Original Article
ISSN (Online): 2582-7472

ShodhKosh: Journal of Visual and Performing Arts
May 2024 5(5), 431–438

ADVANCES IN COMPUTER VISION: NEW HORIZONS AND ONGOING CHALLENGES

Rahul Kumar Majhi[1], Dr. Akhilesh A. Waoo[2]

[1] Department of Computer Science and Information Technology, AKS University, Satna, India

Corresponding Author
Akhilesh A. Waoo,
akhileshwaoo@gmail.com

DOI
10.29121/shodhkosh.v5.i5.2024.1893

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ABSTRACT

Computer vision, a rapidly evolving field at the intersection of computer science and artificial intelligence, has witnessed unprecedented growth in recent years. This comprehensive review paper provides an overview of the advancements and challenges in computer vision, synthesizing the latest research findings, methodologies, and applications. We explore the historical evolution of computer vision and discuss recent advancements in algorithms and techniques, including deep learning models such as convolutional neural networks (CNNs) and generative adversarial networks (GANs). Diverse applications of computer vision across domains such as healthcare, autonomous vehicles, surveillance, and augmented reality are also examined. Despite remarkable progress, computer vision faces significant challenges, including robustness to adversarial attacks, interpretability, ethical considerations, and regulatory compliance. We discuss these challenges in-depth and highlight the importance of interdisciplinary collaboration in addressing them. Additionally, recent trends and future directions in computer vision research, such as self-supervised learning and explainable AI, are identified. By synthesizing insights from academic research and industrial developments, this review paper aims to provide a comprehensive understanding of the current landscape of computer vision and guide future research endeavors.

Keywords: Computer Vision Deep Learning, Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), Image Recognition, Explainable AI, Semantic Segmentation.

1. INTRODUCTION

In recent years, computer vision has emerged as a transformative field at the intersection of computer science, artificial intelligence, and image processing. With the proliferation of digital imagery and the rapid advancement of machine learning techniques, computer vision has revolutionized various sectors, including healthcare, autonomous vehicles, surveillance, and augmented reality. This review paper aims to provide a comprehensive overview of the advancements and challenges in computer vision, synthesizing the latest research findings, methodologies, and applications.

The exponential growth of computational power, coupled with the availability of large-scale annotated datasets, has propelled the development of deep learning models for image recognition, object detection, segmentation, and image generation tasks. Convolutional neural networks (CNNs) have emerged as the cornerstone of modern computer vision systems, achieving remarkable accuracy in various visual recognition tasks. Furthermore, the evolution of generative adversarial networks (GANs)[1] and transformer architectures has enabled novel approaches to image synthesis, style transfer, and image captioning, pushing the boundaries of visual content generation.

While the field of computer vision has witnessed unprecedented progress, it also faces significant challenges that warrant careful consideration. One such challenge is the robustness and generalization of deep learning models in real-world scenarios. Despite achieving impressive performance on benchmark datasets, these models often struggle with adversarial attacks, domain shifts, and variations in lighting, viewpoint, and occlusion. Additionally, ethical concerns surrounding privacy, bias, and accountability in computer vision algorithms have garnered increasing attention from researchers, policymakers, and the public.

Moreover, the deployment of computer vision systems in safety-critical applications, such as autonomous vehicles and medical diagnosis, necessitates robustness, interpretability, and regulatory compliance. Ensuring the reliability and fairness of these systems requires interdisciplinary collaboration between computer scientists, ethicists, domain experts, and policymakers.

In light of these advancements and challenges, this review paper provides a structured examination of the state-of-the-art techniques, methodologies, and applications in computer vision. By synthesizing insights from academic research, industrial developments, and societal implications, this review aims to provide a comprehensive understanding of the current landscape of computer vision and chart a course for future research directions.

2. HISTORICAL DEVELOPMENT

Computer vision, rooted in the ambition to replicate human visual perception through machines, has a rich history marked by significant milestones and breakthroughs. Dating back to the mid-20th century, the field initially focused on basic image processing tasks, gradually evolving into a sophisticated interdisciplinary domain at the nexus of computer science, artificial intelligence, and neuroscience.

Early Foundations (1950s-1960s)

The seminal work of neuroscientists David Marr and Hubel & Wiesel provided foundational insights into the structure and function of the visual cortex, inspiring early computational models of visual processing.

The development of edge detection algorithms by researchers such as Roberts, Sobel, and Canny laid the groundwork for basic image analysis techniques.

Shape Representation and Object Recognition (1970s-1980s)

In the 1970s, researchers began exploring methods for representing and recognizing shapes in images, leading to the development of techniques like the Hough transform for line detection and region-based segmentation methods.

The 1980s saw the emergence of rule-based expert systems for object recognition, such as Rensselaer's AI System (RAIS) and the Generalized Hough Transform[2], which provided rudimentary capabilities for identifying objects in images.

Early Neural Network Approaches (1980s-1990s)

The resurgence of interest in artificial neural networks (ANNs) in the 1980s sparked research into neural network-based approaches to computer vision. Pioneering works by Fukushima (Neocognitron)[3] and LeCun et al. (LeNet)[2] introduced Convolutional neural network (CNN) architectures for image recognition tasks.

The 1990s witnessed further advancements in neural network-based techniques, with breakthroughs in areas such as optical character recognition (OCR) and handwriting recognition.

Rise of Feature-Based Methods (1990s-2000s)

Feature-based methods, including SIFT (Scale-Invariant Feature Transform)[4] and SURF (Speeded-Up Robust Features)[5], gained popularity for their robustness to variations in scale, rotation, and illumination.

The development of bag-of-words models[6] and the introduction of the Visual Geometry Group (VGG)[7] descriptor contributed to significant improvements in object recognition and image classification.

Deep Learning Revolution (2010s-present)

The advent of deep learning, fueled by advances in computational power and the availability of large-scale labeled datasets, revolutionized the field of computer vision.

Breakthrough architectures such as AlexNet[8], VGGNet[9], GoogLeNet[10], and ResNet[11] surpassed traditional methods in image classification accuracy, paving the way for deep learning's dominance in various visual recognition tasks.

Recent years have witnessed the proliferation of deep learning techniques across diverse applications, including object detection, semantic segmentation, image generation, and video analysis, marking a new era of unprecedented progress in computer vision.

3. ADVANCEMENTS IN ALGORITHMS AND TECHNIQUES

In recent years, the field of computer vision has experienced a transformative shift propelled by breakthroughs in algorithms and techniques. This section delves into the forefront of advancements that have revolutionized how machines perceive and understand visual information.

A. **Deep Learning Paradigm Shift:** The advent of deep learning has fundamentally transformed computer vision, enabling unprecedented accuracy and performance across a myriad of tasks. Convolutional Neural Networks (CNNs)[12], inspired by the visual cortex of animals, have emerged as the cornerstone of modern computer vision systems. Through successive layers of convolution, pooling, and nonlinear[13] activations, CNNs can automatically learn hierarchical representations of visual features directly from raw pixel data.

B. **Architectural Innovations:** Architectural innovations have played a pivotal role in enhancing the efficiency and effectiveness of deep learning models. Variants of CNN architectures, such as Residual Networks (ResNets), DenseNets[14], and Inception Networks[15], have introduced novel connectivity patterns and feature aggregation techniques to alleviate the vanishing gradient problem and facilitate feature reuse across network depths. Moreover, attention mechanisms, initially popularized in natural language processing, have been seamlessly integrated into vision models to dynamically focus on salient regions of input data, enhancing discriminative capability and spatial awareness.

C. **Transfer Learning and Pre-trained Models:** Transfer learning has emerged as a crucial strategy for leveraging large-scale pre-trained models to bootstrap performance on downstream tasks with limited labeled data. Pre-trained models, such as ImageNet-trained architectures (e.g., VGG, ResNet, and EfficientNet[16]), serve as powerful feature extractors, capturing rich visual representations that generalize well across diverse datasets and domains. Fine-tuning these models on target tasks allows for rapid convergence and improved generalization performance, significantly reducing the need for extensive annotated data.

D. **Generative Adversarial Networks (GANs) and Image Synthesis:** Generative Adversarial Networks (GANs) have opened new frontiers in image synthesis and manipulation, enabling the generation of photorealistic images from noise vectors. By training a generator network to produce realistic samples that deceive a discriminator network, GANs can produce high-quality, diverse outputs across various domains, including art generation, style transfer, super-resolution, and data augmentation. Conditional GANs and their variants have further extended the capabilities of GANs to controlled image generation tasks, such as image-to-image translation and image inpainting.

E. **Attention Mechanisms and Transformer Architectures:** Inspired by the success of attention mechanisms in natural language processing, transformer architectures have gained prominence in computer vision tasks requiring long-range dependencies and context understanding. Transformers facilitate efficient parallelization and global information aggregation by attending to all input positions simultaneously, allowing for scalable modeling of spatial and temporal relationships in visual data. Vision Transformer (ViT)[17] and its variants have demonstrated competitive performance in image classification, object detection, and semantic segmentation tasks, showcasing the versatility and efficacy of attention-based approaches in computer vision.

4. APPLICATIONS OF COMPUTER VISION

Computer vision has permeated numerous industries and domains, revolutionizing processes, enhancing efficiency, and enabling novel functionalities. This section highlights some of the key applications where computer vision technologies have made significant contributions.

Healthcare

Medical Imaging[18]- Computer vision techniques play a critical role in medical imaging modalities such as X-ray, MRI, CT scans, and ultrasound, aiding in disease diagnosis, treatment planning, and surgical guidance. Automated image analysis algorithms can detect abnormalities, segment organs, and tissues, and assist radiologists in interpreting complex medical images, leading to faster and more accurate diagnosis.

Autonomous Vehicles

Object Detection and Recognition[19]- Computer vision algorithms enable autonomous vehicles to perceive and interpret their surroundings, detecting and recognizing pedestrians, vehicles, cyclists, traffic signs, and road markings in real time. By integrating cameras, LiDAR, and radar sensors, autonomous vehicles can make informed decisions and navigate complex driving scenarios safely.

Scene Understanding[20]- Computer vision systems infer semantic and geometric information from sensor data to understand the context of the driving environment, including lane markings, traffic lights, road signs, and obstacles. Scene understanding facilitates path planning, trajectory prediction, and collision avoidance, ensuring smooth and efficient navigation.

Surveillance and Security

Video Analytics[21,22,]- Computer vision algorithms analyze surveillance footage to detect suspicious activities, identify unauthorized individuals, and monitor crowd behavior in public spaces, airports, and commercial establishments. Video analytics systems can track objects of interest, perform facial recognition, and generate real-time alerts for security personnel, enhancing situational awareness and threat detection.

Perimeter Protection[23]- Computer vision-based perimeter protection systems utilize cameras and image processing algorithms to monitor and secure critical infrastructure, industrial facilities, and sensitive installations. Intrusion detection, object tracking, and anomaly detection algorithms enable proactive threat mitigation and response, safeguarding assets and personnel from potential security breaches.

Augmented Reality (AR) and Virtual Reality (VR)

AR Applications[24]- Computer vision powers augmented reality applications that overlay digital information, graphics, and interactive elements onto the user's view of the physical world. AR experiences enhance gaming, education, retail, marketing, and remote collaboration by seamlessly integrating virtual content with the user's environment, creating immersive and engaging experiences.

VR Environments[25]

Computer vision technologies enable the creation of immersive virtual environments that simulate real-world scenarios and interactions. VR applications in training, simulation, design, and entertainment leverage computer vision algorithms for hand tracking, gesture recognition, and spatial mapping, enabling users to interact naturally with virtual objects and environments.

Industrial Automation and Robotics

Quality Inspection[26]- Computer vision systems inspect manufactured products for defects, anomalies, and deviations from quality standards in industrial assembly lines and production processes. Automated visual inspection reduces manual labor, improves product consistency, and enhances overall manufacturing efficiency.

Robot Vision[27]

Computer vision enables robots to perceive and manipulate objects in dynamic and unstructured environments, such as warehouses, factories, and logistics centers. Robot vision systems use cameras and depth sensors to identify objects, navigate obstacles, and perform tasks such as picking, packing, sorting, and assembly with precision and dexterity.

5. CHALLENGES AND LIMITATIONS

Despite the remarkable progress in computer vision, the field faces several challenges and limitations that warrant careful consideration. Addressing these challenges is crucial for advancing the capabilities, reliability, and societal impact of computer vision systems.

i. **Robustness to Adversarial Attacks:** Adversarial attacks exploit vulnerabilities in computer vision models by subtly perturbing input data to induce incorrect predictions or misclassifications. Adversarial examples can evade detection by human observers but cause significant errors in machine learning models, posing security risks in real-world applications such as autonomous vehicles, biometric authentication, and malware detection. Deep learning models, particularly complex neural networks, are often considered “black-box” systems, making it challenging to interpret their decisions and understand the underlying reasoning processes. Lack of interpretability hinders trust, accountability, and regulatory compliance in safety-critical domains such as healthcare, criminal justice, and finance, where transparent decision-making is paramount.
ii. **Ethical Considerations:** Ethical concerns surrounding privacy, bias, fairness, and accountability in computer vision algorithms have garnered increasing attention from researchers, policymakers, and the public. Biased datasets, algorithmic discrimination, and surveillance technologies raise ethical dilemmas related to data privacy, algorithmic fairness, and societal implications, necessitating ethical frameworks and guidelines for responsible development and deployment of computer vision systems.
iii. **Data Quality and Bias:** The performance and generalization capabilities of computer vision models heavily depend on the quality, diversity, and representativeness of training data. Biases present in training datasets, such as demographic imbalances, cultural stereotypes, and underrepresented minority groups, can propagate into algorithmic decision-making, leading to unfair outcomes and perpetuating societal inequalities.
iv. **Regulatory Compliance and Safety:** Deploying computer vision systems in safety-critical applications, such as autonomous vehicles, medical diagnosis, and surveillance, requires compliance with regulatory standards, industry norms, and ethical guidelines. Ensuring the safety, reliability, and legal compliance of computer vision technologies entails rigorous testing, validation, and certification processes to mitigate risks of accidents, errors, and unintended consequences.
v. **Scalability and Efficiency:** Deep learning models often require significant computational resources, memory bandwidth, and energy consumption, limiting their scalability and applicability in resource-constrained environments, embedded devices, and real-time applications. Addressing scalability and efficiency challenges involves optimizing model architectures, pruning redundant parameters, and exploring hardware accelerators tailored for computer vision workloads.
vi. **Domain Adaptation and Generalization:** Computer vision models trained on specific datasets or environments may struggle to generalize to new domains, unseen conditions, or distribution shifts encountered in real-world scenarios. Domain adaptation techniques aim to mitigate domain gaps and improve model robustness by leveraging unlabeled data, unsupervised learning, and transfer learning strategies tailored for domain shift scenarios.
vii. **Human-AI Collaboration and Trust:** Fostering effective collaboration and trust between human users and AI systems is essential for integrating computer vision technologies into decision-making processes and augmenting human capabilities. Human-centric design principles, user-centric interfaces, and transparent communication strategies can enhance user acceptance, confidence, and trust in computer vision systems, promoting collaborative problem-solving and shared decision-making.

6. RECENT TRENDS AND FUTURE DIRECTIONS

The field of computer vision is dynamic and continuously evolving, driven by emerging trends, technological advancements, and novel research directions. This section explores recent trends and outlines potential future directions shaping the landscape of computer vision.

i. **Self-Supervised Learning -** Self-supervised learning approaches aim to leverage unlabeled data to learn useful representations through pretext tasks, such as image inpainting, colorization, or context prediction. Recent advancements in self-supervised learning have demonstrated promising results in improving model generalization and reducing reliance on large labeled datasets, opening new avenues for unsupervised and semi-supervised learning in computer vision.

ii. **Few-Shot and Meta-Learning-** Few-shot learning techniques focus on training models to recognize new object categories or tasks with limited labeled examples. Meta-learning algorithms aim to learn efficient learning algorithms that can adapt quickly to new tasks or environments. By enabling models to generalize from small datasets and rapidly adapt to new domains, few-shot, and meta-learning approaches hold the potential to enhance the scalability and flexibility of computer vision systems.

iii. **Lifelong and Continual Learning-** Lifelong learning paradigms address the challenge of model degradation and catastrophic forgetting when sequentially learning multiple tasks over time. Continual learning algorithms aim to retain previously learned knowledge while adapting to new tasks or data distributions. By enabling models to continuously learn and accumulate knowledge incrementally, lifelong and continual learning approaches facilitate lifelong adaptation and robustness in evolving environments.

iv. **Multi-Modal and Cross-Modal Learning-** Multi-modal learning frameworks integrate information from multiple modalities, such as images, text, audio, and sensor data, to enhance understanding and reasoning capabilities. Cross-modal learning algorithms aim to learn shared representations across different modalities and bridge semantic gaps between heterogeneous data sources. By leveraging complementary information from diverse modalities, multi-modal and cross-modal learning approaches enable a richer and more comprehensive understanding of the world.

v. **Explainable AI and Interpretability-** Explainable AI techniques aim to enhance the transparency, interpretability, and trustworthiness of computer vision models by providing human-understandable explanations for model predictions and decisions. Interpretability methods such as attention mechanisms, saliency maps, and model introspection facilitate understanding of model behavior and reasoning processes. By enabling users to interpret and trust AI-driven decisions, explainable AI techniques promote responsible and ethical deployment of computer vision systems in real-world applications.

vi. **Federated Learning and Privacy-Preserving AI-** Federated learning frameworks enable collaborative model training across decentralized data sources while preserving data privacy and confidentiality. Privacy-preserving AI techniques such as differential privacy, secure multi-party computation, and homomorphic encryption protect sensitive information during model training and inference. By decentralizing computation and minimizing data exposure, federated learning and privacy-preserving AI approaches address privacy concerns and regulatory requirements in sensitive domains such as healthcare, finance, and personal data processing [28, 29].

These recent trends and future directions in computer vision research hold the promise of advancing the state-of-the-art capabilities, scalability, and societal impact of computer vision systems. By embracing interdisciplinary collaborations, ethical considerations, and responsible innovation, the field of computer vision is poised to tackle complex real-world challenges and unlock new frontiers in intelligent perception and understanding.

7. CONCLUSION

In conclusion, the field of computer vision has witnessed remarkable advancements and transformative developments, revolutionizing how machines perceive, interpret, and interact with visual information. From the early days of handcrafted feature engineering to the recent proliferation of deep learning models, computer vision has evolved into a multidisciplinary field at the forefront of artificial intelligence research and application.

Recent trends in computer vision, such as the widespread adoption of deep learning architectures, the emergence of attention mechanisms and transformers, and the integration of multi-modal and self-supervised learning approaches, have propelled the field to new heights of performance and capability. These advancements have fueled breakthroughs in diverse application domains, including healthcare, autonomous vehicles, surveillance, augmented reality, and industrial automation, unlocking unprecedented opportunities for innovation and societal impact.

However, alongside these advancements come significant challenges and ethical considerations that must be addressed to ensure the responsible development and deployment of computer vision technologies. Challenges such as robustness to adversarial attacks, interpretability, ethical considerations, data bias, regulatory compliance, and scalability pose critical hurdles that require interdisciplinary collaboration, ethical frameworks, and innovative solutions.

Looking ahead, the future of computer vision holds immense promise for further innovation and societal transformation. Future directions in computer vision research may include advancements in explainable AI, fairness and accountability, human-centered design, lifelong learning, and domain adaptation, aimed at addressing current limitations and pushing the boundaries of what is possible.

In this rapidly evolving landscape, collaboration between academia, industry, policymakers, and stakeholders will be essential to navigate the complexities and harness the full potential of computer vision technologies for the benefit of society. By embracing these challenges and opportunities, the field of computer vision is poised to continue driving innovation, empowering human creativity, and shaping the future of artificial intelligence.

CONFLICT OF INTERESTS

None

ACKNOWLEDGMENTS

None

REFERENCES

Pan, Z., Yu, W., Yi, X., Khan, A., Yuan, F., and Zheng, Y., 2019. Recent progress on generative adversarial networks (GANs): A survey. *IEEE Access*, *7*, pp.36322-36333.

Ulrich, M., Steger, C. and Baumgartner, A., 2003. Real-time object recognition using a modified generalized Hough transform. *Pattern Recognition*, *36*(11), pp.2557-2570.

Kukacka, M., Neocognitron: A Survey of a Classical Hybrid Neural Network Model.

Wang, Y., Li, Z., Wang, L. and Wang, M., 2013. A Scale Invariant Feature Transform Based Method. *J. Inf. Hiding Multim. Signal Process.* *4*(2), pp.73-89.

Bay, H., Tuytelaars, T. and Van Gool, L., 2006. Surf: Speeded up robust features. In *Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9* (pp. 404-417). Springer Berlin Heidelberg.

Zhang, Y., Jin, R. and Zhou, Z.H., 2010. Understanding bag-of-words model: a statistical framework. *International journal of machine learning and cybernetics*, *1*, pp.43-52.

Zakaria, N. and Hassim, Y.M.M., 2024. A Review Study of the Visual Geometry Group Approaches for Image Classification. *Journal of Applied Science, Technology and Computing*, *1*(1), pp.14-28.

Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Van Esesn, B.C., Awwal, A.A.S. and Asari, V.K., 2018. The history began from Alexnet: A comprehensive survey on deep learning approaches. *arXiv preprint arXiv:1803.01164*.

Majib, M.S., Rahman, M.M., Sazzad, T.S., Khan, N.I. and Dey, S.K., 2021. Vgg-scnet: A vgg net-based deep learning framework for brain tumor detection on MRI images. *IEEE Access*, *9*, pp.116942-116952.

Anand, R., Shanthi, T., Nithish, M.S. and Lakshman, S., 2020. Face recognition and classification using GoogleNET architecture. In *Soft Computing for Problem Solving: SocProS 2018, Volume 1* (pp. 261-269). Springer Singapore.

Reddy, A.S.B. and Juliet, D.S., 2019, April. Transfer learning with ResNet-50 for malaria cell-image classification. In *2019 International Conference on Communication and Signal Processing (ICCSP)* (pp. 0945-0949). IEEE.

Lindsay, G.W., 2021. Convolutional neural networks as a model of the visual system: Past, present, and future. *Journal of cognitive neuroscience*, *33*(10), pp.2017-2031.

Zoumpourlis, G., Doumanoglou, A., Vretos, N. and Daras, P., 2017. Non-linear convolution filters for cnn-based learning. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 4761-4769).

Iandola, F., Moskowitz, M., Karayev, S., Girshick, R., Darrell, T. and Keutzer, K., 2014. Densenet: Implementing efficient convnet descriptor pyramids. *arXiv preprint arXiv:1404.1869*.

Pouyanfar, S., Chen, S.C. and Shyu, M.L., 2017, July. An efficient deep residual-inception network for multimedia classification. In *2017 IEEE International Conference on Multimedia and Expo (ICME)* (pp. 373-378). IEEE.

Hoang, V.T. and Jo, K.H., 2021, July. Practical analysis on architecture of EfficientNet. In *2021 14th International Conference on Human System Interaction (HSI)* (pp. 1-4). IEEE.

Chen, C.F.R., Fan, Q. and Panda, R., 2021. Crossvit: Cross-attention multi-scale vision transformer for image classification. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 357-366).

Gao, J., Yang, Y., Lin, P. and Park, D.S., 2018. Computer vision in healthcare applications. *Journal of Healthcare Engineering*, *2018*.

Alfahdawi, M.G., Alheeti, K.M.A. and Al-Rawi, S.S., 2021, June. Intelligent Object Recognition System for Autonomous and Semi-Autonomous Vehicles. In *2021 International Conference on Communication & Information Technology (ICICT)* (pp. 227-233). IEEE.

Nadeem, U., Shah, S.A.A., Sohel, F., Togneri, R. and Bennamoun, M., 2019. Deep learning for scene understanding. *Handbook of deep learning applications*, pp.21-51.

Olatunji, I.E. and Cheng, C.H., 2019. Video analytics for visual surveillance and applications: An overview and survey. *Machine Learning Paradigms: Applications of Learning and Analytics in Intelligent Systems*, pp.475-515.

Canedo, D. and Neves, A.J., 2019. Facial expression recognition using computer vision: A systematic review. *Applied Sciences*, *9*(21), p.4678.

Lipton, A.J., Heartwell, C.H., Haering, N. and Madden, D., 2002, October. Critical asset protection, perimeter monitoring, and threat detection using automated video surveillance. In *Proceedings of the Thirty-Sixth Annual International Carnahan Conference on Security Technology*.

Sharma, R. and Molineros, J., 1995, March. Role of computer vision in augmented virtual reality. In *Stereoscopic Displays and Virtual Reality Systems II* (Vol. 2409, pp. 220-231). SPIE.

Kim, W.S., 1999. Computer vision-assisted virtual reality calibration. *IEEE Transactions on Robotics and Automation*, *15*(3), pp.450-464.

Vergara-Villegas, O.O., Cruz-Sánchez, V.G., de Jesús Ochoa-Domínguez, H., de Jesús Nandayapa-Alfaro, M. and Flores-Abad, Á., 2014. Automatic product quality inspection using computer vision systems. *Lean manufacturing in the developing world: Methodology, case studies and trends from Latin America*, pp.135-156.

Kragic, D. and Christensen, H.I., 2005. Advances in robot vision. *Robotics and Autonomous Systems*, *52*(1), pp.1-3.

Gour L. and Waoo A. A., 2018. Implementing Fault Resilient Strategies in Cloud Computing via Federated Learning Approach, Journal of Innovation in Applied Research, Vol.1Issue 1m pp. 1-5.

Lokendra Gour and Akhilesh A Waoo. Fault-tolerant framework with federated learning for reliable and robust distributed system. In THEETAS 2022: Proceedings of The International Conference on Emerging Trends in Artificial Intelligence and Smart Systems, THEETAS 2022, 16-17 April 2022, Jabalpur, India, page 219. European Alliance for Innovation, 2022.