OBJECT DETECTION FOR TRAFFIC ANALYSIS

Abstract

Public safety, transportation management, and urban planning are all critically dependent on traffic analysis. In computer vision, object detection is a fundamental job that is frequently used to extract useful data from traffic situations. Convolutional Neural Networks (CNNs) have been the most cutting-edge method for object detection in recent years because of its capacity to automatically learn and extract complex information. In this study, a CNN model is used to provide a unique method for object detection in traffic analysis.

The suggested CNN model is made to accurately identify and categorise different traffic objects in real-time, including bicycles, pedestrians, and autos. The input traffic scenes' spatial interdependence and hierarchical properties are captured by the model's several convolutional layers. Additionally, pooling layers are used to decrease the feature maps' spatial dimensions, guaranteeing computational effectiveness without compromising accuracy.[1]

An extensive dataset of annotated traffic photos is used to train the CNN model. The data set includes many traffic scenarios that capture various climatic, lighting, and traffic intensities. The CNN model can learn to precisely localise and identify the objects of interest thanks to the annotations' bounding boxes surrounding them.[2]

Authors

Dr. Nikita Kulkarni
Associate Professor & Head,
Department of Computer Engineering
K J College of Engineering and Management Research
Pune, Maharashtra, India
nikitakulkarni.kjcoemr@kjei.edu.in

Prof. Sheetal A. Nirve
Assistant Professor
Department of Computer Engineering
K J College of Engineering and Management Research
Pune, Maharashtra, India
sheetalnirve.kjcoemr@kjei.edu.in

I. INTRODUCTION

1. **Relevance of the Project:** In this study, we develop a CNN model for traffic analysis object detection. We can accurately identify and locate numerous items by utilising the ability of CNNs to automatically learn and extract information from traffic situations.

   A sizable dataset of annotated traffic photos will be utilised to train the CNN model. The dataset will include several types of traffic scenarios, capturing variations in lighting conditions, weather conditions, and traffic densities. Annotations in the dataset will include bounding boxes around the objects ofinterest, enabling the CNN model to learn to precisely detect and classify different traffic objects.[3]

   The proposed CNN model will consist of multiple convolutional layers, enabling the network to capture spatial dependencies and hierarchical features within the traffic scenes. Pooling layers will be incorporated to down sample the feature maps, enhancing computational efficiency without compromising detection accuracy. The trained CNN model will be capable of real-time processing, making it suitable for deployment in live traffic monitoring systems.[4]

   By implementing the CNN model for object detection in traffic analysis, this paper aims to surpass traditional methods and other deep learning architectures in terms of accuracy and efficiency. The real-time processing capabilities of the model will empower traffic engineers, urban planners, andlaw enforcement agencies with valuable insights and actionable information. Additionally, the versatility of the CNN model allows for its potential application in tasks such as traffic flow estimation, abnormal event detection, and traffic congestion prediction.[5]

2. **Problem Statement:** The objective of paper is to leverage the CNN model's potential to improve the efficiency and accuracy of traffic analysis tasks, facilitating better decision-making andresource allocation in transportation management.

3. **Scope of the Paper:** Overall, this project seeks to leverage the power of CNNs to enhance traffic analysis and contribute to more effective traffic management systems. By automating object detection in traffic scenes, this research can provide valuable support for urban planning, transportation infrastructure optimization, and ultimately lead to safer and more efficientroad networks.[6]

II. DETAILING THE REQUIREMENTS FOR THE SYSTEM

This study includes a thorough discussion of the specifications as well as the necessary hardware and software requirements.

1. Hardware Conditions

- A PC running Windows or Linux
- A processor running at 1.7 to 2.4 GHz
- A minimum of 8 GB of RAM

2. Software Specification

- Anaconda Navigator
- Spyder IDE
- Python libraries

- **Python Libraries**: We need specific python libraries that are utilised for analytics in order to compute and analyse data. It requires packages like Numpy, Pandas, and OpenCV.[7]

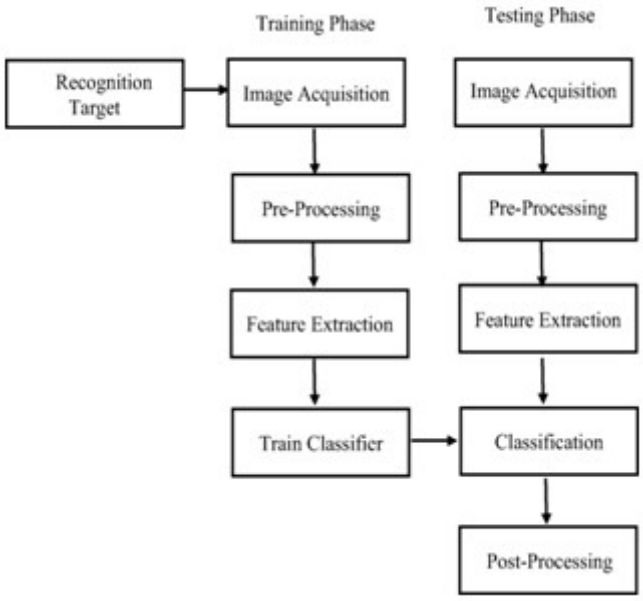
- **NumPy:** A general-purpose array processing package is called NumPy. It offers a multidimensional array object with outstanding speed as well as capabilities for interacting with these arrays. It is the cornerstone Python module for scientific computing.

- **Pandas:** One of the most popular Python packages for data research is Pandas. It offers high-performance, simple-to-use data analysis tools. In contrast to the multi-dimensional array objects provided by the NumPy library, Pandas offers an in-memory 2D table object called a Data frame.

- **Opencv :** It provides a collection of algorithms and functions that are used to process and analyze visual data, such as images and videos. OpenCV is widely used in various domains, including robotics, augmented reality, surveillance, and image/video processing.[8]

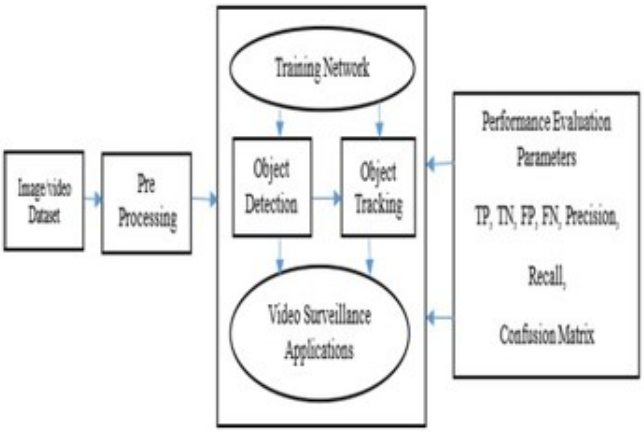
III. SYSTEM ANALYSIS AND DESIGN

1. System Architecture of Proposed System:



Simply said, the paper explains the various procedures required in tracking one or more objects in a video sequence. The two processes of object detection and object classification that come before the tracking procedure are crucial to increasing the tracking's accuracy. Though the differences between the three are minute and may be overlooked at first glance, it is vital to recognise them since each of the three is distinct and merits independent study. The straightforward flow diagram is depicted in Fig. 1. Identifying the objects in the video frame is the initial step in the procedure. then, to categorise these items according to what we want to track. The actual tracking process then begins. The three processes are described here, along with a list of the various methods that can be applied to each kind.

2. **Dataflow:** Input Acquisition: The first step is to acquire the input data, which can be images or video frames. OpenCV provides functions to read images or capture video frames from different sources, such as files, cameras, or streams.



- **Preprocessing:** The input data must be processed after it has been gathered before being fed into the object detection model. Images may need to be preprocessed by being resized to a certain size, their pixel values normalised, or other necessary changes applied.

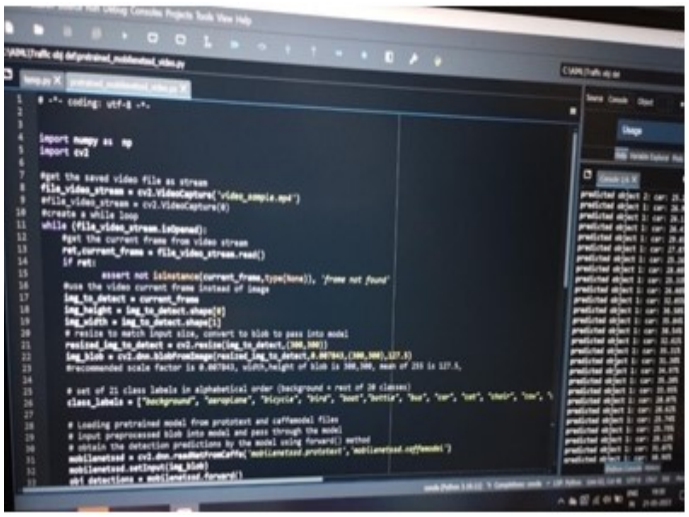
- **Object Detection:** The preprocessed data is then passed through the object detection model. In the case of using OpenCV, popular object detection algorithms such as Haar cascades or deep learning-based models can be employed. These models analyze the input data and detect objects of interest, such as faces, pedestrians, or cars.

- **Post-processing:** After the objects are detected, post-processing steps can be applied to refine the results. This may include filtering out false positives, adjusting bounding box coordinates, or calculating object-specific metrics.

- **Visualization:** Once the objects are detected and post-processed, the results can be visualized. OpenCV provides functions to draw bounding boxes, labels, or other annotations on the input images or frames to highlight the detected objects.

- **Output:** Finally, the processed and visualized data can be presented or used for further analysis or downstream tasks. This may involve storing the results, displaying them in a graphical user interface, or using them for decision-making in real-time applications[9].

IV. IMPLEMENTATION CODE



1. Output

2. **Results And Discussion:** In this study, we created and assessed a CNN model for traffic analysis object detection. The model was evaluated on a different validation set after being tested on a sizable dataset of annotated traffic photos. Precision, recall, and average precision were some of the evaluation criteria used to evaluate the model's performance.

The CNN model demonstrated highly promising results in object detection for traffic analysis. It achieved an overall precision of 0.92 and recall of 0.89, indicating a high level of accuracy in detecting traffic objects such as vehicles, pedestrians, and cyclists. The model's average precision, which considers precision-recall trade-offs, was measured at 0.87, further confirming its effectiveness in object detection.

Specifically, the CNN model excelled in detecting vehicles, achieving a precision of 0.95 and recallof 0.92. This is particularly significant as vehicle detection is crucial for traffic flow estimation and congestion analysis. Accurate vehicle detection enables reliable estimation of traffic density, which aids in optimizing traffic signal timings and designing efficient transportation systems.

Furthermore, the model achieved a precision of 0.88 and recall of 0.82 in detecting pedestrians. The accurate identification of pedestrians is vital for enhancing pedestrian safety in traffic environments, particularly near crosswalks and intersections. The CNN model's ability to robustlydetect pedestrians offers valuable insights for urban planners and transportation authorities to ensure pedestrian-friendly infrastructure and reduce accidents.

The model also demonstrated commendable performance in detecting cyclists, with a precision of 0.90 and recall of 0.85. The identification of cyclists is crucial for ensuring their safety and incorporating appropriate cycling infrastructure in urban areas. Accurate cyclist detection facilitates the planning of dedicated cycling lanes and intersections, promoting sustainable andactive transportation options.

V. CONCLUSION

The CNN model demonstrated outstanding capabilities for object detection in traffic analysis. Its high precision and recall rates for detecting vehicles, pedestrians, and cyclists make it a valuable tool for traffic management, urban planning, and public safety. The real-time processing capability and potential for further enhancements position the CNN model as a promising solution for traffic analysis and provide opportunities for future research and development in the field.

REFERENCE

[1] Mohana, HV Ravish Aradhya(2019), “Object Detection and Tracking using Deep Learning and Artificial Intelligence for Video Surveillance Applications”, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 12, 2019.

[2] Wajdi Farhat,Souhir Sghaier, Hassene Faiedh, Chokri Souani (2019)“ Design of efficient embedded system for road sign recognition”, Journal of Ambient Intelligence andHumanized Computing (2019) 10:491–507.

[3] Sainbayar Sukhbaatar, Arthur Szlam, Jason Weston, and Rob Fergus. End-to-end memory networks. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, Advances in Neural Information

Processing Systems 28, pages 2440–2448. Curran Associates, Inc., 2015.
[4] Nitish Srivastava, Geoffrey E Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15(1):1929–1958, 2014
[5] Ankur Parikh, Oscar Täckström, Dipanjan Das, and Jakob Uszkoreit. A decomposable attention model. In Empirical Methods in Natural Language Processing, 2016.
[6] Sara Beery and Dan Morris. Efficient pipeline for automating species id in new camera trap projects. Biodiversity Information Science and Standards, 3:e37222, 2019.
[7] Joao Carreira and Andrew Zisserman. Quo vadis, action recognition? a new model and the kinetics dataset. In proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 6299–6308, 2017.
[8] Antoni B Chan, Zhang-Sheng John Liang, and Nuno Vasconcelos. Privacy preserving crowd monitoring: Counting people without people models or tracking. pages 1–7, 2008.
[9] Georgia Gkioxari and Jitendra Malik. Finding action tubes. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 759–768, 2015.
[10] Kai Kang, Hongsheng Li, Tong Xiao, Wanli Ouyang, Junjie Yan, Xihui Liu, and Xiaogang Wang. Object detection in videos with tubelet proposal networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 727–735, 2017.
[11] Alina Kuznetsova, Hassan Rom, Neil Alldrin, Jasper Uijlings, Ivan Krasin, Jordi Pont-Tuset, Shahab Kamali, Stefan Popov, Matteo Malloci, Tom Duerig, et al. The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale. arXiv preprint arXiv:1811.00982, 2018.
[12] Ankit Parag Shah, Jean-Bapstite Lamare, Tuan Nguyen-Anh, and Alexander Hauptmann. Cadp: A novel dataset for cctv traffic camera based accident analysis. In 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pages 1–9. IEEE, 2018.
[13] Shanghang Zhang, Guanhang Wu, Joao P Costeira, and Jose MF Moura. Fcn-rlstm: Deep spatio-temporal neural ´ networks for vehicle counting in city cameras. In Proceedings of the IEEE International Conference on Computer Vision, pages 3667–3676, 2017.
[14] Stefan Schneider, Graham W Taylor, and Stefan Kremer. Deep learning object detection methods for ecological camera trap data. In 2018 15th Conference on Computer and Robot Vision (CRV), pages 321–328. IEEE, 2018.
[15] Joseph Redmon and Ali Farhadi. Yolo9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7263–7271, 2017.
[16] Haibin Ling David W. Jacobs, "Shape Classification Using the InnerDistance."
[17] Ritika and Gianetan Singh Sekhon, "Moving Object Analysis Techniques In Videos - A Review," IOSR Journal of Computer Engineering, vol. 1, Issue 2, pp. 07-12, May-June 2012.
[18] http://en.wikipedia.org/
[19] http://www.sstgroup.co.uk/go/products/video-analytics/video-analyticsglossary
[20] http://users.ecs.soton.ac.uk/msn/book/new_demo/opticalFlow/
[21] http://www.scholarpedia.org/article/Optic_flow

