UNIHEZ JOURNAL OF HEALTH, SCIENCES, ENGINEERING, INFORMATION AND COMMUNICATION TECHNOLOGY
VOL. 7 NO. 1 JAN 2026

FACE ALIGNMENT AND NORMALIZATION: FOUNDATIONS, TECHNIQUES AND ADVANCES FOR ROBUST FACE RECOGNITION

AWODELE S. O
Department Of Computer Science,
Babcock University Ilshan-Remo, Ogun State, Nigeria
Awodeles@Babcock.Edu.Ng

CHUKWULOBE I
Department Of Computer Science,
Babcock University Ilshan-Remo, Ogun State, Nigeria
Chukwulobe0408@Pg.Babcock.Edu.Ng

FARUNA J. O
Department Of Computer Science,
Babcock University Ilshan-Remo, Ogun State, Nigeria
Faruna0100@Pg.Babcock.Edu.Ng

FAYEMI T. A
Department Of Computer Science,
Babcock University Ilshan-Remo, Ogun State, Nigeria
Fayemi0197@Pg.Babcock.Edu.Ng

OJUAWO O. O
Department Of Computer Science,
Babcock University Ilshan-Remo, gun State, Nigeria
Ojuawo0687@Pg.Babcock.Edu.Ng

MUSTAPHA M. M
Department Of Computer Science,
Babcock University Ilshan-Remo, Ogun State, Nigeria
Mustapha0219@Pg.Babcock.Edu.Ng

&

OLORUNYOMI O. B
Department Of Computer Science,
Babcock University Ilshan-Remo, Ogun State, Nigeria
Olorunyomi0052@Pg.Babcock.Edu.Ng

Article history:
Received: Jan 2026;
Received in revised form:
27 Jan 2026;
Accepted: 25 Feb 2026;

Keywords:

Abstract

Face alignment and normalization are the two irreplaceable foundations of the modern face recognition systems that cannot be assembled without prerequisite work. The purpose of these underlying processes is to canonicize the geometric arrangement as well as photographic look of facial images and to change extensively diverse, in-the-wild captures into a standard

form that is amenable to dependable feature extraction and matching. This is the extensive paper that involves the in-depth analysis of the history, modern trends, and perspectives in the development of face alignment and normalization techniques. We begin with a clarification of their significant role in mitigating the adverse effect of the essential nuisance factors, such as pose, illumination, expression, and occlusion. The paper continues by giving a systematic taxonomy and a detailed discussion of fundamental methods, including the classical model-based methods such as Active Shape Models, the period of discriminative cascaded regression, and the modern model of deep learning-based holistic alignment and feature-level normalization networks. Much importance is taken on advanced and new issues: 3D-aware alignment to extreme pose correction, adversarial and disentangled representation learning to achieve photometric invariance, and the role of equitable normalization to reduce demographic bias in order to promote algorithmic fairness. More so, we consider how these modules can be strategically integrated into effective end-to-end recognition pipelines and optimized to execute on resource-constrained edge devices. This review summarizes the insights and findings of more than fifty publications that are regarded as seminal with an aim of formulating how advanced alignment and normalization have become more central than peripheral, of facilitating technologies. The development has played a crucial role in the transformation of face recognition as a limited laboratory use to a powerful and efficient biometric technology that can be able to function in the dynamic and challenging real world environment..

Introduction

The consistency, quality and canonical nature of the input data are intrinsically and fundamentally connected to the performance and reliability of automated face recognition (FR) systems [1]. Poses are frontal, and faces in full view in controlled situations, where the lighting is even, high accuracy has consistently been attained. Nevertheless, a range of unconstrained scenarios is bound to be faced by the real-world implementation of FR technology. In this case, faces are represented by enormous and stochastic geometric variations- non frontal pose, non rigid expression changes, and enormous photometric variations due to the variations in the direction, intensity, and color temperature of the illumination and sensor-specific responses [2]. These aspects bring in big intra-class variances, which in a short time can outsmart inter-class variances between distinct individuals, in effect, drastically reducing the discriminatory ability of variables of extracted facial features and resulting in system failures. Face Alignment and Normalization have in turn come out as the central front-end computation operations and activities that are specifically meant to offset these issues. Alignment is largely understood to refer to the mechanical procedure of mapping a identified face onto a pre-defined frame a spatial rotation and a spatial scale of what initially existed in an arbitrary position and scale. It is normally done by localizing a set of fiducial facial points and then using a spatial transformation, which most of the times is projective or non-rigid, relative to these fiducial points [3]. With a wider and more abstract focus, normalization includes both geometric fixes and a set of photometric fixes designed to standardize appearance properties. This involves correcting uneven lighting, removing color casts and at times even flattening textures, to create a more consistent and identity oriented imagery [4]. The history of the development of these methods is closely connected with the development of FR in general. The early systems had little pre-processing and mostly relied on manual cropping or were severely restricted to almost frontal passport-like images. An important progress was achieved with the invention of automated statistical model-based landmark detectors, including Active Shape Models (ASMs) and more expressive versions, Active Appearance Models (AAMs) [5]. These offered a conceptual means of modeling shape and texture change. The next step to discriminative cascaded regression models sharply enhanced speed and accuracy. Nevertheless, the most disruptive innovation was the one related to the emergence of deep learning. The Convolutional Neural Networks (CNNs) have made it possible to achieve impressive improvement in both accuracy and robustness in the alignment, which can now be localized to the landmarks with accuracy in a case of severe occlusion, extreme angles, and adverse lighting conditions [6], [7]. At the same time, the conceptual definition of what is meant by normalization has grown beyond the basic image space filters such as histogram equalization.

The contemporary perspective is furthermore, more concerned with the learning of non-linear, non-linear transformations, which work in the deep feature space and proactively, decouple the fundamental identity representation with confounding nuisance factors [8], [9]. This paper is an attempt to review these classic and innovative methods in detail. Our argument is that the impressive working of modern robust FR systems is not because of strong deep feature extractors such ArcFace [10] or loss functions. Instead, it is also, and obviously, heavily reliant on the advanced alignment and normalization pipelines which feed these networks with canonicalized input. This review will discuss the direct relevance of these pre-processing steps to the fundamental problems of FR, how their methodological development has evolved through classical to deep learning models, as well as the importance of these steps to dealing with new priorities, including algorithmic fairness and efficient edge computing. In this manner, we will highlight their position, not as initial steps, but as pillars in the framework of reliable and competent face recognition.

Statement of the Problem

The large intra-class variation is the main barrier to the attainment of reliable face recognition in unconstrained settings. This is the drastic alteration of face of one single person under the influence of various factors that do not relate to his/her identity. In cases where these non-identity variations are strong they may surpass the appearance variation between two distinct individuals that exists naturally (inter-class variation) and cause false non-matches or false matches. The methods of alignment and normalization are principally intended to strike the greatest of these sources of this nuisance variation: Complex Geometric Deformation: Non-frontal pose will cause one section of the face to be hidden (self-occlusion) and non-linear warping of texture which cannot be corrected by simple 2D translation or scaling.

In the same way, local, non-rigid deformations (such as stretching around the mouth when smiling) are caused by the facial expression and produce changes in the spatial relation of the facial features. It takes complex 2D non-rigid warping or 3D shape model and alignments [11], [12] to do so. Photometric Variance: The changes in the direction (shadows and highlights), intensity (over exposure or dark) and spectral composition (warm vs cool light) of a light source cause changes in pixel intensities drastically. Such alterations can add edges and gradients that are artifacts of illumination as opposed to the facial structure and therefore hide the real features of identity. The effective solutions should not stop at the global brightness/contrast options but should address the problem of intricate, variable lighting effects across the face [13], [14]. The High-Precision Registration Requirement: Face detector output (a bounding box based on maximizing the detection recall) is usually not optimal with regards to recognition (maximizing feature discriminability). The deep neural networks are also highly sensitive to the accurate spatial position of features [15], [16]. Therefore, the smallest deviations of the confining box, which might not be noticeable by humans, can cause massive performance loss. This gap between detection and recognition goals leads to a very serious requirement of high-precision registration as one of the fundamental functions of the alignment stage. Propagation and Amplification of Bias: It is possible to introduce or propagate demographic biases through imperfect or non-robust alignment and normalization in a systematic and inadvertent way. As an example, when a landmark detector is trained on a non-diverse dataset, it might have inferior localization on morphologies of the face popular in a particular ethnic group. Such a worse fit causes poorer normalization, and thus worse feature extraction and worse error rates on those demographics- transforming a technical constraint into a systemic fairness problem [17], [18].

Thus, the main research and engineering challenge is to develop alignment and normalization algorithms that are simultaneously: a) very accurate and precise in the entire variety of geometric and photometric conditions; b) resistant to partial occlusions, low-resolution data and extreme expressions; c) computationally efficient in order to support real-time operation and be implemented on edge devices with limited power; and d) fair and equitable to allow high-performance to be achieved consistently across diverse demographic populations. It is not an augmenting option to address this multilayered issue but rather a required condition that must be met before the implementation of face recognition systems which are indeed robust, trustworthy, and purpose fit to be utilized in various societal contexts.

Objectives

The review article is designed in a way that it attempts to fulfill the following specific objectives:

1. To survey, classify, and critically analyze methods of photometric normalization and illumination invariance
2. To trace the history of evolving classical image-processing filters to current deep learning-based methods that learn invariant representations by adversarial training and disentanglement.
3. To analyze the important intersection of alignment/normalization with the acute problem of demographic bias in FR. This includes the examination of nascent yet crucial methods precisely created to enhance algorithmic impartiality by enhancing, justifiable localization of landmarks and feature space standardization [19], [20].
4. To raise the topic of useful architectural issues of implementing alignment and normalization modules into modern, end-to-end deep FR pipelines. It involves the trade-off analysis between sequential and joint optimization, model compression and acceleration strategies to allow deployment in resource-constrained, edge-computing settings.
5. To integrate the literature reviewed in a coherent approach to the identification of persistent problems, the promising but not researched directions, and informed suggestions about the future research as the interface of computer

vision, geometric modeling, and ethical AI.

Literature Review

Backgrounds: Classical Models to Cascaded Regression.

Statistical shape and appearance models predominated in the first automation of the localization of facial landmarks. The main modes of variation in the spatial structure of the landmarks were learned with Active Shape Models (ASMs) [5] which restricted the search to possible face shapes. Active Appearance Models (AAMs), their extension, added a model of texture (pattern of pixel intensities within the shape) allowing a more powerful yet computationally expensive fit to the image data.

Although basic, these generative models tended to be poor in initializing, lighting and masking. Discriminative, cascaded regression frameworks introduced in the field also created a paradigm shift. The classical example of Kazemi and Sullivan [21] of an Ensemble of Regression Trees (ERT) proved that the direct regression of landmark position based on feature of images in a cascade of fines was much faster and more resistant. This method did not require a specific generative model, but it learns a direct mapping of image patches to shape changes. It influenced the standard of decades because of its attractive combination of speed and precision and laid the groundwork of numerous working libraries (e.g., Dlib).

Deep Learning Revolution of Geometric Alignment.

Deep convolutional neural networks were the essential shift in thinking as they learn data-driven, hierarchical features in localization of landmarks. Initial CNN solutions focused on it as an ordinary regression or heatmap prediction task. One practical development that has seen extensive application was the Multi-task Cascaded CNN (MTCNN) [22], which made it easy to jointly learn face detection, bounding box regression, and landmark localization in a single, lightweight cascaded model, which pointed to the advantages of multi-task learning. Besides cascaded frameworks, FCN-based heatmap regression emerged as a paradigm of high-accuracy alignment. The state of the art performance of networks such as the Face Alignment Network (FAN) and its variations, involved making a likelihood heatmap prediction of each landmark, allowing localization to be localized precisely, and also naturally robust to initialization. The current studies have even gone beyond strength and accuracy. The problem of initialization failure in challenging scenarios (extreme pose, heavy occlusion) is of critical importance, and Shao et al. [28] overcame it through introducing the notion of deep progressive reinitialization mechanism, which allows the network to overcome the poor initial estimates. To ensure very precise alignment which is important in recognition, Wan et al. [29] developed a multi-order hourglass network that learns and fuses features across different granularities producing very precise landmark positions. Lin et al. [30] focused on the structure-coherent deep features explicitly designed to be used in the alignment task to enhance consistency in different appearances.

3D Alignment of Extreme pose invariance.

There are intrinsic limitations to 2D alignment in large yaw or pitch rotations because of self-occlusion. 3D face alignment is used to solve this problem by estimating the 3D structure of the face. This may entail training a 3D Morphable Model (3DMM) [25], a statistical model of 3D face shape and texture, to the 2D image. After the 3D model has been installed, in-plane and out-of-plane rotation can be reversed to the canonical frontal pose. More modern deep learning schemes regress 3MM parameters, or even 3D vertex coordinates, directly regressing off the image using encoder networks [26]. Jabberi et al. [11] looked at the adequacy of landmark counts, and found that 68 points are typically efficient in 3D alignment, but additional points may help in particular tasks.

Photometric Normalization: Image Processing to Feature Disentanglement.

This is aimed at making facial representations insensitive to lighting variations. Classical image processing techniques are global algorithms such as Histogram Equalization (HE), local algorithms such as Self-Quotient Image (SQI) or Multiscale Retinex (MSR), which seek to divide an estimated illumination map and the reflectance image [13]. Dalal and Vishwakarma [13] introduced the optimized adaptive illumination-normalization algorithm with a kernel-based classifier. Iqbal et al. [4] employed the layers extraction and histogram processing method of illumination normalization. The epoch of deep learning marked the beginning of the change in conception: rather than rectify the input image, learn features at the location of lighting suppressed information. This disentanglement has been very successful through adversarial learning.

In a study by Zhang et al. [8], IL-GAN is proposed, in which a generator generates features that are immune to lighting, and a discriminator attempts to determine what the original lighting state was, which encourages the generator to eliminate this information. In the same way, Silva and Farias [9] trained an Adversarial Disentangling Variational Autoencoder (AD-VAE) as a way of disentangling identity-related latent codes and non-identity-related ones such as illumination and expression.

Unified and Feature-Level Normalization Structures.

One of the most popular tendencies is to abandon the notion of alignment and normalization as independent and consecutive pre-processing processes. The new paradigm aims at unified or feature level normalization in which canonicalization is done in the feature extraction pipeline itself. Pang et al. [25] proposed a single multi-domain face normalization system, which is applied to features space, to allow successful cross-domain recognition (e.g., in visible light, near-infrared). Kim et al. [26] were able to show that explicit geometric alignment could be bypassed by directly computing face shape-controllable deep feature alignment on the convolutional feature maps, and thus the network was inherently resistant to spatial misalignment. This co-dependency was formalized by Liu et al. [27] in a framework of joint face normalization and representation learning, the two processes were learned together in a synergetic manner.

Mitigating Fairness and Bias by Better Normalization.

It is increasingly being acknowledged that bias in FR is frequently upstream-based. In turn, enhancing alignment and normalization towards the fairer systems is a straightforward way to the fairer ones [19]. What is imperative is that what it is worth the fairness must be engineered into the initial phases of the FR pipeline, not only at the level of classifier.

Methodology (Analytical Framework).

The review uses a multi-dimensional analytical system that is structured to classify, compare, and assess the ample literature on face alignment and face normalization. The framework will be aimed at clarifying the connections between the various approaches and evaluating their effectiveness in tackling the fundamental issues identified in Section 3. In this category, the data is categorized using a metric that is not ordinal, such as nominal or ratio scales.

Dimensional Categorization:

This is where the data is categorized based on a non-ordinal metric like the nominal or ratio scale. Techniques are mostly categorized under two important dimensions:

- **Primary Objective:** Does the method emphasize Geometric Correction (pose, expression) or Photometric Normalization (illumination, color)? Most contemporary methods cover both, although their main focus may be identified quite often.
- **Implementation Paradigm:** Is the technique part of the Classical/Model-Based paradigm (e.g. ASM/AAM, cascaded regression with hand-crafted features) or the Deep Learning-Based paradigm (e.g. CNN regressors, adversarial networks)? The historical and technological change is taken up in this dimension.

There is a third, cross-cutting dimension which is used to point out to a pressing modern issue:

- **Fairness-Conscious Design:** Does it specifically adopt design elements or objectives that minimize demographic performance differences? [19], [22].

In-Depth Technical Analysis:

In the geometric correction category, one further analyzes it in sub-categories:

- **Landmark-Based 2D Alignment:** Non-enhanced cascaded regression [21], deep holistic networks [6], [23], [24], and occlusion and expression resistance.
- **3D Model-Based Alignment:** An evaluation of the trade-offs between conventional 3DMM fitting [12] and direct deep regression of 3D parameters [21] to emphasize the extremity pose behavior.
- **Feature-Level Canonicalization:** Finding ways to elude explicit landmarks to do spatial transformation in feature space [12], [22], [23].

To perform photometric normalization, the analysis compares:

- **Image-Domain Techniques:** Evaluation of the weaknesses and usefulness of filtering-based approaches such as SQI and MSR [13].
- **Deep Feature-Domain Techniques:** Proper consideration of the functioning and performance of adversarial disentanglement to attain illumination invariance [8], [9].

Metrics and criteria used in the evaluation:

All sets of techniques are compared to a set of common criteria based on the problem statement:

- **Accuracy & Precision:** To be aligned, it is normally quantified by the length of Mean Errors (NME) of landmarks on difficult benchmarks (e.g., 300-W, WFLW). To be normalized, it is the corresponding increase in Face Verification Accuracy (e.g., on LFW, CFP-FP, IJB-C) with regard to changes in lighting.

- **Robustness:** Capability to continue functioning when subjected to stresses such as large pose angles (>60deg yaw), partial occlusions (e.g. glasses, masks), low image resolution and harsh and non-uniform illumination.
- **Efficiency:** The computational cost (measured in Floating-Point Operations (FLOPs)) of a model, model size (number of parameters), and the inference time (milliseconds) on standard hardware (CPU, GPU, mobile processor).
- **Fairness:** Decrease in performance gap, as indicated by such measures as the difference in False Non-Match Rate (FNMR) between demographic subgroups (e.g., on the RFW benchmark) [19].

Integration Analysis:

Lastly, the review examines how these modules can be incorporated into whole FR systems. This includes examining:

- **Sequential Pipelines:** This is the classical model where a detected face is sent to individual, typically pre-trained, alignment and normalization models, followed by feature extraction.
- **End-to-End Joint Training:** New models where alignment/normalization sub-networks are learned together with the feature extractor and classifier, and they can be optimized over the entirety [21].
- **Design towards Edge:** How to streamline or implement them in lightweight FR models to serve IoT and edge devices [17], [18], [19].

Such a complex structure makes it possible to have a systematic examination of methods in various periods and design ideologies and gives a clear picture of how the field has evolved, and the merits/demerits of the existing methods.

Result and Discussion

The literature synthesis process based on the analytical framework presented in Section 6 provides major results in a number of dimensions. Such insights outline the present state of the art and its implications organized as follows:

Synthesis on Implementation paradigm: The Ascendancy of Deep Learning.

DNNs have conclusively established the state of the art in landmark localization error and, more to the point, resilience to adverse, in-the-wild settings. This has made classical model-based models (ASM/AAM) virtually obsolete in high-performance systems, but they still have a conceptual impact.

Synthesis of Primary Objective and Processing Domain: The Canonicalization of Shifting Image-Space to Feature-Space.

One of the most important and consequential tendencies is the shift in the place of geometric and photometric corrections in the pixel/image space towards geometric and photometric corrections in the deep feature space. This deep normalization has the ability to learn task-specific canonicalization which are better than generic, predefined image warps or filters. It is a transition of pre-processing to integrated-processing.

Photometric Disentanglement Synthesis: Invariance Learner Adversarial Learning.

Adversarial learning as demonstrated by Zhang et al. [8] and Silva and Farias [9] has turned out to be an extremely successful model to photometric normalization. This causes the network to learn to factor out such variations by structuring invariance as a minimax game wherein the feature extractor attempts to deceive a discriminator which attempts to determine the nuisance variable (illumination, expression). The principle has become a foundational principle in learning invariant representations and is being generalized to tease out other variables such as age, accessories and even demographic information to minimize bias [20].

Synthesis on Fairness-Aware Design: Removing Bias at the Core.

Much of the algorithmic bias may be remedied during the alignment and normalization phase [19]. Poor normalization and poor features generated by inaccurate landmarking of some structural features of the face amount to a systematic disadvantage. Thus, strengthening, diversifying, and equitable landmark detectors and normalization schemes is a high leverage, critical intervention point in the construction of equitable FR systems. This makes fairness a fundamental need of the front-end modules, rather than a supplement to the classifier.

Efficiency and Integration Synthesis: Trading off between Accuracy and Speed.

There is a longstanding conflict between accuracy of alignment (large quantities of landmarks and model complexity) and speed. Although detailed alignment (e.g. 68+ points as the case in [11]) tends to enhance the accuracy of downstream recognition, it has higher latency and energy usage. The answer of the research community is dual: 1) creating highly efficient but accurate alignment models, such as MobileFAN [12], and 2) creating recognition models which are more resistant to spatial misalignment, which include GhostFaceNets [19] and EdgeFace [22]. It means that the most suitable system design can presuppose the co-design of the alignment and recognition modules where the precision of the former will be adjusted to the tolerance of the latter.

Discussion: Towards Holistic Deep Normalization Networks.

The overlapping of trends is an indicator of the existence of Deep Normalization Networks in the future. They would be end-to-end trainable architectures that implicitly do 3D-aware geometry canonicalization and adversarial photometric disentanglement in one feature-learn feature-learning framework. Synthetic data [23], which is produced through GANs or a 3D-rendering process, will be essential to enable these networks to produce the extensive and controlled variations required to oversee them successfully. The grand dilemma is how to unite all the desirable characteristics such as extreme accuracy, the ability to withstand all the nuisance factors, computational efficiency, and demographic fairness into one practical system.

Summary and Recommendations.

This paper has given detailed overview of the importance of the face alignment and normalization in the creation of strong, precise, and just face recognition systems. We have followed the history of technology since its first statistical models and cascaded regression up to the modern state of deep learning, which has made holistic alignment, feature-level canonicalization, and adversarial disentanglement of nuisance variables possible. These have been identified in the review as not just preliminary steps but are in actual sense building blocks that directly define the highest level of system performance, strength and justice.

Main Recommendations on the Future Research:

1. **Feedback-based End-to-End Joint Optimization:** Future architectures are to transfer past architectures based on pretrained, frozen alignment modules. Future research ought to be done on fully differentiable pipelines in which the alignment/normalization sub-networks are explicitly fed the recognition loss, and therefore can optimize to a canonical form that is not only geometrically correct in a broad way, but also optimal (by the specific feature extractor and task) as well.
2. **Lightweight, Unified, 3D-Aware, Normalizers:** It is necessary to have efficient network designs capable of implicitly learning 3D facial structure of a single 2D image and apply this knowledge to jointly normalize geometric (pose, expression) and photometric features directly in the feature space. These Normalization Nets must be small to deploy the edges.
3. **Explainability and Diagnostic Tools to Failure Analysis:** The community might create more elaborate visualization and diagnostic tools to learn how, and why alignment or normalization did not succeed in a certain situation (e.g. in case of a certain demographic or given a certain lighting condition). This is necessary to debug bias, enhance robustness and develop trust**.**
4. **Generalization of Normalization Principles:** The generalization rules of canonicalization and nuisance factor disentanglement trained on face recognition is generalized to more perceptual problems. Future studies ought to be active and consider further expansion of these methods in order to enhance strength in face expression recognition [24], age prediction, medical phenotyping and other similar areas.
5. **Development of Standardized Benchmarks to Fairly Evaluate:** More complex benchmarks are required to strictly gauge the progress, especially the fairness of this development. These must not only store differing information but also have particular analysis plans and measures of evaluation of the performance and bias of the alignment and normalization phases themselves, without regard to the backend recognizer.

References

[1] J. L. et al., "Cloud versus edge deployment strategies of real-time face recognition inference," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 1, pp. 458–470, 2021, doi: 10.1109/TNSE.2021.3055835.

[2] A. R. S. et al., "Robust face recognition under challenging conditions: a comprehensive review of deep learning methods and challenges," *Applied Sciences*, vol. 15, no. 17, p. 9390, 2025, doi: 10.3390/app15179390.

[3] Y. Xu, W. Yan, G. Yang, J. Luo, T. Li, and J. He, "CenterFace: joint face detection and alignment using face as point," *Scientific Programming*, vol. 2020, Article ID 7845384, 2020, doi: 10.1155/2020/7845384.

[4] "Synthetic data for face recognition: Current state and future prospects," *Image and Vision Computing*, vol. 130, Article ID 104688, 2023, doi: 10.1016/j.imavis.2023.104688.

[5] "A comparative analysis of various methods for attendance framework based on real-time face recognition

technology," in *Proceedings of the International Conference on Communication and Computer Security*, 2022, pp. 401–410, doi: 10.1007/978-981-19-7615-5_40.

[6] "A comprehensive review of face recognition techniques, trends, and challenges," *IEEE Access*, vol. 12, pp. 89882–89920, 2024, doi: 10.1109/ACCESS.2024.3424933.

[7] *Proceedings of the Twelfth International Conference on Digital Image Processing (ICDIP 2020)*, Osaka, Japan, 2020, Proc. SPIE 11519, Art. no. 115191K, 2020, doi: 10.1117/12.2573135.

[8] Y. Zhang, C. Hu, and X. Lu, "IL-GAN: Illumination-invariant representation learning for single sample face recognition," *Journal of Visual Communication and Image Representation*, vol. 59, pp. 501–513, 2019, doi: 10.1016/j.jvcir.2019.02.007.

[9] A. Silva and R. Farias, "AD-VAE: Adversarial disentangling variational autoencoder," *Sensors*, vol. 25, no. 5, p. 1574, 2025, doi: 10.3390/s25051574.

[10] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 10, pp. 5962–5979, Oct. 2021, doi: 10.1109/TPAMI.2021.3087709.

[11] M. Jabberi, A. Wali, B. B. Chaudhuri, et al., "68 landmarks are efficient for 3D face alignment: what about more?," *Multimedia Tools and Applications*, vol. 82, pp. 41435–41469, 2023, doi: 10.1007/s11042-023-14770-x.

[12] H.-I. Kim, K. Yun, and Y. M. Ro, "Face Shape-Guided Deep Feature Alignment for Face Recognition Robust to Face Misalignment," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 4, no. 4, pp. 556–569, Oct. 2022, doi: 10.1109/TBIOM.2022.3213845.

[13] S. Dalal and V. P. Vishwakarma, "A novel approach of face recognition using optimized adaptive illumination-normalization and KELM," *Arabian Journal for Science and Engineering*, vol. 45, no. 11, pp. 9977–9996, 2020, doi: 10.1007/s13369-020-04566-8.

[14] K. Nagumo, T. Kobayashi, K. Oiwa, and A. Nozawa, "Face alignment in thermal infrared images using cascaded shape regression," *International Journal of Environmental Research and Public Health*, vol. 18, no. 4, p. 1776, 2021, doi: 10.3390/ijerph18041776.

[15] H. Wang, R. Cheng, J. Zhou, et al., "Multistage Model for Robust Face Alignment Using Deep Neural Networks," *Cognitive Computation*, vol. 14, pp. 1123–1139, 2022, doi: 10.1007/s12559-021-09846-5.

[16] R. Cheng, H. Wang, X. Liu, X. Yan, and L. Tao, "Deformable hourglass network with face region normalization for robust face alignment," in *Proceedings of SPIE 11373, Eleventh International Conference on Graphics and Image Processing (ICGIP 2019)*, 2020, Art. no. 1137303, doi: 10.1117/12.2557185.

[17] J. A. et al., "Racial bias within face recognition: a survey," *ACM Computing Surveys*, vol. 57, no. 3, Art. no. 61, 2025, doi: 10.1145/3705295.

[18] [S. B. et al., "Review of demographic fairness in face recognition," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 7, no. 1, pp. 15–31, 2025, doi: 10.1109/TBIOM.2025.3601217.

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

[20] M. T. et al., "Mitigating demographic bias in facial recognition through adversarial representation learning: a data-driven study," *International Research Journal of Innovation in Engineering and Technology*, vol. 9, no. 6, Art. no. 035, 2025, doi: 10.47001/IRJIET/2025.906035.

[21] L. Persona, F. Meloni, and A. A. Macedo, "A lightweight method for almost real-time recognition of emotions through facial geometry with normalized landmarks," *Journal of the Brazilian Computer Society*, vol. 31, no. 1, pp. 325–337, 2025, doi: 10.5753/jbcs.2025.4456.

[22] M. Pang, W. Zhang, Y. Lu, Y.-M. Cheung, and N. Zhou, "A Unified Multi-Domain Face Normalization Framework for Cross-Domain Prototype Learning and Heterogeneous Face Recognition," *IEEE Transactions on Information Forensics and Security*, vol. 20, pp. 5282–5295, 2025, doi: 10.1109/TIFS.2025.3570121.

[23] M. Cai, S. Zhang, G. Xiao, and S. Fan, "3D face reconstruction and dense alignment with a new generated dataset," *Displays*, vol. 70, Article ID 102094, 2021, doi: 10.1016/j.displa.2021.102094.

[24] "An efficient machine learning based attendance monitoring system through face recognition," *Research on Intelligent and Artificial*, vol. 38, no. 2, pp. 123–134, 2024, doi: 10.18280/ria.380233.

[25] M. Pang *et al.*, "A unified framework for bidirectional prototype learning from contaminated faces across heterogeneous domains," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 1544–1557, Apr. 2022, doi: 10.1109/TIFS.2022.3164215.

[26] H.-I. Kim, K. Yun, and Y. Ro, "Face shape-guided deep feature alignment for face recognition robust to face misalignment," *IEEE Trans. Biometrics, Behavior, and Identity Sci.*, vol. 5, no. 1, pp. 110-123, Jan. 2023, doi: 10.1109/TBIOM.2022.3213845.

[27] X. Liu, M. Kan, W. Wu, S. Shan, and X. Chen, "VIPLFaceNet: An open source deep face recognition SDK," *arXiv e-prints*, Sept. 2016, Art. no. arXiv:1609.03892, doi: 10.48550/arXiv.1609.03892.

[28] X. Shao, J. Xing, J. Lyu, X. Zhou, Y. Shi, and S. Maybank, "Robust Face Alignment via Deep Progressive Reinitialization and Adaptive Error-driven Learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, pp. 5488–5502, Sept. 2022, doi: 10.1109/TPAMI.2021.3054174.

[29] J. Wan, J. Ma, Y. Ma, S. Wang, H. Hu, and C. Yu, "Robust Face Alignment by Multi-Order High-Precision Hourglass Network," *IEEE Trans. Image Process.*, vol. 30, pp. 120–133, 2021, doi: 10.1109/TIP.2020.3034261.

[30] C. Lin, B. Zhu, Z. Liu, P. Wang, M. Sun, and J. Zhou, "Structure-Coherent Deep Feature Learning for Robust Face Alignment," *IEEE Trans. Image Process.*, vol. 30, pp. 4930–4943, 2021, doi: 10.1109/TIP.2021.3082319.