual makeup experiences tailored to diverse user preferences.

7 Facial Recognition and Security Applications

7.1 Role of Facial Recognition in Security

Facial recognition systems (FRS) are integral to enhancing security in both commercial and governmental contexts, owing to their precision and efficiency in identifying individuals [77]. However, these systems are vulnerable to morphing attacks, which produce synthetic images that resemble

AI-generated, for reference only.

multiple subjects, compromising biometric verification processes [80]. This vulnerability necessitates robust detection mechanisms to protect system integrity.

The susceptibility of FRS to morphing attacks involving lookalikes and identical twins is notably higher than to standard techniques, highlighting the need for improved detection strategies [44]. The TetraLoss method exemplifies efforts to enhance system robustness, significantly improving recognition accuracy under high-security thresholds [111]. This innovation underscores ongoing advancements in fortifying facial recognition against sophisticated attacks.

In video surveillance, FRS effectiveness is a critical focus. Comparative analyses of models under uniform conditions provide insights into performance and vulnerabilities, guiding optimization in security applications [112]. Discrepancies between synthesized and detected highlights have proven effective in distinguishing genuine images from morphing attacks, bolstering biometric security [113].

Benchmarks assessing the impact of morphing techniques on face image quality and exploring unsupervised detection methods are crucial for understanding system limitations and fostering the development of more resilient technologies [78]. Integrating advanced detection mechanisms and evaluation frameworks is vital for enhancing FRS reliability, particularly in countering morphing threats and ensuring accurate identification in varied conditions. Innovations like the TetraLoss function demonstrate significant progress in differentiating morphed images from genuine subjects while maintaining high verification performance. Comparative analyses tailored for diverse scenarios help identify effective approaches to mitigate risks and improve overall system efficacy [13, 112, 111]. Addressing morphing attack vulnerabilities and optimizing performance will ensure these technologies continue to provide reliable identification solutions in an increasingly digital landscape.

7.2 Challenges in Detecting Morphing Attacks

Detecting morphing attacks in facial recognition systems presents numerous challenges due to sophisticated morphing techniques and existing detection methods' limitations. A significant issue is the loss of fine details and digital artifacts during printing and scanning, crucial for identifying morphed images, which hampers current detection methods' effectiveness [82].

The generalization capability of detection methods is another concern, as they often struggle across different datasets and morphing techniques, leading to high error rates in real-world applications [114]. Existing benchmarks frequently fail to differentiate genuine from morphed images effectively, exposing vulnerabilities in security applications [115].

Ethnic bias in Morphing Attack Detection (MAD) techniques is also pressing, as current benchmarks inadequately address this concern, limiting MAD performance understanding across diverse populations and potentially undermining detection systems' fairness and reliability [116].

Additionally, the blending process in morphing attacks merges facial features from multiple images, creating synthetic images that deceive recognition systems. Existing detection systems' inability to incorporate information about fused identities during training further restricts their effectiveness [117]. Creating visually compelling and effective morphed images that evade detection poses a significant threat to security applications [118].

Developing advanced detection methods using depth information and explainable AI is essential for enhancing transparency and reliability. Expanding publicly available datasets to encompass diverse morphing techniques and addressing ethnic biases in benchmarks will facilitate comprehensive evaluations and drive robust detection method development. Addressing these issues can bolster facial recognition systems' resilience against morphing attacks, ensuring their effectiveness in security applications [119].

7.3 Impact of Technological Advancements

Recent technological advancements have profoundly impacted facial recognition systems, enhancing accuracy and robustness while introducing new challenges. Notably, face embeddings, particularly from models like MagFace, have proven effective in pre-selecting morphing pairs and improving morphing attack detection [79]. This dual functionality underscores embeddings' potential to bolster facial recognition systems against sophisticated attack vectors.

Advanced morphing attack detection methods have contributed to reducing Detection Equal Error Rates (D-EER), as recent studies outperform baseline methods [120]. These methods leverage innovative algorithms and improved data processing techniques to enhance detection capabilities, ensuring greater reliability in security applications.

Experiments indicate specific alignment settings significantly influence face morphing detection effectiveness, emphasizing precise alignment in enhancing recognition accuracy [119]. This finding highlights the need for continuous refinement of alignment techniques to maintain high detection performance.

Recent advancements in facial recognition technology signify a notable evolution, driven by the need to counter increasingly sophisticated morphing attacks and enhance detection accuracy. As morphing techniques become more accessible and realistic, integrating deep learning models for detecting such threats has intensified. Challenges remain regarding model interpretability. Innovations like IDistill offer state-of-the-art performance while providing insights into identity separation in morph samples. Ongoing research into facial landmark detection highlights annotation noise's impact on accuracy and stability, leading to new metrics and solutions that enhance facial recognition capabilities [117, 36]. By integrating advanced embeddings, optimizing alignment processes, and developing innovative detection methods, these systems can better withstand morphing attack challenges and other security threats, ensuring their effectiveness in diverse applications.

8 Challenges and Future Directions

8.1 Current Challenges in Facial Landmark Detection

Facial landmark detection is impeded by challenges such as model dependency on specific poses and environmental factors, which affect detection accuracy [6]. The reliance on distinct models for varying poses necessitates the development of unified models capable of handling diverse pose variations. Additionally, 3D Morphable Models (3DMM) often assume linear relationships between anthropometric measurements and eigenvector weights, which may not always be accurate, calling for more sophisticated models that capture non-linear relationships [72].

Environmental factors like lighting variations and occlusions pose significant challenges, as current methodologies show diminished accuracy under such conditions. Techniques like 3D face modeling and shadow removal have been developed to address these issues, yet their effectiveness in enhancing landmark detection remains under study [13, 109, 36, 90, 87]. Furthermore, the quality of input images significantly impacts detection performance, emphasizing the importance of preprocessing and enhancement techniques.

The reliance on initial landmark locations and human annotations introduces potential noise, affecting the reliability of landmark localization. Handling extreme variations in facial appearances, particularly with occlusions or non-frontal views, presents a significant challenge, as existing methods struggle under uncontrolled conditions and diverse demographics [13, 109, 23, 17, 121].

Challenges also arise from the limited number of landmarks and dataset quality, particularly in applications involving animals, where accurate models depend on diverse, high-quality datasets. Research indicates instability in facial landmarks across domains, including animal affective computing, often due to inconsistent labeling quality, complicating cross-dataset applications [51, 7, 36]. Misalignment issues can lead to lower quality scores in face image assessments, necessitating further exploration.

To address these challenges, continuous innovation and the development of robust methodologies are essential to enhance accuracy and adaptability across conditions and applications. This includes deploying deep learning models on resource-constrained systems, ensuring performance across diverse ethnicities and expressions, and improving dataset representation. Recent advancements, such as knowledge distillation techniques, show promise in creating lightweight yet powerful models capable of accurate real-time detection under varying conditions. Ongoing research emphasizes understanding different algorithmic approaches, including holistic, Constrained Local Model (CLM), and regression-based methods, to leverage their strengths and improve performance [8, 9].

8.2 Enhancing Robustness and Generalization

Enhancing robustness and generalization in facial landmark detection is crucial for effective deployment across diverse scenarios. Integrating facial expression data with head pose information can refine 3D face reconstruction accuracy, improving model performance [89]. Future research could focus on optimizing adaptive attention mechanisms and incorporating temporal information to enhance performance in dynamic settings [18].

Leveraging diverse annotations from multiple datasets enhances model training and generalization, highlighting the potential of varied data sources to improve adaptability [51]. Fortifying model robustness against occlusions and exploring applications beyond facial landmarks could provide valuable insights [43].

Optimizing architectural elements, such as the temperature parameter in models like RePFormer, and investigating additional enhancements could further improve detection accuracy [62]. Future work should also explore uncertainty estimation techniques and enhance model robustness to various occlusion scenarios [122].

Capturing structural relationships between landmarks is a key advantage of certain methods, providing improved robustness under challenging conditions [123]. Future research should focus on enhancing generalization capabilities, particularly in real-world applications, while addressing dataset bias observed in cross-dataset evaluations [39]. Refining alignment techniques and their application to 3D face models can address challenges in extreme face misalignment cases [50].

8.3 Optimization for Real-Time and Mobile Applications

Optimizing facial landmark detection for real-time and mobile applications requires addressing computational efficiency, robustness, and adaptability across diverse environments. Enhancing detection algorithms' robustness against expression variations and occlusions, and exploring emotion-invariant multimodal identification approaches, are key areas for future research [15]. Integrating minimal 3D supervision could enhance landmark detection accuracy, especially in mobile applications with limited computational resources [67].

Developing models that handle occlusions and improve expression recognition accuracy for smaller facial features remains a priority [59]. Research should enhance real-time tracking algorithms, focusing on robustness against challenges such as illumination changes and low-resolution images [3].

Exploring factors affecting face image quality and methods to mitigate misalignment impacts is essential for optimizing mobile detection systems [91]. Joint training across multiple datasets and integrating additional information, such as depth data, could enhance performance in complex scenarios [56].

Enhancing model robustness to challenging conditions and refining the optimization process are crucial for high performance in real-time applications [5]. Further exploration of the Mixture of Expert Layer and its application in real-time tracking could significantly improve detection accuracy and efficiency [57].

Future research could focus on expanding datasets, such as the CatFLW dataset, and improving model performance across different species, highlighting potential cross-domain applications [7]. By addressing these challenges and exploring innovative optimization strategies, researchers can advance detection technologies suitable for real-time and mobile applications.

8.4 Integration of Advanced Techniques and Data Sources

Integrating advanced techniques and diverse data sources is crucial for enhancing facial landmark detection accuracy and robustness. Lightweight models that maintain high accuracy with fast inference speeds are essential for real-time applications [26]. Unsupervised learning techniques and improved GAN architectures offer promising avenues for enhancing face data augmentation and model generalization [30].

Adaptive models that generalize across different environments are vital for addressing challenges posed by occlusions and complex scenarios [1]. Optimizing the discrimination network and exploring hybrid loss functions could enhance landmark predictions in such conditions [4].

Future research should focus on developing methods for better disentangling facial features, enhancing 3D face reconstruction methods' performance and generalizability [85]. Exploring new deep learning architectures and data augmentation techniques is critical for improving detection system robustness.

Further optimizations in quantization techniques could be explored, broadening the scope of facial landmark detection technologies [58]. Enhancing model robustness to occlusions and refining landmark predictions in complex scenarios remain crucial research areas [6].

Utilizing cutting-edge methodologies and broadening data resources can significantly enhance advancements in facial analysis and recognition technologies, ensuring efficacy in applications such as facial anonymization, aesthetic assessment, and data augmentation. Developing facial anonymization algorithms through average face morphing addresses ethical concerns, while machine learning techniques in evaluating attractiveness highlight the intersection of technology and social perceptions. Data augmentation strategies improve deep learning models' robustness in face-related tasks, leading to more reliable outcomes [13, 17, 30].

8.5 Expanding Applications and Domains

The potential for expanding facial landmark detection applications is immense, driven by advancements and demand for robust, versatile solutions. Integrating landmarking methods with face frontalization techniques can enhance pose-invariant 3D face recognition systems, improving accuracy and reliability in diverse scenarios [124].

In security, morphing significantly impacts face image quality, indicating potential for using quality measures as indicators for unsupervised morphing attack detection [78]. Future research should increase dataset sizes, explore morphing in various scenarios, and develop effective detection techniques for morphing attacks [44]. Investigating identity vectors' reconstruction capabilities and quantifying morphing percentages based on vector intensities could enhance recognition systems' robustness [117].

Expanding detection applications to non-human subjects is promising. Few-shot model adaptation to non-human facial images could broaden these technologies' scope, enabling use in veterinary science and wildlife monitoring. Enhancing face alignment accuracy for side views and exploring 3D interaction capabilities will improve user experience in applications like FaceAtlasAR, localizing facial acupuncture points [12].

Expanding datasets and improving detection methods for mobile environments are critical for addressing current benchmarks' limitations [40]. Future research could refine methods to handle low-quality video inputs and integrate additional supervisory signals to enhance detection accuracy, optimizing systems for mobile platforms.

8.6 Addressing Dataset and Annotation Limitations

The effectiveness of facial landmark detection systems is significantly influenced by dataset and annotation quality and diversity. A major issue is the lack of scientific validation across various 3D face acquisition systems, hindering standardized benchmark establishment [23]. This gap underscores the need for comprehensive datasets to support robust detection algorithm development. Table 5 provides a detailed overview of various benchmarks crucial for advancing facial recognition and morphing attack detection, addressing the challenges of dataset and annotation limitations.

New datasets, such as MERL-RAV, with over 19,000 face images labeled for visibility, advance face alignment research [122]. These datasets provide valuable resources for training and evaluating models, particularly in improving landmark visibility under various conditions.

Despite advancements, challenges remain in ensuring detection algorithms' robustness under extreme conditions, such as significant occlusions or unusual lighting [26]. Addressing these challenges requires developing adaptable models capable of maintaining high accuracy across diverse environments.

Benchmark	Size	Domain	Task Format	Metric
MAD-Q[78]	4,000	Face Recognition	Morphing Attack Detection	MagFace, CNNIQA
MAD-Bench[79]	55,134	Face Recognition	Morphing Attack Detection	Morphing Attack Potential, Product Average Mated Morph Presentation Match Rate
FPN[125]	2,600,000	Face Alignment	Face Recognition	Recognition Accuracy
VSB[112]	33,630	Video Surveillance	Face Identification	CMC, Separation Performance
MABench[126]	2,715	Biometrics	Morphing Attack Detection	MMPMR, FMR
SynMorph[127]	141,000	Face Morphing Attack Detection	Morphing Attack Detection	Morphing Attack Potential, Face Image Quality Assessment
face.evoLVe[128]	5,080,000	Face Recognition	Classification	Accuracy, F1-score
3DMM[129]	1,000,000	Facial Expression Analysis	Dimensional Emotion Recognition	CCC, ICC

Table 5: This table presents a comprehensive overview of representative benchmarks utilized in facial recognition and related domains. It details the size, domain, task format, and evaluation metrics of each benchmark, providing insights into their application in morphing attack detection, face alignment, and facial expression analysis.

Reliance on specific datasets can limit detection systems' generalization capabilities, particularly when facing varying morphing techniques. Expanding benchmarks to include a wider variety of morphing techniques and integrating additional quality measures is crucial for enhancing morphing attack detection systems' robustness [78].

Enhancing training data diversity is essential for improving landmark detection systems' reliability, as inconsistent labeling quality causes instability. Incorporating a broader range of expressions, ethnicities, and conditions allows better training for accurate feature location and tracking across varied contexts, increasing robustness and performance [8, 60, 36]. Generating diverse, comprehensive datasets equips detection models to handle complex conditions, advancing facial landmark detection.

8.7 Future Directions and Ethical Considerations

The future of facial landmark detection technologies promises substantial advancements, focusing on enhancing model robustness, accuracy, and applicability across diverse applications. Optimizing hyperparameters and exploring fully quantum annealing methods could significantly enhance detection system effectiveness [102]. Exploring landmark-free methods in facial analysis and improving algorithms for extreme expressions could pave new innovation pathways [86].

Future research should emphasize refining cross-format training strategies and utilizing additional datasets to bolster model robustness and accuracy. Enhancing the ENiqab-V1 dataset and exploring new model architectures for occluded faces are crucial steps [95]. Integrating additional student networks and self-supervised techniques could leverage unused unlabeled samples, enhancing performance [60].

Developing strategies for end-to-end training and improving transform estimation processes are critical for advancing deep alignment networks [130]. Reducing dependence on labeled data and extending model applicability to areas like human pose estimation will broaden facial landmark detection technologies' scope [131].

Enhancing landmark detection algorithms and improving blending techniques to minimize artifacts and preserve identity are vital future exploration areas [118]. Enhancing realism in style modeling and extending frameworks to high-level vision tasks beyond facial landmark detection are promising research directions [2].

Ethical considerations remain paramount in developing facial landmark detection technologies. Face recognition systems' vulnerability to morphing attacks underscores benchmarks' importance as critical resources for understanding and improving system resilience [126]. Addressing algorithm robustness against extreme noise and exploring additional domains' applications are essential for ensuring ethical standards in deploying these technologies [109].

9 Conclusion

This survey has provided a detailed exploration of facial landmark detection, highlighting its transformative impact on computer vision. By integrating diverse methodologies, such as hybrid models, the field has achieved notable improvements in detecting facial keypoints under challenging conditions like heavy makeup and occlusions. The development of advanced frameworks, including those optimized for motion-blurred video, underscores the potential for enhanced robustness and accuracy in varied scenarios.

Innovative techniques, such as the 2D Wasserstein loss, have addressed limitations in existing metrics, improving the robustness of heatmap regression models. Approaches like the Attentive One-Dimensional Heatmap Regression have further advanced accuracy and efficiency, minimizing errors and outperforming existing methods. Memory-efficient strategies, exemplified by the Global Refinement method, have optimized resource usage while maintaining high accuracy in face alignment tasks. The incorporation of human examiner decisions into face manipulation detection has also significantly improved system performance.

The joint multi-view approach has proven effective in enhancing facial landmark localization and tracking, paving the way for more comprehensive detection systems. As these technologies continue to evolve, they are set to drive innovation across computer vision applications, broadening their applicability in diverse domains. The ongoing refinement of these methodologies promises to advance the development of robust, efficient, and versatile systems capable of meeting the demands of modern applications.

References

[1] Nannan Wang, Xinbo Gao, Dacheng Tao, and Xuelong Li. Facial feature point detection: A comprehensive survey, 2014.

[2] Shengju Qian, Keqiang Sun, Wayne Wu, Chen Qian, and Jiaya Jia. Aggregation via separation: Boosting facial landmark detector with semi-supervised style translation, 2019.

[3] Yue Wu and Qiang Ji. Robust facial landmark detection under significant head poses and occlusion, 2017.

[4] Chih-Fan Hsu, Chia-Ching Lin, Ting-Yang Hung, Chin-Laung Lei, and Kuan-Ta Chen. A detailed look at cnn-based approaches in facial landmark detection, 2020.

[5] Zhiwei Liu, Xiangyu Zhu, Guosheng Hu, Haiyun Guo, Ming Tang, Zhen Lei, Neil M. Robertson, and Jinqiao Wang. Semantic alignment: Finding semantically consistent ground-truth for facial landmark detection, 2019.

[6] Jiankang Deng, George Trigeorgis, Yuxiang Zhou, and Stefanos Zafeiriou. Joint multi-view face alignment in the wild, 2017.

[7] George Martvel, Ilan Shimshoni, and Anna Zamansky. Automated detection of cat facial landmarks, 2024.

[8] Zong-Wei Hong and Yu-Chen Lin. Improving facial landmark detection accuracy and efficiency with knowledge distillation, 2024.

[9] Yue Wu and Qiang Ji. Facial landmark detection: a literature survey, 2018.

[10] Yue Wu, Tal Hassner, KangGeon Kim, Gerard Medioni, and Prem Natarajan. Facial landmark detection with tweaked convolutional neural networks, 2016.

[11] Yongzhe Yan, Stefan Duffner, Priyanka Phutane, Anthony Berthelier, Christophe Blanc, Christophe Garcia, and Thierry Chateau. Facial landmark correlation analysis, 2020.

[12] Menghe Zhang, Jurgen Schulze, and Dong Zhang. Faceatlasar: Atlas of facial acupuncture points in augmented reality, 2022.

[13] Joar Midtun. Automatic facial anonymisation using average face morphing. an explorative attempt at designing a tool for investigating the normative views of journalism practices in dealing with manipulation and usage of facial press photographies. Master's thesis, The University of Bergen, 2017.

[14] S L Happy and Aurobinda Routray. Automatic facial expression recognition using features of salient facial patches, 2015.

[15] Sk Rahatul Jannat, Diego Fabiano, Shaun Canavan, and Tempestt Neal. Subject identification across large expression variations using 3d facial landmarks, 2020.

[16] Diego Fabiano, Manikandan Jaishanker, and Shaun Canavan. Impact of multiple modalities on emotion recognition: investigation into 3d facial landmarks, action units, and physiological data, 2020.

[17] Pr EL HASSOUNI Mohammed. 3d facial aesthetic quality: Analysis and enhancement system. 2020.

[18] Zhiwen Shao, Zhilei Liu, Jianfei Cai, and Lizhuang Ma. J$\hat{a}$a-net: Joint facial action unit detection and face alignment via adaptive attention, 2020.

[19] Andreas Steger, Radu Timofte, and Luc Van Gool. Failure detection for facial landmark detectors, 2016.

[20] Yongzhe Yan, Stefan Duffner, Priyanka Phutane, Anthony Berthelier, Christophe Blanc, Christophe Garcia, and Thierry Chateau. 2d wasserstein loss for robust facial landmark detection, 2020.

[21] Prathima Dileep, Bharath Kumar Bolla, and Sabeesh Ethiraj. Revisiting facial key point detection: An efficient approach using deep neural networks, 2022.

[22] Yuhang Wu, Le Anh Vu Ha, Xiang Xu, and Ioannis A. Kakadiaris. Convolutional point-set representation: A convolutional bridge between a densely annotated image and 3d face alignment, 2018.

[23] Pradeep Singh, Michael M Bornstein, Richard Tai-Chiu Hsung, Deepal Haresh Ajmera, Yiu Yan Leung, and Min Gu. Frontiers in three-dimensional surface imaging systems for 3d face acquisition in craniofacial research and practice: an updated literature review. *Diagnostics*, 14(4):423, 2024.

[24] Mohammad Amin Haghpanah, Ehsan Saeedizade, Mehdi Tale Masouleh, and Ahmad Kalhor. Real-time facial expression recognition using facial landmarks and neural networks, 2022.

[25] Xin Jin and Xiaoyang Tan. Face alignment in-the-wild: A survey, 2016.

[26] Kostiantyn Khabarlak and Larysa Koriashkina. Fast facial landmark detection and applications: A survey, 2022.

[27] Yangyu Huang, Xi Chen, Jongyoo Kim, Hao Yang, Chong Li, Jiaolong Yang, and Dong Chen. Freeenricher: Enriching face landmarks without additional cost, 2022.

[28] Christian Rathgeb, Robert Nichols, Mathias Ibsen, Pawel Drozdowski, and Christoph Busch. Crowd-powered face manipulation detection: Fusing human examiner decisions, 2022.

[29] Jagmohan Meher, Hector Allende-Cid, and Torbjörn E. M. Nordling. A survey and classification of face alignment methods based on face models, 2023.

[30] Xiang Wang, Kai Wang, and Shiguo Lian. A survey on face data augmentation. *arXiv preprint arXiv:1904.11685*, 2019.

[31] Lucas Faria Porto, Laise Nascimento Correia Lima, Marta Flores, Andrea Valsecchi, Oscar Ibanez, Carlos Eduardo Machado Palhares, and Flavio de Barros Vidal. Automatic cephalometric landmarks detection on frontal faces: an approach based on supervised learning techniques, 2019.

[32] Xiangyu Zhu, Xiaoming Liu, Zhen Lei, and Stan Z. Li. Face alignment in full pose range: A 3d total solution, 2018.

[33] Heng Yang, Xuhui Jia, Chen Change Loy, and Peter Robinson. An empirical study of recent face alignment methods, 2015.

[34] Adrian Bulat and Georgios Tzimiropoulos. How far are we from solving the 2d 3d face alignment problem? (and a dataset of 230,000 3d facial landmarks), 2017.

[35] Ethan Grooby, Chiranjibi Sitaula, Soodeh Ahani, Liisa Holsti, Atul Malhotra, Guy A. Dumont, and Faezeh Marzbanrad. Neonatal face and facial landmark detection from video recordings, 2023.

[36] Xiaoyu Xiang, Yang Cheng, Shaoyuan Xu, Qian Lin, and Jan Allebach. The blessing and the curse of the noise behind facial landmark annotations, 2020.

[37] Yue Wu, Chao Gou, and Qiang Ji. Simultaneous facial landmark detection, pose and deformation estimation under facial occlusion, 2017.

[38] Vishnu Naresh Boddeti, Myung-Cheol Roh, Jongju Shin, Takaharu Oguri, and Takeo Kanade. Face alignment robust to pose, expressions and occlusions, 2017.

[39] Roberto Valle, José M. Buenaposada, Antonio Valdés, and Luis Baumela. Face alignment using a 3d deeply-initialized ensemble of regression trees, 2019.

[40] Na Zhang. Facial landmark detection evaluation on mobio database, 2023.

[41] Chandrasekhar Bhagavatula, Chenchen Zhu, Khoa Luu, and Marios Savvides. Faster than real-time facial alignment: A 3d spatial transformer network approach in unconstrained poses. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3980–3989, 2017.

[42] Wayne Wu, Chen Qian, Shuo Yang, Quan Wang, Yici Cai, and Qiang Zhou. Look at boundary: A boundary-aware face alignment algorithm, 2018.

[43] Marcin Kopaczka, Justus Schock, and Dorit Merhof. Super-realtime facial landmark detection and shape fitting by deep regression of shape model parameters, 2019.

[44] Raghavendra Ramachandra, Sushma Venkatesh, Gaurav Jaswal, and Guoqiang Li. Vulnerability of face morphing attacks: A case study on lookalike and identical twins, 2023.

[45] Xu Sun, Zhenfeng Fan, Zihao Zhang, Yingjie Guo, and Shihong Xia. A backbone replaceable fine-tuning framework for stable face alignment, 2020.

[46] Brandon M. Smith and Charles R. Dyer. Efficient branching cascaded regression for face alignment under significant head rotation, 2016.

[47] Yuting Zhang, Kui Jia, Yueming Wang, Gang Pan, Tsung-Han Chan, and Yi Ma. Robust face recognition by constrained part-based alignment, 2015.

[48] Jihua Huang and Amir Tamrakar. Ace-net: Fine-level face alignment through anchors and contours estimation, 2022.

[49] Huabin Wang, Rui Cheng, Jian Zhou, Liang Tao, and Hon Keung Kwan. Multistage model for robust face alignment using deep neural networks, 2020.

[50] Huawei Wei, Peng Lu, and Yichen Wei. Balanced alignment for face recognition: A joint learning approach, 2020.

[51] Shizhan Zhu, Cheng Li, Chen Change Loy, and Xiaoou Tang. Transferring landmark annotations for cross-dataset face alignment, 2014.

[52] Feng Liu, Qijun Zhao, Xiaoming Liu, and Dan Zeng. Joint face alignment and 3d face reconstruction with application to face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 42(3):664–678, 2018.

[53] Amin Jourabloo and Xiaoming Liu. Pose-invariant 3d face alignment, 2015.

[54] Vítor Albiero, Xingyu Chen, Xi Yin, Guan Pang, and Tal Hassner. img2pose: Face alignment and detection via 6dof, face pose estimation, 2021.

[55] Jun Wan, Jun Liu, Jie Zhou, Zhihui Lai, Linlin Shen, Hang Sun, Ping Xiong, and Wenwen Min. Precise facial landmark detection by reference heatmap transformer, 2023.

[56] Jun Wan, He Liu, Yujia Wu, Zhihui Lai, Wenwen Min, and Jun Liu. Precise facial landmark detection by dynamic semantic aggregation transformer, 2024.

[57] Amir Zadeh, Tadas Baltrušaitis, and Louis-Philippe Morency. Convolutional experts constrained local model for facial landmark detection, 2017.

[58] Nenad Markuš, Ivan Gogić, Igor S. Pandžić, and Jörgen Ahlberg. Memory-efficient global refinement of decision-tree ensembles and its application to face alignment, 2018.

[59] Bindita Chaudhuri, Noranart Vesdapunt, and Baoyuan Wang. Joint face detection and facial motion retargeting for multiple faces, 2019.

[60] Xuanyi Dong and Yi Yang. Teacher supervises students how to learn from partially labeled images for facial landmark detection, 2019.

[61] Xuanyi Dong, Yan Yan, Wanli Ouyang, and Yi Yang. Style aggregated network for facial landmark detection, 2018.

[62] Jinpeng Li, Haibo Jin, Shengcai Liao, Ling Shao, and Pheng-Ann Heng. Repformer: Refinement pyramid transformer for robust facial landmark detection, 2022.

[63] Hongwen Zhang, Qi Li, Zhenan Sun, and Yunfan Liu. Combining data-driven and model-driven methods for robust facial landmark detection, 2018.

[64] Prashanth Chandran, Gaspard Zoss, Paulo Gotardo, and Derek Bradley. Infinite 3d landmarks: Improving continuous 2d facial landmark detection, 2024.

[65] Luan Tran and Xiaoming Liu. Nonlinear 3d face morphable model. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7346–7355, 2018.

[66] Hao Zhang, Tianyuan Dai, Yu-Wing Tai, and Chi-Keung Tang. Flnerf: 3d facial landmarks estimation in neural radiance fields, 2023.

[67] Soubhik Sanyal, Timo Bolkart, Haiwen Feng, and Michael J Black. Learning to regress 3d face shape and expression from an image without 3d supervision. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 7763–7772, 2019.

[68] Jialiang Zhang, Lixiang Lin, Jianke Zhu, and Steven C. H. Hoi. Weakly-supervised multi-face 3d reconstruction, 2021.

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

[70] Ilja Gubins. Automated generation of animated 3d facial meshes: A photogrammetry and deformation transfer-based model. Master's thesis, 2017.

[71] Luan Tran and Xiaoming Liu. On learning 3d face morphable model from in-the-wild images, 2019.

[72] Donya Ghafourzadeh, Sahel Fallahdoust, Cyrus Rahgoshay, Andre Beauchamp, Adeline Aubame, Tiberiu Popa, and Eric Paquette. Local control editing paradigms for part-based 3d face morphable models. *Computer Animation and Virtual Worlds*, 32(6):e2028, 2021.

[73] Erroll Wood, Tadas Baltrušaitis, Charlie Hewitt, Matthew Johnson, Jingjing Shen, Nikola Milosavljević, Daniel Wilde, Stephan Garbin, Toby Sharp, Ivan Stojiljković, et al. 3d face reconstruction with dense landmarks. In *European Conference on Computer Vision*, pages 160–177. Springer, 2022.

[74] Cho-Ying Wu, Qiangeng Xu, and Ulrich Neumann. Synergy between 3dmm and 3d landmarks for accurate 3d facial geometry, 2024.

[75] Michael Zollhöfer, Justus Thies, Pablo Garrido, Derek Bradley, Thabo Beeler, Patrick Pérez, Marc Stamminger, Matthias Nießner, and Christian Theobalt. State of the art on monocular 3d face reconstruction, tracking, and applications. In *Computer graphics forum*, volume 37, pages 523–550. Wiley Online Library, 2018.

[76] Feng Liu, Qijun Zhao, Xiaoming Liu, and Dan Zeng. Joint face alignment and 3d face reconstruction with application to face recognition, 2018.

[77] Jag Mohan Singh and Raghavendra Ramachandra. 3d face morphing attack generation using non-rigid registration, 2024.

[78] Biying Fu and Naser Damer. Face morphing attacks and face image quality: The effect of morphing and the unsupervised attack detection by quality, 2022.

[79] Roman Kessler, Kiran Raja, Juan Tapia, and Christoph Busch. Towards minimizing efforts for morphing attacks – deep embeddings for morphing pair selection and improved morphing attack detection, 2024.

[80] Clemens Seibold, Anna Hilsmann, and Peter Eisert. Style your face morph and improve your face morphing attack detector, 2020.

[81] Muhammad Hamza, Samabia Tehsin, Mamoona Humayun, Maram Fahaad Almufareh, and Majed Alfayad. A comprehensive review of face morph generation and detection of fraudulent identities. *Applied Sciences*, 12(24):12545, 2022.

[82] Matteo Ferrara, Annalisa Franco, and Davide Maltoni. Face morphing detection in the presence of printing/scanning and heterogeneous image sources, 2021.

[83] Luan Tran and Xiaoming Liu. Nonlinear 3d face morphable model, 2018.

[84] Yuxiang Wu, Zehua Cheng, Bin Huang, Yiming Chen, Xinghui Zhu, and Weiyang Wang. Foxnet: A multi-face alignment method, 2019.

[85] Zihao Jian and Minshan Xie. Realistic face animation generation from videos, 2021.

[86] Feng-Ju Chang, Anh Tuan Tran, Tal Hassner, Iacopo Masi, Ram Nevatia, and Gerard Medioni. Expnet: Landmark-free, deep, 3d facial expressions, 2018.

[87] Lan Fu, Qing Guo, Felix Juefei-Xu, Hongkai Yu, Wei Feng, Yang Liu, and Song Wang. Benchmarking shadow removal for facial landmark detection and beyond, 2021.

[88] Sahil Sharma and Vijay Kumar. 3d face reconstruction in deep learning era: A survey. *Archives of Computational Methods in Engineering*, 29(5):3475–3507, 2022.

[89] Jaehyun So and Youngjoon Han. 3d face alignment through fusion of head pose information and features, 2023.

[90] Qiulin Chen and Jan P. Allebach. Face image lighting enhancement using a 3d model, 2022.

[91] Eren Onaran, Erdi Sarıtaş, and Hazım Kemal Ekenel. Impact of face alignment on face image quality, 2024.

[92] Oncel Tuzel, Tim K. Marks, and Salil Tambe. Robust face alignment using a mixture of invariant experts, 2016.

[93] Qingtian Wu and Liming Zhang. A real-time multi-task learning system for joint detection of face, facial landmark and head pose, 2023.

[94] Hyung-Il Kim, Kimin Yun, and Yong Man Ro. Face shape-guided deep feature alignment for face recognition robust to face misalignment, 2022.

[95] Zieb Rabie Alqahtani, Mohd Shahrizal Sunar, and Abdelmonim M Artoli. Comparative analysis of pre-trained deep learning models for facial landmark localization on enhanced dataset of heavily occluded face images. *Journal of Advances in Information Technology*, 15(11), 2024.

[96] Jinheng Xie, Jun Wan, Linlin Shen, and Zhihui Lai. Think about boundary: Fusing multi-level boundary information for landmark heatmap regression, 2020.

[97] Hui Li, Zidong Guo, Seon-Min Rhee, Seungju Han, and Jae-Joon Han. Towards accurate facial landmark detection via cascaded transformers, 2022.

[98] Xiangyu Zhu, Zhen Lei, Xiaoming Liu, Hailin Shi, and Stan Z. Li. Face alignment across large poses: A 3d solution, 2015.

[99] Amin Jourabloo, Mao Ye, Xiaoming Liu, and Liu Ren. Pose-invariant face alignment with a single cnn, 2017.

[100] Rongye Meng, Sanping Zhou, Xingyu Wan, Mengliu Li, and Jinjun Wang. Teacher-student asynchronous learning with multi-source consistency for facial landmark detection, 2020.

[101] Yangyu Huang, Hao Yang, Chong Li, Jongyoo Kim, and Fangyun Wei. Adnet: Leveraging error-bias towards normal direction in face alignment, 2022.

[102] Archismita Dalal, Mohsen Bagherimehrab, and Barry C. Sanders. Quantum-assisted support vector regression for detecting facial landmarks, 2021.

[103] Sota Kato, Kazuhiro Hotta, Yuhki Hatakeyama, and Yoshinori Konishi. Lite-hrnet plus: Fast and accurate facial landmark detection, 2023.

[104] Haibo Jin, Jinpeng Li, Shengcai Liao, and Ling Shao. When liebig's barrel meets facial landmark detection: A practical model, 2021.

[105] Hyungjoon Kim, Hyeonwoo Kim, and Eenjun Hwang. Real-time shape tracking of facial landmarks, 2018.

[106] Zhaoyi Wan, Haoran Chen, Jie An, Wentao Jiang, Cong Yao, and Jiebo Luo. Facial attribute transformers for precise and robust makeup transfer. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 1717–1726, 2022.

[107] Christian Rathgeb, Kevin Bernardo, Nathania E. Haryanto, and Christoph Busch. Effects of image compression on face image manipulation detection: A case study on facial retouching, 2021.

[108] Guanhua Zhao, Yu Gu, Xuhan Sheng, Yujie Hu, and Jian Zhang. Face2face: Label-driven facial retouching restoration, 2024.

[109] Stylianos Moschoglou, Evangelos Ververas, Yannis Panagakis, Mihalis Nicolaou, and Stefanos Zafeiriou. Multi-attribute robust component analysis for facial uv maps, 2017.

[110] Wei Wei, Edmond SL Ho, Kevin D McCay, Robertas Damaševičius, Rytis Maskeliūnas, and Anna Esposito. Assessing facial symmetry and attractiveness using augmented reality. *Pattern Analysis and Applications*, pages 1–17, 2022.

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

[112] Eker Onur and Bal Murat. A comparative analysis of the face recognition methods in video surveillance scenarios, 2022.

[113] Clemens Seibold, Anna Hilsmann, and Peter Eisert. Reflection analysis for face morphing attack detection, 2018.

[114] Juan Tapia and Christoph Busch. Face feature visualisation of single morphing attack detection, 2023.

[115] Iurii Medvedev, Joana Pimenta, and Nuno Gonçalves. Fused classification for differential face morphing detection, 2023.

[116] Raghavendra Ramachandra, Kiran Raja, and Christoph Busch. Algorithmic fairness in face morphing attack detection, 2021.

[117] 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.

[118] Qiaoyun He, Zongyong Deng, Zuyuan He, and Qijun Zhao. Optimal-landmark-guided image blending for face morphing attacks, 2024.

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

[120] Juan Tapia and Christoph Busch. Single morphing attack detection using feature selection and visualisation based on mutual information, 2021.

[121] Azin Asgarian, Shun Zhao, Ahmed B. Ashraf, M. Erin Browne, Kenneth M. Prkachin, Alex Mihailidis, Thomas Hadjistavropoulos, and Babak Taati. Limitations and biases in facial landmark detection – an empirical study on older adults with dementia, 2019.

[122] Abhinav Kumar, Tim K. Marks, Wenxuan Mou, Ye Wang, Michael Jones, Anoop Cherian, Toshiaki Koike-Akino, Xiaoming Liu, and Chen Feng. Luvli face alignment: Estimating landmarks' location, uncertainty, and visibility likelihood, 2020.

[123] Lisha Chen, Hui Su, and Qiang Ji. Deep structured prediction for facial landmark detection, 2020.

[124] Janez Križaj, Peter Peer, Vitomir Štruc, and Simon Dobrišek. Simultaneous regression and feature learning for facial landmarking, 2019.

[125] Fengju Chang, Anh Tuan Tran, Tal Hassner, Iacopo Masi, Ram Nevatia, and Gerard Medioni. Faceposenet: Making a case for landmark-free face alignment, 2017.

[126] Eklavya Sarkar, Pavel Korshunov, Laurent Colbois, and Sébastien Marcel. Vulnerability analysis of face morphing attacks from landmarks and generative adversarial networks, 2020.

[127] Haoyu Zhang, Raghavendra Ramachandra, Kiran Raja, and Christoph Busch. Synmorph: Generating synthetic face morphing dataset with mated samples, 2024.

[128] Qingzhong Wang, Pengfei Zhang, Haoyi Xiong, and Jian Zhao. Face.evolve: A high-performance face recognition library, 2022.

[129] Mani Kumar Tellamekala, Ömer Sümer, Björn W. Schuller, Elisabeth André, Timo Giesbrecht, and Michel Valstar. Are 3d face shapes expressive enough for recognising continuous emotions and action unit intensities?, 2023.

[130] Marek Kowalski, Jacek Naruniec, and Tomasz Trzcinski. Deep alignment network: A convolutional neural network for robust face alignment, 2017.

[131] Jun Wan, Hui Xi, Jie Zhou, Zhihui Lai, Witold Pedrycz, Xu Wang, and Hang Sun. Robust and precise facial landmark detection by self-calibrated pose attention network, 2021.

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.