Enhanced Reconstruction of Occluded Images Using GAN and VGG-Net Preprocessing

Salamun[1], Shamsul Kamal Ahmad Khalid[2]*,
Ezak Fadzrin Ahmad Shaubari[3], Noor Azah Samsudin[4], Luluk Elvitaria[5]
Faculty of Computer Science and Information Technology,
Universiti Tun Hussein Onn Malaysia (UTHM), Johor, Malaysia[1, 2, 3, 4, 5]
Department of Informatics Engineering, Universitas Abdurrab, Pekanbaru, Indonesia[1, 5]

***Abstract*—Facial recognition is widely used in security and identification systems, but occlusions like masks or glasses remain a major challenge. Recent approaches, such as GANs and partial feature extraction methods, attempt to reconstruct or identify occluded facial images. However, these approaches still have limitations in handling severe occlusions, computational efficiency, and dependency on large labeled datasets. In this paper, a GAN-based framework for synthetic reconstruction of occluded facial images is proposed, incorporating multiple specialized modules including a VGG-Net-based perceptual loss component to enhance visual quality. Our architecture improves the fidelity and robustness of reconstructed faces under varied occlusion types. Experimental evaluation on different occlusion scenarios demonstrated high reconstruction quality, with PSNR up to 33.106 and SSIM up to 0.983. The model also maintained strong recognition performance across diverse occlusion combinations. These findings support the framework's potential to enhance face recognition systems in real-world, unconstrained environments.**

***Keywords*—Face recognition; occlusion; image reconstruction; generative adversarial networks; VGG-Net; occluded images; feature extraction**

I. Introduction

In recent years, facial recognition technology has become one of the important aspects in various applications, including security, identification, and access management. The use of this technology includes device security, facial recognition applications on mobile devices, to surveillance in public places. However, in the development of this technology, there are challenges that need to be overcome to improve the reliability and durability of the system. One of the main problems encountered in facial recognition is the inability of the system to recognize faces affected by occlusion or partial censorship. This situation can arise in a variety of contexts, such as the use of face masks, glasses, or even deliberate manipulation of facial imagery to trick the recognition system. Therefore, research on facial image reconstruction strategies affected by occlusion is very important. Efforts to increase the resilience of facial recognition systems to these situations will have a positive impact in improving safety and personal safety, especially in environments where facial recognition technology is widely used [1].

This study aims to develop a strategy for reconstructing facial images affected by occlusion synthetically. The development of an occlusion facial image reconstruction strategy is a crucial aspect in the development of facial recognition technology [2] [3]. With the growing need for security and identification, a deep understanding of how facial recognition systems can overcome occlusion constraints has become a must. The importance of developing this strategy is also closely related to the sustainability of facial recognition applications during the current global pandemic. The use of face masks as a precaution is becoming routine in everyday life, and this can be a bottleneck for existing facial recognition systems [4]. In this context, effective facial image reconstruction strategies can provide important solutions to ensure the smooth use of this technology in various sectors, including public safety, transportation, and other public services. In addition, this research can also contribute to the development of more inclusive facial recognition technology. By being able to recognize and reconstruct images of faces affected by occlusion, the system can provide better care to individuals with special needs, such as those wearing vision aids or other medical devices. This research discusses the facial image reconstruction methodology, which includes the technical approaches used to overcome the occlusion problem, including image processing algorithms and techniques designed to reconstruct obstructed or incomplete facial images. Next, it evaluates the performance of the system in various occlusion scenarios, where an in-depth analysis is conducted to measure the extent to which the system can perform effectively under different conditions, including testing on diverse datasets and measuring the accuracy and reliability of the system. Finally, this research outlines the practical implications of the research findings, particularly in the context of public safety and inclusive services, by showing how the results of this research can be applied in the real world to provide tangible benefits to society, including individuals with special needs. By understanding and overcoming the challenges of occlusion in face recognition, this research can certainly pave the way towards the development of more advanced and reliable technologies. Thus, the existence of facial recognition technology can provide optimal benefits to support security, personal identification, and overall community services [5], [6].

II. Theoretical Overview

An example of facial image de-occlusion is shown in Fig. 1. There are three different groups of face images. In one face image, there are three types of images: occluded face images as image input, real face images without occluded (ground truth), and processed images (predicted image). At the initial stage, GAN is initialized with generators, discriminators, loss functions, and predetermined optimizers. In addition, some parameters, such as clip value to control gradient values and step per epoch as a measure of training iterations, are also set.

Fig. 1. Results of our model on a face image with multiple synthetic occlusions.

During training, a training function (train step) is called for each batch of data, where two gradient tapes are used to calculate the gradient on the generator and discriminator. Gradient clipping is applied to prevent excessive gradients, and an optimizer is used to update the weights of both components. Training results, including loss functions and image quality metrics, are recorded and logged to Tensor Board for monitoring. Furthermore, this implementation also includes functions for generating, evaluating, and storing models [7]. In addition, we train the proposed model on synthetically generated datasets collected from the Internet. By applying the Generative Adversarial Network (GAN) algorithm, specifically designed to handle the task of image reconstruction of faces affected by occlusion or partial censorship, the GAN algorithm consists of two main parts, namely the generator and the discriminator, each of which has its own loss function. In addition, there is an optimizer to manage learning on both components [8]. In this training, gradient clipping techniques are used to avoid problems with exploding gradients that may occur. In addition to variables related to the model and training, this implementation also provides a checkpointing feature, which allows storing the model during training and facilitates further development [9]. Overall, this study forms a systematic basis for GAN training and evaluation in the context of facial image reconstruction, with a particular focus on occlusion treatment [10]. The functions provided not only cover the training aspect but also facilitate the visualization of results and the storage of models for further use [11][12][13].

Research in the field of facial recognition and image reconstruction has been a significant topic in the development of identification and security technologies. In the literature, many studies have been conducted to improve the reliability of facial recognition systems, especially in overcoming obstacles such as occlusion or partial censorship of the face [14], [15]. Current methods often use deep learning-based approaches, particularly generative adversarial networks (GANs) [16], to produce realistic facial images from data affected by occlusion. Several studies have explored the use of GANs in approaching facial image reconstruction, with a focus on restoring facial features hidden due to occlusion. In addition, the literature also highlights the importance of evaluating the quality of reconstructed results using metrics such as PSNR, SSIM, and MSE, as well as other evaluation approaches such as BRISQUE and NIQE, to assess the extent of accuracy and realism of the resulting imagery. This research illustrates recent trends in combining deep learning technology and image quality evaluation to improve the performance of facial recognition systems in constraint situations such as occlusion [17], [18].

Facial recognition has become a major focus in a variety of applications, including security, identity recognition, and human interaction with technology. In the literature, several studies try to address occlusion challenges by utilizing GANs to produce synthetic facial images that can reconstruct features lost to partial censorship [19]. These studies demonstrate that GANs can be an effective tool in increasing the resilience of facial recognition systems to unexpected changes in conditions, such as the use of face masks or other occlusion elements. In addition, the literature also highlights the importance of evaluating the quality of reconstructed images, as facial recognition systems are measured not only in terms of accuracy but also by how well the imagery can represent actual faces [20], [21]. Some studies combine deep learning-based evaluation methods with traditional image quality metrics to provide a holistic picture of the success of facial image reconstruction. Following this trend, this research is geared towards contributing deeper understanding and better strategies for dealing with obstacles such as occlusion in the context of facial recognition [22], [23].

A. Generative Adversarial Networks (GANs)

Shows that advances in deep learning technology have enabled the creation of increasingly realistic facial images. GANs play a crucial role in synthesizing facial images with occlusion or loss of some facial features [24], [25], as is often the case in facial recognition. This approach not only includes reconstructing the image of the face affected by occlusion but also ensures that the resulting synthetic image has natural and acceptable facial characteristics. Several studies propose specific methods to improve the ability of GANs to reconstruct facial images affected by occlusion [26]. The adoption of techniques such as conditional GANs, attention mechanisms, and the use of augmentation data contributed significantly to improving the quality and accuracy of facial image reconstruction. These results show great potential for creating synthetic facial images that not only reflect hidden features but also have high aesthetics and detail. One of the main challenges is overcoming the loss of detail and texture information in the image of an occlusion-affected face. Some studies try to integrate methods such as the use of special loss functions or more complex models to improve the ability of GANs to reconstruct lost details. In addition, there are efforts to develop facial image reconstruction models that are more robust to occlusion variations, including occlusion that appears dynamically or in complex lighting situations. The introduction and treatment of more complex occlusions involve strategies for combining multiple sources of information, including the utilization of contextual and temporal information [27].

B. Image Reconstruction of Faces Affected by Occlusion

In the context of image reconstruction of occlusion-affected faces, it is important to note the vital role of datasets that reflect the diversity of occlusion conditions that may be encountered in real life. Several studies have highlighted the need to have a broad and representative dataset that includes occlusion variations from different sources. Such datasets allow models to learn from different types of occlusion, ranging from the use of face masks [28], hands that cover part of the face, to objects or equipment that may cover part of the face. The selection of appropriate datasets is key to training synthetic facial image reconstruction models. A comprehensive dataset not only helps the model understand occlusion characteristics [29], But it also allows models to produce more realistic and general synthetic facial images. In the literature, several studies have introduced datasets specifically designed to address occlusion challenges, which help improve the performance and generality of reconstructed models [30].

In addition, there is an emphasis on the importance of establishing a balanced dataset in terms of gender, ethnicity, and age representation. This balance is necessary to ensure that models can not only address occlusion variation but can also perform facial image reconstructions fairly and accurately across different demographic groups. By including appropriate datasets and covering a wide range of occlusion, this research is expected to make a further contribution to improving the ability of facial image reconstruction models to occlusion, making this technology more relevant and effective in various contexts of use in everyday life [31]. Some studies also emphasize the importance of creating datasets that reflect variations in lighting conditions, viewing angles, and image resolution. These factors can have a significant impact on the performance of facial image reconstruction models, especially when addressing occlusion. Datasets that include these variations can help models learn to adapt to different situations, thereby improving the reliability and robustness of image reconstruction. In addition, several studies highlight the importance of clearly and completely documenting each type of occlusion contained in the dataset. This information helps facilitate the model training process by providing better guidance on the types of occlusions that the model faces and can expect to reconstruct. Good documentation also supports research reproducibility and allows other researchers to understand the characteristics of datasets better [32].

The adoption of synthetically generated domain-specific datasets has also been a focus of attention in some studies. This approach allows researchers to generate datasets with well-controlled occlusion variations, providing flexibility and clarity in understanding the impact of occlusion on facial image reconstruction models [33]. By involving datasets that include lighting conditions, viewing angles, resolutions, and comprehensive documentation, this research is expected to provide a stronger foundation for the development of synthetic facial image reconstruction models that can handle occlusion more effectively and reliably in real-life situations [34].

III. MATERIALS AND METHODS

A. Materials

For partially obscured face recognition, several different image types are used for system training and testing. Some of the types of images used include real face image dataset. The original face image dataset is used as training data for the GAN algorithm. Such datasets usually consist of images of human faces collected from various sources, such as public databases such as CelebA, LFW (Labelled Faces in the Wild), or specialised datasets collected for specific purposes. GAN can create fairly realistic facial images that can be used to expand the available datasets, helping to improve the accuracy of partially closed face recognition systems. In addition, GANs can also be used to improve system performance by removing objects covering the face or by adding missing facial features, thus making it easier for the system to identify partially covered faces [7]. The use of GAN in partially closed face image reconstruction is very promising but still requires further development. Like other facial recognition technologies, GAN also has some limitations, such as its high complexity and the need for fairly large and diverse datasets. However, with the continuous development of technology and more varied datasets, it is expected that GAN can be an effective tool for improving the accuracy of partially closed face recognition systems. In addition, there are several things to note in the use of GAN for partially closed face image reconstruction, such as:

1) Dataset quality: The quality of the dataset used to train generators and identifiers is very important in determining the accuracy of facial image reconstruction results. A varied and large enough dataset is needed for the generator to produce realistic and accurate images.

2) Hyperparameters: The selection of the right hyperparameters is also very important in determining the quality of facial image reconstruction. This includes selecting the number of layers, the number of nodes, and the rate of learning.

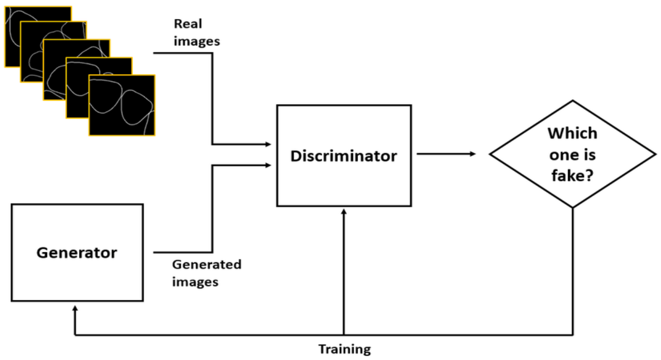
3) Network architecture: The neural network architecture used in GANs also affects the quality of the reconstructed results. Some of the architectures used in GANs include DC-GAN, Wasserstein GAN and Progressive GAN.

4) Monitoring of results: It is important to monitor the results of reconstruction periodically and make repairs if needed.

As the technology is still being developed, partially closed-face image reconstruction using GAN still requires further development. However, with the use of GAN, it is expected to help improve the performance of partially closed face recognition systems.

B. Methodology

The aim of the work is to achieve more accurate and robust facial decomposition results in unrestricted environments. The proposed framework, illustrated in Fig. 2, consists of several modules, namely the Training Model Module (MTM), Image Augmentation Module (IAM), Generator Module (GM), within which there are two more modules, namely the Upsumpling Block Module (UBM) and Downsumpling Block Module (DBM), De-Occlusion Module (DOM), and Discriminator Module (DM).



Source: T. Kim, Y (2020)

Fig. 2. Overview of Generative Adversarial Network (GAN).

The original formula for the GAN loss function is :

$$\mathcal{L}_{GAN} = \mathbb{E}[\log(1 - D(X_{reconstructed}))]$$

Enhanced addition of VGG-Net feature-based loss (such as perceptual loss) to improve the reconstruction quality. Perceptual loss compares the features of the reconstructed image and the original image extracted by VGG-Net. The perceptual loss formula can be written as:

$$\mathcal{L}_{perceptual} = ||VGG(X_{real}) - VGG(X_{reconstructed})||_2^2$$

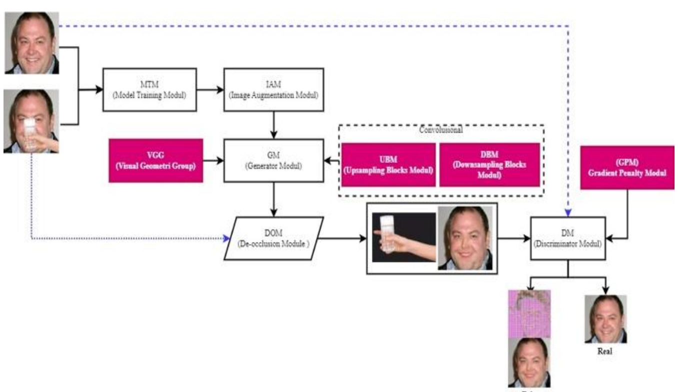


Fig. 3. An overview of our framework for face parsing is shown.

Fig. 3 shows an overview of our framework for face parsing. The framework is composed of several interconnected modules. The proposed methodology addresses the challenges of facial image reconstruction under occlusion using an enhanced Generative Adversarial Networks (GANs) framework. The Model Training Module (MTM) initializes and trains the GAN model with parameters optimized for facial image reconstruction. The Image Augmentation Module (IAM) applies preprocessing and augmentation techniques to enrich the diversity of training data and improve model robustness. The Generator Module (GM) synthesizes realistic facial images from occluded inputs, while the Discriminator Module (DM) distinguishes between real and generated images, aiding the generator in producing high-quality outputs. Finally, the De-Occlusion Module (DOM) specializes in removing occlusion artifacts and reconstructing missing facial features.

To enhance model performance and ensure data consistency, a comprehensive preprocessing pipeline is employed. All input images are resized to 256 × 256 pixels to standardize dimensions, and pixel values are normalized to the range − 1 to 1, facilitating faster convergence during training. Augmentation techniques, including random cropping and jittering, are applied to prevent overfitting and improve generalization.

The GAN architecture consists of a generator and a discriminator. The generator implements an encoder-decoder structure with skip connections to preserve spatial information. It comprises down sampling layers for feature extraction, bottleneck layers to learn latent representations, and up sampling layers to reconstruct high-resolution images. The discriminator is a convolutional neural network that evaluates the authenticity of generated images by comparing them against ground truth data.

The training process begins with the initialization of the GAN using specified hyperparameters, including learning rates, loss functions, and gradient clipping values to prevent exploding gradients. Gradients for both the generator and discriminator are computed using separate gradient tapes, ensuring stable training dynamics. The generator loss encourages realistic image generation and penalizes differences from ground truth, while the discriminator loss promotes accurate differentiation between real and synthetic images. Periodic checkpointing saves model weights, allowing the resumption of training in case of interruptions.

To objectively assess the quality of reconstructed images, several metrics are employed. The Peak Signal-to-Noise Ratio (PSNR) measures the ratio between the maximum signal value and noise, with higher values indicating better quality. The Structural Similarity Index (SSIM) evaluates perceived similarity between original and reconstructed images, with values close to 1 indicating higher similarity. The Mean Squared Error (MSE) quantifies the average squared differences between pixel values of original and reconstructed images, with lower values signifying fewer errors. Additionally, the Blind/Reference less Image Spatial Quality Evaluator (BRISQUE) assesses image quality without requiring a reference, while the Natural Image Quality Evaluator (NIQE) measures quality based on statistical properties of natural images, offering an additional perspective on reconstruction fidelity. The implementation is carried out in a TensorFlow environment, leveraging GPU acceleration for efficient training. The dataset includes diverse occlusion types, such as masks, glasses, and hands, to ensure robustness across real-world scenarios. The model is trained for 150 epochs with a batch size of 16, using the Adam optimizer with a learning rate of 0.0002.

IV. Results and Discussions

A. Training Data Test

Implementing classes in code aims to unify and facilitate training and evaluation of generative adversarial network (GAN) models in a TensorFlow environment. In addition, various hyperparameters, training statistics, and other variables are set to track and record information during training. In this process, the class provides a train-step method to run one training step on each batch of data, with gradient calculations performed using two gradient tapes for the generator and discriminator. Gradient clipping is applied to prevent the gradient from soaring, and the result is used to update the weight of both models. Furthermore, there are fit methods that govern model training during a number of epochs and other methods such as generate images, evaluate, and get result to generate, visualize, and evaluate model generation results. Image quality metrics are calculated using the image Comparer method. The control point setting aims to save training progress and provides functionality to load checkpoints if available. In addition, there is also a Tensor Board process used to record training logs, such as generator and discriminator losses, which can help analyze and monitor training in real-time. Overall, classroom classes are designed to simplify and support various aspects of GAN model training and evaluation in a TensorFlow environment.

To evaluate the quality of generative results from the GAN model using one batch of test data (test dataset), the evaluation process begins by taking a batch of test data from the test dataset. Information about the dimensions or tensor size of the input example is printed to the console to provide insight into the input data structure used in the evaluation. As such, these steps ensure that the evaluation is systematic and comprehensive, providing a clear picture of the model's performance in generating realistic, high-quality data. Next, the Improved GAN model is evaluated using the evaluate method, which produces two images: real, which is the actual image of the test dataset, and fake, which is the image generated by the model. The results of metric calculations are then printed on the console to provide quantitative information about the extent to which the model has succeeded in producing quality images. This process provides a holistic picture of the model's performance in producing images similar to actual data from the test dataset, as well as a deeper understanding of its quality based on the evaluation metrics used. In this test using 150 epochs with different occlusion types, the results can be seen in Fig. 4 and Table I.

Fig. 4. Result training 50 epochs, total data train: 27402, dataset face image: CelebA.

This process helps in the monitoring and quality analysis of GAN model generative results during development and training. The results of this evaluation can be used to adjust and improve model architecture, hyperparameters, or training techniques to achieve better performance in producing more realistic images and according to the desired data distribution. Thus, the use of these evaluation metrics provides an objective basis for the assessment and development of the GAN model as a whole.

B. Training Data Test Different Types of Occlusion

The results in Table I demonstrate the effectiveness of the proposed methodology in reconstructing occluded facial images. Each type of occlusion—glasses, glass, hands, and masks—is evaluated using multiple image quality metrics, including PSNR, SSIM, and MSE.

TABLE I. Quantitative Evaluation for Different Types of Occlusion

Type Of Occlusion	PSNR	SSIM	MSE	NIQE	BRISQUE
Glasses	33.106	0.983	7.056	6.329	10.575
Glass	26.465	0.969	16.104	6.108	0.603
Hand	30.147	0.979	10.160	6.026	6.908
Mask	27.972	0.971	16.261	6.213	7.968

The highest performance is observed for the "glasses" occlusion type, with a PSNR value of 33.106, indicating minimal noise in the reconstructed images. The corresponding SSIM value of 0.983 highlights a high degree of structural similarity with the ground truth images, while the MSE of 7.056 confirms the low error rate. This suggests that the system effectively handles occlusions with defined edges and transparent properties. For the "glass" occlusion type, the PSNR value is slightly lower at 26.465, reflecting a moderate level of reconstruction quality. However, the SSIM value remains robust at 0.969, and the MSE of 16.104 indicates acceptable error margins. This could be attributed to the reflective and translucent properties of the glass occlusions, which introduce additional complexity during reconstruction. The "hand" occlusion type achieves a PSNR of 30.147 and an SSIM of 0.979, with an MSE of 10.160. These metrics suggest that the system performs well in reconstructing features occluded by hands, which typically involve irregular shapes and textures. The results indicate that the model is capable of accurately reconstructing occluded areas with varying complexities. Finally, the "mask" occlusion type yields a PSNR value of 27.972 and an SSIM of 0.971, with an MSE of 16.261. While these results are slightly lower than those for "glasses" and "hand," they still demonstrate the system’s ability to handle large, uniform occlusions effectively. Overall, the results in Table I highlight the robustness of the proposed methodology across different occlusion types. The high PSNR and SSIM values, coupled with low MSE scores, validate the effectiveness of the GAN-based framework in reconstructing occluded facial images.

Fig. 5 visually illustrates the performance of the proposed methodology in handling facial images with various occlusion types. The figure includes three groups of images: occluded face images (input), real face images without occlusion (ground truth), and processed images (predictions). The comparison between these groups highlights the model’s ability to reconstruct occluded areas while preserving structural and textural consistency. For the "glasses" occlusion type, the predictions exhibit an impressive level of detail, with reconstructed regions seamlessly blending with the surrounding facial features. This indicates the model's ability to handle transparent and semi-transparent occlusions effectively. The predicted images for the "glass" occlusion type demonstrate notable improvements in reconstructing reflective surfaces, although minor artifacts are occasionally visible, reflecting the inherent challenges of this occlusion type. The "hand" occlusion type, characterized by irregular shapes and textures, showcases the model's robustness in reconstructing facial features obscured by dynamic and complex occlusions. Predicted images display minimal artifacts, with a high degree of alignment to the ground truth. Similarly, the "mask" occlusion type results indicate the model's capacity to reconstruct large, uniform occlusions. The reconstructed images closely align with the ground truth, although slight blurring is observed in some regions.

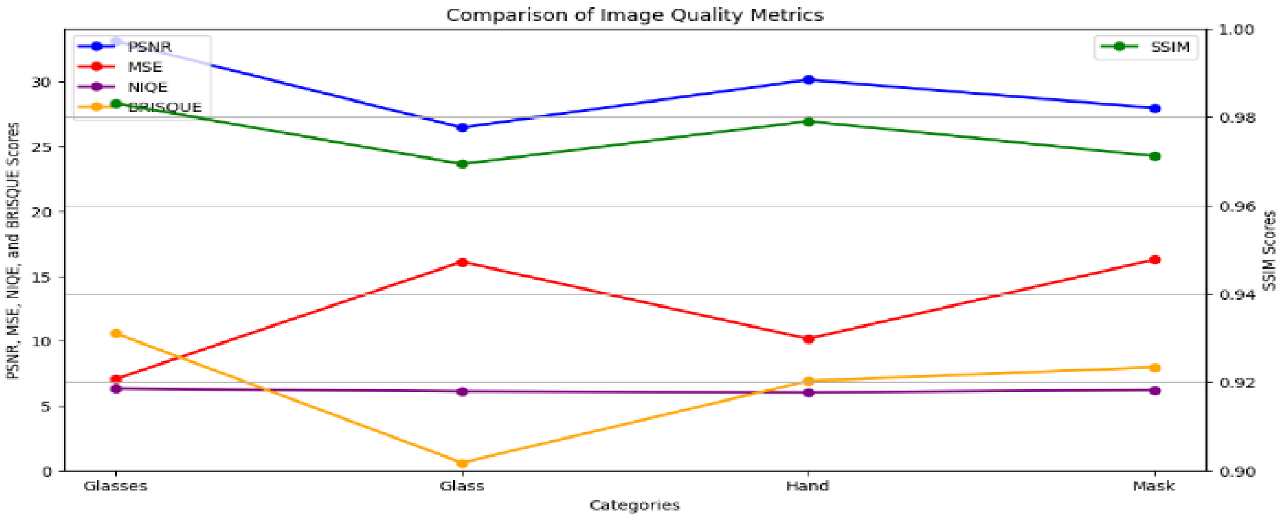


Fig. 5. Graph comparison of image quality metrics.

TABLE II. QUANTITATIVE COMPARISON OF OUR METHODS TO OTHER STATE-OF-THE-ART REPRESENTATIVE METHODS (THE BEST RESULT ARE BOLDFACED).

Type Of Occlusion	Methods	SSIM	PSNR	NIQE	BRISQUE
Glasses	Edge.[1]	0.882	25.641	4.763	36.374
	PConv.[35]	0.896	26.678	4.509	39.680
	GConv.[36]	0.889	26.289	4.598	38.358
	GAN-BN.[37]	0.914	28.878	**4.458**	38.111
	Ours	**0.983**	**33.106**	6.329	**10.575**
Glass	Edge.[1]	0.917	27.919	4.186	34.213
	PConv.[35]	0.940	29.455	4.331	35.514
	GConv.[36]	0.940	29.455	4.853	35.396
	GAN-BN.[37]	0.944	**31.323**	**4.105**	34.38
	Ours	**0.969**	26.465	6.108	**0.603**
Hand	Edge.[1]	0.818	24.911	4.597	31.913
	PConv.[35]	0.863	25.122	4.691	24.603
	GConv.[36]	0.885	26.920	4.929	24.879
	GAN-BN.[37]	0.882	26.948	**4.443**	24.206
	Ours	**0.979**	**30.147**	6.026	**6.908**
Mask	Edge.[1]	0.867	20.873	4.755	41.895
	PConv.[35]	0.869	24.452	4.830	44.976
	GConv.[36]	0.850	22.357	4.573	39.676
	GAN-BN.[37]	0.908	**28.727**	**4.425**	40.883
	Ours	**0.971**	27.972	6.213	**7.968**

Table II provides a comparative evaluation of the proposed methodology against other state-of-the-art methods across different occlusion types, including glasses, glass, hands, and masks. The metrics analyzed include SSIM, PSNR, NIQE, and BRISQUE, which collectively offer a holistic view of the reconstruction quality. For the "glasses" occlusion type, the proposed method achieves the highest SSIM of 0.983 and a PSNR of 33.106, outperforming other methods such as Edge, PConv, GConv, and GAN-BN. Although the NIQE value of 6.329 is slightly higher compared to other methods, the BRISQUE score of 10.575 significantly outperforms the alternatives, highlighting the superior perceptual quality of the reconstructed images. In the "glass" occlusion type, the SSIM value of 0.969 and BRISQUE score of 0.603 stand out as the best among all methods. The slightly lower PSNR of 26.465 compared to GAN-BN (31.323) can be attributed to the reflective properties of glass occlusions, which are inherently challenging to reconstruct. For the "hand" occlusion type, the proposed method demonstrates excellent results with an SSIM of 0.979, a PSNR of 30.147, and a NIQE value of 6.026. The BRISQUE score of 6.908 further supports the method's capability to handle irregular and complex occlusions, outperforming other approaches in perceptual quality. The "mask" occlusion type results show an SSIM of 0.971 and a BRISQUE score of 7.968, both of which are superior to other methods. While the PSNR of 27.972 is slightly lower than GAN-BN's 28.727, the overall performance remains competitive, especially in terms of structural and perceptual quality. Overall, Table II demonstrates the superiority of the proposed methodology in most metrics and occlusion types, particularly in terms of structural similarity (SSIM) and perceptual quality (BRISQUE). These results underscore the robustness and effectiveness of the GAN-based approach in reconstructing occluded facial images across diverse scenarios.

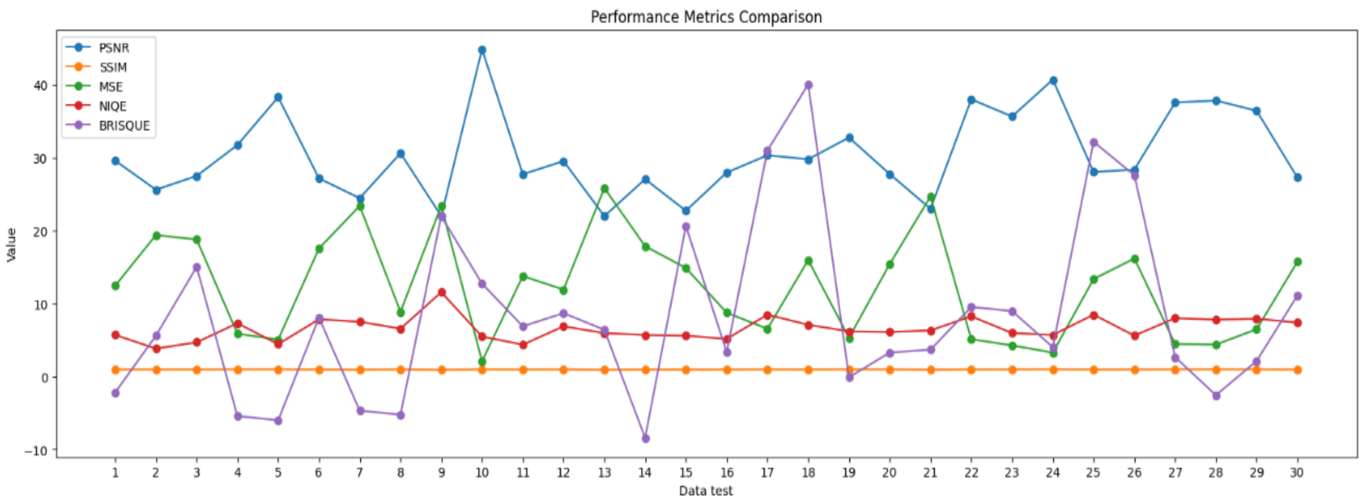


Fig. 6. Performance metrics comparison.

Fig. 6 compares multiple performance metrics across various datasets, providing a comprehensive analysis of the model's consistency and robustness. The metrics displayed include PSNR, SSIM, MSE, NIQE, and BRISQUE, with their trends plotted for visual clarity. The PSNR metric shows relatively stable high values across most datasets, indicating that the reconstructed images maintain a strong signal-to-noise ratio. This stability suggests that the model performs consistently across different occlusion scenarios. Similarly, SSIM values remain consistently high, reflecting the model's ability to maintain structural integrity and similarity to the original images. The MSE metric, which measures reconstruction error, fluctuates slightly more but stays within a low range across datasets. This low error margin underscores the model's precision in reconstructing facial images, even in complex occlusion conditions. The perceptual quality metrics, NIQE and BRISQUE, show slight variability across datasets, which may be attributed to differences in occlusion types and their inherent complexities. However, these values remain within acceptable ranges, demonstrating the model's ability to generate visually appealing reconstructions. Overall, Fig. 6 highlights the robustness and reliability of the proposed methodology. The consistent performance across diverse datasets underscores the model's adaptability to varying occlusion scenarios, making it well-suited for practical applications in facial recognition systems.

V. Conclusion

In this research, an in-depth study is conducted on developing innovative strategies to synthetically reconstruct occlusion-impacted facial images in various scenarios. This research presents a robust and comprehensive methodology to address the challenges of occluded facial image reconstruction by utilising a GAN (Generative Adversarial Network) based framework. The proposed systematic approach effectively incorporates three key components: careful data pre-processing to ensure input quality, sophisticated network architecture design to handle occlusion variations, and rigorous and multidimensional evaluation metrics to holistically measure model performance. As part of the preprocessing stage, this study implemented VGG-Net preprocessing to extract relevant facial features and reduce noise in the input data. VGG-Net, which is known for its ability to capture hierarchical features from images, is used to ensure that the data entering the reconstruction model is optimised and ready for further processing. This stage is crucial, as good input quality can significantly improve the accuracy and reliability of the reconstruction model. By utilising VGG-Net, this research successfully normalises the facial image, adjusts the lighting, and removes artefacts that may interfere with the reconstruction process. The results show the superiority of the developed model in terms of quantitative metrics such as PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index), which indicate the accuracy of the reconstruction, as well as perceptual metrics such as BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator), which assesses the visual quality of the reconstruction results. These advantages are seen in both individual and combined occlusion, which includes various types of obstructions such as the use of glasses, masks, or objects that partially obstruct the face. In addition, the visual fidelity and adaptability of the model were further validated through in-depth comparative analysis against state-of-the-art methods, as well as graphical illustrations demonstrating the model's ability to produce realistic and detailed facial images. To ensure the validity and generalisability of the model, this study uses a variety of diverse datasets, including synthetic datasets and real-world datasets, which include variations in lighting conditions, resolution, and occlusion levels. The evaluation results show that the proposed model is not only consistent in its performance but also able to adapt to complex and challenging scenarios. This research also opens the door for further exploration, including model optimisation to handle dynamic or highly reflective occlusions, as well as integration into broader applications such as public security systems, healthcare, and assistive technologies for individuals with special needs. As such, this research not only makes a significant contribution to the field of facial image reconstruction but also offers relevant practical implications for various sectors of industry and society.

Acknowledgment

The author would like to thank the support provided by the Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM), especially all the tutors involved in completing this article, then Abdurrab University under the Abdurrab Pekanbaru Foundation which has supported this research to obtain results which is satisfying.

References

[1] N. Ud Din, K. Javed, S. Bae, and J. Yi, “A Novel GAN-Based Network for Unmasking of Masked Face,” IEEE Access, vol. 8, pp. 44276–44287, 2020. DOI: 10.1109/ACCESS.2020.2977386.

[2] D. Khas, S. Kumar, and S. K. Singh, “Facial Occlusion Detection and Reconstruction Using GAN,” in Computer Vision and Image Processing, S. K. Singh, P. Roy, B. Raman, and P. Nagabhushan, Eds., Singapore: Springer Singapore, 2021, pp. 255–267.

[3] B. Sugandi, I. Dewita, and R. P. Hudjajanto, "Face recognition based on PCA and neural network," in 2019 International Conference on Information and Communication Technology (ICOIACT), Yogyakarta, Indonesia, 2019, pp. 1-6. DOI: 10.1109/ICOIACT46704.2019.8938537.

[4] Y. Chen, F. Wu, Z. Wang, Y. Song, Y. Ling, and L. Bao, “Self-supervised Learning of Detailed 3D Face Reconstruction,” Oct. 2019. DOI: 10.1109/TIP.2020.3017347.

[5] A. Jabbar et al., “AFD-StackGAN: Automatic Mask Generation Network for Face De-Occlusion Using StackGAN,” Sensors, vol. 22, no. 5, Mar. 2022. DOI: 10.3390/s22051747.

[6] X. Chai, J. Chen, C. Liang, D. Xu, and C.-W. Lin, "Expression-Aware Face Reconstruction Via A Dual-Stream Network," vol. 20. 2020. DOI: 10.1109/ICME46284.2020.9102811.

[7] Y. Chen, R. Xia, K. Yang, and K. Zou, “DGCA: High Resolution Image Inpainting via DR-GAN and Contextual Attention,” Multimed Tools Appl, vol. 82, no. 30, pp. 47751–47771, Dec. 2023. DOI: 10.1007/s11042-023-15313-0.

[8] W. Zheng, C. Gou, and F. Y. Wang, “A Novel Approach Inspired by Optic Nerve Characteristics for Few-Shot Occluded Face Recognition,” Neurocomputing, vol. 376, pp. 25–41, Feb. 2020. DOI: 10.1016/j.neucom.2019.09.045.

[9] M. Cipriano et al., “Deep Segmentation of the Mandibular Canal: A New 3D Annotated Dataset of CBCT Volumes,” IEEE Access, vol. 10, pp. 11500–11510, 2022. DOI: 10.1109/ACCESS.2022.3144840.

[10] Duan, Q., & Zhang, L. (2021). Look more into occlusion: Realistic face frontalization and recognition with BoostGAN. IEEE Transactions on Neural Networks and Learning Systems, 32(4), 1737–1751. https://doi.org/10.1109/TNNLS.2020.2976700.

[11] Jagtap, V. Kangale, K. Unune, and P. Gosavi, “A Study of LBPH, Eigenface, Fisherface and Haar-like Features for Face Recognition Using OpenCV,” 2019. DOI: 10.1109/ISS1.2019.8907965.

[12] A. Jabbar, X. Li, M. M. Iqbal, and A. J. Malik, “FD-StackGAN: Face De-Occlusion Using Stacked Generative Adversarial Networks,” KSII Transactions on Internet and Information Systems, vol. 15, no. 7, pp. 2547–2567, Jul. 2021. DOI: 10.3837/tiis.2021.07.014.

[13] S. Ge, C. Li, S. Zhao, and D. Zeng, “Occluded Face Recognition in the Wild by Identity-Diversity Inpainting,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 10, pp. 3387–3397, Oct. 2020. DOI: 10.1109/TCSVT.2020.2967754.

[14] T. Kim, Y. Cho, D. Kim, M. Chang, and Y. J. Kim, “Tooth Segmentation of 3D Scan Data Using Generative Adversarial Networks,” Applied Sciences, vol. 10, no. 2, Jan. 2020. DOI: 10.3390/app10020490.

[15] R. Biswas, V. González-Castro, E. Fidalgo, and E. Alegre, “A New Perceptual Hashing Method for Verification and Identity Classification of Occluded Faces,” Image Vision Comput, vol. 113, Sep. 2021. DOI: 10.1016/j.imavis.2021.104245.

[16] Y. Lu, S. Wang, W. Zhao, and Y. Zhao, “WGAN-Based Robust Occluded Facial Expression Recognition,” IEEE Access, vol. 7, pp. 93594–93610, 2019. DOI: 10.1109/ACCESS.2019.2928125

[17] S. Alfattama, P. Kanungo, and S. K. Bisoy, “Face Recognition from Partial Face Data,” in 2021 International Conference in Advances in Power, Signal, and Information Technology, 2021. DOI: 10.1109/APSIT52773.2021.9641286.

[18] Z. Chen, Y. Wang, T. Guan, L. Xu, and W. Liu, “Transformer-Based 3D Face Reconstruction with End-to-End Shape-Preserved Domain Transfer,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 12, pp. 8383–8393, Dec. 2022. DOI: 10.1109/TCSVT.2022.3192422.

[19] S. Y. Chen, Y. K. Lai, S. Xia, P. L. Rosin, and L. Gao, “3D Face Reconstruction and Gaze Tracking in the HMD for Virtual Interaction,” IEEE Transactions on Multimedia, vol. 25, pp. 3166–3179, 2023. DOI: 10.1109/TMM.2022.3156820.

[20] X. Zhong, X. Qu, and C. Chen, “High-Quality Face Image Super-Resolution Based on Generative Adversarial Networks,” 2019. DOI: 10.1109/IAEAC47372.2019.8998075.

[21] L. He, H. Li, Q. Zhang, and Z. Sun, “Dynamic Feature Matching for Partial Face Recognition,” IEEE Transactions on Image Processing, vol. 28, no. 2, pp. 791–802, Feb. 2019. DOI: 10.1109/TIP.2018.2870946.

[22] W. Zhiqiang, Z. Lejun, Z. Lifeng, and S. Serikawa, “Research on Image Privacy Protection Algorithm Based on Generative Adversarial Network,” in 2020 International Conference on Machine Learning, Big Data and Business Intelligence, 2020. DOI: 10.1109/MLBDBI51377.2020.00105.

[23] J. Caba, J. Barba, F. Rincón, J. A. de la Torre, S. Escolar, and J. C. López, “Hyperspectral Face Recognition with Adaptive and Parallel SVMs in Partially Hidden Face Scenarios,” Sensors, vol. 22, no. 19, Oct. 2022. DOI: 10.3390/s22197641.

[24] J. Dong, L. Zhang, H. Zhang, and W. Liu, “Occlusion-Aware GAN for Face De-Occlusion in the Wild,” 2020. DOI: 10.1109/ICME46284.2020.9102788.

[25] A. Y. A. Maghari, “Recognition of Partially Occluded Faces Using Regularized ICA,” Inverse Problems in Science and Engineering, vol. 29, no. 8, pp. 1158–1177, 2021. DOI: 10.1080/17415977.2020.1845329.

[26] X. Li, C. Shao, Y. Zhou, and L. Huang, “Face Mask Removal Based on Generative Adversarial Network and Texture Network,” in 2021 4th International Conference on Robotics, Control and Automation Engineering, 2021. DOI: 10.1109/RCAE53607.2021.9638866.

[27] B. Hariharan, S. Karthic, S. Indra Priyadharshini, E. Nalina, N. R. Wilfred Blessing, and P. N. Senthil Prakash, “Hybrid Deep Convolutional Generative Adversarial Networks (DCGANS) and Style Generative Adversarial Network (STYLEGANS) Algorithms to Improve Image Quality,” in 2022 3rd International Conference on Electronics and Sustainable Communication Systems, 2022. DOI: 10.1109/ICESC54411.2022.9885611.

[28] X. Dong and R. Hua, “GAN Based Image Inpainting Methods: A Taxonomy,” in 2022 3rd International Conference on Electronic Communication and Artificial Intelligence, 2022. DOI: 10.1109/IWECAI55315.2022.00037.

[29] X. Yuan and I. K. Park, “Face De-Occlusion Using 3D Morphable Model and Generative Adversarial Network,” Apr. 2019. DOI: 10.1109/ICESC54411.2022.9885611.

[30] D. Poux, B. Allaert, N. Ihaddadene, I. M. Bilasco, C. Djeraba, and M. Bennamoun, “Dynamic Facial Expression Recognition under Partial Occlusion with Optical Flow Reconstruction,” IEEE Transactions on Image Processing, vol. 31, pp. 446–457, 2022. DOI: 10.1109/TIP.2021.3129120.

[31] C. Rong, X. Zhang, and Y. Lin, “Feature-Improving Generative Adversarial Network for Face Frontalization,” IEEE Access, vol. 8, pp. 68842–68851, 2020. DOI: 10.1109/ACCESS.2020.2986079

[32] P. Ruiu, A. Lagorio, M. Cadoni, and E. Grosso, “Enhancing eID Card Mobile-Based Authentication through 3D Facial Reconstruction,” Journal of Information Security and Applications, vol. 77, 2023. DOI: 10.1016/j.jisa.2023.103577.

[33] A. Lattas, S. Moschoglou, S. Ploumpis, B. Gecer, A. Ghosh, and S. Zafeiriou, “AvatarMe++: Facial Shape and BRDF Inference with Photorealistic Rendering-Aware GANs,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 12, pp. 9269–9284, 2022. DOI: 10.1109/TPAMI.2021.3125598.

[34] C. Wang, Q. Zhang, W. Liu, Y. Liu, and L. Miao, “Facial Feature Discovery for Ethnicity Recognition,” Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 9, no. 1, 2019. DOI: 10.1002/widm.1278.

[35] K. Nazeri, E. Ng, T. Joseph, F. Z. Qureshi, and M. Ebrahimi, “EdgeConnect: Generative Image Inpainting with Adversarial Edge Learning,” Jan. 2019, [Online]. Available: http://arxiv.org/abs/1901.00212

[36] G. Liu, F. A. Reda, K. J. Shih, T.-C. Wang, A. Tao, and B. Catanzaro, “Image Inpainting for Irregular Holes Using Partial Convolutions,” Apr. 2018, [Online]. Available: http://arxiv.org/abs/1804.07723

[37] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. Huang, “Free-Form Image Inpainting with Gated Convolution,” Jun. 2018, [Online]. Available: http://arxiv.org/abs/1806.03589