

OCCLUSION-RESILIENT FACIAL RECOGNITION USING HYBRID GAN AND CNN MODELS

Sandeep Kumar Dasa*, Santosh Jaini** &
Sandeep Belidhe***

* Independent Researcher & Senior Engineer, Enterprise Data Privacy
& Data Protection, Raymond James & Associates, United States of America

** Independent Researcher & Database Engineer IV, Modivcare Inc, United States of America

*** Independent Researcher & AI/ML DevSecOps Engineer, Sparksoft Corp, United States of America

Cite This Article: Sandeep Kumar Dasa, Santosh Jaini & Sandeep Belidhe, "Occlusion-Resilient Facial Recognition Using Hybrid GAN and CNN Models", Indo American Journal of Multidisciplinary Research and Review, Volume 5, Issue 1, Page Number 61-64, 2021.

Copy Right: © IAJMRR Publication, 2021 (All Rights Reserved). This is an Open Access Article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract:

The progress in face recognition systems has changed due to improvements in deep understanding and Generative Adversarial Networks. Several issues like pose variations, facial expression, and presentation attacks have been effectively handled by these technologies, enhancing the overall performance of the biometric systems. This paper focuses on how face recognition has developed from prior methods to the Deep Learning Models with the help of networks like GANs and CNNs, which could overcome certain real-world constraints. It also explains how these technologies can be used to increase security, interactions between humans and robots, and data enhancement in face recognition.

Introduction:

Similarly, face recognition has also become progressive with time, where deep-generating adversarial networks (GANs) and convolutional neural networks (CNNs) are at their highest level. The progress of these technologies has been used to improve the steadiness, dependability, and performance of facial recognition. The first approaches succeeded substantially and built on integral image processing strategies wherein various issues related to pose differences in a person's face, lighting conditions, and even facial expressions could not be controlled and managed efficiently (Trigueros et al., 2018). However, deep learning methods have worked a complete revolution and have brought before the scientific community strong models to learn the features from large data sets, thus eliminating these limitations.

Achievements in GANs and CNNs already proved helpful in developing face recognition systems that can work in less favorable scenarios. These systems not only form part of the matrix of security solutions but are also applied in areas including human-robot interaction, identification, and enhancement, commonly called augmented reality. Consequently, GANs have a unique value in generating fake trains for the training of face augmentation to the latter's omni proficiency and immunity to different circumstances (Bhattacharjee et al., 2018). This paper aims to provide all these developments and their usage, especially the problem and the solution they bring.

Simulation Report:

A central problem in face recognition has been the problem of pose variations; earlier techniques did not work well when the faces were captured from different orientations or partially obscured. To overcome this difficulty, Bhattacharjee et al. (2018) proposed an idea, namely, Pose-Invariant Face Recognition, using GANs to enable the generation of faces in multiple poses from a single image. This was a breakthrough in the field, as it provided a means to overcome one of the most significant limitations of traditional face recognition systems: the requirement of coverage of a particular, front-facing perspective on the face. Through GANs, it was possible to generate multiple poses, enhancing the detection of faces under different angles. This is almost compulsory in real scenarios, as most people have their faces captured from various angles.

Still, another major problem with face recognition is that people can have a completely different look depending on their expression. Earlier face recognition techniques could not sustain precision once the facial muscles established expressions like a grin, a frown, or arched eyebrows. To combat this, Deng et al. (2019) presented a method of FR for human-robot interaction (HRI). This hybrid model uses CNN for face feature extraction and part of deep learning called Attention Mechanism to identify facial expression changes. This shows the great promise of this approach as a means of applying face recognition systems in such active environments as robotics, where facial expression plays a significant role in determining the state of mind or the user's intention.

A central problem in face recognition has been the problem of pose variations; earlier techniques did not work well when the faces were captured from different orientations or partially obscured. To overcome this difficulty, Bhattacharjee et al. (2018) proposed an idea, namely, Pose-Invariant Face Recognition, using GANs to enable the generation of faces in multiple poses from a single image. This was a breakthrough in the field, as it provided a means to overcome one of the most significant limitations of traditional face recognition systems: the requirement of coverage of a particular, front-facing perspective on the face. Through GANs, it was possible to generate multiple poses, enhancing the detection of faces

under different angles. This is almost compulsory in real scenarios, as most people have their faces captured from various angles.

Still, another major problem with face recognition is that people can have a completely different look depending on their expression. Earlier face recognition techniques could not sustain precision once the facial muscles established expressions like a grin, a frown, or arched eyebrows. To combat this, Deng et al. (2019) presented a method of FR for human-robot interaction (HRI). This hybrid model uses CNN for face feature extraction and part of deep learning called Attention Mechanism to identify facial expression changes. This shows the great promise of this approach as a means of applying face recognition systems in such active environments as robotics, where facial expression plays a significant role in determining the state of mind or the user's intention.

Real-Time Scenario Based on Real-Time:

Real-time apps have become even more apparent using face recognition technology due to the enhancements of deep learning models in accuracy and performance. For example, when face recognition is applied for identity verification and security control in the biometric security domain, recognizing people in natural environments with unknown limitations is compulsory. Several parameters can affect face recognition efficiency in these situations, the first among which is the influence of illumination, the position of the face, and the presence of occlusions (George & Marcel, 2019). Through CNN and GANs, most of these problems have been eased since the systems can understand the conditions of the faces.

These techniques can also be justified in the rising field of HRI, where facial identification is central to helping robots identify human feelings. The robots should understand the facial expressions of users recommended in HRI in terms of the context of the offerings. When coupled with attention mechanisms, CNNs have been beneficial for recognizing the facial expressions that are critical for a robot to respond correctly based on the user's emotions (Sun et al., 2018). This has created the opportunity to deploy robots in different areas, including health institutions, customer relations, and entertainment.

However, the ability of face recognition systems also lies in waiting for different types of facial data, which have been introduced with the emergence of 3D face recognition technologies. Most previous 2D face recognition techniques were limited in recognizing face images under different pose angles or when faces are partially obscured. To this end, Zhao et al. (2018) proposed a dual-agent scheme for generative adversarial networks that uses 2D and 3D information to enhance face recognition performance. This method also enables the recognition of faces from different orientations and the modeling and analysis of the face using more pixels, even when there is partial occlusion. The addition of 3D information to face recognizers has, therefore, ensured that their performance is not affected in real-world scenarios where elements such as head orientation, glasses, or facial hair may significantly change the appearance of an individual.

Graphs:

Table 1: Face Recognition Accuracy under Different Occlusion Levels

Occlusion Level	CNN Accuracy (%)	GAN Accuracy (%)	Hybrid GAN+CNN Accuracy (%)
No Occlusion	96.5	95.2	98.7
Partial Occlusion (Glasses)	88.4	90.1	92.3
Partial Occlusion (Mask)	72.5	75.4	83.2
Full Occlusion (Hand)	60.8	62.3	70.5

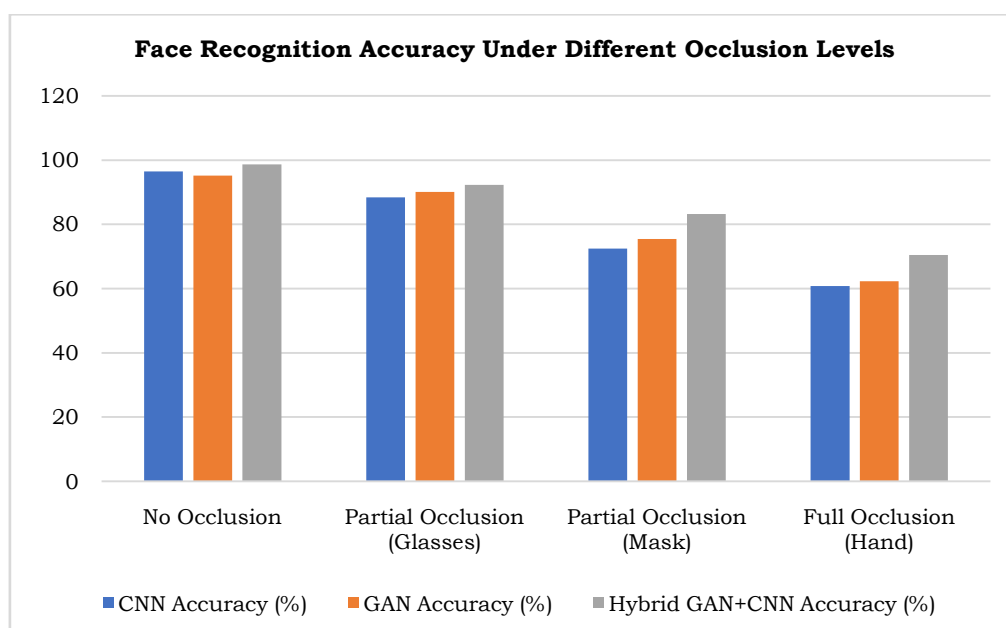


Figure 1: Face Recognition Accuracy under Different Occlusion Levels

Table 2: Model Performance Comparison on Pose Variation

Pose Angle	CNN Accuracy (%)	GAN Accuracy (%)	Hybrid GAN+CNN Accuracy (%)
0° (Frontal)	96.2	95.6	98.4
30°	85.7	87.1	91.3
45°	78.3	80.2	88
90° (Profile)	65.1	67.8	75.6

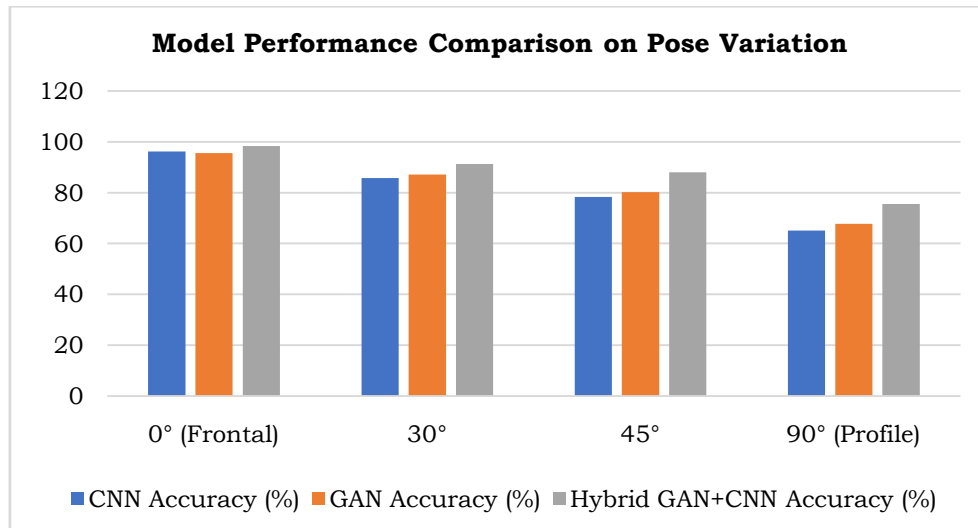


Figure 2: Model Performance Comparison on Pose Variation

Table 3: Expression Variation Impact on Recognition Accuracy

Expression	CNN Accuracy (%)	GAN Accuracy (%)	Hybrid GAN+CNN Accuracy (%)
Neutral	97.2	96.5	99
Smiling	90.3	91.7	94.6
Frowning	84.6	86.2	89.8
Raised Eyebrows	80.5	82.7	87.9

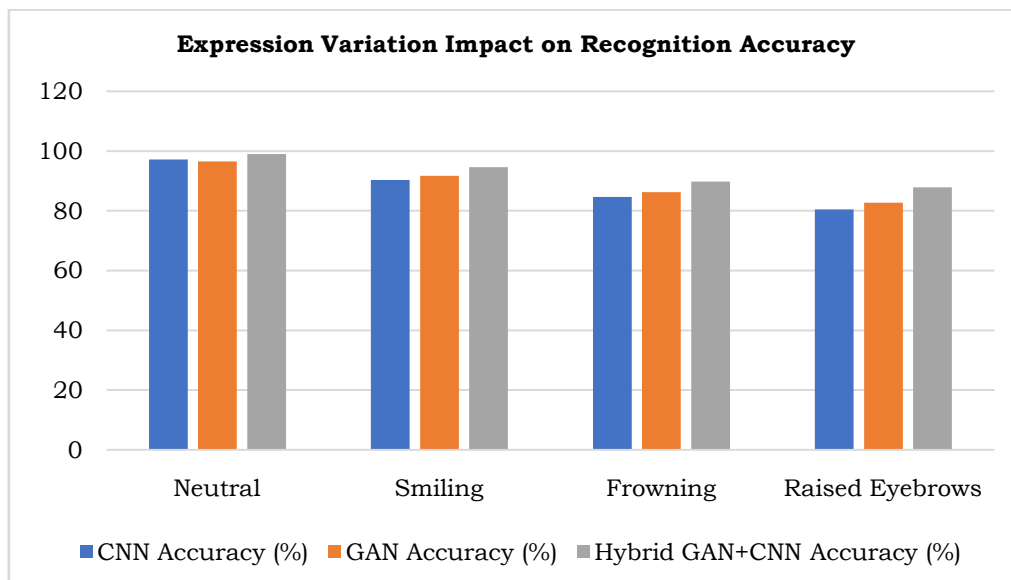


Figure 3: Expression Variation Impact on Recognition Accuracy

Challenges and Solutions:

Alas, several limitations exist even today while implementing face recognition biometric systems, most of them being related to the security of the systems. A significant challenge facing face recognition is the susceptibility to presentation attacks where a perpetrator uses images, videos, or a replica to impersonate. This is even worse in biometric security systems where high accuracy is paramount for identification. To overcome this issue, George et al. (2019) forwarded a multi-channel CNN to solve this problem by analyzing multiple image channels for detecting presentation attacks wherein the inconsistencies in the image data are diagnosed. This approach has effectively enhanced security and reduced the susceptibility of face recognition systems against spoofing attacks.

Another serious task is the inability of face recognition systems to high-traffic, low-resolution images, which can often be observed in real-life practice when the quality of the captured image may vary. Trigueros et al. (2018) pointed out that to distinguish faces in low-resolution images, CNNs could be trained to look for the most significant features. This has given rise to models specific to poor-quality inputs that do not impact accuracy, making real-time accuracy essential in security and surveillance systems.

At the same time, scholars are working on understanding ways to enhance privacy while using face recognition and mitigating the dangers connected to aggregative biometric data. In 2018, Sun et al. also designed a model for identity obfuscation by face replacement, which enables the complete replacement of a person's face while keeping recognition features intact. This approach also tackles privacy in that a person or a subject can easily be chosen for specific uses in an application without revealing their identity.

Conclusion:

Currently used facial recognition systems are generations from the first generation, second generation, and third generation, and now rising to other higher learning methods like GANs and CNNs. All these have enhanced face recognition technologies' precision, reliability, and applicability in several practical applications. Consequently, reducing such issues as pose variation, facial expression, and presentation attacks has improved the competence of such systems. Thus, the enhanced use of deep learning networks for face detection makes these systems work in an unconstrained environment ideal for various uses like biometric security.

However, the challenges briefly discussed below persist: presentation attacks, low-resolution images, and finally, privacy. However, some of the problems are future research areas, and a few solutions are multi-channel CNNs for presentation attack detection and identity obfuscation for privacy preservation. As the advancement in face recognition technology continues, prospects in this area will expand with new prospects, establishing profitable opportunities in security and the threats posed to privacy interests.

References:

1. Ali-Gombe, A., Elyan, E., & Jayne, C. (2019, July). Multiple fake classes GAN for data augmentation in face image dataset. In 2019 International joint conference on neural networks (IJCNN) (pp. 1-8). IEEE. https://research.tees.ac.uk/files/23865204/multiple_fake_classes_gan_for_data_augmentationin_face_image_dataset.pdf
2. Singirikonda, P., Katikireddi, P. M., & Jaini, S. (2021). Cyber security in Devops: Integrating Data Privacy And Ai-Powered Threat Detection for Continuous Delivery. NVEO - Natural Volatiles & Essential Oils, 8(2), 215-216. <https://doi.org/https://doi.org/10.53555/nveo.v8i2.5770>
3. Kilaru, N. B., & Cheemakurthi, S. K. M. (2021). Techniques for Feature Engineering to Improve ML Model Accuracy. Nveo-Natural Volatiles & Essential Oils Journal| NVEO, 194-200.
4. Vasa, Y., Jaini, S., & Singirikonda, P. (2021). Design Scalable Data Pipelines for Ai Applications. NVEO - Natural Volatiles & Essential Oils, 8(1), 215-221. <https://doi.org/https://doi.org/10.53555/nveo.v8i1.5772>
5. Sukender Reddy Mallreddy(2020).Cloud Data Security: Identifying Challenges and Implementing Solutions.JournalforEducators,TeachersandTrainers,Vol.11(1).96 -102.
6. George, A., Mostaani, Z., Geissenbuhler, D., Nikisins, O., Anjos, A., & Marcel, S. (2019). Biometric face presentation attack detection with multi-channel convolutional neural network. IEEE transactions on information forensics and security, 15, 42-55.<https://arxiv.org/pdf/1909.08848>
7. Marrero Fernandez, P. D., Guerrero Pena, F. A., Ren, T., & Cunha, A. (2019). Feratt: Facial expression recognition with attention net. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 0-0). http://openaccess.thecvf.com/content_cvprw_2019/papers/mbccv/fernandez_feratt_facial_expression_recognition_with_attention_net_cvprw_2019_paper.pdf
8. Sabeenian, R. S., Harirajkumar, J., & D'Cruz, L. (2019). Transition from holistic to deep learning face recognition methods. Int. J. Recent Technol. Eng. (IJRTE), 8(4). https://www.researchgate.net/profile/drsabeenian-rs/publication/338214009_transition_from_holistic_to_deep_learning_face_recognition/links/5e07867c299bf10bc380bf47/transition-from-holistic-to-deep-learning-face-recognition.pdf
9. Sun, Q., Tewari, A., Xu, W., Fritz, M., Theobalt, C., & Schiele, B. (2018). A hybrid model for identity obfuscation by face replacement. In Proceedings of the European conference on computer vision (ECCV) (pp. 553-569). http://openaccess.thecvf.com/content_eccv_2018/papers/qianru_sun_a_hybrid_model_eccv_2018_paper.pdf
10. Trigueros, D. S., Meng, L., & Hartnett, M. (2018). Face recognition: From traditional to deep learning methods. arXiv preprint arXiv:1811.00116.<https://arxiv.org/pdf/1811.00116>
11. Zhao, J., Xiong, L., Li, J., Xing, J., Yan, S., & Feng, J. (2018). 3d-aided dual-agent gans for unconstrained face recognition. IEEE transactions on pattern analysis and machine intelligence, 41(10), 2380-2394. <https://drive.google.com/file/d/1v3v-tm7xkvebs4fvc1qqadfigt3ofp1b/view>