Image reconstruction by using the occluded scenes

[1] CHETANA TUKKOJI, [2] VEERA SAIVA SAI LAKSHMI KALYAN , [3]PENJURI LOHITH SAI , [4]T PRADEEP KUMMAR REDDY , [5]B M SRINIVAS
[1]Assistant Professor, GITAM Deemed to be University, Bangalore
2, 3, 4,5, B.Tech students of GITAM Deemed to be University, Bangalore

1. Abstract:

In computer vision and image processing, picture reconstruction of obscured situations is an essential procedure. In order to greatly improve scene understanding, this research focuses on creating sophisticated algorithms and techniques that can precisely rebuild buried or obscured areas of an image. Our goal is to improve object detection and categorization by utilizing machine learning and computer vision techniques to fill in the gaps in visual information and produce a more complete picture of the scene. This work tackles practical issues where scenes could be partially concealed, offering possible uses in fields such as robotics, autonomous systems, healthcare, and surveillance. By improving the quality and interpretability of photos, the results of this study hope to have a substantial impact on a number of industries and advance technologies that rely on visual data.

Keywords: Computer Vision, Object recognition, Image reconstruction, Machine Learning.

2. Introduction:

In the quest to develop resilient computer vision systems, the imperative task of reconstructing high-quality images from partially obscured scenes takes precedence. Occlusions, arising from diverse factors such as foreground objects, fluctuating lighting conditions, or limitations in sensors, present a formidable challenge to conventional image reconstruction methods. Successfully overcoming these challenges is pivotal for applications such as object recognition, scene understanding, and 3D reconstruction, where achieving a comprehensive and accurate representation of the environment holds paramount importance.

The impetus for this research arises from the discerned limitations within existing image reconstruction techniques when confronted with scenes containing occlusions. Contemporary methodologies often grapple with the effective handling of occlusions, resulting in reconstructions that are either incomplete or distorted. This shortfall impedes the deployment of computer vision systems in real-world scenarios where occlusions are pervasive. Addressing this critical gap is essential to unleash the full potential of computer vision technologies across diverse applications.

3. Literature survey:

3.1 Xu, J., Li, et.al [1], “Occluded Scene Classification Via Cascade Supervised Contrastive Learning”.
In this paper author introduced a novel approach named cascade supervised contrastive learning to tackle the challenge of occluded scene classification. Leveraging this technique, they significantly improve classification accuracy even in environments with visual obstructions. The paper's focus on occlusion is essential for remote sensing

applications where scenes may often be partially obscured. The cascade supervised contrastive learning approach demonstrates promise in enhancing the robustness and reliability of automated scene classification, vital for numerous applications in Earth and environmental sciences.

Research gap: this work holds significance in advancing classification methodologies for remote sensing imagery, contributing to a better understanding and analysis of complex scenes in Earth observations.

3.3 **Title:** Occlusion-Free Scene Recovery via Neural Radiance Fields.
Authors: Zhu, C., Wan, R., Tang, Y., & Shi, B.
Year: 2023
Summary of the paper: Zhu, Wan, Tang, and Shi (2023) present a cutting-edge approach named Neural Radiance Fields (NRF) to address occlusion challenges in scene recovery. Their work, published in the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, focuses on occlusion-free scene recovery, a fundamental aspect in computer vision. The utilization of NRF in scene recovery represents a significant advancement, showcasing potential in overcoming occlusion complexities that can hinder accurate scene reconstruction.

Research gap: This research contributes to the realm of computer vision by enhancing the reconstruction capabilities, offering a pathway towards more comprehensive scene understanding, crucial in various applications such as robotics, augmented reality, and 3D modeling.

3.3 **Title:** AccNet: occluded scene text enhancing network with accretion blocks.
Authors: Gong, Y., Zhang, Z., Duan, G., Ma, Z., & Xie, M.
Year: 2023
Summary of the paper: In their paper "AccNet: Occluded Scene Text Enhancing Network with Accretion Blocks" published in Machine Vision and Applications, Gong, Zhang, Duan, Ma, and Xie (2023) introduce a novel approach to enhance occluded scene text. The proposed AccNet leverages accretion blocks, showcasing a promising solution to mitigate challenges posed by occlusions in scene text. This research is a notable contribution to the field of machine vision, particularly in improving the legibility and accuracy of scene text recognition, which is fundamental in applications like optical character recognition (OCR), autonomous vehicles, and visual assistance systems.
Research gap: The utilization of accretion blocks demonstrates its potential in addressing occlusion-related issues, signifying a step forward in advancing the robustness and efficacy of text-related vision tasks.
3.4 **Title:** A large depth-of-field virtual measurement network for non-cooperative 6DOF pose estimation in occlusion scenes.
Authors: Deng, Z., & Li, A.
Year: 2023
Summary of the paper: In their recent work titled "A Large Depth-of-Field Virtual Measurement Network for Non-cooperative 6DOF Pose Estimation in Occlusion Scenes" published in Measurement, Deng and Li (2023) present a novel approach to tackle the challenges of occlusion in 6DOF (Six Degrees of Freedom) pose estimation. The proposed large depth-of-field virtual measurement network offers a significant advancement, enabling accurate pose estimation in non-

cooperative settings, even amidst occlusion. This innovation has the potential to impact various domains, including robotics, computer vision, and automation, where precise pose estimation is critical for task execution and decision-making.
Research gap: By focusing on addressing occlusion challenges and enhancing accuracy in pose estimation, this research contributes to the broader goal of improving the performance and reliability of vision-based systems in complex real-world scenarios.

3.5 **Title:** Occlusion-aware human mesh model-based gait recognition.
Authors: Xu, C., Makihara, Y., Li, X., & Yagi, Y.
Year: 2023
Summary of the paper: In their recent publication "Occlusion-aware Human Mesh Model-based Gait Recognition" in IEEE Transactions on Information Forensics and Security, Xu, Makihara, Li, and Yagi (2023) propose an innovative approach to gait recognition. Leveraging occlusion-aware human mesh models, the authors aim to enhance the accuracy and robustness of gait recognition systems. The introduction of human mesh models provides a promising avenue for more accurate gait analysis, leading to potential advancements in security and biometric authentication systems.
Research gap: This work is pivotal in pushing the boundaries of gait recognition technology, making strides towards more reliable and effective recognition methods that can operate effectively in various complex and dynamic environments.

3.6 **Title:** 3D integral imaging depth estimation of partially occluded objects using mutual information and Bayesian optimization.
Authors: Wani, P., & Javidi, B.
Year: 2023
Summary of the paper: In their recent study published in Optics Express titled "3D Integral Imaging Depth Estimation of Partially Occluded Objects using Mutual Information and Bayesian Optimization," Wani and Javidi (2023) present an innovative approach to depth estimation in 3D integral imaging. The research focuses on accurately estimating the depth of partially occluded objects within the integral imaging framework. By integrating mutual information and Bayesian optimization techniques, the proposed method aims to enhance the accuracy and robustness of depth estimation, particularly in challenging scenarios where objects are partially obscured.
Research gap: This work contributes significantly to advancing the field of 3D imaging and computational methods, offering potential applications in various domains, including computer vision, robotics, and medical imaging.

3.7 **Title:** Occlusion-aware light field depth estimation with view attention.
Authors: Wang, X., Tao, C., & Zheng, Z.
Year: 2023
Summary of the paper: In their recent paper "Occlusion-aware Light Field Depth Estimation with View Attention" published in Optics and Lasers in Engineering (2023), Wang, Tao, and Zheng propose an innovative approach to depth estimation in light field imagery. The study focuses on addressing occlusion challenges by incorporating view attention mechanisms into the depth estimation process. By utilizing light field data and considering occlusions, their approach enhances the accuracy of depth estimation.
Research gap: This work holds promise for various applications such as virtual reality, robotics, and 3D scene reconstruction, where accurate depth perception is essential.

3.8 **Title:** I See-Through You: A Framework for Removing Foreground Occlusion in Both Sparse and Dense Light Field Images.
Authors: Hur, J., Lee, J. Y., Choi, J., & Kim, J.
Year: 2023
Summary of the paper: In their paper "I See-Through You: A Framework for Removing Foreground Occlusion in Both Sparse and Dense Light Field Images" presented at the IEEE/CVF Winter Conference on Applications of Computer Vision (2023), Hur, Lee, Choi, and Kim propose an innovative framework to address the challenge of foreground occlusion in light field imagery. The framework is designed to effectively remove occlusions from both sparse and dense light field images.
Research gap: This work has implications in various domains, including augmented reality, virtual reality, and computational photography, where occlusion-free visualization is crucial.

3.9 **Title:** Self-supervised learning of scene flow with occlusion handling through feature masking.
Authors: Xiang, X., Abdein, R., Lv, N., & Yang, J.
Year: 2023
Summary of the paper: In their recent research titled "Self-supervised Learning of Scene Flow with Occlusion Handling through Feature Masking" proposed an innovative approach to self-supervised learning in the context of scene flow estimation. Addressing the challenge of occlusions in scene flow analysis, the authors introduce feature masking as a mechanism to handle occluded regions effectively. By leveraging self-supervised learning techniques, the proposed model learns to estimate scene flow while considering occlusions, contributing to more accurate and robust results.
Research gap: This work is significant in the field of computer vision and scene analysis, enhancing the understanding and interpretation of dynamic scenes by accounting for occlusions, a critical aspect in real-world applications such as autonomous driving and robotics.

3.10 **Title:** 3D object detection through fog and occlusion: passive integral imaging vs active (LiDAR) sensing.
Authors: Usmani, K., O'Connor, T., Wani, P., & Javidi, B.
Year: 2023
Summary of the paper: In their recent study "3D Object Detection Through Fog and Occlusion: Passive Integral Imaging vs Active (LiDAR) Sensing" published in Optics Express (2023), Usmani, O'Connor, Wani, and Javidi investigate 3D object detection under challenging environmental conditions, including fog and occlusion. They compare passive integral imaging with active LiDAR sensing, aiming to assess the performance of these approaches in adverse scenarios.
Research gap: The research sheds light on the effectiveness of both methods in capturing 3D information of objects despite challenging factors, emphasizing the potential of passive integral imaging in foggy and occluded settings.

4. Objectives of the project:

4.1 Enhanced Scene Understanding: Develop algorithms to accurately reconstruct occluded scenes in images, improving the understanding of complex scenes even when parts are obscured.

4.2 Improved Visual Representations: Create techniques that enhance visual representations by filling in missing or occluded regions, resulting in a more complete and informative image.

4.3 Robust Object Recognition: Build models that aid in object recognition and classification, even in scenarios where objects may be partially hidden or occluded.

4.4 Enhanced Computer Vision Applications: Enable better performance for computer vision applications such as object tracking, augmented reality, and autonomous navigation by providing a clearer and more complete visual input.

4.5 Real-world Applications and Impact: Apply the reconstructed images to various domains like healthcare (medical image analysis), surveillance, robotics, and autonomous vehicles, ultimately contributing to safer and more efficient systems in these domains.

5. Steps for image reconstruction:

5.1 Data Collection:
Collect a diverse dataset of occluded scenes, ensuring it covers different types of occlusions and varying levels of complexity.

5.2 Preprocessing:
Implement preprocessing steps to clean and enhance the collected data, which may include noise reduction, contrast enhancement, and resolution adjustment.

5.3 Feature Extraction:
Extract relevant features from the preprocessed data that are vital for understanding occlusions and scene context.

5.4 Occlusion Detection:
Develop algorithms to detect occluded regions within the image, employing techniques like edge detection, color analysis, or deep learning-based approaches.

5.5 Scene Understanding:
Utilize machine learning models to understand the un-occluded regions and overall scene structure, considering context and semantic information.

5.6 Reconstruction Algorithms:
Implement reconstruction algorithms that use the detected occlusions and scene understanding to fill in the occluded parts accurately.

5.7 Integration of Deep Learning:
Incorporate deep learning architectures like convolutional neural networks (CNNs) for robust feature extraction and scene reconstruction, utilizing their ability to learn complex patterns.

5.8 Testing and Validation:
Evaluate the performance of the reconstruction system using appropriate metrics, ensuring accurate restoration of occluded regions.

5.9 Optimization:
Optimize the algorithms and processes to enhance the speed and efficiency of image reconstruction, crucial for real-time applications.

5.10 Scalability and Adaptability: Design the system to be scalable and adaptable to varying occlusion types, scene complexities, and input data formats, allowing for broader applicability across different domains and use cases.

6. System Architecture: Data flow diagram:

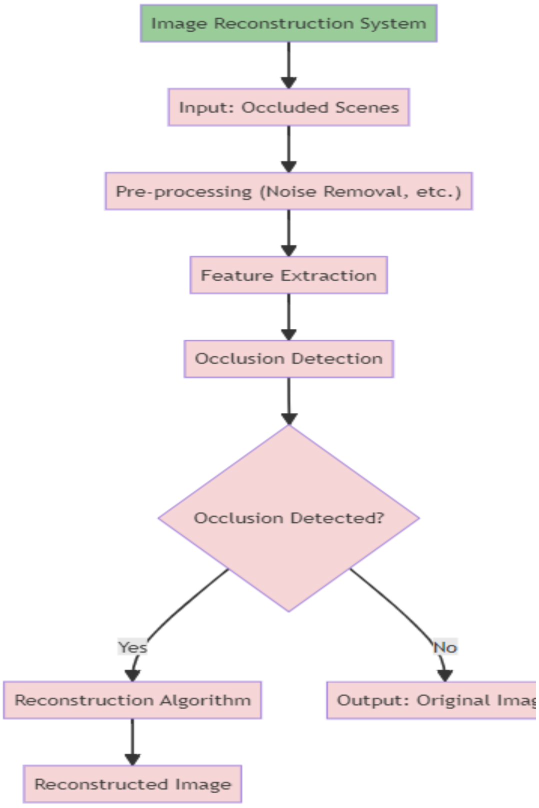


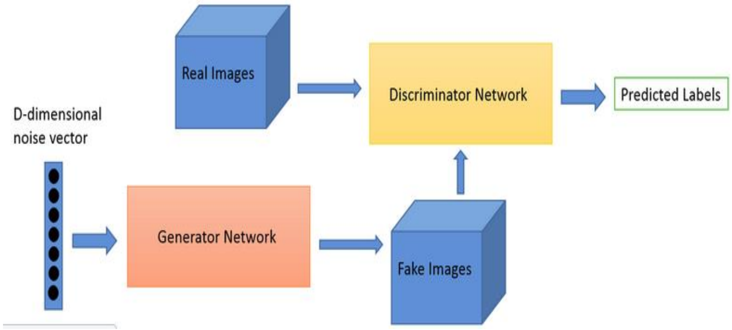
Figure 1:- Architecture diagram

The above flowchart illustrates the working architecture of "IMAGE

7. METHODOLOGY:

Generative Adversarial Networks (GANs) were developed in 2014 by Ian Goodfellow and his teammates. GAN is basically an approach to generative modeling that generates a new set of data based on training data that look like training data. GANs have two main blocks (two neural networks) which compete with each other and are able to capture, copy, and analyze the variations in a dataset. The two models are usually called Generator and Discriminator which we will cover in Components on GANs

FLOW CHART OF WORKING PROCESS OF GAN's ALGORITHM:



Steps to Implement Basic GAN:

1. Importing all libraries
2. Getting the Dataset
3. Data Preparation – It includes various steps to accomplish like preprocessing data, scaling, flattening, and reshaping the data.
4. Define the function Generator and Discriminator.
5. Create a Random Noise and then create an Image with Random Noise.
6. Setting Parameters like defining epoch, batch size, and Sample size.
7. Define the function of generating Sample Images.
8. Train Discriminator then trains Generator and it will create Images.
9. Will see what clarity of Images is created by Generator.

8. RESULTS:

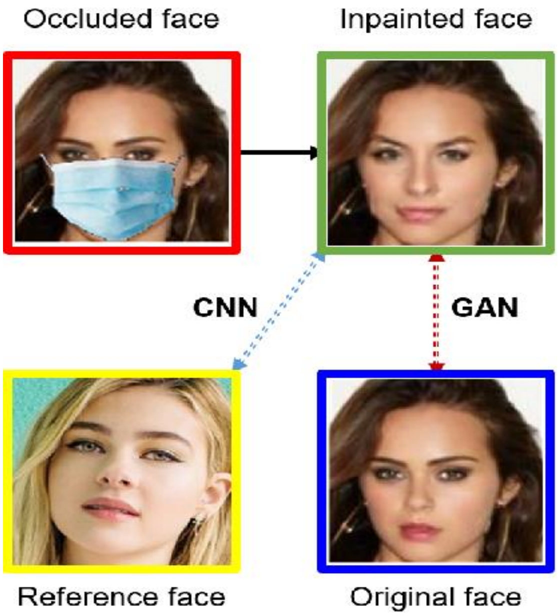


Figure 2:-IMAGE RECONSTRUCTION

Figure 3:- INCOMPLETE IMG WITH RECONSTRUCTION

9. CONCLUSION:

In summary, the implementation of "Generative Adversarial Network (GAN)" can improve image reconstruction employing obscured scenes and provide approximate research outcomes compared to other types of Deep Neural Networks.

As a probabilistic model, the GAN algorithm improves the reconstruction model's correctness and raises the likelihood of producing the right images for use in the fields of health care, surveillance, and automobile industries, among others.

10. References:

[10.1] Xu, J., Li, Y., Shi, Q., & He, L. (2023). Occluded Scene Classification Via Cascade Supervised Contrastive Learning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

[10.2] Zhu, C., Wan, R., Tang, Y., & Shi, B. (2023). Occlusion-Free Scene Recovery via Neural Radiance Fields. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 20722-20731).

[10.3] Gong, Y., Zhang, Z., Duan, G., Ma, Z., & Xie, M. (2023). AccNet: occluded scene text enhancing network with accretion blocks. Machine Vision and Applications, 34(1).

[10.4] Deng, Z., & Li, A. (2023). A large depth-of-field virtual measurement network for non-cooperative 6DOF pose estimation in occlusion scenes. Measurement, 113111.

[10.5] Xu, C., Makihara, Y., Li, X., & Yagi, Y. (2023). Occlusion-aware human mesh model-based gait recognition. IEEE transactions on information forensics and security, 18, 1309-1321.

[10.6] Wani, P., & Javidi, B. (2023). 3D integral imaging depth estimation of

partially occluded objects using mutual information and Bayesian optimization. Optics Express, 31(14), 22863-22884.
[10.7] Wang, X., Tao, C., & Zheng, Z. (2023). Occlusion-aware light field depth estimation with view attention. Optics and Lasers in Engineering, 160, 107299.
[10.8] Hur, J., Lee, J. Y., Choi, J., & Kim, J. (2023). I See-Through You: A Framework for Removing Foreground Occlusion in Both Sparse and Dense Light Field Images. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 229-238).
[10.9] Xiang, X., Abdein, R., Lv, N., & Yang, J. (2023). Self-supervised learning of scene flow with occlusion handling through feature masking. Pattern Recognition, 139, 109487.
[10.10] Usmani, K., O'Connor, T., Wani, P., & Javidi, B. (2023). 3D object detection through fog and occlusion: passive integral imaging vs active (LiDAR) sensing. Optics Express, 31(1), 479-491.
[10.11] J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.
[10.12] M. S. Bartlett, G. Littlewort, M. G. Frank, C. Lainscsek, I. Fasel, and J. R. Movellan, "Automatic Recognition of Facial Actions in Spontaneous Expressions," Journal of Multimedia, 2006.
[10.13] I. J. Goodfellow, D. Erhan, P. L. Carrier, A. Courville, M. Mirza, B. Hamner, W. Cukierski, Y. Tang, D. Thaler, D.-H. Lee, et al., "Challenges in Representation Learning: A report on three machine learning contests," Neural Networks, 2013.
[10.14]. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, 2015.
[10.15]. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv:1409.1556, 2014.
[10.16]. A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), 2016.
[10.17]. X. Liu, J. Song, Y. Wang, Y. Zhou, S. Wang, and T. Tan, "Deeply Learning Deformable Facial Action Parts Model for Dynamic Expression Analysis," IEEE Transactions on Image Processing, 2018.
[10.18]. Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep Learning Face Attributes in the Wild," Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015.
[10.19]. D. Kollias, P. Tzirakis, S. Zafeiriou, and M. Pantic, "A Multi-Task Learning Approach for Facial Expression Recognition," Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2019.
[10.20]. M. M. Kacem, B. Fertil, and A. Benazza-Benyahia, "A Hybrid Deep Architecture for Facial Expression Recognition," Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), 2017.