Mechatronics and Intelligent Transportation Systems
https://www.acadlore.com/journals/MITS

Advanced Vehicle Detection and License Plate Recognition via the Kanade-Lucas-Tomasi Technique

Egina Nyati[1*], John Sabelo Mahlalela[2]

[1] Fuels and Energy Department, Midlands State University, ZW170407 Gweru, Zimbabwe
[2] Electrical and Electronic Engineering Department, University of Eswatini, M200 Kwaluseni, Eswatini

[*] Correspondence: Egina Nyati (nyatie@staff.msu.ac.zw)

Received: 10-18-2023 **Revised:** 10-25-2023 **Accepted:** 11-02-2023

Citation: E. Nyati ahd J. S. Mahlalela, "Advanced vehicle detection and license plate recognition via the Kanade-Lucas-Tomasi technique," *Mechatron. Intell Transp. Syst.*, vol. 2, no. 4, pp. 191–200, 2023. https://doi.org/10.56578/mits020401.

Abstract: The optimization of traffic flow, enhancement of safety measures, and minimization of emissions in intelligent transportation system (ITS) pivotally depend on the Vehicle License Plate Recognition (VLPR) technology. Challenges predominantly arise in the precise localization and accurate identification of license plates, which are critical for the applicability of VLPR across various domains, including law enforcement, traffic management, and both governmental and private sectors. Utilization in electronic toll collection, personal security, visitor management, and smart parking systems is commercially significant. In this investigation, a novel methodology grounded in the Kanade-Lucas-Tomasi (KLT) algorithm is introduced, targeting the localization, segmentation, and recognition of characters within license plates. Implementation was conducted utilizing MATLAB software, with grayscale images derived from both still cameras and video footage serving as the input. An extensive evaluation of the results revealed an accuracy of 99.267%, a precision of 100%, a recall of 99.267%, and an F-Score of 99.632%, thereby surpassing the performance of existing methodologies. The contribution of this research is significant in addressing critical challenges inherent in VLPR systems and achieving an enhanced performance standard.

Keywords: Intelligent transportation system (ITS); MATLAB; Vehicle detection; Kanade-Lucas-Tomasi (KLT) algorithm

1 Introduction

The transformative impact of technological advancements on numerous facets of daily life is undeniable, and the realm of transportation management stands as a prime example [1]. Within the ambit of ITS, the development of VLPR systems has been identified as crucial for achieving optimal traffic management and bolstering security measures. These systems are adept at capturing and deciphering vehicle license plate numbers, thereby offering a broad spectrum of applications encompassing access control in parking lots, crime prevention, and traffic analysis. The origins of VLPR can be traced back to the 1970s, marking the commencement of efforts to automate the reading of license plates [2]. Yet, it was the advent of digital cameras and image processing techniques that propelled the widespread adoption of VLPR systems. Early iterations of these systems predominantly employed rule-based approaches, utilizing manually crafted features to discern license plate characters [3, 4]. These initial systems, however, demonstrated limitations in adapting to variances in lighting conditions, image quality, and license plate formats. A paradigm shift towards machine learning-based approaches for VLPR has been observed in recent years [5–7]. Such algorithms, capable of learning from data, exhibit enhanced robustness to fluctuations in image quality and conditions. VLPR systems, it is acknowledged, occupy a central role across diverse sectors, finding application in scenarios ranging from toll collection and traffic enforcement to border control and vehicle tracking [8, 9]. The present research is committed to contributing to the evolution of VLPR technology, with the objective of forging a system characterized by accuracy, robustness, and efficiency [10, 11]. Emphasis will be placed on image enhancement, license plate detection, character segmentation, character recognition, and the optimization of the overall system.

This paper's structure is as follows: Section 2 provides a comprehensive review of the VLPR literature. In Section 3, the research methodology designed to address existing challenges in VLPR and improve its performance is outlined, with innovative techniques integrated to enhance accuracy, robustness, and efficiency. Empirical findings are presented in Section 4, where performance metrics, comparative studies, and practical insights are explored. The paper concludes in Section 5, emphasizing the contributions of this work to the advancement of VLPR systems and the assurance of safer and more efficient transportation networks.

2 Literature Review

The review encompasses an exploration of various methodologies pertinent to object tracking, with a particular emphasis on applications within traffic monitoring and surveillance contexts. Techniques such as video analytics, vehicle detection, and motion tracking are scrutinized. The evaluation encompasses a range of algorithms, including ne-Class Support Vector Machine (OC-SVM) and Convolutional Neural Network (CNN)-based approaches, with a focus on augmenting accuracy and mitigating false alarm occurrences. The overarching aim is articulated as the enhancement of object tracking efficacy in real-time scenarios, inclusive of challenging environmental conditions. In the work of Velazquez-Pupo et al. [12], a stationary camera is utilized in a video analytics context, serving multifarious functions such as vehicle detection, occlusion handling, vehicle counting, tracking, and classification. Within this context, the application of OC-SVM with an RBF Kernel is highlighted, having demonstrated superior performance, particularly in the classification of midsize vehicles, yielding a F-measure of 98.190% and 99.051% respectively. It is underscored that SVM is acknowledged as the optimal classifier in this scenario. Furthermore, the research conducted by Qu et al. [13] is brought into focus, advocating for the implementation of an accurate moving vehicle detector. This encompasses the incorporation of techniques such as candidate target recognition, CNN-based vehicle screening, and the application of motion sensors with image normalization for real-time scenarios, aiming for a high detection rate. Empirical studies employing diverse datasets underscore the effectiveness of moving vehicle detection, achieving up to 90% detection performance for automobiles, while maintaining an average false alarm rate below 10%.

In the work presented by Sarcevic and Pletl [14], a novel technique has been introduced for the filtration of false alarms. Regulations have been constructed separately, based on various data types derived from the signals, serving as the foundation for the filtration process. The parameters exerting the most significant influence were subjected to independent examination across each data type. Subsequently, these parameters were amalgamated into sophisticated algorithms to yield more precise outcomes. Optimization of the model parameters was achieved through the application of evolutionary algorithms. Results garnered from this approach indicate that 97% of false detections could be successfully eliminated, with a negligible loss of 0.3% in accurate detection systems, when rules are meticulously crafted. It was observed that even the application of a singular parameter could facilitate this process.

In a separate study conducted by Guo et al. [15], an augmented Single Shot MultiBox Detector (SSD) method has been proposed, aimed at addressing the shortcomings associated with low accuracy and missing detections in existing SSD methodologies for object tracking. The backbone of the proposed SSD network is ResNet50, selected for its capability to extract intricate details pertaining to vehicle features. The Feature Fusion Model, designed to enhance the accuracy of small target vehicle recognition, amalgamates positional data from shallow features with semantic information from feature representation. The incorporation of a Squeeze-and-Excitation (SE) block within the feature extraction layer further augments the model's performance, enabling more comprehensive feature extraction and a reevaluation of the channel's significance. Experimental findings attest to the efficacy of the modified approach, as evidenced by an average accuracy of 83.09% on a dataset comprising home-built vehicles, surpassing the accuracy of the preceding algorithm by 3.23% . The work of Ma et al. [16] introduces the Partial Anchors based Detection Network (PADeN), advocating for the identification and subsequent removal of incomplete anchors on vehicles to expedite the object detection process significantly. Contextual information is utilized within PADeN to discern and discard unnecessary anchors, enhancing the efficiency of object detection in images. The integration of the centerness mask branch into the network is highlighted as a pivotal enhancement to PADeN's performance. Results from this study indicate a Mean Average Precision (mAP) of 76.9%, positioning PADeN as a superior method in comparison to previous object tracking methodologies.

In another study, Barnouti et al. [17] propose the utilization of the KLT tracker in conjunction with the Two-Dimensional Principal Component Analysis (2DPCA) tracker for the purpose of monitoring and recognizing facial features within video sequences. The initial phase employs the Viola-Jones face identification technique for face detection in images or video sequences, followed by the application of the KLT method for face tracking. The KLT tracker maintains a long-term tracking capability of facial objects across successive frames, ensuring continuity even in instances of facial appearance and disappearance. The 2DPCA feature extraction method is utilized for noise reduction and enhancement of face recognition through a distance classifier. The proposed methodology undergoes validation using the Face94 database and webcam images. Experimental results confirm the efficacy of the Viola-Jones method in frontal face detection, the proficiency of the KLT system in face tracking across diverse webcam-shot videos, and the successful face recognition capabilities of 2DPCA in both the Face94 dataset and computer webcam video series.

In the work of Yue [18], a recursive tracking system oriented towards Augmented Reality (AR) for human motion tracking is introduced. This system leverages the positional relationship between consecutive frames, employing the KLT approach in tandem with Oriented Rapid and Rotated Brief (ORB) feature descriptors. The KLT tracking technique is applied to track the ORB feature descriptor, matching the first frame image and the reference image, while concurrently tracking the feature descriptor from the preceding frame in the current frame. This approach effectively mitigates the phenomenon of virtual object jitter. Comparative analysis reveals that the recursive tracking method surpasses the detection tracking strategy in terms of both speed and accuracy. Nevertheless, the study acknowledges the existence of challenges, particularly the inability to develop a feature tracking technique with enhanced accuracy and extended tracking longevity to diminish or mitigate the effects of cumulative error.

In the work conducted by Ramakrishnan et al. [19], an investigation into the optimization of the window size in the KLT tracking algorithm was presented, emphasizing the necessity of adapting the window size to mitigate the impact of distortions surrounding each feature point. The researchers introduced an adaptive window size technique, employing the iterations of the KLT algorithm as a metric to assess the quality of the tracks and consequently determine the optimal window sizes. Experimental results from well-established tracking datasets indicated that this adaptive approach exhibits enhanced robustness in comparison to the conventional fixed-window KLT, and offers a comparable level of robustness to the affine KLT, all the while achieving an average runtime speedup of seven-fold.

The system "Traffic Sensor" was introduced by Fernández et al. [20], employing deep learning techniques for the automatic detection and classification of vehicles on highways, utilizing a stationary, calibrated camera. The models were trained on a novel traffic image dataset, inclusive of images captured under sub-optimal lighting and weather conditions, as well as low-resolution images. The system is comprised of two principal modules: the first responsible for vehicle detection and classification, and the second for vehicle tracking. Extensive evaluation and comparison of various neural models were conducted for the first module, culminating in the selection of a network based on YOLOv3 or YOLOv4, trained on the new traffic dataset. The second module integrates a straightforward spatial association technique with the more intricate KLT tracker for the tracking of moving vehicles. Validation of the system was undertaken through numerous tests on challenging traffic videos, demonstrating the system's capability to effectively and real-time detect, track, and classify vehicles on highways.

Yin et al. [21] detailed the development of an optical flow target tracking system based on the KLT algorithm, implemented on the OpenCV platform and evaluated in the context of a water pipeline intelligent inspection competition. The technique leverages the optical flow method, aiming to achieve high detection certainty and rapid operational speed for frame differentiation, with a particular focus on underwater target detection and localization. The system ensures the stable control error of an underwater vehicle's motion through the application of incremental Proportional-Integral-Derivative (PID) control.

Several limitations have been identified in the prevailing state-of-the-art algorithm employed for the tasks of vehicle detection and tracking. These constraints are primarily attributed to the algorithm's inherent complexity, its operation within a confined frequency range for feature extraction, and its reliance on the minimum enclosing rectangle (MER) as a mechanism for object detection. The adoption of a SVM for the task of classification, particularly in scenarios characterized by high traffic volumes and noisy datasets, has been observed to yield suboptimal performance [22, 23]. The algorithm's effectiveness is further compromised by its dependence on a fixed threshold value, a factor that serves to impede its adaptive capabilities.

The proposed approach distinguishes itself through several innovative facets:

•The employment of Haar features in conjunction with the KLT algorithm is central to the development of a vehicle detection and tracking algorithm, which is anticipated to demonstrate both computational efficiency and robustness.

•A deep learning model is integrated with the aim of enhancing the accuracy of the proposed algorithm, particularly when applied to extensive datasets and those characterized by the presence of noise.

•The introduction of a novel thresholding technique is proposed, with the objective of rendering the algorithm less susceptible to variations in threshold values.

In alignment with these innovative aspects, the study sets forth several key research objectives:

•A methodological framework is to be established for the processing and analytical examination of real-time video data, with a particular focus on the accurate localization of vehicle license plates. This effort is expected to significantly contribute to the location and retrieval of lost automobiles.

•The capabilities of Haar features and the KLT algorithm are to be harnessed for the detection and tracking of vehicles within video streams.

•A rigorous comparative analysis is planned, wherein the proposed methodology will be evaluated against existing approaches. This evaluation will utilize a comprehensive set of performance metrics, including but not limited to accuracy, efficiency, recall, and precision, ensuring a thorough assessment of the methodologies in question.