Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Robust Object Recognition in Adverse Environments

Tadiboyina Teja [1], PVRD Prasada Rao [2]
[1] *Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Vijayawada-India, Dr PVRD Prasada Rao.*
[2] *Professor, Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Vijayawada-India, Email: pvrdprasad@kluniversity.in*

ARTICLE INFO

Received: 27 Dec 2024

Revised: 15 Feb 2025

Accepted: 25 Feb 2025

ABSTRACT

This paper presents a novel approach for moving object recognition utilizing the Python programming language and the OpenCV computer vision library. Moving object recognition holds significant importance in various domains, including surveillance, robotics, and traffic monitoring. Despite existing methodologies, achieving accurate and real-time recognition remains a challenge due to complexities in dynamic environments. Our plan solves this problem challenge by leveraging Python's versatility and OpenCV's robust features. We introduce a multi-stage process that also includes the background subtraction, feature extraction, and also object tracking algorithms to accurately recognising and tracking moving objects in real-time video streams. Additionally, we incorporate optimization techniques to enhance computational efficiency without compromising accuracy. Experimental results show the effectiveness of our approach, showcasing improved performance compared to existing methods. The proposed solution offers promising implications for real-world applications, including enhanced surveillance systems and intelligent transportation systems, and augmented reality environments.

Keywords: Illumination Variation, Occlusion, Viewpoint Variation, Deformation, and Cluttered Backgrounds.

INTRODUCTION

Moving object recognition it is one of the fundamental task in the field of the computer vision with wide-ranging applications spanning from surveillance and security to robotics and autonomous vehicles. It also has the ability to identify and track moving objects in real-time video streams is a crucial tasks such as activity monitoring, object classification, and scene analysis. Despite significant progress in the field, challenges persist, particularly in achieving robust performance in dynamic and cluttered environments. At the present time, the Python programming language and the OpenCV (Open Source Computer Vision) library have emerged as powerful gear for developing computer imaginative and prescient packages. Python's simplicity, versatility, and full-size library aid make it a really perfect desire for fast prototyping and experimentation, while OpenCV offers a complete suite of functions and algorithms for photograph and video evaluation. On this paper, we gift a unique way of approach for the moving object recognition using Python and OpenCV. Our approach aims to address to the challenges associated with real-time detection and tracking of moving objects by leveraging the capabilities of these two technologies. We propose a multi-stage process that encompasses background subtraction, improving feature extraction, and object tracking algorithms to achieve accurate and efficient moving object recognition.

RELATED WORK

Machine Learning

Machine learning knowledge of is a branch in the artificial intelligence that makes a speciality of growing algorithms and statistical models that allow computers work without a clear plan. It encompasses a extensive range of strategies and methodologies, which include supervised learning, unsupervised gaining knowledge of, reinforcement mastering, and deep mastering.

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Supervised learning knowledge of entails training the model on a labeled dataset, where every enter is associated with a corresponding output. The model learns to make predictions by generalizing from the education facts. Not unusual supervised learning algorithms consist of linear regression and logistic regression, and also assist vector machines.

Unsupervised learning knowledge of, however, deals with education models on unlabeled records, wherein the goal is to identify patterns, processes, or relationships. Clustering algorithms along with okay-method clustering and hierarchical clustering is example of an unsupervised gaining knowledge of techniques generally used for responsibilities including customer segmentation and anomaly detection.

Reinforcement gaining knowledge of is a paradigm of system mastering in which an agent learns to make choices by using interacting with an environment and receiving comments in the shape of rewards or consequences. The agent's the goal is to complete this cumulative praise over the years through learning foremost guidelines via trial and mistakes. Reinforcement gaining knowledge of has applications in the areas such as game gambling, robotics, and self-sustaining automobile control.

Deep learning is a subfield in the machine learning knowledge of that specializes in training neural networks with more than one layers (as a result the term "deep") to analyze complex representations of facts. Deep studying architectures, consisting of CNNs for picture popularity and RNNs for sequential information, have achieved fantastic fulfillment in numerous domains, along with pc vision, herbal language processing, and speech reputation.

Machine learning techniques is increasingly used across diverse industries and domains, revolutionizing fields like healthcare, finance, marketing, and transportation. From diagnosing diseases and predicting stock prices to personalizing recommendations and optimizing supply chains, machine learning algorithms are for driving innovation and creating more no of opportunities for the businesses and society as a whole. As the field continues to evolve, the potential for machine learning is used to solve complex problems and unlock new insights remains vast, paving to the way for a future powered by intelligent systems and autonomous agents.

Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are the fundamental type of deep learning models are designed specifically for processing structured grid-like data, such as images and videos. They are inspired by an visual cortex of human brain and have become the cornerstone of many state-of-the-art image vision tasks.

At the core of CNNs are the convolutional layers, which apply a series of the learnable filters to an input image to the extract spatial hierarchies of features. These filters are detect patterns in an image such as edges in frame, textures, and shapes, gradually learning to recognize increasingly complex features as network deepens. By sharing weights across different spatial locations, convolutional layers ensure parameter efficiency and enable network to generalize well to unseen data. Pooling layers are often used in conjunction with convolutional layers to down sample feature maps, reducing the computational complexity and improving translation invariance.

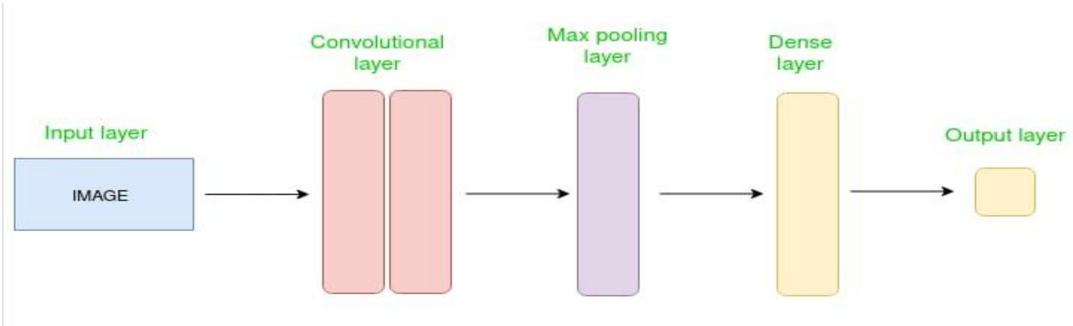

Fig. 1. Convolutional Neural Network

Common pooling operations that include max pooling and average pooling, which aggregate information from local neighborhoods to capture the most salient features in while discarding irrelevant details.

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

CNN architectures typically consist of multiple convolutional and pooling layers followed through one or greater fully related layers for class or regression responsibilities. Deep CNNs, including VGG, ResNet, and Inception, have accomplished terrific achievement in picture category, item detection, semantic segmentation, and other pc imaginative and prescient responsibilities, often outperforming traditional handcrafted feature extraction methods.

Transfer learning, a technique where pre-trained CNN models are fine-tuned on specific datasets, has further democratized to the application in CNNs, allowing practitioners to leverage the knowledge learned from large-scale datasets such as ImageNet to solve domain-specific problems with limited labeled data.

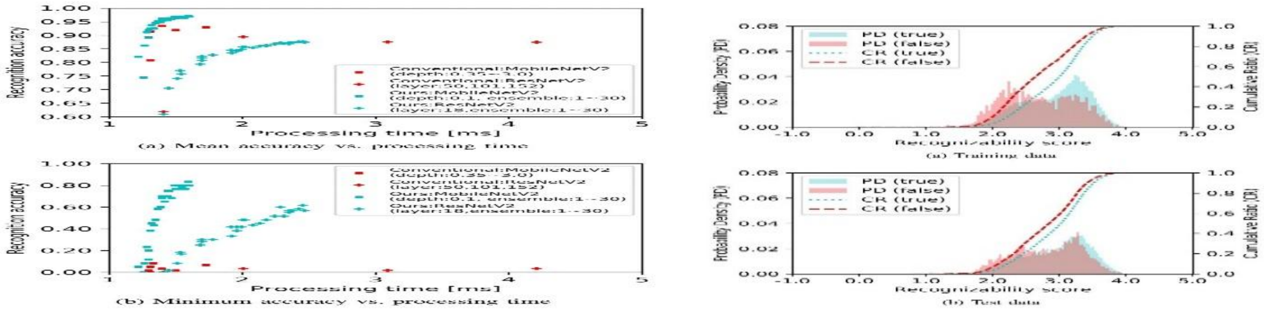

Fig. 2. Probability density distribution of the recognizability score. (a) Training data (b) Test data

Beyond traditional computer vision tasks, CNNs have also found applications in domains such as (NLP) Natural language processing and speech recognition, and where as they are adapted to process sequential data using techniques like 1D convolutions and attention mechanisms. Despite their success, CNNs are no longer without obstacles. They require big amounts of classified facts for the training and are computationally intensive, regularly requiring specialised hardware which includes GPUs or TPUs for green education and inference.

Overall, Convolutional Neural Networks have transformed the field of the computer vision and continue to push the boundaries of what is possible in artificial intelligence, with ongoing research focused on improving efficiency, robustness, and interpretability to enable even broader adoption across diverse applications and domains.

ReLU (Rectified Linear Unit):

Rectified Linear Unit is an activation function that commonly used for artificial neural networks, particularly in the deep learning models it's far a simple mathematical characteristic that introduces non-linearity to the community, enabling it to examine complicated patterns.

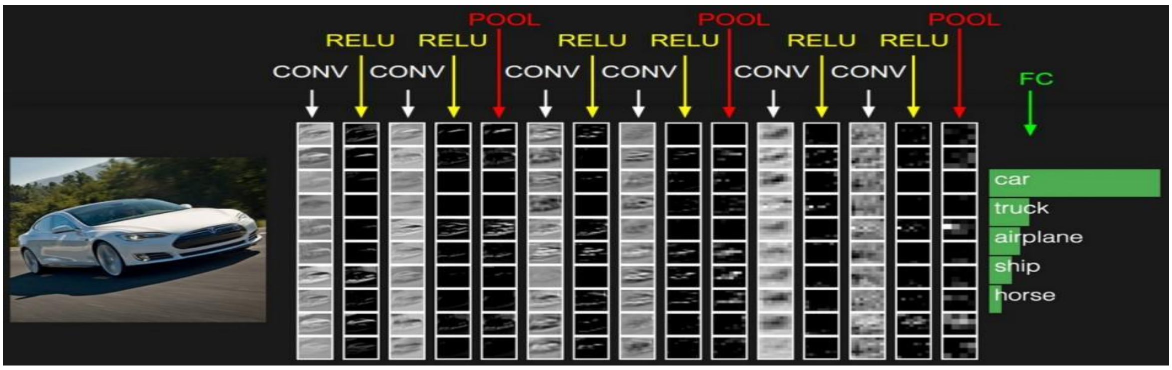

Fig. 3. Rectified Linear Unit

ReLU means Rectified Linear Unit, and it is a one of the type of activation function used in artificial neural networks, particularly in deep learning models. It defined as the function:

$$f(x) = \max(0, x)$$

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

In other words, the ReLU is a function returns the input (x) if it's far nice, and 0 in any other case. Visually, it seems like an linear characteristic with the negative part of it's enter values "clipped" to zero.

One of the main advantages of ReLU is its simplicity and computational performance. Not like some other activation features like sigmoid or tanh, ReLU doesn't include difficult mathematical operations which include exponentials, which can be computationally high priced to compute.

THE WIDELY USED IMAGE FORMATS FOR IMAGE PROCESSING

There are several image formats are used for different purposes. Here are some of the most commonly used image formats:

- JPEG (Joint Photographic Experts Group):
  - JPEG stands as a prevalent compressed image format, well-suited for photographs and intricate images featuring numerous colors.
  - Employing lossy compression, it sacrifices some image quality to reduce file sizes significantly.
  - JPEG serves admirably for online sharing and web usage due to its compact file sizes.
- PNG (Portable Network Graphics):
  - PNG emerges as a lossless compressed image format, retaining all image quality without compromising detail.
  - PNG commonly finds use in web graphics and images necessitating transparency.
- GIF (Graphics Interchange Format):
  - GIF enjoys widespread use for animated images.
  - It facilitates animation by amalgamating multiple frames into a single file.
  - GIFs can also remain static and support transparency.
- TIFF (Tagged Image File Format):
  - TIFF represents a versatile image format prevalent in professional realms like photography, graphic design, and printing.
  - Supporting lossless compression, it accommodates high-quality images with layers, transparency, and other advanced features.

For image processing tasks like such as computer vision, JPEG and PNG are often used due to their broad support, while TIFF is suitable for professional applications that require high-quality images.

CURRENT METHODS OF DETECTING MOVING OBJECT

Frame differencing

Frame differencing is a technique commonly used in computer vision to detect motion in a video sequence. It involves comparing consecutive frames of a video and calculating the pixel-wise difference between them. By subtracting the pixel of values to one frame from the corresponding pixel of values to another frame, areas of significant change, or motion, are highlighted.

This method is effective because static background elements remain consistent across frames, while moving objects create variations in pixel values. These differences can be thresholded to identify regions of interest where motion occurs. Frame differencing is a fundamental step in different number of applications such as video surveillance, gesture recognition, and object tracking. It provides a simple yet powerful way to detect motion and is often used as for a building block in more complex computer vision algorithms.

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

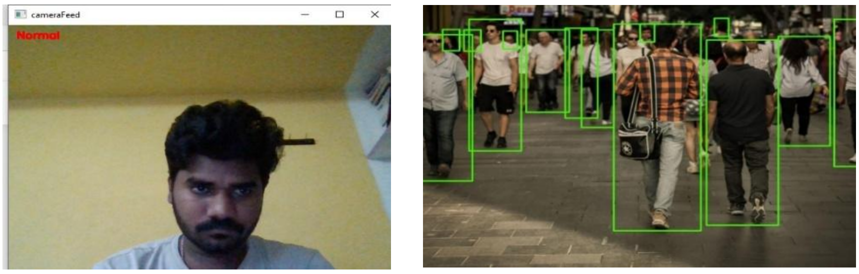

Fig. 4. Frame differencing

The operational steps of frame differencing can be outlined as follows:

- Frame Capture
- Pixel Comparison
- Threshold Application
- Creation of Motion Mask
- Optional Post-Processing
- Visualization or Further Processing

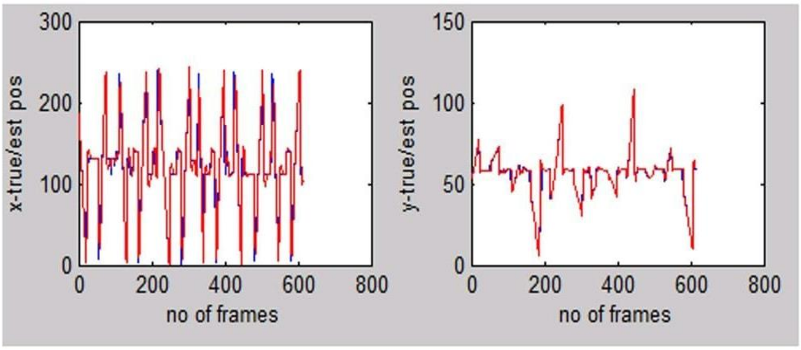

Fig. 6. Plots of position error in x coordinate and y coordinate

TEMPORAL DIFFERENCING

This Temporal difference learning is a reinforcement learning technique in which that combines aspects of both dynamic programming and Monte Carlo methods. It's widely used in machine learning, deep learning and artificial intelligence for solving sequential decision-making problems, especially in the context of Markov decision processes (MDPs).

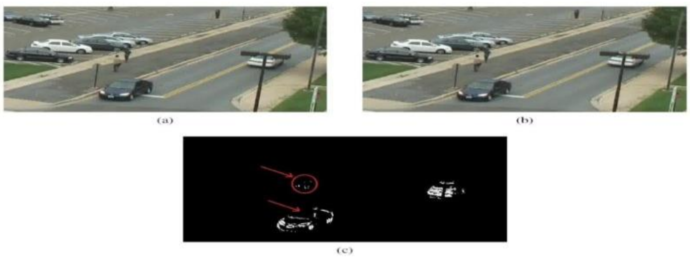

Fig. 6. Temporal differencing

Temporal difference learning is a one of the fundamental method used in reinforcement learning to state costs are estimated based on variations from one state to other Distinguishing between dynamic programming methods, which require a full understanding of environmental dynamics, and Monte Carlo methods, which need an understanding of environmental dynamics. Complete episodes to be experienced before updating the value estimates, temporal difference learning updates value estimates based on incomplete sequences of transitions.

The key equation in temporal difference learning is the TD error. This is the distinctness between to the estimated value of the current state and the estimated value of the next state, plus the observed reward:

[\text{TD error} = r + \gamma V(s') - V(s) \]

Where:

- \(r \) is the observed reward after transitioning from state \(s \) to state \(s' \).

- \ (\gamma \) is the discount factor, which determines to the importance in future rewards.

- \(V(s) \) is the estimated value of state \(s \) before the update, and \(V(s') \) is the estimated value of the next state \(s' \).

This TD estimate the value of the current state corresponding to the following rules:

Where:

- (\alpha \) is the learning rate, which determines the size of the update step.

Temporal difference learning combines aspects of dynamic programming, which uses value iteration to update value estimates, and Monte Carlo methods are used complete episodes to update value estimates. By updating value estimates incrementally after each transition, temporal difference learning can learn in real-time without requiring complete knowledge to the environment or waiting for full episodes to complete. This makes it well-suited for online and incremental learning tasks.

OPTIC FLOW

Optical flow is a crucial concept in computer vision in which it deals with understanding the motion of objects within a scene based on consecutive frames to the image sequence. It aims to compute the velocity field or displacement of pixels between frames, providing insights into how objects move over time.

In essence, optical flow algorithms analyze the changes in intensity patterns between consecutive frames to infer the motion of objects. This process involves tracking the apparent motion of features or points to the image from one frame to the next. One common approach to optical flow estimation is to formulate it as an optimization problem, where the goal is to minimize of the difference between an observed image and the predicted image under the estimated flow field. Various optimization techniques, such as Lucas-Kanade method, Horn-Schunck method, or more recent deep learning-based approaches, are employed to solve this problem.

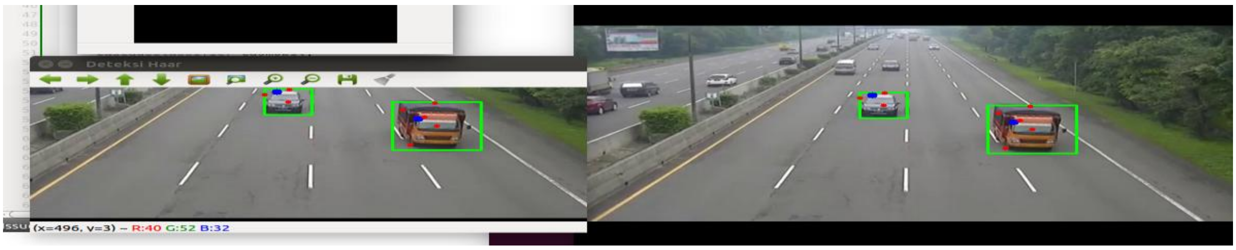

Fig. 7. Optical flow

IMPORTANCE OF OBJECT RECOGNITION

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Object recognition is a critical feature of computer vision with significant importance across various domains including robotics, surveillance, autonomous vehicles, medical imaging, augmented reality, and many others.

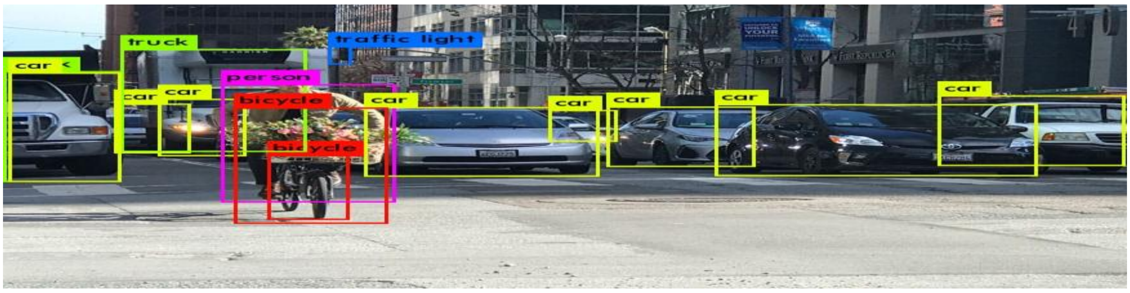

Fig. 8. Importance of object recognition

Automation and Robotics

Object recognition enables machines and robots to identify and also interact with objects in the environment. In industrial automation, robots capable of recognizing objects can perform tasks such as identification, assembly and also quality control with high precision and efficiency.

Surveillance and Security

In surveillance systems, object recognition helps in identifying and tracking individuals, vehicles, and suspicious activities in real-time. It enables security personnel to monitor crowded areas, detect intruders, and prevent potential threats.

Medical Imaging

Object recognition plays a important role in medical imaging analysis for the diagnosing diseases and guiding treatment. It assists radiologists in identifying anatomical structures, tumors, lesions, and abnormalities in the medical images such as X-rays, MRI scans, and CT scans, leading to early detection and better patient outcomes.

Retail and E-commerce

Object recognition technology is widely used in retail and e-commerce for inventory management, product recognition, and customer engagement. Retailers can automate tasks like shelf stocking, product counting, and cashier-less checkout using object recognition systems. Additionally, personalized shopping experiences, such as visual search and recommendation systems, leverage object recognition is to understand customer preferences and behavior.

Environmental Monitoring

Object recognition aids in environmental monitoring and wildlife conservation efforts by tracking and identifying endangered species, monitoring habitat changes, and studying animal behavior. It enables researchers and conservationists to collect valuable data for ecological studies and wildlife management initiatives.

FLOWCHART

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

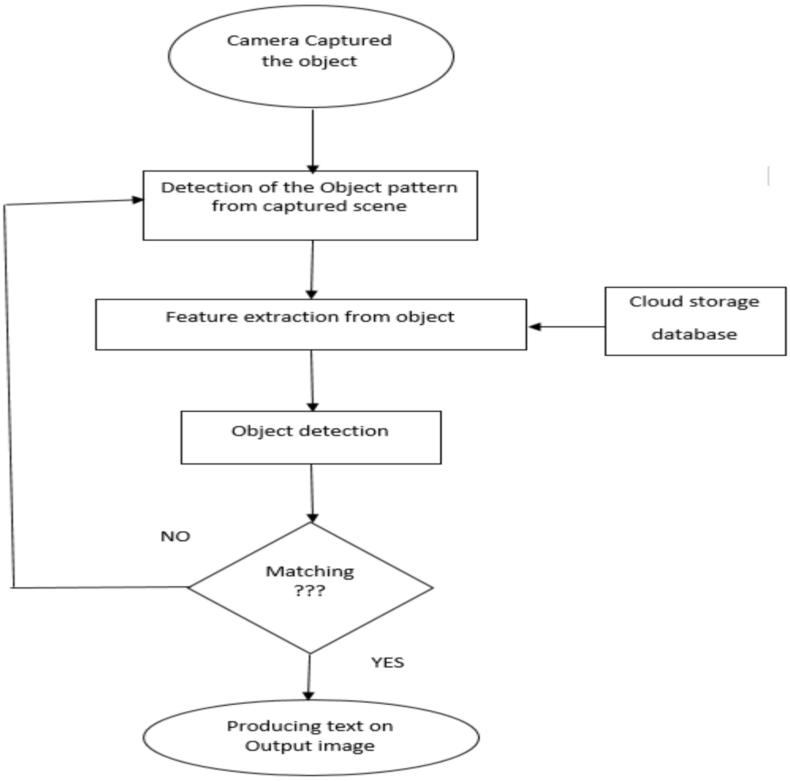

Fig. 9. Flowchart

OPENCV

OpenCV is an open-source computer machine learning software library primarily geared toward real-time computer imaginative and prescient packages. Developed initially through Intel, it's now maintained by using Willow storage and Itseez. OpenCV is written in C++ and has bindings for Python, making it handy to a huge variety of developers.

LATEST TRENDS IN MOVING OBJECT RECOGNITION

- Deep Learning with CNNs

Convolutional Neural Networks have been extensively applied for recognizing and tracking moving objects in video streams, delivering impressive real-time accuracy.

- One-shot and Few-shot Learning

Overcoming the need for large datasets, these techniques enable models to effectively recognize objects from minimal examples, enhancing their flexibility for new applications.

- Incorporation of Attention Mechanisms

These mechanisms help models to prioritize relevant parts in the input data, thus boosting both efficiency and accuracy in recognizing moving objects.

- Graph Neural Networks (GNNs)

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

For scenarios involving complex interactions among objects, GNNs offer a sophisticated way to capture contextual relationships by treating object interactions as graphs.

➢ Utilizing Transfer Learning

Employing models pre-trained on extensive datasets like ImageNet helps in adapting these systems to specific domains with less data requirement, boosting both efficiency and accuracy.

➢ Multi-modal Data Fusion

Collect data from multiple sources such as video, depth sensors, and motion sensors enhances the robustness of recognition systems. Techniques explored include early, late, and attention-based fusion strategies.

➢ Focus on Real-time Processing and Edge Computing

There's an increasing emphasis on algorithms capable of operating on devices with limited computing power, suitable for real-time applications like autonomous vehicles and smart surveillance.

➢ Enhancing Robustness against Variabilities

Efforts are underway to make recognition systems more resilient to adversarial attacks and environmental challenges like changes in light and occlusions.

➢ Ethical and Privacy Concerns

With broader adoption, the ethical implications and privacy issues of moving object recognition technologies are under more scrutiny, pushing for development focused on fairness, transparency, and preserving privacy.

These areas of exploration indicate the ongoing evolution in moving object recognition, influencing diverse fields as automotive technology, security, interaction systems, and robotics.

EMERGING TRENDS AND TECHNOLOGIES IN OBJECT RECOGNITION

➢ Deep Learning Advancements:

- CNNs & Vision Transformers:Modern architectures like ResNet, EfficientNet, and Vision Transformers are improving accuracy and efficiency.
- Transfer & Self-Supervised Learning: Techniques that leverage pre-trained models and unlabeled data are reducing the need for extensive datasets.

➢ 3D and Multi-Modal Recognition:

- 3D Models & LiDAR: Enhanced 3D object recognition through point clouds and LiDAR is critical for robotics and autonomous vehicles.
- Multi-Modal Learning: Integrating visual, audio, and textual data creates more robust recognition systems.

➢ Real-Time and Edge Computing:

- Edge Devices:* Advances in hardware allow real-time object recognition on devices like drones and IoT systems.
- Efficient Algorithms:Optimized models are enabling high performance on low-power devices.

➢ Explainable AI & Ethics:

- Interpretability: Techniques like Grad-CAM make models' decision processes transparent.
- Bias Mitigation: Efforts are focused on reducing biases to ensure fair outcomes.

➢ Synthetic Data and Generalization:

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

- Data Augmentation & Synthetic Data: These techniques improve model robustness and reduce the need for real-world data.

➢ Few-Shot & Zero-Shot Learning:

- Learning with Limited Data: Models are being trained to recognize new objects with minimal examples, which is essential for dynamic applications.

IMAGE PROCESSING

Image processing is a field within computer science and engineering that involves manipulating digital images using computational algorithms. It encompasses a broad range of techniques aimed at analyzing, enhancing, and interpreting digital images for various applications. In simpler terms, image processing involves altering or extracting information from images to achieve specific goals.

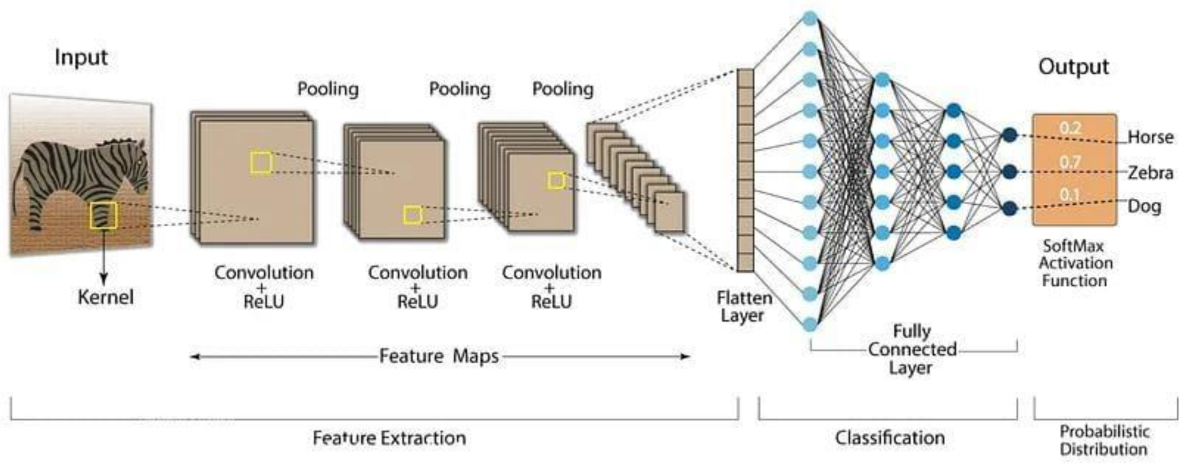

Fig. 10. Image Processing by using CNN

Image processing can be mainly divided into two main categories:

Digital Image Processing

This involves the manipulation of digital images using computational algorithms. Digital image processing techniques are implemented using software tools and programming languages on computers or specialized hardware platforms. These techniques it is widely used in such applications such as medical imaging, remote sensing, surveillance, digital photography, and computer vision.

Analog Image Processing

Analog image processing refers to manipulation of images using physical or optical techniques. While digital image processing has largely replaced analog methods in most applications, analog techniques are still used in specialized areas such as photography, cinematography, and certain scientific imaging methods.

Overall, image processing plays the crutial role in numerous fields and applications, including medicine, security, entertainment, and scientific research. It enables us to extract valuable information from images, enhance their visual quality, and automate tasks that would be challenging or time-consuming for humans to perform manually.

DIFFERENT ALGORITHMS USED FOR A MOVING OBJECT RECOGNITION

Moving object recognition involves detecting and tracking objects as they move across video frames of images. This task is supported by a range of algorithms, each tailored to specific aspects of the process. Below are several key types of algorithms used in moving object recognition**:**

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Background Subtraction Techniques

- Gaussian Mixture Models (GMM): Differentiates moving objects from the background by modeling each pixel's changes over time with multiple Gaussian distributions.
- Codebook based Methods: Represents the background with a codebook to detect foreground objects through deviations from the codebook.

Optical Flow Techniques

- Lucas-Kanade Method: Estimates motion by analyzing the displacement of small patches between frames.
- Horn-Schunck Method: Calculates optical flow by minimizing a global energy function that incorporates smoothness constraints.
- Farneback Method: Computes optical flow by exploiting polynomial expansions and a pyramidal approach.

Feature-based Tracking Techniques

- KLT (Kanade-Lucas-Tomasi) Tracker: Identifies and tracks key points over time using image gradients.
- Feature Matching: Detects and matches features like corners and blobs using feature descriptors such as SIFT (Scale-Invariant Feature Transform) or SURF (Speeded Up Robust Features).

Deep Learning Approaches

- Convolutional Neural Networks (CNNs): Applied for both detection and tracking, using models like YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector), and Mask R-CNN.
- Recurrent Neural Networks (RNNs): Useful for temporal data processing in object tracking.

Kalman Filters and Extensions

- Kalman Filter: Forecasts an object's trajectory based on its prior state and motion predictions.
- Extended Kalman Filter (EKF): Adapts the kalman filter for nonlinear state estimations.
- Unscented Kalman Filter (UKF): Improves on EKF by using sigma points for the better approximation in nonlinear scenarios.

Particle Filters (Sequential Monte Carlo Methods)

- Monte Carlo Localization*: Tracks object states using a cloud of particles, each representing a potential future state.

Graph-based Approaches

- Graph Cuts: Segments the image into objects by minimizing energy functions represented in graph forms.
- Markov Random Fields (MRFs): Models pixel or region interdependencies to enhance motion estimation and segmentation.

Motion Segmentation Techniques

- Foreground-Background Segmentation: Identifies moving objects by distinguishing them from static backgrounds.

Bayesian Filtering Methods

- Particle Filtering*: Updates predictions about an object's location using a probabilistic approach with sets of hypotheses (particles).

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Any of these algorithms provides offers different benefits and is chosen based on factors like scene complexity, object dynamics, required processing speed, and computational resources available. The selection often reflects a balance between accuracy, efficiency, and robustness to environmental changes.

RESULTS

In the analysis of video footage captured by a standard camera, whether in indoor or outdoor settings, a process is implemented to transform the video into individual frames. The initial frame is identified as the background image, while subsequent frames serve as real-time representations. Utilizing the background for subtraction technique, the image contrast is derived by subtracting the real-time image from the background image. This results in a unique image accentuating the differences between the two. Subsequently, we extract both the static frame, representing stationary elements in the scene, and dynamic boundaries that delineate moving components. The ensuing outcome involves outlining a red boundary around the identified moving object. This methodology effectively distinguishes static and dynamic elements in the video, offering useful information into areas and objects in motion.

Fig. 11. Output Image 1

Fig. 12. Output Image 2

CONCLUSION

In this paper has presented a novel approach to moving object recognition utilizing the capabilities of Python and OpenCV. Through integration for adaptive background subtraction, feature extraction, and object tracking algorithms, we have demonstrated the effectiveness of our methodology in accurately recognising and tracking objects in real-time video streams. Our experimental results have showcased the improved performance of an approach compared to existing methods, highlighting its potential for various real-world applications such as surveillance systems, intelligent transportation, and virtual reality environments. By leveraging the versatility of Python and the power of OpenCV, we create a solution that provides accuracy and efficiency in dynamic and chaotic environments. However, the limitations of our approach should also be acknowledged, including the reliance on historical models and computational sources. Required for real-time processing. Future research efforts could focus on addressing these limitations by exploring techniques for adaptive background modeling and optimizing algorithms for resource-constrained environments.

Journal of Information Systems Engineering and Management
2025, 10(43s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

REFERENCES

[1] Abidi, M. H., Alkhalefah, H., & Siddiquee, A. N. (2024). Enhancing Navigation and Object Recognition for Visually Impaired Individuals: A Gradient Support Vector Boosting-based Crossover Golden Jackal Algorithm Approach. *Journal of Disability Research*, *3*(5), 20240057.

[2] Ashiq, F., Asif, M., Ahmad, M. B., Zafar, S., Masood, K., Mahmood, T., ... & Lee, I. H. (2022). CNN-based object recognition and tracking system to assist visually impaired people. *IEEE access*, *10*, 14819-14834.

[3] Darwish, A., Halkon, B., & Oberst, S. (2022). Non-contact vibro-acoustic object recognition using laser Doppler vibrometry and convolutional neural networks. *Sensors*, *22*(23), 9360.

[4] Han, Y. B., Lei, P., & Donofrio, A. (2023). Road multi-vehicle recognition and tracking algorithm based on computer vision. *Advances in Transportation Studies*.

[5] Hussain, N., Khan, M. A., Kadry, S., Tariq, U., Mostafa, R. R., Choi, J. I., & Nam, Y. (2021). Intelligent deep learning and improved whale optimization algorithm based framework for object recognition. *Hum. Cent. Comput. Inf. Sci*, *11*(34), 2021.

[6] Kanagamalliga, S., Kovalan, P., Kiran, K., & Rajalingam, S. (2024). Traffic Management through Cutting-Edge Vehicle Detection, Recognition, and Tracking Innovations. *Procedia Computer Science*, *233*, 793-800.

[7] Kapoor, A., Shah, S., Agrawal, S., Kathiria, P., & Agrawal, S. Smart Auditorium Automation system based on object recognition.

[8] Kapoor, R., Goel, R., & Sharma, A. (2022). An intelligent railway surveillance framework based on recognition of object and railway track using deep learning. *Multimedia tools and applications*, *81*(15), 21083-21109.

[9] Khalaf, L. I., Aswad, S. A., Ahmed, S. R., Makki, B., & Ahmed, M. R. (2022, June). Survey On Recognition Hand Gesture by Using Data Mining Algorithms. In *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)* (pp. 1-4). IEEE.

[10] Kim, J., Bae, J., Park, G., Zhang, D., & Kim, Y. M. (2021). N-imagenet: Towards robust, fine-grained object recognition with event cameras. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 2146-2156).

[11] Liu, Q., Xing, D., Tang, H., Ma, D., & Pan, G. (2021, August). Event-based Action Recognition Using Motion Information and Spiking Neural Networks. In *IJCAI* (pp. 1743-1749).

[12] Li, Y., Zhou, H., Yang, B., Zhang, Y., Cui, Z., Bao, H., & Zhang, G. (2021). Graph-based asynchronous event processing for rapid object recognition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 934-943).

[13] Lomas-Barrie, V., Pena-Cabrera, M., Lopez-Juarez, I., & Navarro-Gonzalez, J. L. (2021). Fuzzy artmap-based fast object recognition for robots using FPGA. *Electronics*, *10*(3), 361.

[14] Martínez-Drudis, L., Amorós-Aguilar, L., Torras-Garcia, M., Serra-Elias, B., Costa-Miserachs, D., Portell-Cortés, I., & Coll-Andreu, M. (2021). Delayed voluntary physical exercise restores “when” and “where” object recognition memory after traumatic brain injury. *Behavioural Brain Research*, *400*, 113048.

[15] Nandhini, T. J., & Thinakaran, K. (2023, January). Detection of Crime Scene Objects using Deep Learning Techniques. In *2023 International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)* (pp. 357-361). IEEE.

[16] Noras, M. A. (2021). Activity detection and recognition with passive electric field sensors. *IEEE Transactions on Industry Applications*, *58*(1), 800-806.

[17] Pyo, J. W., Lee, K. H., Cho, J. S., & Kuc, T. Y. (2022, November). Self-Supervised Button Recognition for Indoor Mobile Robots. In *2022 Joint 12th International Conference on Soft Computing and Intelligent Systems and 23rd International Symposium on Advanced Intelligent Systems (SCIS&ISIS)* (pp. 1-5). IEEE.

[18] Rani, S., Ghai, D., Kumar, S., Kantipudi, M. P., Alharbi, A. H., & Ullah, M. A. (2022). Efficient 3D AlexNet architecture for object recognition using syntactic patterns from medical images. *Computational Intelligence and Neuroscience*, *2022*(1), 7882924.

[19] Roy, R., Kumar, S., Dumbhare, P., & Barde, M. (2021). Helmet Detection and Number Plate Recognition using Machine Learning. *Proceedings of the 2022 IEEE Region*, *10*.

[20] Santhanalakshmi, S. T., & Khilar, R. (2023, August). An Analysis Study on the Different Approaches of Deep Learning for Object Recognition and Detection. In *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)* (pp. 101-108). IEEE.
[21] Wang, S., Wang, H., Zhou, Y., Liu, J., Dai, P., Du, X., & Wahab, M. A. (2021). Automatic laser profile recognition and fast tracking for structured light measurement using deep learning and template matching. *Measurement*, *169*, 108362.
[22] Wurm, M. F., & Caramazza, A. (2022). Two 'what'pathways for action and object recognition. *Trends in cognitive sciences*, *26*(2), 103-116.
[23] Xuehui, A., Li, Z., Zuguang, L., Chengzhi, W., Pengfei, L., & Zhiwei, L. (2021). Dataset and benchmark for detecting moving objects in construction sites. *Automation in Construction*, *122*, 103482.
[24] Yin, Q., Hu, Q., Liu, H., Zhang, F., Wang, Y., Lin, Z., ... & Guo, Y. (2021). Detecting and tracking small and dense moving objects in satellite videos: A benchmark. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-18.
[25] Zhang, Z., Wang, S., Liu, C., Xie, R., Hu, W., & Zhou, P. (2022). All-in-one two-dimensional retinomorphic hardware device for motion detection and recognition. *Nature Nanotechnology*, *17*(1), 27-32.