International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 13, Issue 8, August 2024)

A Review of AI Techniques and Applications in Computer Vision

Devika Sahu[1], Gousia Hazra Anjum Khan[2], Mrs. Arjoo Pandey[3], Vijaya Sahu[4]
[1,2,3]Assistant Professor, Department of CSE, Government Engineering College, Raipur, India
[4]Assistant Professor, Department of CSE, Shantiniketan College, Raipur, India

***Abstract*— The rapid advancements in artificial intelligence (AI) have significantly transformed the field of computer vision, leading to the development of powerful techniques that enable machines to interpret and understand visual data. This review explores the evolution, current state, and future directions of AI techniques in computer vision, emphasizing key methodologies such as deep learning, convolutional neural networks (CNNs), and reinforcement learning. We discuss their applications across various domains, including image recognition, object detection, and video analysis. The review highlights the challenges faced in the field, such as data scarcity, computational complexity, and ethical considerations. Finally, it provides insights into emerging trends and potential future research avenues that could further enhance the capabilities of AI in computer vision.**

***Keywords*— AI, CNN, RNN, Computer Vision.**

I. INTRODUCTION

The field of computer vision, a subdomain of artificial intelligence (AI), is concerned with enabling machines to interpret and understand the visual world. By emulating human vision, computer vision systems have the potential to perform complex tasks such as image recognition, object detection, and video analysis with unprecedented accuracy and efficiency. The evolution of AI, particularly the advent of deep learning, has revolutionized computer vision, allowing for the development of sophisticated algorithms that can process vast amounts of visual data and extract meaningful insights.

Historically, computer vision relied heavily on handcrafted features and traditional machine learning techniques. Early methods were limited by their dependence on human expertise for feature extraction, which often resulted in suboptimal performance, especially in complex and dynamic environments. The introduction of deep learning, particularly convolutional neural networks (CNNs), marked a paradigm shift in the field. CNNs, with their ability to automatically learn hierarchical features from raw data, have significantly outperformed traditional methods, setting new benchmarks in various computer vision tasks.

The success of deep learning in computer vision can be attributed to several factors, including the availability of large-scale labeled datasets, advances in computational hardware (such as GPUs), and the development of novel architectures like ResNet and GANs. These advancements have enabled the deployment of AI-powered computer vision systems in a wide range of applications, from autonomous vehicles and medical imaging to facial recognition and augmented reality.

Despite these achievements, the field of computer vision still faces several challenges. The scarcity of annotated data in certain domains, the high computational cost of training deep networks, and the need for explainability in AI decisions are some of the critical issues that researchers are striving to address. Moreover, the ethical implications of deploying AI in sensitive areas, such as surveillance and law enforcement, have sparked significant debate, necessitating a careful consideration of privacy, fairness, and bias in AI systems.

This review aims to provide a comprehensive overview of the AI techniques that have shaped the field of computer vision, examining both the fundamental principles and the latest advancements. We will explore the various applications of these techniques across different domains, discuss the ongoing challenges, and identify potential future research directions that could further enhance the impact of AI in computer vision. Through this review, we hope to offer valuable insights to researchers and practitioners alike, contributing to the continued growth and innovation in this dynamic and rapidly evolving field.

II. LITERATURE SURVEY

I. Ahmed et al.,[1] presents the transformative potential of AI and computer vision in the field of commodity supply chain management. The capability of AI to reduce yield loss and enhance supply chain efficiency is a growing trend and vision-based commodity defect monitoring can be useful in this regard. explored the employment of real-time computer vision techniques in supply chain flaw management, which include Detection Transformer (DETR), a type of Vision Transformer (ViT), and compared its performance with the You Only Look Once (YOLO) and other AI models. Computational feasibility is assessed, encompassing various computer vision and AI models, by using a dataset comprising images of commodity items used to substantiate our findings. The obtained results have shown the improved performance of DETR with a detection and classification accuracy of 96%, directly correlating with improved supply chain management. On the other hand, the higher computational burden imposed by DETR makes it less feasible for the higher constrained embedded applications.

G. A. Senthil et al.,[2] Among the causes of traffic deaths are increased traffic volumes, excessive speeding, careless and intoxicated driving, driver weariness, inadequate road infrastructure, and the presence of animals on roadways in forest locations. Road accident fatalities as a percentage of all deaths globally have grown by 2.2%, according to the World Health Organization (WHO). Every year, traffic accidents claim the lives of almost 1.35 million individuals. Emergency medical help is often delayed in traffic incidents that lead to the deaths of people. The sensor camera captures images through IoT sensor networks and computer vision. The image resizing method is employed in data preparation to prepare the datasets. Here, deep learning (DL) algorithms will be used to train these datasets, and a model file will be produced.

S. S. Ittannavar et al.,[3] Significant strides have been achieved in the use of deep learning to computer vision, which has changed the way that computers process and respond to visual data. The authors of this study apply a thorough approach that includes data collecting, model construction, training, assessment, and ethical concerns in their investigation of the many facets of picture categorization using CNNs. The study shows how these techniques might be used in the actual world, namely in the fields of healthcare and autonomous systems. Ethical concerns highlight the significance of justice and accountability, and transfer learning emerges as a beneficial technique for optimizing model performance. Future prospects include researching advanced architectures and multimodal fusion, tackling real-world difficulties, and enhancing ethical and explainable AI, as well as reinforcing models against adversarial assaults.

B. Sathish et al.,[4] present a mobile application designed to assist visually impaired individuals in navigating their daily lives. The application incorporates several features, including text-to-speech, currency detection, calculator, object detection using trained machine learning models, and reminders. These features are all integrated with voice command functionality and can be accessed with a single button touch. The text-to-speech feature allows users to listen to text on their mobile devices from documents. The currency detection feature enables users to recognize all Indian currency through their phone camera, which is particularly helpful for those who may struggle with identifying money. The calculator feature helps users to perform simple calculations with ease. Using the mobile camera, the object detection feature detects objects in real-time using the trained ML models, both indoors and outdoors.

H. Lee et al.,[5] established a sustainable and efficient parking system at seaports with a comprehensive framework based on deep learning and transfer learning techniques. Through an in-depth literature review, we determined the benefits of combining AI and CV to optimize parking facility management. Then, data gathering, quality evaluation, and practical applications of CV and AI in parking spaces and lanes were conducted. The significance of AI and CV technologies in developing efficient and sustainable transportation systems was verified by enhancing container drayage efficiency in seaports. The proposed system allowed AI and CV technologies to improve traffic management, optimize parking space allocation, and streamline container movement within seaports. By using deep learning techniques,

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 13, Issue 8, August 2024)

particularly transfer learning, the proposed framework accurately identified parking spaces and lanes considering the crane position. CCTV camera data were used for real-time monitoring and efficient utilization of parking spaces in the seaport.

S. V. Mahadevkar et al.,[6] Computer applications have considerably shifted from single data processing to machine learning in recent years due to the accessibility and availability of massive volumes of data obtained through the internet and various sources. Machine learning is automating human assistance by training an algorithm on relevant data. Supervised, Unsupervised, and Reinforcement Learning are the three fundamental categories of machine learning techniques. Discussed the different learning styles used in the field of Computer vision, Deep Learning, Neural networks, and machine learning. Some of the most recent applications of machine learning in computer vision include object identification, object classification, and extracting usable information from images, graphic documents, and videos. Some machine learning techniques frequently include zero-shot learning, active learning, contrastive learning, self-supervised learning, life-long learning, semi-supervised learning, ensemble learning, sequential learning, and multi-view learning used in computer vision until now.

I. A. Idowu et al.,[7] A safety inspection is an on-site walk-through to identify potential hazards to occupants and personnel and options for remedial action. This computer vision-based approach automates the inspection of an interior building hallway (exit access) for an obstruction that may be a potential fire hazard. Our approach is important because it will mitigate the risk of a fire hazard to the building occupants by sensing and alerting the safety officer before a situation turns into an emergency. The performance of our proposed approach, the benefits, and the implementation challenges, were evaluated through a case study. The result demonstrates that our proposed Dual Temporal Buffer Differencing (DTBD) method can detect a potential fire hazard in the building exit access effectively and continuously. As a result, the approach can facilitate safety in the building and allow safety inspectors to plan more trained human resources.

R. Pitale et al.,[8] AI and DL become supreme problem-solving strategies in many areas of research and industrial applications. Deep learning is one of the most efficient, accurate and cost efficient algorithms. Applications of Deep learning via computer vision in Bio medical and healthcare, security, education and latest trends in technologies are stated here. This paper focuses on giving an overview of work done by many authors on Deep learning and computer vision applications and algorithms. At last common findings from the papers are reviewed.

III. CHALLENGES

Challenges in AI Techniques for Computer Vision-

1. **Data Scarcity and Quality**: While deep learning models thrive on large datasets, many domains lack sufficient annotated data, limiting model performance. Data augmentation techniques help, but generating high-quality, labeled data remains a challenge, especially in specialized areas like medical imaging.
2. **Computational Complexity**: Training deep neural networks, particularly those with complex architectures like CNNs and GANs, demands significant computational resources. High costs associated with hardware (GPUs/TPUs) and energy consumption can be prohibitive, especially for smaller research institutions.
3. **Generalization and Robustness**: AI models in computer vision often struggle with generalization across different environments or domains. Factors such as variations in lighting, occlusion, and viewpoint changes can drastically affect model accuracy. Ensuring robustness against adversarial attacks, where subtle changes in input can lead to misclassification, is also a significant concern.
4. **Explainability and Interpretability**: The black-box nature of deep learning models poses challenges in understanding how decisions are made, which is critical in applications like medical diagnosis or autonomous driving. Lack of transparency hinders

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 13, Issue 8, August 2024)

trust and acceptance, particularly in high-stakes scenarios.

5. **Real-Time Processing**: Many computer vision applications, such as autonomous driving and video surveillance, require real-time processing capabilities. Achieving low-latency inference while maintaining high accuracy is challenging, especially on resource-constrained devices.
6. **Scalability**: As AI models grow in complexity, scaling them for deployment across different platforms, including edge devices and cloud environments, becomes increasingly difficult. Efficiently managing model size, inference time, and energy consumption is critical for practical deployment.
7. **Integration with Other Technologies**: Combining AI-based computer vision systems with other technologies, such as IoT and robotics, presents integration challenges. Ensuring seamless communication, data exchange, and system synchronization requires sophisticated engineering solutions.

These challenges highlight the ongoing need for innovation and interdisciplinary collaboration to further advance the field of computer vision and realize its full potential across diverse applications.

IV. Comparative Analysis

This analysis will cover key methodologies such as Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), Reinforcement Learning (RL), and Transformer models, providing insights into their performance across tasks like image recognition, object detection, and video analysis.

1. Convolutional Neural Networks (CNNs)

- **Strengths**: CNNs are highly effective in extracting spatial hierarchies from visual data, making them the standard for image recognition and classification tasks. Their ability to automatically learn features from raw images has led to state-of-the-art performance in many applications.
- **Limitations**: CNNs require large amounts of labeled data and significant computational resources for training. They also struggle with tasks that require understanding long-range dependencies due to their limited receptive fields.
- **Applications**: Widely used in image recognition, object detection (e.g., YOLO, Faster R-CNN), and facial recognition systems.

2. Generative Adversarial Networks (GANs)

- **Strengths**: GANs excel in generating realistic images and data augmentation. They have been pivotal in tasks like image synthesis, super-resolution, and style transfer, where the goal is to create or enhance visual content.
- **Limitations**: GANs are notoriously difficult to train, with challenges such as mode collapse and instability during training. They also require careful tuning of hyperparameters and architectures.
- **Applications**: Used in image generation, deepfake technology, and enhancing image quality in applications like medical imaging and video game development.

3. Reinforcement Learning (RL)

- **Strengths**: RL techniques are powerful for sequential decision-making tasks and environments where actions influence future states. In computer vision, RL is used for tasks like robotic vision, where agents learn to interact with their environment to achieve specific goals.
- **Limitations**: RL often requires a significant amount of training data and time, as well as complex reward engineering. It can be less efficient compared to supervised learning methods for certain vision tasks.
- **Applications**: Applied in autonomous navigation, robotic manipulation, and dynamic object tracking in video analysis.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 13, Issue 8, August 2024)

V. CONCLUSION

AI techniques have profoundly transformed computer vision, enabling advancements in image recognition, object detection, and image generation. Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), Reinforcement Learning (RL), and Transformer models each offer unique strengths, contributing to the field's evolution. However, challenges such as data scarcity, computational demands, and ethical concerns remain critical. Continued innovation and interdisciplinary collaboration are essential to overcoming these obstacles and unlocking the full potential of AI in computer vision, paving the way for more robust and ethical applications across diverse domains.

REFERENCES

1. I. Ahmed, M. Alkahtani, Q. S. Khalid and F. M. Alqahtani, "Improved Commodity Supply Chain Performance Through AI and Computer Vision Techniques," in IEEE Access, vol. 12, pp. 24116-24132, 2024, doi: 10.1109/ACCESS.2024.3361756.
2. G. A. Senthil, R. V. Lakshmi Priya, S. Geerthik, G. Karthick and R. Lavanya, "Safe Road AI: Real-Time Smart Accident Detection for Multi-Angle Crash Videos using Deep Learning Techniques and Computer Vision," 2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2024, pp. 617-622, doi: 10.1109/ICAAIC60222.2024.10575074.
3. S. S. Ittannavar, B. P. Khot, M. Mahadevaswamy and R. H. Havaldar, "Deep Learning for Computer Vision: Recent Breakthroughs and Emerging Trends," *2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, Chennai, India, 2023, pp. 1-5, doi: 10.1109/ICSES60034.2023.10465486.
4. B. Sathish, B. R, S. BS, T. Popli, S. R and R. R. Hebbar, "An Advanced Computer Vision Techniques for Object Recognition and Navigation Assistance for the Visually Impaired," *2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN)*, Vellore, India, 2023, pp. 1-6, doi: 10.1109/ViTECoN58111.2023.10157346.
5. H. Lee, I. Chatterjee and G. Cho, "Enhancing Parking Facility of Container Drayage in Seaports: A Study on Integrating Computer Vision and AI," *2023 IEEE 6th International Conference on Knowledge Innovation and Invention (ICKII)*, Sapporo, Japan, 2023, pp. 384-387, doi: 10.1109/ICKII58656.2023.10332699.
6. S. V. Mahadevkar *et al.*, "A Review on Machine Learning Styles in Computer Vision—Techniques and Future Directions," in *IEEE Access*, vol. 10, pp. 107293-107329, 2022, doi: 10.1109/ACCESS.2022.3209825.
7. I. A. Idowu, K. Nyarko and O. Toutsop, "Computer Vision Method in Means of Egress Obstruction Detection," *2022 IEEE World AI IoT Congress (AIIoT)*, Seattle, WA, USA, 2022, pp. 167-173, doi: 10.1109/AIIoT54504.2022.9817354.
8. R. Pitale, H. Kale, S. Kshirsagar and H. Rajput, "A Schematic Review on Applications of Deep Learning and Computer Vision," *2021 Asian Conference on Innovation in Technology (ASIANCON)*, PUNE, India, 2021, pp. 1-6, doi: 10.1109/ASIANCON51346.2021.9544941.