A New Method for Revealing Traffic Patterns in Video Surveillance using a Topic Model

Yao Wang*
Henan Industry and Trade, Vocational College, Zhengzhou Henan, 450053, China

***Abstract*—Research on video surveillance systems, for instance, in intelligent transportation systems, has advanced due to the growing requirement for monitoring, control, and intelligent management. One of the next issues is extracting patterns and automatically classifying them, given the volume of data produced by these systems. In this study, a theme approach was utilized to translate visual patterns into visual words in order to reveal and extract traffic patterns at crossings. The supplied video is first cut up into segments. The optical flux technique is then used to determine the clips' optical flux characteristics, which are based on a lot of local motion vector data, and translate them into visual words. The thin-group thematic coding method is then used to teach traffic patterns to the proposed system using a non-probable thematic model. By responding to a behavioral query like "Where is a vehicle going?" these patterns convey observable motion that can be utilized to characterize a scene. The results of applying the suggested method to the QM_UL video database demonstrated that the suggested method can accurately identify and depict significant traffic patterns such as left turns, right turns, and intersection crossings.**

Keywords*—Group thin topic coding; QM_UL video; optical flux; traffic patterns

I. INTRODUCTION

The deployment of surveillance cameras is one of the most common strategies for boosting security and managing public spaces better. The majority of public spaces and intelligent transportation networks virtually always have these cameras in situ [1]. These cameras are crucial in video surveillance because they may give decision-makers access to all the details and activities occurring in the monitored area [2]. For instance, in smart transportation systems, these binoculars can be used to monitor and control the volume and kind of traffic, vehicle offenses, and traffic patterns at crossings. On the one hand, the use of these video surveillance cameras has become quite widespread [3]. On the other hand, as their use has increased, a vast amount of multimedia data has been produced, making it virtually impossible for people to monitor these cameras, necessitating the need for a system. Intelligent monitoring, automatic pattern detection, and pattern extraction are all clearly audible [4]. There are two sorts of studies that are accessible for the examination of traffic scenes: a) studies based on the route. A training dataset for machine learning is created using paths that have been observed over a lengthy period of time [5] and [6]. In fact, such a study is still unreliable in challenging circumstances because of the lack of steady and dependable multipurpose tracking algorithms [7]. Rapid response to unexpected shifts in traffic [8] also commonly causes analysis problems. b) Analysis that does not involve tracking and is based solely on the low-level motion vector Optical flux [9] is the most commonly used technique and contains a lot of data about local movement. Using this simple motion feature, more complex models, like thematic ones, might be built for evaluating complex traffic scenarios [10]. Some examples of probabilistic topic models are the Probabilistic Latent Structure Analysis [11], the Probabilistic Conceptual Latent Analysis [12], the Dirichlet Latent Allocation LDA [13], and the Hierarchical Dirichlet Process HDP [14]. were originally developed to unearth concealed headers in a massive corpus of text documents before they were adopted by academics for video analysis. In [15], a fully sparse topic model (TM) (FSTM) is proposed as a simple variant of LDA and PLSA. According to research [16], thin topic coding (STC) using a non-probabilistic topic model (NPTM) can be used to learn hierarchical hidden representations of enormous data sets. Group Thin Topic Coding (GS-TC), a brand-new non-probabilistic topic model, is suggested in [17] for learning thin hidden representations of big text document sets. An effort has been made in [18] to highlight odd and uncommon actions, such as the fire engine stopping the regular flow of traffic. Thin group thematic coding has been enhanced for this purpose, and the thin light flux method has also been applied to extract movement patterns. When compared to previous similar works, the key distinction between our work and others is the utilization of dense light flux in this article to extract movement pattern features and the thematic coding of the primary thin group. Even though using dense light flux requires a larger computation area than using thin light flux, more movement patterns are still recovered. The approaches used to convert the prevalent visual traffic movement patterns into descriptive phrases allowed us to zero down on eight key traffic patterns within the dataset. These eight traffic patterns are: a) Crossing the intersection from the eastern side to the western side; b) Crossing the intersection from the southern side to the northern side; c) Turning right from the eastern side to the northern side; d) Turning left from the western side to the northern side; d) Turning left from the west to the northern side; e) Crossing the intersection from west to east; f) Turning right from south to east; g) Crossing the intersection from north to south; and h) Finding traffic patterns directly results in a useful scene model and streamlines scene analysis. Despite the fact that there have been numerous studies in this area, machine vision systems still struggle with this problem. The automatic detection of traffic patterns in this paper is accomplished using the group thin thematic coding framework (GS-TC). First, a series of separate, non-overlapping clips are created from the surveillance camera video. Then, the discrete visual stream words are translated into discrete light flux characteristics for each pair of subsequent frames in the clips. The words in the visual stream are then translated into words contained in each video clip, which is then treated as a document. To put it another way, visual elements become visual words. In order to find hidden patterns that reveal the distribution of common motion in the scene, the GS-TC approach is then used. The results demonstrate that useful traffic patterns, such as left turn, right turn, and crossing the intersection, can be retrieved from intersections using the proposed method when applied to real footage. The reasons for choosing the "proposed method" that we consider suitable for dealing with such problems are the following:

1) Growth in video surveillance data: This study confirms the increasing use of surveillance cameras in various applications, leading to significant growth in video data. This increasing amount of data makes manual monitoring impractical and necessitates the development of automatic pattern recognition methods. The proposed method offers a solution to this problem by automatically identifying and classifying traffic patterns.

2) Communication with intelligent transportation systems: surveillance cameras are important in intelligent transportation systems. These systems need to monitor and control traffic patterns at intersections, which makes the proposed method particularly relevant to this field.

3) Using optical flux features: This method uses optical flux features extracted from video frames. Optical flux data contains information about local movement, which is very important for analyzing traffic scenarios. This research discusses the use of optical flux and its potential to build more complex models, such as thematic models, to assess traffic patterns.

4) Thematic coding approach: In this work, a thematic coding approach is chosen, which is usually used in natural language and text processing. Applying this approach to video data, they aim to transform visual traffic movement patterns into descriptive expressions that enable the identification and classification of traffic patterns.

5) Identifying common traffic behaviors: This article also emphasizes the importance of recognizing common traffic behaviors such as turning left, turning right, and moving straight through intersections. These behaviors are considered significant traffic patterns and the proposed method is designed to identify them. The authors' overall contributions to this paper can be summarized as follows:

- Providing a non-surveillance way for video surveillance to identify traffic trends.
- Using the thematic paradigm, drivers at crossings are instructed on traffic patterns.

The overview of related studies follows in Section II. Section III provides background information on the theory behind flow detection and traffic patterns. The traffic dataset, the execution strategy, and the useful outcomes are all covered in Section IV. Section V concludes with recommendations for additional research.

II. BACKGROUND

In Europe, road traffic noise is a big problem, exposing almost 20% of the population to dangerously high noise levels. In order to facilitate decision-making techniques aimed at controlling or reducing noise exposure, effective monitoring and measurement of sound levels in sensitive regions is essential [31]. However, expensive equipment and maintenance duties are needed for spatiotemporal measurements with continuous range and lengthy duration. Thus, the goal of the research [32] is to create an intelligent mobility approach that is affordable and can be used to estimate traffic noise levels using roadside video images. The created method consists of an algorithm that uses dynamic microscopic models to evaluate noise and extract traffic volume, identify vehicle classes, and estimate each vehicle's speed from video records. These later models are based on the already available noise emission models (NEMs) for estimating source sound power levels and a sound emission model that can estimate the corresponding A-weighted sound pressure levels given any road vehicle speed as input slow. The created method is distinguished by its modular structure, which makes it simple to add new variables to the sound propagation model and/or replace the NEM. The process is put to the test under various service levels on a medium-sized city's rural roads. The findings reveal that the noise estimation errors are less than 1 dBA, indicating great accuracy. The issue of traffic congestion in large cities is getting worse due to the swift growth in both the population living in metropolitan areas and the quantity of motor vehicles. The paper aims to perform cluster analysis for daily traffic congestion index curves in order to discover patterns of traffic congestion and examine their spatial-temporal changes. Initially, the coefficient of variation is used to apply weights in order to improve the K-means clustering method because the significance of sample points varies slightly over different time segments. A better K-means clustering algorithm is suggested to find patterns of traffic congestion. Second, changes in traffic congestion patterns over time and space are analyzed using the paired t-test method.

A. Exposing Traffic Patterns

Fig. 1 displays an example of traffic monitoring scenarios at an intersection using a surveillance camera positioned above the intersection. Traffic movement patterns are regular and circular traffic situations, such as "straight passage" and "right turn". The technologies and algorithms for detection, identification, tracking, and categorization of items suffer substantially as a result of the complexity of many of these settings. In these situations, topic models can be utilized to locate and highlight patterns in order to map words, documents, and subjects to certain pattern recognition ideas.

B. Coding of Thin Group Topic

Consider a collection of documents $D = \{w_1, \dots, w_D\}$ that includes N terms from the vocabulary collection V. A document is only a vector | I| after which $D = \{w_1, \dots, w_{|I|}\}$, appears, where I is a set of m word indices. The nth entry (n ∈I) in w_n indicates the number of occurrences of the desired word in a particular document. We consider the parameter $\beta \in R^{K*N}$ as a unitary 2 distribution in V is used to describe a dictionary with K bases, each of which is taken to be the foundation of the subject [19].

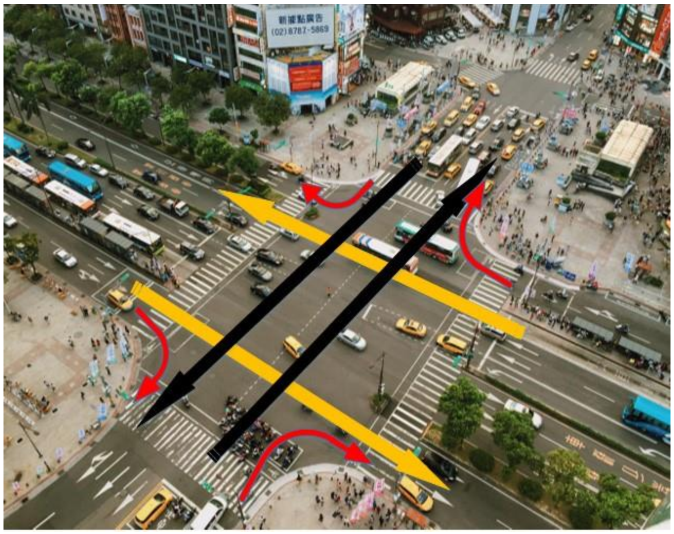

Fig. 1. An intersection traffic scene includes typical traffic movements including left turns, right turns, and crossing the intersection.

For the dth document (w_d), the GS-TC method maps it to a semantic space assigned by a set of auto-learned β addresses and directly encodes the non-normal word $s_{d,m} \in R^K$ determines for each specific word in the document w_d.

Then the mixture ratio of the whole document w can be derived from the code word $s = \{s_{.1}, \dots, s_{|I|}\}$ and the headings β. The optimization issue is resolved by the GS-TC approach in accordance with (1). The first part (1) is the non-normal difference KL between the observed words $w_{d,n}$ and their reconstruction $\beta_{.n}\ s^T_{d,m.}$.

$$\min \sum_{d=1}^{D} \sum_{m=1}^{|I_d|} \left(\sum_{k=1}^{K} S_{d,kn}\ \beta_{kn} - w_{d,m} \ln\left(\sum_{k=1}^{K} S_{d,km}\ \beta_{km} \right) \right) + \lambda \sum_{d=1}^{D} \sum_{k=1}^{K} \left|\left| S_{d,k.} \right|\right|_2 + C \quad (1)$$

$$\text{s.t}\ s_{d,n} \geq 0\ \forall d, m$$

$$\sum_{k=1}^{K} \beta_{km} = 1, \forall k$$

$$\theta_k = \frac{\sum_{m=1}^{|I|} S_{km}\ \beta_{km}}{\sum_{m=1}^{|I|} \sum_{l=1}^{K} S_{\mathrm{Im}}\ \beta_{\mathrm{Im}}} \quad (2)$$

The second step involves using the LASSO approach to apply a variety of norms for the reconstruction coefficients matrix, which results in proportionate thin coding at the document level. The symbols used in (1) are defined in Table I. It should be noted that the word codes can be used to determine the mixture ratio of the document surface. The variable θ_k is used in (2) [20] to indicate the contribution vector of the k_{th} title in the document w.

TABLE I. THE DEFINITION OF THE SYMPTOMS AND VARIABLES UTILIZED IN (1) AND OPTIMIZATION OF THE GS-TC TECHNIQUE

Signs	Description
$m = 1, \dots, N$	Index_of words
θ_d	Percentage_of titles in the dth _document
$d = 1, \dots, D$	Document_index
β	Titles_Dictionary
$S_{d,m}$	mth-word representation in dth-document
$w_{d,m}$	How many times the word n appears_in the given document
$k = 1, \dots, K$	Index of _titles
I_d	Set of index words _that appear in document dth

C. Related Works

Most methods for analyzing traffic can be classified into two broad categories. In the first category, the target object, such as a car or a person, is first recognized, followed by tracking, and the paths gleaned are then used for additional analysis. Undoubtedly, the grouping of traces is a straightforward technique that makes it possible to identify anomalies [21]. Regardless, the accuracy and dependability of these approaches heavily rely on the detection and tracking techniques, which are susceptible to errors brought on by obstruction, noise, shifting lighting conditions, and changing weather. Additionally, the grouping of routes necessitates a comparison of all samples' similarity, which can be computationally taxing [22]. In the second type, information about motion and appearance is gleaned from video frames without the aid of detection and tracking methods. The strategy utilized in these articles [7] and [23] is to directly develop a model of movements and activities using these extracted attributes. A sizable number of research projects have recently concentrated on the use of topic models (TM). For instance, utilizing probabilistic conceptual latent analysis (PLSA), anomaly detection and motion-based scene segmentation have been carried out in [24]. Using the two-level Dirichlet process (dp), normal and aberrant activity were distinguished in [25]. In [26], a two-level LDA topic model was applied to extracting rules and recognizing anomalies after it was utilized to identify anomalies and learn behavior for the efficient display of clips with a dispersed set of motion patterns. The author in [19] presents an unsupervised method for anomaly detection that uses the thin thematic group coding (GS-TC) framework to learn movement patterns. In [27], it is planned to use license plate number data obtained from video surveillance cameras to determine urban road vehicle density, city-wide regional vehicle density, and hot routes. To improve the precision of the visualization's impact, this article used a method for detecting outliers based on Dixon's detection approach and Internet crawling technologies during data analysis and processing. This study developed an urban road vehicle traffic index for the visualization map design in order to visually and numerically depict the region's traffic operating state. An experiment was carried out in Guiyang using the information from the road video surveillance camera system to confirm the method's viability. From three visualization maps, a number of geographical and temporal characteristics of urban traffic are clearly and effectively identified. The outcomes demonstrate the suggested framework's satisfactory performance in terms of visual analysis, which helps with traffic management and operations. The author in [28] introduces DeWiCam, a smartphone-based detection system that is small and powerful. Utilizing the natural traffic patterns of wireless camera streams is the core concept behind DeWiCam. DeWiCam is more difficult to utilize than conventional traffic pattern analysis because it cannot access data that has been packed in data packets. DeWiCam, on the other hand, circumvents this issue and can accurately find neighboring Wi-Fi cameras. A human-assisted recognition approach is suggested to further determine whether a camera is interested in a certain room. Image resolution and audio channel inference are two additional DeWiCam add-on functions that are available to enhance further security. DeWiCam is put to use on the Android operating system and assessed using in-depth tests on 20 cameras. DeWiCam can identify cameras with 99% accuracy in 7:2 seconds, according to test data.

III. PROPOSED METHOD FOR DETECTING TRAFFIC PATTERNS

A. *Process of the Proposed Method*

Fig. 2 and 3 depict the general and specific steps of the suggested strategy that makes use of the topic model. When considering an input video, the video is first momentarily split into *D* non-overlapping clips, with each clip being treated as a separate w_d document. The scene is initially divided into $C_x * C_y$ square cells, each of which covers *p∗p* pixels, in order to construct stream words. After that, using the optical flux technique, motion vectors are extracted for each pair of subsequent frames. In order to eliminate noise and preserve dependable flows, a motion vectors are subjected to a threshold value, t. The other motion vectors $s_i = (x, y, u, v)$, whose positions (x and y) are fixed in a grid with a distance of p pixels, are sampled to create flow words. The examples of motion vectors include then split into an O-number of directions based on their displacement (u, v). Finally, a group of fixed words is generated, $N = C_x \times C_y \times O$ and $V = \{1, \ldots N\}$, each of which has two content aspects: "position information" and "direction information" of movement. The frames include a collection of the video clips' stream words. Then, a video clip is shown as a vector with the form $w = (\omega_1, \ldots, \omega_N)$, where ω_n denotes how many times the n_{th} word appears in the clip. The symptoms are depicted in Table II along with the video's counterparts.

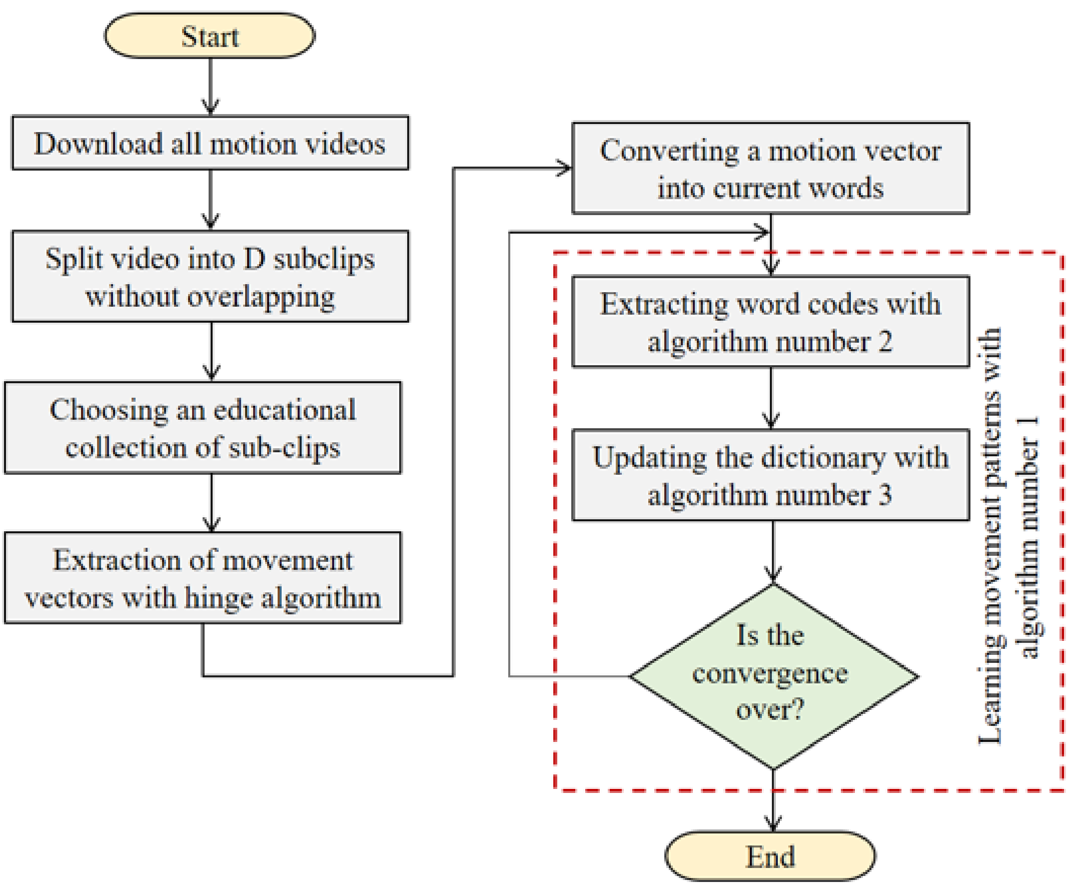

Fig. 2. The general process of the proposed method of detecting normal traffic patterns of vehicles.

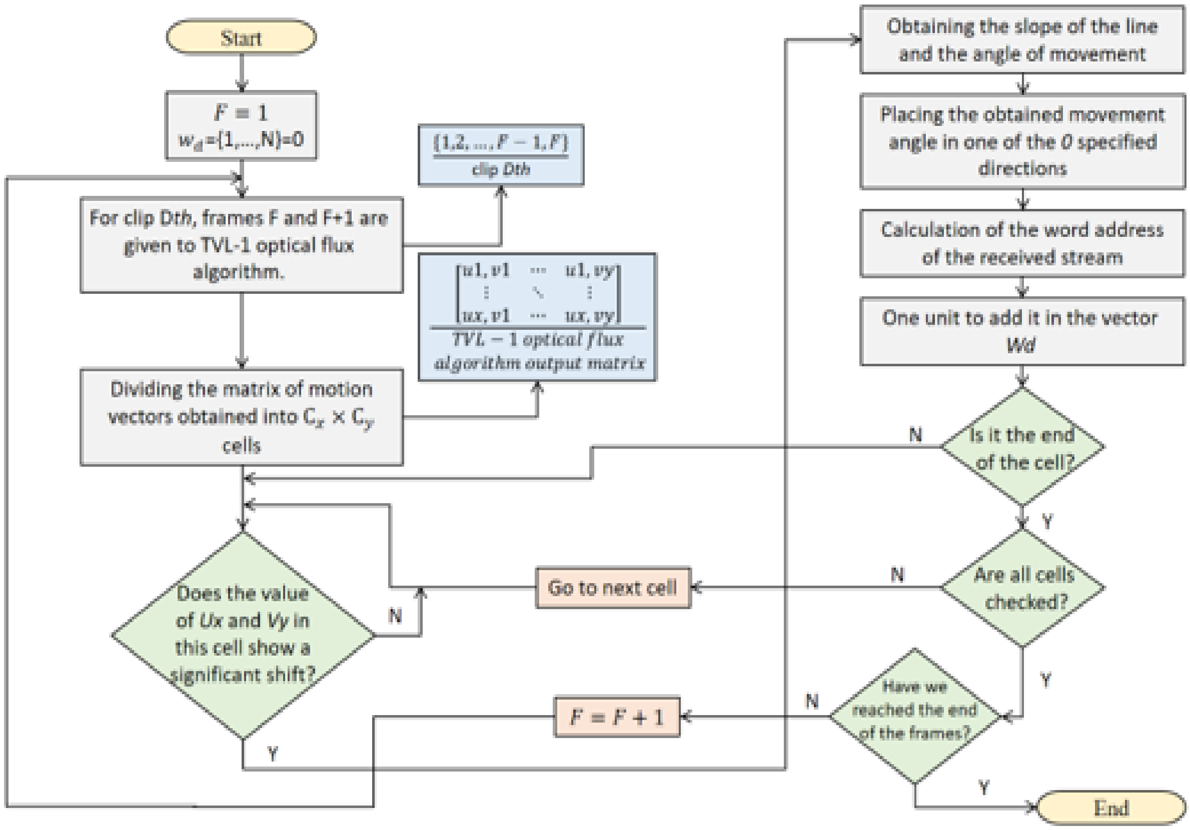

Fig. 3. Partial flowchart of the proposed method, video display with the subject model in detecting normal traffic patterns of vehicles at intersections. Identifying normal traffic patterns of vehicles.

TABLE II. SYMBOLS AND VARIABLES USED IN THE VIDEO, ALONG WITH THEIR EQUIVALENTS

Signs	Description
$\boldsymbol{d = 1, \ldots, D}$	Index of clips
$\boldsymbol{k = 1, \ldots, K}$	Index of movement patterns
$\boldsymbol{n = 1, \ldots, N}$	A Word-Flow Index
$\boldsymbol{I_d}$	Index set of flow words that happened in clip dth
$\boldsymbol{w_{d,n}}$	number of occurrences _ word nth clip dth
β	A Movement Pattern Dictionary
$\boldsymbol{\theta_d}$	Clip share dth of templates
$\boldsymbol{S_{d,n}}$	How significant the nth stream word was in the dth clip

B. Learning Traffic Patterns using GS-TC

Eq. (3) provides a summary of the general GS-TC formula. The GS-TC formula in [18], which is based on reducing the variance of KL, does not apply to this equation because the 2 soft reduces reconstruction errors to a minimum. The optimization problem can be solved more easily by using soft minimization 2 [21]. The objective function in (3) is convex, which means that while one of the two is constant, it is convex around $\{S_d\}_{d=1}^{D}$ and β. The CDA algorithm [29], which alternately optimizes on $\{S_d\}_{d=1}^{D}$ and β and is depicted in Algorithm 1, is a common solution.

Algorithm 1: Intelligent transportation systems' learning algorithms for understanding typical vehicle traffic patterns

Input: training video clips$\{w_d\}_{d=1}^{D}$, hyper parameter λ, and topic count
Output: dictionary β, word codes s
$\beta \in R^{K\times N}$ a positive-definite random matrix.
Initialize $\{S_d\}_{d=1}^{D} \in R^{D\times K\times N}$to random matrices with positive elements
l^{old} = Using Eq. (3), figure out the cost function.
repeat
for $d = 1$: D
calculate $S_d$with Algorithm 2
end for
Apply Algorithm 3 to Dictionary β
l = A cost function can be calculated using formula (3).
If ($\boldsymbol{l - l^{old} < \varepsilon}$)then
Break
Else
$l = l^{old}$ until the conditions for termination are met, usually convergence.

$$\min_{\{S_d\}_{d=1}^{D},\beta} \sum_{d=1}^{D} \left(||w_d - diag(S_d^T\,\beta)||_2^2 + \lambda\,||S_d||_{2,1}\right) \tag{3}$$

$$\text{s.t.}\, s_d \geq .0\,\forall d\,, \beta_k \geq .,\, ||\beta_{k.}||_1 \geq 1\,, \forall k$$

where $S_d \in R^{K\times N}$ and parameter λ is a non-negative parameter for thin control.

C. Extracting Word Codes using Thin Coding Algorithm

In the following, the word codes $\{S_d\}_{d=1}^{D}$ will be obtained by optimization according to Eq. (4) with the assumption of constant β dictionary.

Due to conditional independence, this step can be carried out separately for each document by addressing the optimization problem. Since the word codes are organized into titles and each title is unique, this optimization makes use of the BCD technique by solving the issue for each S_{kn}.

The difficulty of locating the roots of a complex quadratic equation is the end result of this approach. By expressing the optimization problem in accordance with Eq. (5) and (6), it is possible to acquire the entire matrix s without having to calculate each S_{kn} individually.

Algorithm 2: Thin coding algorithm

$P = diag(w_{d,1}, \dots, w_{d,N})$
l^{old} = Cost function calculation per formula (5)
Repeat
$R = diag(\frac{1}{||s_{d,1.}||_2 + \varepsilon}, \dots, \frac{1}{||s_{d,k.}||_2 + \varepsilon})$
$S_d = (\beta\beta^T + \lambda R)^{-1}\beta P$
l = Cost function calculation per formula (5)
If ($l - l^{old} < \varepsilon$)then
Break
Else
$l = l^{old}$
Until the convergence occurs or the end point is reached.

$$\min_{\{S_d\}_{d=1}^{D},\beta} \sum_{d=1}^{D}\left(||w_d - diag(S_d^T\beta)||_2^2 + \lambda\,||S_d||_{2,1}\right) = \sum_{d=1}^{D}\left(\sum_{n=1}^{N}\left(W_{d.n} - S_{d,.n}^T\beta_{.n}\right)^2 + \lambda\sum_{k=1}^{K}||S_{d,k.}||_2\right) \quad \text{s.t.}\, s_d \geq .0, \forall d \tag{4}$$

$$\min_s ||w_d - diag(s_d^T\beta)||_2^2 + \lambda||S_d||_{2,1} \quad \text{s.t. } s > 0 \tag{5}$$

$$\min_s trace(PP + s^T\beta\beta^T s - 2Ps^T\beta) + \lambda\, trace(s^T Rs) \quad \text{s.t. } s > 0 \tag{6}$$

that in (6), $R = diag(\frac{1}{||s_{10}||_2} + \varepsilon, \dots, 1/||S_{k.}||_2)$ and $P = diag(w_1, \dots, w_N)$ is. We consider the value of ε much smaller than the non-zero values of s and add it to the denominator to prevent division by zero. Since the norms of l_p for $p \geq 1$ are convex functions, according to (5) and its equivalent, Eq. (6), the problem is the optimization of convex functions with respect to the matrix s. Therefore, assuming zero slope can be a closed-form solution for calculating the ss matrix according to Eq. (7). Algorithm 2 shows the thin coding algorithm.

$$\beta\beta^T s - \beta P + \lambda Rs = 0 \qquad s = (\beta\beta^T + \lambda R)^{-1}\beta P \tag{7}$$

D. Dictionary Update Using Dictionary Learning Algorithm

After discovering each of the set's hidden word codes, the optimization problem (8) is solved to update the β dictionary. By locating the slope's root, Eq. (8)'s convex optimization problem can be successfully solved. Instead of calculating each β_{kn} separately to solve the dictionary learning problem, a general solution can be suggested to acquire the full β matrix by altering (8) to (9).

$$\min_\beta \sum_{d=1}^{D}\left(||w_d - diag(s_d^T\beta)||_2^2\right) = \sum_{d=1}^{D}\sum_{n=1}^{N}(w_{d,n} - s_{d,n.}^T.\beta_{.n})^2 \quad \text{s.t. } \beta \geq 0, \sum_{n=1}^{N}\beta_{kn} = 1, \forall k \tag{8}$$

$$\min_\beta \sum_{d=1}^{D}\left(||w_d - diag(s_d^T\beta)||_2^2\right) = \sum_{d=1}^{D} trace(P^T P + s_d^T\beta\beta^T s_d - 2P^T s_d^T\beta) \tag{9}$$

which is $P = diag(w_{d,1}, \dots, w_{d,N})$. Setting the slope to zero yields the value of β per Eq. (10).

Algorithm 3: Dictionary learning algorithm

$P = diag(w_{d,1}, \dots, w_{d,N})$
$\beta = (\sum_{d=1}^{D} s_d\, s_d^T)^{-1}(\sum_{d=1}^{D} s_d\, P)$
for $k = 1 : K$
for $n = 1 : N$
$\beta_{kn} = \max(\beta_{kn}, 0)$
end
end
for $k = 1 : K$
$\beta_{k.} = \frac{\beta_{k.}}{||\beta_{k.}||_1}$
End

$$\sum_{d=1}^{D}(S_d S_d^T\beta - s_d P) = 0 \qquad \beta = (\sum_{d=1}^{D} S_d S_d^T)^{-1}(\sum_{d=1}^{D} s_d P) \tag{10}$$

The proposed method is depicted in detail in Algorithm 3 of the dictionary learning process in Fig. 2.

E. Discussion

In this article, a new method for revealing traffic patterns in video surveillance using the topic model is presented. This method focuses on extracting and classifying traffic patterns from video footage, with specific application in intelligent transportation systems. The proposed method involves segmenting video clips, extracting optical flux features, and using a topic coding approach to train the system about traffic patterns. The results of applying this method in the video data set are promising to identify and depict different traffic patterns. The importance of the article begins by emphasizing the increasing use of surveillance cameras in various applications, including intelligent transportation systems. The growth of video data is a challenge for manual monitoring and necessitates the development of automatic pattern recognition methods. This research deals with a practical problem and the identification of traffic patterns is very important for traffic management and analysis. It briefly discusses the available approaches for analyzing traffic scenes. It mentions the challenges associated with tracking-based methods and the potential of optical flux data to create more complex models. The introduction provides valuable background to the proposed method and provides context for its importance. This paper also discusses the identification of eight key traffic patterns in the dataset and provides a list of these patterns. However, it would be useful to include more information on how these patterns are determined and how they relate to real-world traffic scenarios. In addition, the paper notes that two traffic patterns were indistinguishable due to camera positioning, which raises questions about the method's limitations and potential improvements. In summary, the paper introduces an attractive approach to traffic pattern detection in video surveillance, but could benefit from a more detailed description of the method, a deeper analysis of the results, and a discussion of its broader significance and potential limitations.

IV. Implementation of the Proposed Method and Results

A. Implementation Environment

To acquire the optical flux of the clips, the proposed method was implemented in the C++ programming environment using open_CV functions. The methods were implemented using the C++ linear algebra package Armadillo [27], and the software was run on a computer with an Intel Core i4790 7 processor and 8 GB of memory.

B. Total Data and their Characteristics

The video images used in this article are from the QMUL1 dataset, which is shown in Fig. 4 as an example.

The activity analysis and behavior comprehension applications of this traffic data collection are particularly beneficial [28]. The tough video images in this bank have a resolution of 288 x 360, a frame rate of 25 frames per second, and a total of one hour (90,000 frames). It should be noted that the topic model has swiftly adopted this image bank as a go-to resource. Fig. 5 displays the eleven typical traffic patterns from the QMUL dataset, and Table III lists each one's description. They are as follows: Turning left from the south side to the west is traffic pattern number one. Crossing the intersection from the east to the west is traffic pattern number two. Turning right from the north side to the west is traffic pattern number four. Turning left from the west to the north is traffic pattern number five. Turning right from the east side to the north is traffic pattern number six. Turning left from the north side to the east is traffic pattern number seven. It should be noted that the camera's angle and height off the ground have a significant impact on how accurately the optical flux algorithm detects motion vectors. Also, Fig. 6 shows examples of traffic patterns in this collection.

Fig. 4. QMUL dataset containing one hour (90,000 frames) of high-volume traffic video collected at the intersection.

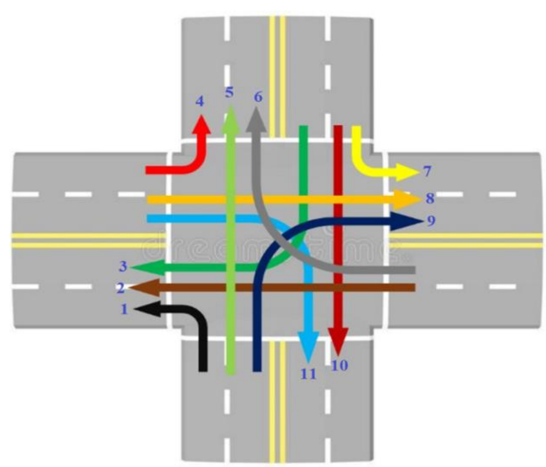

Fig. 5. Common traffic patterns in the QMUL data set (the description and numbering of the patterns are given in Table III).

TABLE III. TRAFFIC PATTERNS THAT ARE TYPICAL AND FREQUENT IN THE QUML DATASET

Traffic pattern number	name of the movement pattern	Recognizability
1	From the south side, turn left and head west.	□
2	moving from the east side of the intersection to the west side	□
3	From the north side, turn right to reach the west side.	□
4	From the west side, make a left turn to the north side.	□
5	moving from the south side of the crossroads to the north side	□
6	From the east side, make a right turn to the north side.	□
7	From the north side, make a left turn to the east side.	□
8	moving from the west to the east across the intersection	□
9	On the south side, turn right toward the east.	□
10	from the north side to the south side of the intersection	□
11	From the west side, turn right to go to the south side.	□

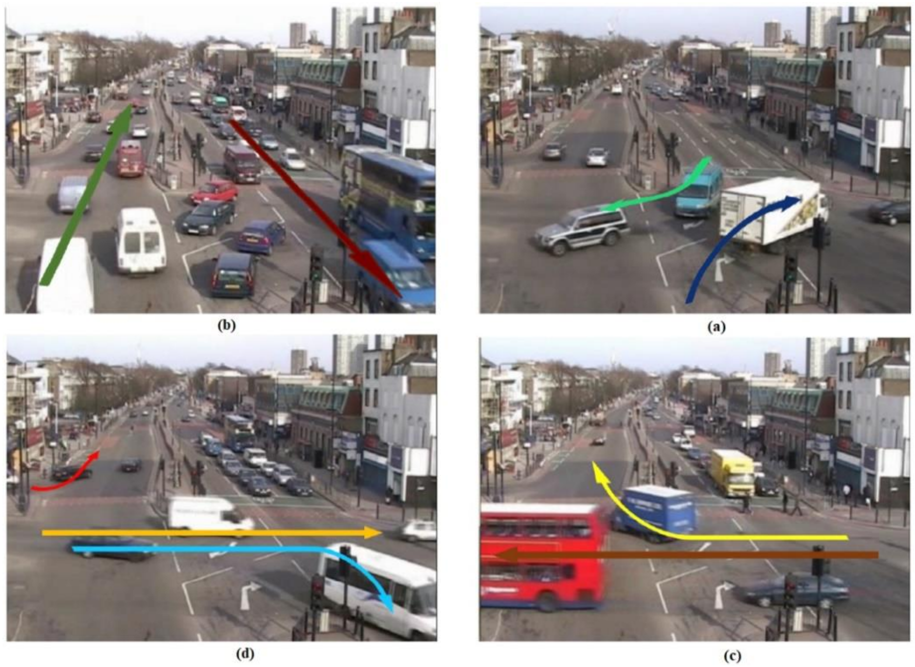

Fig. 6. Examples of traffic patterns in the QMUL image bank, (a) pattern 3 and 9, (b) pattern 5 and 10, (c) pattern 2 and 6 and (d) pattern 4, 8 and 11.

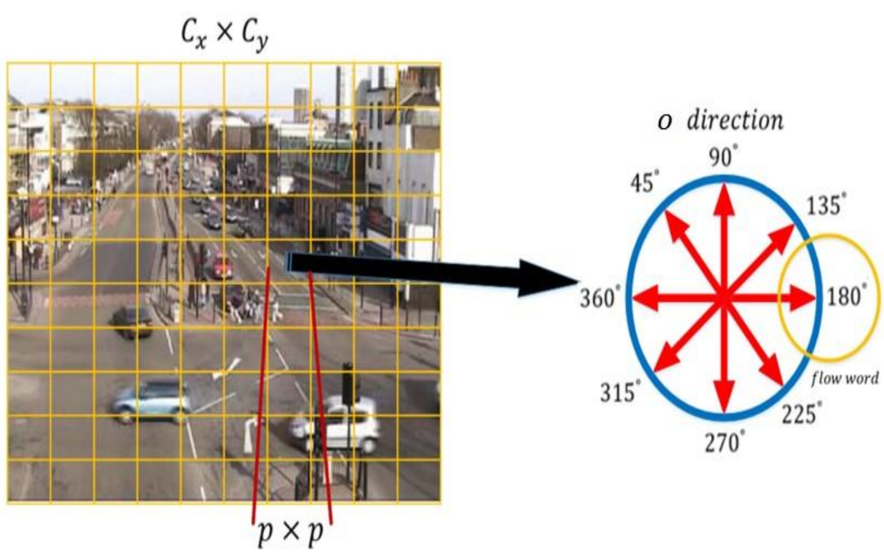

Fig. 7. How to make stream words.

C. Implementation Processes of the Proposed Method

In implementing the proposed method, the input video was first divided into three-second sub-clips. Then the optical flux characteristics were extracted for each pair of consecutive frames in each clip using the TVL-1 optical flux algorithm [29]. After that, as shown in Fig. 7, the output of the optical flux algorithm was divided into 10 x 10 cells. According to the resolution of each frame, there are 29 x 36 cells, and each of the motion vector values in each cell is one of 8 directions: 45º, 90º, 135º, 180º, 225º, 270º, 315º, and 360º. Then this process is repeated for each clip, and finally each clip is displayed with a vector length of $N = 8 \times 36 \times 29 = 8352$. Each member of the vector represents the number of occurrences of the current word. We experimentally set the λ parameter equal to 0.1 and the K value equal to 25. We considered the maximum number of repetitions of the program to be 20, and the value of ε for the convergence of the algorithms was 0.01. In this experiment, 200 clips were used for training.

D. Implementation Results

After running the proposed algorithm on the QMUL bank, traffic patterns were extracted. As can be seen in Fig. 8, three common movement behaviors—turning left [Fig. 8(a)], turning right [Fig. 8(b)], and crossing the intersection [Fig. 8(c)]—have been extracted. Also, in Fig. 8(d) to 8(i), two traffic patterns have been extracted simultaneously. As mentioned before, due to the importance of the location of the camera in the detection of traffic patterns, traffic patterns one and seven are undetectable in the QMUL bank and are not observed in the detected traffic patterns. Finally, according to the number and timing of red lights and traffic routines of cars in this data set, the proposed method has been able to correctly extract eight meaningful flows and patterns from the nine existing traffic patterns (accuracy 88.8%). The use of local movements, which is the movement of pixels between two frames, as features of "direction of flow detection" and "motion patterns", increases the possibility of misdiagnosis. For example, the presence of pedestrians or other moving objects increases the possibility of detecting meaningless patterns.

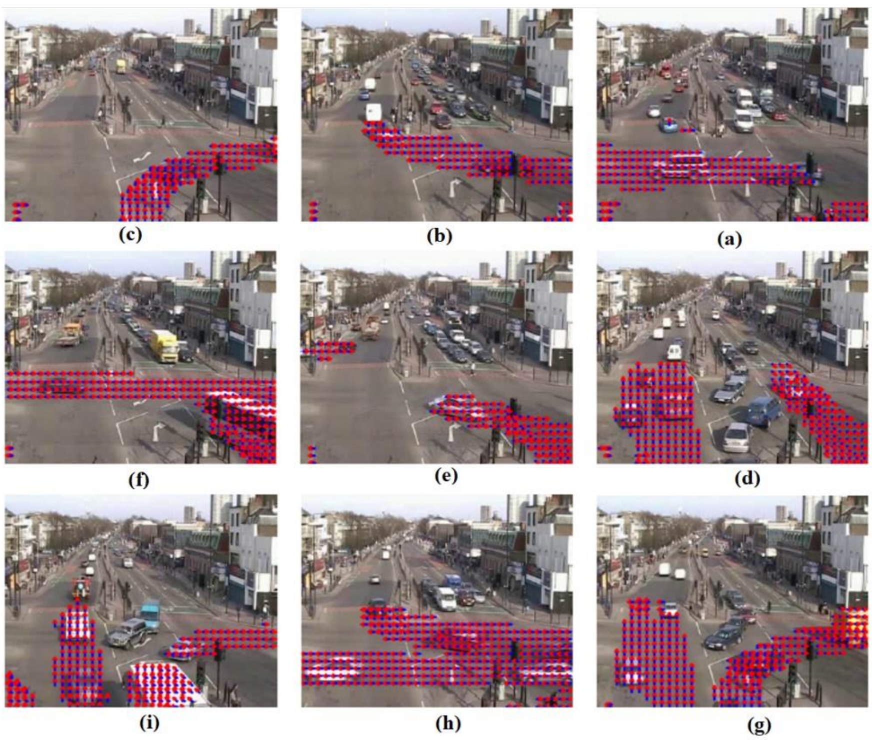

Fig. 8. Traffic flows and patterns obtained by the proposed method, (a) going left (b) turning north (c) a right turn from south to east, (d) crossing the intersection (e)) a right turn from west to south (f) going right from the west side to the south side (g) crossing the intersection from the south side to the north side (h) turning right from the east side to the north (i) turning to Right from south to east.

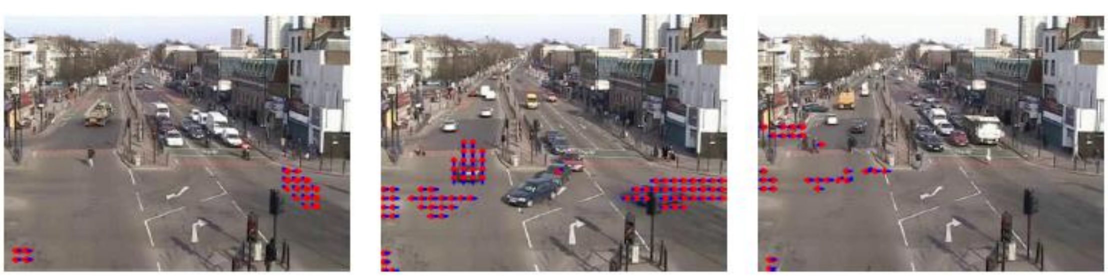

Fig. 9. Traffic patterns extracted from QMUL bank that do not have a specific meaning.

Naturally, scientists have employed these patterns to identify anomalies and peculiar movements [1]. There are other traffic patterns in the bank, as seen in Fig. 9 taken from QMUL, that have no clear significance. Therefore, it will be challenging to analyze the scenario, uncover the rules, and apply them using the patterns gleaned from the thematic model. Some traffic flows and patterns recovered have no particular significance because of the topic model's nature, which was primarily created and utilized for processing natural language and text and does not account for difficulties such as abrupt changes in illumination, image size, or camera viewing angle.

This paper has several significant contributions in the field of traffic pattern detection in video surveillance, which can be summarized as follows:

1) Unsupervised Traffic Pattern Recognition: The main contribution of this work is the development of an unsupervised method to identify and classify traffic patterns in video surveillance data. Traditional approaches often rely on predefined training datasets or tracking techniques, which may be limited by the need for extensive manual labeling or tracking accuracy issues. The proposed method takes a different approach by using a topic coding model to automatically extract traffic patterns without the need for a previous training dataset. This non-supervised nature of the method makes it adaptable to a wide range of traffic scenarios, including scenarios with diverse and evolving patterns.

2) Topic coding with optical flux features: This paper introduces the use of topic coding, a technique commonly used in natural language processing, in the context of video analysis. This method transforms visual traffic movement patterns into descriptive expressions that enable the identification and categorization of significant traffic patterns. To achieve this goal, the authors use optical flux features extracted from video frames that contain detailed information about local motion. This combination of thematic coding and optical flux features provides a new and promising approach for traffic pattern recognition. The contributions of this work extend to the development of an automated system to recognize common traffic behaviors at intersections, such as left turns, right turns, and crossing intersections. These contributions have significant relevance for applications in intelligent transportation systems, traffic management, and public safety, where accurate and automatic traffic pattern detection is essential for efficient monitoring and decision-making.

V. Conclusion

In urban traffic, crossroads play a significant role, and surveillance cameras are frequently utilized to regulate and control public spaces. The traditional method of video surveillance does not function well due to the vast volume of video data and the unreliability of humans; therefore, a system that automatically collects traffic flows and patterns is necessary. A non-supervisory strategy to identify traffic patterns in video surveillance was put forth in this paper. First, the optical flux algorithm is used to determine the traffic flow in each frame. The Car was taught using the theme model of these traffic patterns because behaviors like turning left, turning right, and traveling straight through an intersection are considered to be common significant traffic patterns in the images observed by the camera. According to the position of the camera in the QMUL bank, the implementation results revealed that the suggested method was able to accurately calculate eight significant motion patterns out of a total of nine conceivable patterns. Turning left, turning right, and passing straight are frequent and permitted in areas where traffic is controlled, such as junctions, according to traffic regulations and driving laws. Now, if there are movements that deviate from these authorized and typical patterns, they can first be identified as unexpected events before suitable choices, like recording a violation, can be taken into account. In other words, future studies may focus on the identification of violations in terms of atypical movement. A future study might examine how well the suggested algorithm holds up to environmental elements in the image, such as changes in light intensity, camera position, and difficulties similar to those seen in image processing. Some limitations and potential challenges in this study are as follows:

Reliability of tracking-based methods, one class of which involves first tracking objects (e.g. cars or people) and then using the tracked trajectories for further analysis: This indicates that these methods may suffer from challenges related to the accuracy and reliability of object detection and tracking.

Inadequate anomaly detection, where some existing methods may not effectively detect anomalies or unusual traffic behavior: This shows that the proposed approach can potentially detect unexpected events or violations by identifying patterns that deviate from the permitted and normal traffic behaviors. This indicates that existing methods may have limitations in detecting and classifying anomalies in traffic patterns.

Environmental variability, where the location and angle of surveillance cameras can significantly affect the accuracy of optical flux-based motion vector detection: This suggests that existing methods may be limited by environmental factors, such as changes in lighting conditions, changes in camera positions, and other challenges commonly encountered in video surveillance.

Acknowledgment

This research was supported by Key Technology Research and Development Program of Henan Province China (Grant No.232102210194).

References

[1] H. Zou, K. Cao, and C. Jiang, “Spatio-temporal visual analysis for urban traffic characters based on video surveillance camera data,” ISPRS Int J Geoinf, vol. 10, no. 3, p. 177, 2021.

[2] L. Song, F. Jiang, Z. Shi, R. Molina, and A. K. Katsaggelos, “Toward dynamic scene understanding by hierarchical motion pattern mining,” IEEE Transactions on Intelligent Transportation Systems, vol. 15, no. 3, pp. 1273–1285, 2014.

[3] I. Tvoroshenko and M. Dziubenko, “Modern methods of analysis of the movement scheme using video detection of vehicles,” 2020.

[4] Y. Cheng, X. Ji, T. Lu, and W. Xu, “On detecting hidden wireless cameras: A traffic pattern-based approach,” IEEE Trans Mob Comput, vol. 19, no. 4, pp. 907–921, 2019.

[5] J. Varadarajan, R. Emonet, and J.-M. Odobez, “A sequential topic model for mining recurrent activities from long term video logs,” Int J Comput Vis, vol. 103, no. 1, pp. 100–126, 2013.

[6] M. C. R. Murça, R. J. Hansman, L. Li, and P. Ren, “Flight trajectory data analytics for characterization of air traffic flows: A comparative analysis of terminal area operations between New York, Hong Kong and Sao Paulo,” Transp Res Part C Emerg Technol, vol. 97, pp. 324–347, 2018.

[7] N. Apthorpe, D. Reisman, and N. Feamster, “A smart home is no castle: Privacy vulnerabilities of encrypted iot traffic,” arXiv preprint arXiv:1705.06805, 2017.

[8] X. Ji, Y. Cheng, W. Xu, and X. Zhou, “User presence inference via encrypted traffic of wireless camera in smart homes,” Security and Communication Networks, vol. 2018, 2018.

[9] F. Duarte and C. Ratti, “What urban cameras reveal about the city: The work of the Senseable City Lab,” Urban Informatics, pp. 491–502, 2021.

[10] N. Apthorpe, D. Reisman, S. Sundaresan, A. Narayanan, and N. Feamster, “Spying on the smart home: Privacy attacks and defenses on encrypted iot traffic,” arXiv preprint arXiv:1708.05044, 2017.

[11] Y. Cheng, X. Ji, T. Lu, and W. Xu, “Dewicam: Detecting hidden wireless cameras via smartphones,” in Proceedings of the 2018 on Asia Conference on Computer and Communications Security, 2018, pp. 1–13.

[12] J. Barthélemy, N. Verstaevel, H. Forehead, and P. Perez, “Edge-computing video analytics for real-time traffic monitoring in a smart city,” Sensors, vol. 19, no. 9, p. 2048, 2019.

[13] M. C. R. Murca and R. J. Hansman, “Identification, characterization, and prediction of traffic flow patterns in multi-airport systems,” IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 5, pp. 1683–1696, 2018.

[14] Khezri, E., Zeinali, E., & Sargolzaey, H. (2023). SGHRP: Secure Greedy Highway Routing Protocol with authentication and increased privacy in vehicular ad hoc networks. Plos one, 18(4), e0282031.

[15] H. Luo, J. Liu, W. Fang, P. E. D. Love, Q. Yu, and Z. Lu, “Real-time smart video surveillance to manage safety: A case study of a transport mega-project,” Advanced Engineering Informatics, vol. 45, p. 101100, 2020.

[16] Khezri, E., & Zeinali, E. (2021). A review on highway routing protocols in vehicular ad hoc networks. SN Computer Science, 2, 1-22.

[17] Zhang, H., Zou, Q., Ju, Y., Song, C., & Chen, D. (2022). Distance-based support vector machine to predict DNA N6-methyladenine modification. Current Bioinformatics, 17(5), 473-482.

[18] Cao, C., Wang, J., Kwok, D., Cui, F., Zhang, Z., Zhao, D., ... & Zou, Q. (2022). webTWAS: a resource for disease candidate susceptibility genes identified by transcriptome-wide association study. Nucleic acids research, 50(D1), D1123-D1130.

[19] Zhao, H., Wang, H., Xu, N., Zhao, X., & Sharaf, S. (2023). Fuzzy approximation-based optimal consensus control for nonlinear multiagent systems via adaptive dynamic programming. Neurocomputing, 553, 126529.

[20] M. Samiei, A. Hassani, S. Sarspy, I. E. Komari, M. Trik, and F. Hassanpour, “Classification of skin cancer stages using a AHP fuzzy technique within the context of big data healthcare,” J Cancer Res Clin Oncol, pp. 1–15, 2023.

[21] J. Sun, Y. Zhang, and M. Trik, “PBPHS: a profile-based predictive handover strategy for 5G networks,” Cybern Syst, pp. 1–22, 2022.

[22] M. Trik, H. Akhavan, A. M. Bidgoli, A. M. N. G. Molk, H. Vashani, and S. P. Mozaffari, “A new adaptive selection strategy for reducing latency in networks on chip,” Integration, vol. 89, pp. 9–24, 2023.

[23] Trik, M., Pour Mozafari, S., & Bidgoli, A. M. (2021). An adaptive routing strategy to reduce energy consumption in network on chip. Journal of Advances in Computer Research, 12(3), 13-26.

[24] Yue, S., Niu, B., Wang, H., Zhang, L., & Ahmad, A. M. (2023). Hierarchical sliding mode-based adaptive fuzzy control for uncertain switched under-actuated nonlinear systems with input saturation and dead-zone. Robotic Intelligence and Automation, 43(5), 523-536.

[25] Wang, T., Zhang, L., Xu, N., & Alharbi, K. H. (2023). Adaptive critic learning for approximate optimal event-triggered tracking control of nonlinear systems with prescribed performances. International Journal of Control, 1-15.

[26] M. Trik, A. M. N. G. Molk, F. Ghasemi, and P. Pouryeganeh, “A hybrid selection strategy based on traffic analysis for improving performance in networks on chip,” J Sens, vol. 2022, 2022.

[27] Khezri, E., Zeinali, E., & Sargolzaey, H. (2022). A novel highway routing protocol in vehicular ad hoc networks using VMaSC-LTE and DBA-MAC protocols. Wireless Communications and Mobile Computing, 2022.

[28] Cao, Y., Xu, N., Wang, H., Zhao, X., & Ahmad, A. M. (2023). Neural networks-based adaptive tracking control for full-state constrained switched nonlinear systems with periodic disturbances and actuator saturation. International Journal of Systems Science, 54(14), 2689-2704.

[29] C. Yang, K. Clarke, S. Shekhar, and C. V. Tao, “Big Spatiotemporal Data Analytics: A research and innovation frontier,” International Journal of Geographical Information Science, vol. 34, no. 6. Taylor & Francis, pp. 1075–1088, 2020.

[30] Wang, Z., Jin, Z., Yang, Z., Zhao, W., & Trik, M. (2023). Increasing efficiency for routing in internet of things using Binary Gray Wolf Optimization and fuzzy logic. Journal of King Saud University-Computer and Information Sciences, 35(9), 101732.

[31] Li, X., Gui, J., & Liu, J. (2023). Data-driven traffic congestion patterns analysis: A case of Beijing. Journal of Ambient Intelligence and Humanized Computing, 14(7), 9035-9048.

[32] Pascale, A., Macedo, E., Guarnaccia, C., & Coelho, M. C. (2023). Smart mobility procedure for road traffic noise dynamic estimation by video analysis. Applied Acoustics, 208, 109381.