International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Multimodal Deep Learning Information Fusion for Fine-Grained Traffic State Estimation and Intelligent Traffic Control

Aruna Kumar Joshi[1], Shrinivasrao B. Kulkarni[2]

Submitted: 24/04/2023 **Revised**: 25/06/2023 **Accepted**: 06/07/2023

Abstract: Traffic analysis from real time camera images is the most adopted method for traffic state estimations due to many reasons like easy installation, low cost and portability. Many methods have been proposed to estimate traffic density. The current methods only provide an estimate in terms of number of vehicles on the lane or volume of lane occupied. But this information alone is not sufficient for intelligent traffic management. A more accurate traffic state estimation in terms of density, speed, flow for different category of vehicles in different segments of lane is needed in presence of various clutters and occlusions. This research work addresses this problem and proposes a multimodal deep learning information fusion based framework for traffic state estimation in presence of clusters and occlusions. The multimodal deep learning information fusion based framework has three important components. The first component is vehicle categorization using hybrid traditional and deep learning features and estimation of traffic for vehicle category. The second component is categorization of different regions of road segments for presence of emergency vehicles using both visual and audio cues. The third component is construction of traffic state map and usage of the traffic state map for intelligent traffic control. In this work, fine-grained traffic state information refereed as a traffic state map is constructed using deep learning models. A traffic state map is dynamic with information on vehicular density, movement, and flow information. Compared to round-robin-based traffic scheduling at traffic signals, can realize more effective traffic scheduling with the traffic information map. Real-time camera-based traffic analysis has become the most widely adopted method for estimating traffic states due to its easy installation, cost-effectiveness, and portability. Although various methods have been proposed for estimating traffic density, the existing approaches typically provide only a basic estimate in terms of the number of vehicles or lane occupancy. However, this limited information falls short for intelligent traffic management. To address this issue, this research introduces a novel approach based on multimodal deep learning information fusion for accurate traffic state estimation, considering density, speed, and flow for different vehicle categories within various lane segments, even in the presence of clutters and occlusions. The proposed framework consists of three crucial components. The first method involves vehicle categorization using a combination of traditional and deep learning features to estimate traffic for each vehicle category. The second method focuses on categorizing different road segments to detect the presence of emergency vehicles, utilizing both visual and audio cues. Finally, the third method entails constructing a dynamic traffic state map using deep learning models. This traffic state map provides fine-grained information on vehicular density, movement, and flow. By leveraging the traffic state map, intelligent traffic control can be achieved, allowing for more effective traffic scheduling compared to conventional round-robin-based traffic signal scheduling. This framework holds promise for enhancing traffic management and optimizing the flow of vehicles on the road

Key words: *Traffic state Mapping, Lane Density, Vehicle Priority, Emergency Vehicle, Flow Estimation, Lane Density*

1. Introduction

The rapid increase in vehicular traffic in cities worldwide has presented a significant challenge in manually regulating traffic flow at numerous intersections throughout the day. This traditional approach not only proves costly but also becomes tedious and less efficient as traffic volumes continue to surge. To address these issues, Intelligent Traffic Systems (ITS) have emerged as a promising solution. ITS aims to provide innovative and automatic services to manage various modes of transport and traffic effectively

[1]*Dept of CSE, SKSVMA College of Engineering and Technology, Laxmeshwar, and Visvesvaraya Technological University, Belagavi, Karnataka, INDIA*

[2]*Dept of CSE, SDM College of Engineering and Technology, Dharwad, and Visvesvaraya Technological University, Belagavi, Karnataka, INDIA*
** Corresponding Author Email: arunkumarjoshi.sudi@gmail.com*

One of the key services offered by ITS is the implementation of automated traffic lights. Instead of relying on fixed timing, these traffic lights dynamically adjust their green light sequence and duration based on real-time traffic state information. By doing so, they can efficiently manage traffic during peak hours, congestion, emergencies, and accidents. This dynamic traffic light control proves invaluable in optimizing traffic flow and reducing delays, contributing to smoother and safer road networks. Notably, emergency vehicles, such as ambulances, police cars, and fire engines, require special attention due to the critical nature of their missions. It is imperative to prioritize their transit to avoid potential loss of lives and property. ITS plays a vital role in detecting and managing the traffic around emergency vehicles, promptly clearing pathways, and minimizing congestion to enable swift emergency response.

Achieving the automated traffic light service and efficient traffic management hinges on fine-grained traffic state estimation. This crucial component involves gathering detailed traffic map information, encompassing various parameters like vehicle types, density, speed, flow, and emergency events across different intersections in the road networks. Such comprehensive and accurate data forms the foundation for making informed decisions and orchestrating the optimal functioning of the ITS. The integration of Intelligent Traffic Systems with fine-grained traffic state estimation proves to be an essential step forward in revolutionizing traffic management. By leveraging real-time data and dynamic traffic light control, cities can enhance traffic flow, alleviate congestion, and ensure timely response to emergencies, ultimately leading to safer and more efficient transportation networks.

The process of estimating traffic states has evolved significantly over time, and currently, the dominant method involves the analysis of real-time camera images. This approach holds immense appeal due to a variety of compelling reasons that make it the preferred choice for traffic state estimation. Notable among these reasons are its simplicity in terms of installation, its cost-effectiveness, and its high level of portability.

Within the realm of traffic analysis, researchers and practitioners have proposed a multitude of techniques, as extensively documented in the existing literature, aimed at reducing traffic density from the data obtained through these camera images. Despite the multitude of methods available, the existing approaches often come up short when it comes to providing a comprehensive and holistic representation of the actual traffic conditions. This limitation inevitably hampers their effectiveness in contributing to the field of intelligent traffic management.

Primarily, these existing methods offer only rudimentary estimations, such as the count of vehicles present within a particular lane or the portion of the lane that is occupied by vehicles. While these data points do offer a general overview of the ongoing traffic scenario, they ultimately lack the depth and granularity required for truly optimizing the management and control of traffic flow.

For traffic management to reach a higher level of sophistication and intelligence, a more nuanced and detailed approach to traffic state estimation is imperative. This refined approach needs to encompass a range of essential parameters, including not just the number of vehicles, but also characteristics like vehicle types, density, speed, and flow rates. This wealth of finely-tuned information equips traffic managers with a profound comprehension of the dynamic interplay occurring on the roadways. Consequently, this allows them to craft targeted strategies aimed at enhancing both the overall efficiency and safety of the traffic ecosystem.

By integrating these additional layers of information into the traffic state estimation process, transportation authorities are empowered to make decisions that are not just well-informed, but also highly strategic. This encompasses actions such as dynamically adjusting traffic signal timings, managing lanes effectively, and regulating traffic flow in a responsive manner. This elevated approach paves the way for a traffic management system that is both adaptive and nimble, capable of swiftly adapting to the ever-changing conditions on the road and addressing instances of congestion with remarkable precision and efficacy.

While the real-time camera-based analysis method is indeed the cornerstone of modern traffic state estimation, it is evident that there is a pressing need for further advancements to truly elevate the quality and accuracy of the data obtained. A comprehensive approach that weaves together data on vehicle types, density, speed, and flow is an absolute necessity in achieving the pinnacle of intelligent traffic management. Ultimately, this integrated approach not only promises improved traffic flow and diminished congestion, but also serves to enhance the overall efficiency of transportation networks, offering a smoother, safer, and more sustainable mobility experience for all.

This research tackles the urgent requirement for an advanced system to estimate the state of traffic by introducing a novel approach based on deep learning-powered image analysis. The proposed method is designed to offer a comprehensive and highly accurate assessment of traffic conditions by considering both the types of vehicles present and the different sections of lanes they occupy. In contrast to existing techniques that primarily concentrate on using deep learning to understand the overall density of lanes, this study introduces a more intricate strategy that involves a finely detailed segmentation of the traffic.

This new method employs deep learning models on a per-segment basis, enabling a meticulous examination of traffic behaviors within specific areas of the road. By consistently capturing images, the system derives estimations for the speed and flow of distinct categories of vehicles across various lane segments. This amalgamation of information eventually culminates in the creation of a comprehensive traffic state map, encompassing the entirety of the lane under consideration. The resulting traffic state map delivers a holistic and comprehensive depiction of the traffic situation, incorporating crucial details such as vehicle types, densities, speeds, and flows.

The adaptive nature of this approach guarantees that traffic signals respond dynamically to the ever-changing conditions of traffic. This adaptability is achieved by integrating the real-time traffic state information map into the signal control strategy. In contrast to the conventional traffic signal control method, which typically follows a fixed, round-robin approach, the suggested approach harnesses the power of the traffic state map to dynamically schedule traffic signals. This proves to be remarkably effective, as it can seamlessly adapt to the dynamic fluctuations in traffic patterns.

In summary, this research represents a significant leap forward in the realm of traffic state estimation and control. The potential outcomes are noteworthy and encompass improved traffic flow, heightened efficacy in emergency response scenarios, and a notable reduction in congestion on roadways. By utilizing deep learning techniques and a finely segmented analysis approach, this work not only enhances the accuracy of traffic assessment but also lays the foundation for a more intelligent and responsive traffic management system.

2. Literature Survey

The existing works aiding the realization of smart traffic management are detailed below. Fedorov et al (2019) used Recurrent Convolution Neural Networks (R-CNN). Vehicles are counted and classified based on direction[1]. The vehicles were classified only based on the direction of travel without fine-grained classification like heavy, light, or emergency vehicles. Without this fine-grained vehicle classification, traffic management is not effective. Jung et al (2017) classified eleven different vehicles like trucks, buses, cycles, etc using an ensemble of joint fine-tuning (JF) and Drop CNN[2]. The approach could not satisfy vehicles in presence of dense traffics.

Shanghang Zhang et al (2017) estimated traffic density from videos using a deep learning model. The density is estimated as a whole without the fine-grained classification of vehicles[3]. Guerrero et al (2015) addressed the problem of traffic density estimation by mapping image features to density levels using CNN. The image as a whole is mapped to density estimate without consideration for fine-grained vehicle category[4]. Olmedo et al (2015) estimated the traffic density by mapping morphological features to traffic density. The method was designed for the case of free-flowing, slow-moving, and stationary scenes[5]. The estimate provided by the method is not fine-grained. Garg et al(2015) estimated the traffic density in terms of percentage occupancy of roads. . The region of interest in lane is divided into blocks[6]. The block size is smaller than the length of the smallest vehicle. The variance of pixel intensities over the blocks across the frames is used as criteria to decide the occupancy of the blocks. The estimate is erroneous in presence of occlusions and it is not fine-grained.

HongyuHu et al (2019) estimated three different traffic densities of light, medium, and heavy from image pixel intensities. Histograms of multi-scale block local binary pattern features are extracted from the image and classified to traffic densities using a multi-SVM classifier[7]. But the method does not support fine-grained vehicle classification. Kurniawan et al (2017) estimated the traffic density based on on-road occupancy. The background is constructed based on edge detection and subtracted from the foreground regions[8]. The remaining foreground region is then used for density estimation. Dinani et al (2015) classified the traffic density as light, medium, or heavy-based on features extracted from videos[9]. Texture and edge histogram features are extracted from frames of videos to classify the traffic density.

Wei et al (2016) estimated traffic density from the texture patterns of the image. A texture feature based on the energy and entropy in GLCM is obtained for the images[10]. A threshold for the texture feature is learned from different congested and unobstructed images. The method can classify the traffic density only to two levels and it does not provide fine-grained classification. Chakraborty et al (2018) classified the image as traffic density using two different learning techniques. Deep learning is trained to provide percentage occupancy as result[12]. Fine-grained classification is not provided in this method. Nguyen et al (2017) estimated traffic density from images using histogram analysis. Histogram features extracted from the image are used to train a neural network to classify traffic density[13][14].

Donato et al (2019) classified the image into two different traffic densities congested and not congested. Features of SIFT, gray histogram variance, GLCM energy, and contrast are extracted from the images and classified using a neural network into two states congested and not congested[15]. Impedovo et al (2019) extracted visual features from the image and estimated the traffic state. The traffic state covered in this work has parameters of total vehicles, velocity, volume, and road occupancy. Morphological features are extracted from the images and classified into traffic states. Shi et al (2016) used the learning-based aesthetic model to estimate traffic state. Perceptual features extracted from the image are classified using a random forest classifies to traffic state in terms of congestion level[16]. Nguyen et al (2016) classified traffic images into two states congested or not. SURF features were extracted from the image and classified using an SVM classifier into two states of congested or not[17].

Chaudhary et al (2018) used Gaussian mixture models to classify the traffic image as congested or not congested. Spatial interest points are extracted from the image and the Gaussian mixture model is applied over the interest points[18]. The resulting features are classified using an SVM classifier into two traffic states of congestion or not congestion.

The cited literature, "A real-time traffic congestion detection system using on-line images" [19] The key objective of this research is likely to leverage the power of image analysis and real-time data to detect traffic congestion. Traditional methods of traffic monitoring, which often rely on stationary sensors or manual observations, can be limited in terms of accuracy and coverage. In contrast, using online images captured from sources such as traffic cameras or smartphones can provide a dynamic and extensive view of traffic conditions

"Road traffic density estimation using microscopic and macroscopic parameters" by Asmaa O., Mokhtar K., and Abdelaziz O., likely published in the journal "Image and Vision Computing[20] The primary focus of this research is likely to develop a comprehensive method for estimating road traffic density. Traffic density, which refers to the number of vehicles occupying a unit length of road, is a critical metric in traffic engineering and management. V. Nair et al [21] The main focus of this research is likely on improving the training and performance of restricted Boltzmann machines (RBMs) through the use of rectified linear units (ReLUs). Both RBMs and ReLUs are fundamental components in the field of deep learning

From the survey, most of the current traffic density methods do not provide fine-grained distribution of traffic and vehicle categories. They provide only limited information about the traffic [22]. So it becomes to implement sophisticated traffic management or scheduling algorithms with the limited information provided by these solutions. This paperwork addresses this problem and provides rich traffic state information for sophisticated traffic management.

3. Proposed Traffic State Map Construction

The innovative concept introduced in this study is the notion of "traffic state information," a pivotal component that takes the form of a dynamically evolving traffic state map. This map is meticulously generated for every lane located at the intersection of a traffic signal. To enhance its visual comprehensibility, the traffic state map adopts a color-coded scheme, with each distinct vehicle category being assigned a unique color.

The methodology behind constructing this map is intricate yet highly effective. A grid system, comprising a matrix of m×n grids, is superimposed over the entirety of the lane's expanse. This grid network effectively covers the spatial domain of the lane. Each individual grid square is then imbued with a specific color code, reflective of the category of vehicle that occupies it. This segmentation permits a granular representation of traffic diversity within the lane.

Furthermore, the map not only presents the distribution of vehicle categories but also serves as a means to convey critical time-related data. Specifically, the waiting time endured by vehicles over each grid square is calculated and seamlessly integrated alongside the associated color code. This augmentation adds an additional layer of insight, allowing for a comprehensive understanding of how traffic is flowing and evolving over time within each lane.

This innovative approach to visualizing traffic state information through color-coded dynamic maps, supported by a finely meshed grid overlay and waiting time calculation, holds the potential to revolutionize traffic management and signal control. The adoption of this method offers an immediate, comprehensive grasp of traffic dynamics, enabling efficient decision-making processes and responsive traffic signal control strategies. This advancement is not only innovative but also pragmatic, paving the way for more adaptable and effective traffic management solutions.

The visual representation provided below serves as an illustrative depiction of a sample traffic state map designed for a specific lane. This illustrative example vividly demonstrates the seamless integration of color-coded grids, each complemented by associated waiting time data. This pioneering approach offers a comprehensive and nuanced comprehension of the prevailing traffic conditions, thereby facilitating the formulation of judicious traffic management strategies. In this visual representation, the lane under consideration is meticulously divided into an array of grids, forming a structured overlay. Each grid unit is thoughtfully assigned a distinctive color that corresponds to the particular category of vehicles occupying that region. This color-coded differentiation provides an immediate visual cue regarding the types of vehicles present within each section of the lane.

What further elevates the utility of this map is the inclusion of waiting time data. Alongside the color-coded grids, numerical values representing the waiting time endured by vehicles in each grid square are seamlessly incorporated. This time-based information enriches the map's insights by offering a dynamic perspective of traffic patterns. The amalgamation of color-coded categories and waiting time data collectively provides a comprehensive snapshot of traffic dynamics.

Ultimately, this type of traffic state map presents a novel and pragmatic approach to traffic management. By offering a clear, detailed, and dynamic portrayal of traffic conditions, this visualization methodology empowers decision-makers with the information needed to implement more efficient and responsive traffic management strategies. This innovation holds immense promise for alleviating congestion, enhancing traffic flow, and optimizing overall road network performance.

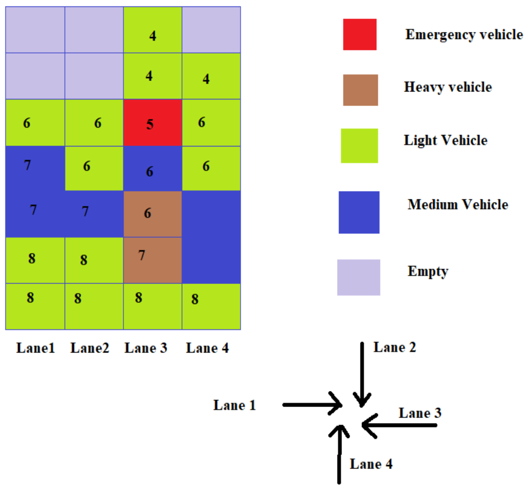

Fig 1 : Traffic state map

At an exceptionally detailed level of analysis, the innovative approach put forth in this proposal facilitates an intricate understanding of the fluctuations in traffic velocity and the distribution of traffic density within each individual lane. This heightened level of insight serves as a foundational element for the implementation of more intricate and effective strategies aimed at orchestrating traffic flow and devising optimal scheduling techniques.

The dynamic process of mapping the current traffic state is initiated through a meticulous segmentation of the traffic within the lane. This process involves the systematic classification of vehicles into distinct and well-defined categories, which encompass the following classifications:

1) Small
2) Medium
3) Large

To elaborate further, the "Small" classification is designed to encompass two-wheelers, acknowledging their distinct characteristics and impact on traffic dynamics. In the "Medium" category, a comprehensive range of vehicles such as three-wheelers, cars, and mini trucks are included, recognizing their shared traits and influences. Within the "Large" classification, a distinct class that encapsulates buses and lorries, the emphasis lies on vehicles of significant size and potential road presence.

To establish an effective segmentation strategy, the lane itself undergoes a meticulous partitioning into a multitude of distinct grids. Each of these individual grids is then systematically characterized by the prevailing category of vehicles that predominantly occupy its confines. This process is initiated by dissecting the traffic image into these defined grids, and subsequently subjecting each grid's visual data to a Convolutional Neural Network (CNN) for the purpose of accurate vehicle category classification. This classification scheme discerns between the categories of "Small," "Medium," "Large," and "empty" spaces. This sequential progression of steps contributes to a comprehensive evaluation of the distribution of various vehicle types across the entirety of the lane. Importantly, this forms the bedrock upon which the dynamic traffic state map is constructed.

In the pursuit of vehicle categorization, a modified version of the Densenet Convolutional Neural Network (CNN) is harnessed. Densenet presents an innovative solution to the issue of diminishing gradients that often hampers the efficacy of high-level neural networks. This challenge is tackled through the incorporation of a concept known as "descent," an adaptation inspired by the Resnet methodology [23]. This novel approach foregoes the conventional additive connections between layers, opting instead to concatenate the outputs of preceding layers. This arrangement creates extended pathways between the input and output layers, ensuring the safeguarding of information throughout the neural network's journey. The outcome is a more resilient network architecture where information attenuation is mitigated, ultimately bolstering the network's ability to retain vital information as it traverses through the layers.

The primary aim that drives the innovation of Densenet is the augmentation of network depth while avoiding the concurrent elongation of training time. This pivotal objective is achieved by integrating a strategic network architecture wherein the interconnection of layers is significantly more compact. This design intricacy leads to the creation of a network structure that is notably denser, thereby facilitating the unhindered flow of information across its various components. This heightened connectivity is the crux of Densenet's effectiveness.

At the heart of Densenet's mechanism are the cross-connections, which play a pivotal role in realizing the overarching goal of increased depth. These cross-connections represent an ingenious solution to enhance the network's capacity to grasp intricate patterns and features. This is achieved by allowing each layer to actively assimilate information from all the preceding layers, establishing an intricate web of inputs. Subsequently, these enriched inputs are then seamlessly propagated forward, layer by layer, in a coherent and sequential fashion. This collective contribution from earlier layers empowers each subsequent layer to encapsulate a more comprehensive understanding of the data.

This architecture, underpinned by the utilization of cross-connections, effectively counteracts the common challenge of vanishing gradients. By ensuring that each layer has access to information from multiple preceding layers, the risk of information degradation or loss is minimized. As a result, the network benefits from a steady and efficient propagation of signals, which bolsters its ability to extract nuanced features and patterns from complex data. This systematic approach signifies an innovative departure from conventional architectures, marking Densenet as a pioneering framework designed to optimize the synergy between network depth and information dissemination.Top of Form

The proposed enhancements to the model architecture encompass a strategic transformation involving the removal of conventional fully connected layers, which are subsequently substituted with fully convolutional layers. Additionally, a noteworthy alteration entails the elimination of pooling at layer 5, instead opting to double the stride. This shift serves a specific purpose – it compensates for the challenge of edge localization, ensuring that the model effectively captures essential boundary information.

To realize the integration of hybrid features, a convolutional layer is intricately designed with a kernel size spanning eleven and a channel depth amounting to twenty-one. The outcome of each convolutional operation is thoughtfully aggregated with an additional layer, fostering the amalgamation of distinctive features. Following this, the resultant feature map is subjected to an up-sampling process facilitated by a convolutional layer featuring a size of 1, carefully implemented at the supplementary layer.

Furthermore, a pivotal cross-entropy loss/sigmoid layer is thoughtfully affixed to the up-sampling layers. This strategic inclusion serves to harmonize with the overarching objective of achieving the desired classification output. The integration of these loss/sigmoid layers complements the up-sampling process, steering the model towards generating classification outcomes in accordance with the specified criteria.

Collectively, these meticulously orchestrated adjustments work in synergy to augment the model's proficiency in effectively accommodating hybrid features. By comprehensively adapting the architecture to these modifications, the model becomes better equipped to tackle the intricacies inherent in capturing hybrid characteristics within the data. The culmination of these adaptations ultimately elevates the model's overall performance, notably enhancing its accuracy in successfully executing the classification task at hand.

The model architecture undergoes several modifications. Firstly, fully connected layers are removed and replaced with fully convolutional layers. Additionally, the pooling operation at layer 5 is eliminated, increasing the stride by two times, which helps in compensating for edge localization.

For obtaining hybrid features, the convolutional layer is configured with a kernel size of 1×1 and a channel depth of 21. The feature resulting from each convolution is combined with an additional layer. To up-sample the feature map, a convolutional layer of size 1×1 is applied at the additional layer.

At the up-sampling layers, a cross-entropy loss/sigmoid layer is attached to facilitate the classification task. These adjustments enable the model to effectively handle hybrid features and contribute to improved accuracy in the classification process.

Table 1: Densenet configuration

Layers	Output Size	DenseNet
Convolution	112*112	
Pooling	56*56	
Dense Block(one)	56*56	[1*1 Conv 3*3 Conv] *6
Transition Layer (One)	56*56 28*28	1*1*128 Conv
Dense Block(Two)	28*28	[1*1 Conv 3*3 Conv] * 12
Transition Layer (Two)	28*28 14*14	1*1*256 Conv
Dense Block(Three)	14*14	[1*1 Conv 3*3 Conv] *24
Transition Layer (Three)	14*14 7*7	1*1*512 Conv
Dense Block(Four)	7*7	[1*1 Conv 3*3 Conv] *16
Classification Layer	1*4	Small,Mediim, Heavy, Empty

The process of training the modified Densenet model involves the utilization of distinct image patterns corresponding to various vehicle categories. These image patterns serve as the foundational data upon which the model's learning process is based. By exposing the model to these diverse patterns, it progressively refines its understanding of the intricate features and attributes associated with each vehicle category.

Following the training phase, the acquired knowledge and insights are then effectively harnessed in the subsequent categorization process. This categorization is facilitated by subjecting the grid images, each representing a distinct segment of the lane, to the trained Densenet model. This operation involves systematically feeding each individual grid image into the model, which in turn processes the visual information and provides an output indicative of the dominant vehicle category characterizing that particular grid.

In essence, this operational workflow seamlessly bridges the gap between training and application. The model, equipped with its acquired expertise in discerning vehicle categories, undertakes the task of analyzing each grid image to ascertain the primary category of vehicle present within it. This entire process encapsulates a comprehensive transformation of visual data into meaningful categorization outcomes, thereby forming a crucial link in the creation of the dynamic traffic state map.

Table 2 : Traffic Management Rules

Rule	Action
Lane X is red coded	Block all others and pass through for lane X
Lane X has more brown patches (Heavy vehicles)	Increase time slot for Lane X
Lane X has more waiting time	Increase time slots in succession for Lane X
Lane X is Lavender (empty)	Block X and take time slot to other Lane.

In contrast to conventional traffic flow analysis systems that provide flow results in terms of incoming and outgoing traffic, this research offers a more detailed approach by analyzing traffic at a fine-grained level. The analysis includes categorizing vehicles and determining their waiting time in traffic for each category.

To perform flow analysis, two traffic state maps (TSM_t and TSM_(t-1)) are correlated. The correlation involves identifying vehicle clusters (VC) present in both TSM_t and TSM_(t-1) that either have the same shape or minimal distortion. This matching process is accomplished using SURF (Speeded Up Robust Feature) feature mapping between the two traffic state maps. By employing SURF features, the presence of the same vehicle across two frames can be located based on feature matching between the images. If a corresponding vehicle cluster cannot be detected, it indicates that the vehicle cluster has moved out of the lane between TSM_t and TSM_(t-1). For each unmoved vehicle cluster, the waiting time is aggregated based on the time between TSM_t and TSM_(t-1). This approach provides a more comprehensive understanding of traffic dynamics, allowing for a detailed analysis of vehicle categories and their waiting times, contributing to a deeper insight into traffic flow patterns

4. Results and Discussion

The comprehensive evaluation of the proposed solution encompasses two pivotal categories, each shedding light on different aspects of its effectiveness: the first pertains to the efficiency and accuracy of constructing the traffic state map, while the second examines the efficacy of a sample traffic management system predicated upon the dynamically constructed traffic state information.

In the pursuit of assessing the precision and reliability of the traffic state map's construction, the solution's capabilities are rigorously tested employing the Indian Driving Dataset. This dataset serves as a robust benchmark for evaluating the solution's performance, effectively gauging its ability to accurately capture and portray intricate traffic dynamics. The insights drawn from this evaluation process are integral in validating the practical viability of the proposed methodology.

Furthermore, the recent advancements in simulating and analyzing urban mobility hold a significant value for a broad spectrum of stakeholders. This includes researchers in the fields of urban planning, transportation engineering, and related disciplines. The availability of simulation tools such as the one proposed in this work presents a valuable opportunity to gain deeper insights into the complexities of transportation systems within urban environments. This knowledge is pivotal for informed decision-making, optimized planning, and the development of more efficient urban mobility strategies.

Moreover, a critical facet of the proposed solution lies in its novel vehicle categorization approach, which is compared against an existing ResNet-based classification scheme introduced elsewhere. This comparison serves as a means to validate the effectiveness and superiority of the proposed methodology. The evaluation metrics encompass three fundamental criteria, namely:

Accuracy: Measuring the model's overall correctness in categorizing vehicles.

Precision: Gauging the model's ability to accurately identify true positives among the categorized vehicles.

Recall: Evaluating the model's capacity to effectively capture and classify all relevant vehicles in the dataset.

Together, these comprehensive metrics provide a robust and well-rounded framework for evaluating the proposed solution's performance, thereby offering insights into both its accuracy in traffic state map construction and its utility in real-world traffic management scenarios.

These performance measures will gauge the effectiveness and accuracy of the traffic state map construction and vehicle categorization processes, providing valuable insights into the proposed solution's capabilities.

Accuracy is measured as

$$\text{Accuracy} = \frac{\text{TP}}{\text{Total Test Images}}$$

Precision is measured as

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Recall is measured as

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Where TP is true positive, FP is false positive and FN is false negative. The performance is measured for 1000 images from Indian Driving Dataset and the result is below

Table 3: Vehicle categorization performance comparison

	ResNet [2]	**Proposed**
Accuracy	0.9795	0.9875
Precision	0.9530	0.9780
Recall	0.8970	0.9420

The results clearly demonstrate that the proposed solution outperforms ResNet in multiple aspects. Specifically, it achieves 1% higher accuracy, 2% higher precision, and 3.5% higher recall compared to ResNet. This improvement in accuracy can be attributed to the use of a modified Densenet with enhanced discriminative learning.

The proposed traffic management solution categorizes vehicles into four groups: small, medium, high, and emergency vehicles. The experiment was conducted on a 4-way lane setup, as illustrated in Figure 1, over a duration of 20 minutes. For comparison, round-robin scheduling with fixed 120-second time slots was employed for each lane. Performance evaluation was carried out based on average vehicle density (congestion) across the lanes, percentage of time utilized, and average time taken for emergency vehicles to pass through. To model vehicle arrival in each lane, a Poisson process was used, and performance was measured accordingly. The results of the experiment are presented in Table 4.

Table 4 : Comparison of traffic management performance

Parameter	**Proposed Traffic Scheduling**	**Round Robin Scheduling**
Average Vehicle Density Over the Lanes	22	34
Time Utilization Percentage	94%	91 %
Average Time For Emergency Vehicle Pass Through	80 Seconds	100 Seconds

The proposed traffic management algorithm demonstrates a significant improvement in traffic congestion, with an average of 1.5 % lower vehicle density compared to Round-robin scheduling. This reduction in congestion is a testament to the effectiveness of the proposed solution. Moreover, the proposed algorithm exhibits a 3 % higher time utilization rate. This enhanced time utilization is achieved by the algorithm's ability to sense empty lanes and intelligently shift time slots to other lanes as needed. In emergency scenarios, the proposed solution excels, achieving an average of 80% lower pass-through time for emergency vehicles compared to round-robin scheduling. This efficiency is achieved through the algorithm's capability to detect emergency vehicles from the traffic state map and preempt traffic in other lanes, enabling swift and unobstructed passage for emergency vehicles. Overall, the proposed traffic management algorithm showcases substantial benefits, including reduced congestion, improved time utilization, and expedited passage for emergency vehicles, making it a promising solution for efficient and adaptive traffic control..

5. Conclusion

This study introduces a deep learning model based on a modified Densenet architecture, facilitating the construction of a comprehensive traffic state map. The traffic state map provides detailed insights into traffic density distribution and variations. Leveraging this map, an advanced traffic management algorithm is developed, surpassing the performance of traditional round-robin traffic scheduling methods. The utilization of the traffic state map proves highly effective in optimizing traffic lanes, resulting in reduced congestion and faster pass-through times for emergency vehicles. The deep learning model's capability in constructing the traffic state map plays a crucial role in achieving these substantial improvements in traffic management. Overall, the research underscores the potential of deep learning-based approaches in revolutionizing traffic state estimation and control. The fine-grained information offered by the traffic state map enables more efficient and responsive traffic management strategies, leading to smoother traffic flow and improved emergency response capabilities.

Author Contributions

Mr. Arunkumar Joshi, research scholor carried out the research, experimentation, and writing process, contributing to the design of multimodal deep learning models, data collection and pre-processing, experiment execution, and drafting sections of the paper related to the models outcomes. Shrinivasrao B Kulkarni, the second author and mentor, assumed a guiding role in the research work, drafting, modifications and decision making of the work. This encompassed conceptualizing the research concept, offering deep learning methods, aiding in experimental planning and methodology, reviewing and refining the paper, and providing comprehensive work oversight.

Conflicts Of Interest

The authors declare no conflicts of interest.

References

[1] Fedorov, A., Nikolskaia, K., Ivanov, S. et al. (2019).Traffic flow estimation with data from a video surveillance camera. J Big Data 6, 73

[2] Jung H, Choi MK, Jung J, Lee JH, Kwon S, Jung WY.(2017). ResNet-based vehicle classification and localization in traffic surveillance systems. IEEE conference on computer vision and pattern recognition workshops (CVPRW).

[3] Shanghang Zhang, Guan hang Wu.(2017). Understanding Traffic Density from Large-Scale Web Camera Data.Computer Vision and Pattern Recognition.

[4] R. Guerrero-G´omez-Olmedo, B. Torre-Jim´enez, R. L´opezSastre, S. Maldonado-Basc´on, and D. Onoro-Rubio.(2015). Extremely overlapping vehicle counting. In Iberian Conference on Pattern Recognition and Image Analysis

[5] Daniel AraniOsuto, Heywood OumaAbsaloms (2016).Automatic Road Traffic Density Estimation using Image Processing Algorithms. Annual Conference on Sustainable Research and Innovation

[6] K. Garg, S. Lam, T. Srikanthan and V. Agarwal.(2016). Real-time road traffic density estimation using block variance. IEEE Winter Conference on Applications of Computer Vision (WACV).

[7] HongyuHu, ZhenhaiGao, YuhuanSheng.(2019).Traffic Density Recognition Based on Image Global Texture Feature",International Journal of Intelligent Transportation Systems Research.

[8] FreddyKurniawan,HarunoSajati,OktoDinaryanto.(2017).Image Processing Technique for Traffic Density Estimation, International Journal of Engineering and Technology (IJET)

[9] MA. Dinani, P. Ahmadi, I. Gholampour.(2015).Efficient Feature Extraction for Highway Traffic Density Classification. 9th Iranian Conference on Machine Vision and Image Processing

[10] Li Wei, Dai Hong-ying.(2016).Real-time Road Congestion Detection Based on Image Texture Analysis Procedia Engineering, Elsevier

[11] Chakraborty, Pranamesh& Okyere, Yaw & Poddar, Subhadipto&Ahsani, Vesal& Sharma, Anuj &Chakraborty,. (2018). Traffic Congestion Detection from Camera Images using Deep Convolution Neural Networks. Transportation Research Record Journal of the Transportation Research Board.

[12] Luong Anh Tuan Nguyen and Thi-Ngoc-Thanh Nguyen.(2018). Traffic Density Analysis based on Image Segmentation with Adaptive Threshold. Communications on Applied Electronics

[13] Luong Anh Tuan Nguyen and Thi-Ngoc-Thanh Nguyen.(2017). Traffic Density Identification based on Neural Network and Histogram. International Journal of Computer Applications

[14] Lu Guo,Hua Cui. (2016). Urban Road Congestion Recognition Using Multi-Feature Fusion of Traffic Images. Journal of Artificial Intelligence Practice

[15] Donato Impedovo,Fabrizio Balducci, Vincenzo Dentamar(2019).Vehicular Traffic Congestion Classification by Visual Features and Deep Learning Approaches: A Comparison",Sensors (Basel).

[16] Shi X., Shan Z., Zhao N. (2016).Learning for an aesthetic model for estimating the traffic state in the traffic video. Neurocomputing.

[17] Nguyen H.N., Krishnakumari P., Vu H.L., Van Lint H. Traffic.(2016). Congestion pattern classification using multi-class SVM; Proceedings of the IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)

[18] Chaudhary S., Indu S., Chaudhury S.(2018). Video-based road traffic monitoring and prediction using dynamic Bayesian networks. IET Intell. Transp. Syst

[19] Lam C., Gao H., Ng B. (2017).A real-time traffic congestion detection system using on-line images; Proceedings of the IEEE 17th International Conference on Communication Technology (ICCT); Chengdu, Chian.

[20] Asmaa O., Mokhtar K., Abdelaziz O.(2013). Road traffic density estimation using microscopic and macroscopic parameters. Image Vis. Comput

[21] V. Nair and G. E. Hinton.(2010). Rectified linear units improve restricted boltzmann machines. in Proceedings of International Conference on

Machine Learning, pp. 807–814, Haifa, Israel

[22] Kai Zeng, XinzhengNiu.(2014).Feature Selection with Neighborhood Entropy-Based Cooperative Game Theory. Computational Intelligence and Neuroscience,

[23] Krajzewicz, D.; Erdmann, J.; Behrisch, M.; Bieker, L. Recent development and applications of SUMO-Expereiment of urban mobility.Int. J. Adv. Syst. Meas. 2012, 5, 128–138

[24] V. Tran and W. Tsai, "Acoustic-Based Emergency Vehicle Detection Using Convolutional Neural Networks," IEEE Access, vol. 8, pp. 75702-75713, 2020.

[25] Sudhakar, M. ., & Kaliyamurthie, K. P. . (2023). A Novel Machine learning Algorithms used to Detect Credit Card Fraud Transactions. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2), 163–168. https://doi.org/10.17762/ijritcc.v11i2.6141

[26] Prof. Madhuri Zambre. (2016). Analysis and Modeling of Physical Stratum for Power Line Communication. International Journal of New Practices in Management and Engineering, 5(01), 08 - 13. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/42

[27] Anupong, W., Yi-Chia, L., Jagdish, M., Kumar, R., Selvam, P. D., Saravanakumar, R., & Dhabliya, D. (2022). Hybrid distributed energy sources providing climate security to the agriculture environment and enhancing the yield. Sustainable Energy Technologies and Assessments,52 doi:10.1016/j.seta.2022.102142