IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 23 March 2022, accepted 6 May 2022, date of publication 30 May 2022, date of current version 5 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3178714

Research on Expressway Traffic Event Detection at Night Based on Mask-SpyNet

XIANGLUN MO, CHUANPENG SUN, CHENYU ZHANG, JINPENG TIAN, AND ZHUSHUAI SHAO
Department of Transportation, China University of Mining and Technology, Xuzhou 221116, China

Corresponding author: Chuanpeng Sun (sunchuanpeng9999@163.com)

This work was supported in part by the China University of Mining and Technology, in part by the National Undergraduate Training Program for Innovation and Entrepreneurship under Grant 202010290060, and in part by the Jiangsu Undergraduate Training Program for Innovation and Entrepreneurship under Grant 20200013cx.

ABSTRACT Expressway traffic event detection at night is essential for improving rescue efficiency and avoiding secondary accidents. Most expressways in China have built a complete Expressway video monitoring system. However, at night, the expressway traffic event detection still adopts manual detection, which is inefficient. In this dissertation, the strategy of expressway traffic event detection at night has been analysed first. On this basis, by combining the Mask method and SpyNet deep learning, this study develops a night highway vehicle detection deep learning network with a dense optical flow formed by night vehicle light flow as the detection object. Finally, the Deepsort algorithm is used to track and measure the velocity of the detected target. The measured data are used to compare the background difference method, classical optical flow method, YOLO-v3 and proposed method in this paper. The results show that the proposed method has the advantages of high detection accuracy and fast detection speed.

INDEX TERMS Expressway traffic event detection at night, image processing, mask algorithm, SpyNet.

I. INTRODUCTION

Expressway has a fast driving speed, few obstacles, and many straight lines. This causes drivers to short emergency response time when faced with traffic events. The severity of traffic accidents caused by traffic accidents is also much higher than that of ordinary highway transportation. Expressway traffic events generally refer to events that reduce the road capacity or interfere with the operation of expressway vehicles, such as traffic accidents, temporary roadside parking, traffic congestion. When a highway traffic incident occurs, if it cannot be confirmed and handled in time, it may lead to more serious traffic accidents, resulting in substantial economic losses and casualties. Especially at night, affected by natural light and lighting conditions, the concealment of expressway traffic events at night is stronger. The expressway management department sets up a video camera every 2 kilometers along the Expressway to enhance the detection efficiency of expressway traffic incidents. It constructs a video monitoring system with complete coverage of the Expressway. However, the detection of traffic events in the system still needs to be completed by watching the video manually, and the automatic alarm of traffic events can not be carried out according to the monitoring video. There are many highway surveillance cameras, which rely entirely on manual traffic event detection, which requires a lot of human resources. At present, the only manual rolling camera can be used for inspection, and the adequate detection coverage is low. As a result, after most traffic events, if there is no alarm or is not found and handled in time by patrol inspection, more severe traffic events will be caused.

It is necessary to develop a method that can detect traffic incidents in real time. The detection of traffic events is closely related to the detection and recognition of vehicle states. Vehicle breakdown, rear-end collision, and illegal parking are all changes in vehicle motion status, which can be judged according to vehicle speed. Therefore, the vehicle state information can be effectively identified and analyzed by evaluating the speed parameters. Based on the analysis of nighttime environment characteristics and highway hardware conditions, this paper proposes a nighttime traffic incident detection method based on Mask-SpyNet. The method can effectively detect the real-time status of vehicles in various environments of highways at night, which is conducive to timely feedback and alarming of abnormal traffic events,

The associate editor coordinating the review of this manuscript and approving it for publication was Gangyi Jiang.

avoids subsequent secondary safety accidents, and fully guarantees the safety of highway transportation.

Vehicle recognition is the key in vehicle speed detection by image processing at night. The earliest target recognition method by image processing at night is based on the principle of the frame difference method. Li combines the adjacent frame difference method and the inter-frame difference method to obtain the cumulative frame difference method for night vehicle detection [1]. Wang *et al.* first extracted the region of interest by frame difference method, then calculated the optical flow of each pixel in the region, and then used the optical flow field for vehicle tracking [2]. Gu uses the dynamic background difference method to adapt to different night environments [3]. N. Arunprasath combines the dynamic background change with the threshold setting range to obtain a moving target detection method [4]. O'malleyr uses a low static camera to reduce exposure and obtain a clearer headlight area. Then the threshold segmentation method and the left-right symmetry of the lamp are used for detection and pairing [5]. Finally, the perspective correction method is used to correct the image of lamp asymmetry caused by vehicle turning and road bending. Kuo used the multi-level image processing algorithm to obtain the region of interest of the tail lamp and then used the tail lamp clustering algorithm to identify the vehicle [6]. Du *et al.* selected tail lamps' R-G colour difference for image segmentation and set the judgment conditions of tail lamp adhesion area using non-maximum suppression [7]. To solve the problem of vehicle occlusion, he also proposed the vehicle bounding box and set the maximum overlap rate. Hu *et al.* applied the gradient threshold to the frame difference method to obtain the binary image of the moving region and then calculated the optical flow of the regional feature points. Sincan *et al.* [8]. Transformed the background image of optical flow into target detection under static background according to the motion vector of optical flow [9]. When dealing with moving vehicle targets at night, Choi *et al.* Can not use the optical flow method to detect the continuous traffic flow in the video [10]. Yolo algorithm is a target detection and recognition algorithm based on the regression method. The algorithm simplifies the detection process, uses global features, significantly improves the speed, and has good real-time performance [11]. SSD algorithm is different from the detection after the full connection layer of the Yolo algorithm. This method uses a convolutional neural network to detect the target directly [12]. RCNN is a target detection algorithm based on candidate regions. The algorithm uses a convolutional neural network in deep learning to learn and extract the features of sample images automatically, so this algorithm's accuracy was greatly improved at that time [13]. Fast RCNN still adopts the selective search algorithm, but the difference is that a whole sample image is an input into the convolution neural network for learning [14]. After obtaining the complete feature map of each image, the candidate regions obtained in the previous step are scaled and mapped into the feature map to obtain the features of each candidate region; the features are normalised by the pooling layer of the region of interest, and finally classified by softmax. Fastrcnn algorithm not only improves the accuracy but also improves the speed. However, this method adopts the selective search method, and the redundant information will increase due to the superposition of search times and will become the main bottleneck of speed improvement. Compared with the Fast-RCNN algorithm, Faster-RCNN algorithm makes full use of a convolutional neural network [15]. It uses a regional recommendation network (RPN) to extract image candidate regions in the initial stage, directly predicting candidate frames, and realise high-speed network processing. Some representative algorithms such as R-FCN and FPN have become the latest candidate region-based methods, and the detection accuracy has been improved [16], [17]. However, the detection process is generally similar to that of the R-CNN algorithm, and the detection speed can not be effectively improved. The first attempt to estimate the optical depth flow with end-to-end CNN was flownet1.0, proposed by Dosovitskiy *et al.* of Freiburg University in Germany in 2015 [18]. Because of the poor accuracy of flownet1. 0, the team of Freiburg University in Germany proposed flownet2. 0 in 2017, which has greatly improved the accuracy and can reach the same accuracy as the traditional method [19]. In 2018, sun *et al.* Proposed a PwC net model based on the CNN model and achieved high accuracy on MPISintel and KITTI2015 data sets [20]. IRRNet is an improvement of PWC-Net [21]. It makes full use of the idea of iteration. The difference from flownet2.0 is that iteration is placed inside the network structure, and parameter sharing is realised. The number of parameters is reduced, and the accuracy is improved. Selflow algorithm is the latest research method. This method uses multidimensional data sets for learning and greatly improves the algorithm's accuracy combined with image information [22]. In 2019,chen et al propose a Multi-modal Neural Network (MMN) to process sensor observations and social media texts simultaneously and detect traffic events [23].

Although the traditional method is simple and easy to operate, it is difficult to extract features due to the low definition of vehicle images at night, and the traditional method cannot adapt to the complex changes of the night environment, so the vehicle detection accuracy is often far away. For the detection method of deep learning, the previous model needs to be fully trained to adapt to the complex background environment at night, so as to achieve a high detection accuracy. However, the detection algorithm of deep learning cannot adapt to high-intensity environmental changes well for the interference of strong light from opposing vehicles and reflected light from objects at night. Although pre-image processing work is required, the over-exposure of the image will result in very few feature points and strong interference, so it cannot meet the requirements of high accuracy. Therefore, we construct a Mask-SpyNet-based optical flow detection method for vehicle optical flow at night. This method can effectively eliminate the interference of unfavorable factors at night. Combined with the Deepsort tracking algorithm, it can complete the calculation of vehicle speed at night, and judge whether the vehicle has abnormal traffic events such as overspeed, abnormally low speed, and parking. The method proposed in this paper combines optical flow, the most significant feature at night, to realize the detection of traffic incidents in a strong interference environment at night, which can well complement the method and theoretical system of nighttime traffic incident detection.

II. DETECTION PROCESS

As described above, the core of expressway traffic incident detection at night is to complete vehicle detection, track, and calculate vehicle speed. Whether the vehicle breaks down, rear-end collision, or illegal parking, it belongs to the change of vehicle motion state caused by the vehicle's speed. Therefore, this paper mainly detects traffic events at night to detect the vehicle speed parameters. The light at night is insufficient. At the same time, it is difficult to identify the vehicle body due to the interference of lamplight. The vehicle position can be determined by identifying the light flow of the lamp. In this study, a Mask-SpyNet deep learning network for light flow detection is constructed for network training, and vehicle recognition and speed detection are completed combined with Deepsort. The detection flow chart of this paper is shown in Fig1.

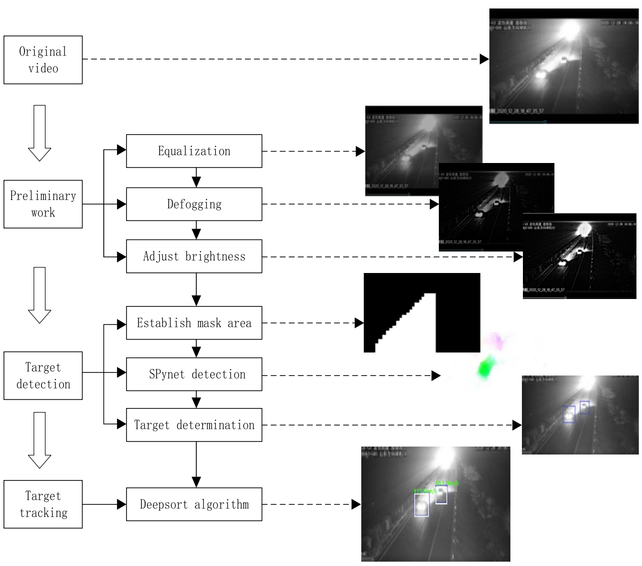

FIGURE 1. Detection process.

III. IMAGE PREPROCESSING

Before training the data set, image preprocessing is usually needed to ensure better robustness of the algorithm. On the one hand, preprocessing aims to enhance the data set and ensure the correctness of subsequent processing. On the other hand, it is to prepare for the follow-up comparison with traditional methods. Night images have the characteristics of local solid light interference and a dark environment. Video images are preprocessed to increase contrast and brightness. Firstly, the histogram equalisation of the image is carried out to weaken the overexposed area and keep the brightness of the enhanced image near the median of the gray level range. In the dark environment at night, when light propagates in fog, haze, and other media, the image information collected by the camera sensor is seriously degraded due to the scattering effect of particles, resulting in significant detection errors. The dark channel a priori algorithm will eliminate the influence of fuzzy environment on image quality and improve image visualisation. Finally, for the problem of insufficient brightness at night, a gamma correction algorithm is used to improve image brightness and contrast.

A. HISTOGRAM EQUALISATION

Histogram equalisation is an automatic image contrast adjustment algorithm widely used because of its simplicity and effectiveness [24]. This method makes the histogram of the transformed image evenly distributed, and the brightness of the enhanced image is usually maintained near the median of the gray level range, which can effectively deal with the contrast of solid light sources in the dark environment so that the image can achieve the effect of overexposure and light reduction. The specific optimisation steps are as follows:

Set the image x size to M × N. The image after histogram equalisation is y, and the gray level range is set to [0255]. The histogram equalisation steps are as follows:

① Statistical image X gray histogram h_r.

Calculate the histogram distribution probability p_f according to the following formula:

$$p_f = \frac{1}{M \times N} \cdot h_f \tag{1}$$

③ Calculate the histogram probability cumulative distribution probability p_a according to the following formula:

$$p_a = \sum_{k=0}^{k} p_f \tag{2}$$

④ According to the mapping function f in the following formula, the gray level k of the input image X is mapped to obtain the input image y.

$$f(k) = 255 \cdot p_a(k) \tag{3}$$

B. DARK CHANNEL ALGORITHM

He *et al.* proposed the classical dark channel prior theory by analysing and summarising the characteristics of fog-free images [25]. In the dark channel prior theory, it is mentioned that in sunny weather, for most clear images taken outdoors, after removing the sky area of the image, there is a pixel value with low intensity tending to 0 in at least one channel in the RGB image. Therefore, any clear image $J(x)$ can be expressed in the following form:

$$J^{dark}(x) = \min_{y \in \Omega(x)} (\min_{c} J^c(y)) \tag{4}$$

$$J^{dark} \rightarrow 0 \tag{5}$$

where J^c is the image of a color channel in RGB color space in the precise image $J(x)$, $\Omega(x)$ is a neighborhood-centered on pixel x in the image $J(x)$. The formula (5) in the dark channel prior. He *et al.* realize image defogging based on the dark channel prior and atmospheric scattering model. The general steps of the DCP algorithm are as follows:

① Estimate and refine the transmittance value;

② Estimate the atmospheric illumination value;

③ Restore the fog-free clear image according to the estimated transmittance and atmospheric illumination values.

Because the radiation illumination of the scene is usually darker than the atmospheric illumination, the image after defogging is darker. Therefore, this paper adjusts the brightness of the image after defogging.

C. GAMMA CORRECTING

In the dark channel algorithm image preprocessing, the brightness of the grey image is too low, so it needs to be adjusted. The contrast of critical features increases through image gamma correction, and the display effect improves. In order to overcome the nonlinear input characteristics of image acquisition and display, a nonlinear compensation function is introduced. The mathematical formula of gamma correction is:

$$s = cr^{\gamma} \tag{6}$$

c and γ are constant. When $\gamma > 1$, the small range grey value and the extensive range high grey value will be interchanged. The small range low grey value will transition to the extensive range high grey value, and the large range high grey value will transition to the small range low grey value. When $\gamma < 1$, on the contrary, a small range of high grey values will transition to an extensive range of low grey values, and a large range of low grey values will transition to a small range of high grey values. Change power value (γ) can change the gamma curve of CRT. Experienced, $\gamma = 1$ is the calculated boundary value, and c is 0.5 [26].

IV. CONSTRUCTION AND TRAINING OF DEPTH LEARNING MODEL FOR DENSE OPTICAL FLOW VEHICLE SETECTION BASED ON MASK-SpyNet

Optical flow estimation and detection can be divided into the sparse optical flow and the dense optical flow. The difference between sparse optical flow and dense optical flow mainly depends on the sparse point of the image, that is, the feature saliency and gradient of the point. Dense optical flow is a point-to-point matching method for optical flow images. For convenience, different colours and brightness represent the size and direction of optical flow. Unlike sparse optical flow, only several feature points on optical flow images are processed. Therefore, dense optical flow calculates the offset of all points on the image to form a dense optical flow field, which is more conducive to determining the image feature information and effectively avoiding the interference of individual feature points when sparse optical flow calculates the night image. This paper combines Mask image processing and SpyNet target extraction method to develop a Mask-SpyNet vehicle recognition method for the dense optical flow of Expressway at night. Mask-SpyNet adds a branch of the predicted split mask based on SpyNet, and its structure is shown in Fig 2. Mask-SpyNet reduces the impact of the environment on night highway optical flow by generating high-quality segmentation masks for each instance.

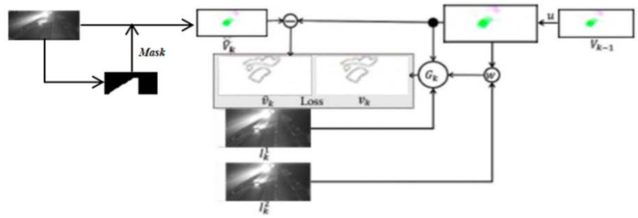

FIGURE 2. Mask-SpyNet model structure.

FIGURE 3. Original image.

FIGURE 4. Establishment of original image mask.

A. MASK PROSSING IN THE IMAGE

Mask processing in image processing comes from lithography in the semiconductor manufacturing process. The image is occluded with specific graphics to control the processing area of the whole image. The process of optical image processing is mostly film, filter. The process of digital image processing is mostly matrix array and multivalued image.

The main functions of Mask processing in the image are as follows.

① Control function. Mask image processing can control the area outside the target area and shield the area outside the control area to prevent the interference of non-characteristic areas.

② Feature structure extraction. The structural features similar to the mask in the image are detected and extracted by similarity variables or image matching methods.

③ Feature region extraction. The image is occluded with a specific figure to obtain the image of the characteristic area. The image value in the characteristic area remains unchanged, while the image value outside the area is 0.

Before detecting the target vehicle, this paper extracts the ROI from the image data in the video and determines the ROI area by establishing the image mask to eliminate the interference of other factors outside the ROI area, such as night opposite optical flow reflected light. The original Fig 3 mask is processed to obtain Fig 4. Establishing an image mask lays a good foundation for subsequent vehicle detection and tracking.

B. THE DEEP LEARNING NETWORK STRUCTURE IN MASK-SpyNet

The convolution network $\{G_0, \cdots\cdots, G_k\}$ is trained in an independent and serialized way. Through a given input $(I_k^1, w(I_k^2, u(V_{k-1})), u(V_{k-1})$, the residual v_k is calculated in layer k through the optical flow field sampled on the previous layer and the target optical flow field, but $u(V_{k-1})$ is obtained from the network trained on the previous layer. Training data set: MPI Sintel data set is adopted, which solves the limitations of the existing optical flow benchmark. It provides natural video sequences, which is challenging for current methods. It aims to encourage people to study long-distance motion, motion blur, multi-frame analysis, and non-rigid motion. The dataset contains flow field, motion boundary, mismatch region, and image sequence. Render image sequences at different levels of difficulty. Here, it modifies the frame slice in many ways for optical flow evaluation. Five convolution networks are trained using the PyTorch machine learning library, each with five convolution layers. After each layer (except the last layer), ReLU is activated. Each layer is trained at different resolutions of each layer, in which 24 * 36 down sampled images are used for training; The picture size of the next layer is a multiple of the size of the previous layer, and the training picture size is 1184 * 608; The training network adopts 7 * 7 convolution kernel, which is better than small convolution kernel. The number of convolution cores of each convolution network is eight channels are output, and two optical flow fields (x direction, Y direction) are output. Network initialization: the parameters used during network training are initialization parameters.

C. DATA SET

The training data set used in this paper consists of two parts, a total of 2045 pictures. Part of the training data comes from MPI Sintel data set. MPI Sintel is an open-source number composed of pairs of animated images collected in the movie data set, selecting 1628 frames with 35 scene amounts from 18000 movie clips. These frames have different motion scenes. Most still or similar motion scenes are screened out in the selection process, and too simple small displacement and motion scenes that humans cannot distinguish are deleted. The MPI Sintel dataset is also copied into clean and final, in which the clean image contains shadows, while the image data in final includes more complex situations such as motion blur, caustics blur, and atmospheric effect in addition to shadows. The MPI Sintel dataset image contains many problems such as occlusion, large displacement, and detail presentation. A complex problem of optical flow prediction expected to be solved in this paper is one of the most commonly used data sets in optical flow prediction.

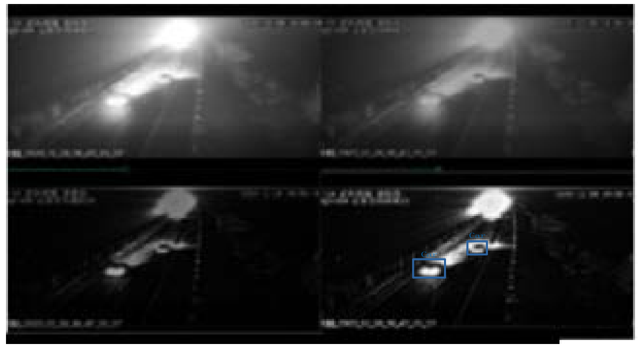

FIGURE 5. Data initial processing and annotation process.

Another part of the training data is annotated and produced by ourselves. We intercepted 417 images containing vehicles in the obtained night highway video. After initial image processing, these images are annotated. This part of the image as a data set has particular pertinence. It can improve the model's accuracy.The data initial processing and annotation process is shown in Fig 5.

D. NETWORK TRAINING PROCESS

Optical flow estimation and detection can be divided into the

1) SPATIAL SAMPLING

Downsampling is represented by d (x). Assuming that the original figure is I (m ∗ n), then the size of d (I) is (m/2 ∗ n/2). The exact process also operates the optical flow field results. Upsampling is represented by

Assuming that the image size before upsampling is I (m ∗ n), then the size u (I) of is (2m ∗ 2n). The exact process also operates the optical flow field results. The image distortion warp is represented by w (I, V). According to the optical flow field V, the picture I is distorted by two-line interpolation to obtain a picture.

2) MODELING CALCULATION

$$v_k = G_k(I_k^1, w(I_k^2, u(V_{k-1})), u(V_{k-1}))) \tag{7}$$
$$V_k = u(V_{k-1}) + v_k \tag{8}$$

where: $\{G_0, \cdots\cdots, G_k\}$ represents a group of trained convolutional neural network models;

v_k represents the optical flow residual of layer k, which the model calculates according to the sampling results on the video frame $\{I_k^1, I_k^2\}$ of layer k and the optical flow field of the upper layer.

$w(I_k^2, u(V_{k-1}))$ represents a distorted new picture and distorts the second picture of layer k from the sampling results of the optical flow field of the previous layer;

V_k represents the optical flow field result of the k layer, which is obtained by adding the upper sampling result of the optical flow field of the previous layer and the optical flow residual of the current layer, that is, formula (8).

As shown in Figure 6, the specific flow chart is as follows.

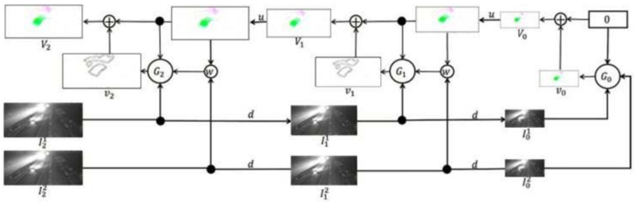

FIGURE 6. Three-layer pyramid network inference.

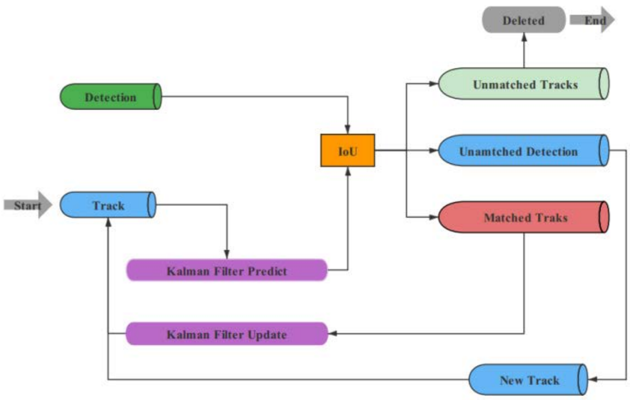

FIGURE 7. Deepsort algorithm framework flow.

Starting from the bottom sampled picture $\{I_0^1, I_0^2\}$, the top layer of the pyramid. Initialize the residual of the optical flow field of the 'upper layer' to 0, calculate the residual v_0 through the network G_0. Calculate the optical flow field V_0 of the G_0 network layer by using formula (8). Transfer the optical flow field result V_0 calculated by G_0 to G_1, calculate the residual V_1 through $\{I_1^1, w(I_1^2, u(V_0))\}$, and calculate the optical flow field V_1 of G_1 network layer by formula (8). Repeat the previous operation, calculate the residual V_2, and finally calculate the optical flow field V_2 (the network has three layers).

V. DEEPSORT VEHICLE TRACKING ALGORITHM

Deepsort tracking algorithm belongs to feature-based and model-based tracking, mainly used for pedestrian target tracking and detection. Due to its excellent performance in pedestrian detection, we use this method for vehicle tracking [27].

Deepsort tracking algorithm is mainly divided into two parts: Hungarian algorithm and Kalman filter. The Hungarian algorithm can recognize and detect the moving target in the current frame and judge whether the target is the same as the target in the previous frame. Kalman filter is mainly used to predict the target position. That is, it predicts the position of the current frame based on the position of the previous frame of the target and updates the target state in time through the tracker to lock the position of the target. Firstly, the algorithm uses the dense optical flow vehicle detection algorithm of mask SpyNet as the detector to detect the bbox coordinates of the current frame and then converts the detected bbox coordinates into feature detection matching target detection.

The Kalman filter is used to predict the specific position and state of the tracks in the previous frame in the current frame. The Mahala Nobis distance is used to calculate the cost matrix of Tracks and Detection, the Hungarian algorithm is used for IoU matching and cascade matching, all matching sequences between the current frame and the previous frame are found, and finally, the detection corresponding to each successfully matched track is updated through the Kalman filter. When the target is a new target, the unsaved track will match it. Then, a new track will be created for it, and matching authentication will be carried out to ensure that it is not false detection, and the tracks that are not matched successfully will be disposed of at the same time.

The main flow chart of the specific Deepsort algorithm architecture is as Fig 7.

VI. EXPERIMENTAL VERIFICATION

In order to verify the effectiveness of this algorithm, the night video images collected in the Lian-Xu section of the Beijing-Taiwan expressway are selected as the experimental data, and the video format is MP4. The algorithm is developed on Python 3.6 and Open CV 4.5.1. The computer is configured as Intel Core i5-9400F 2.90GHz, 16GB system memory, and 64-bit operating system. In order to increase the computing speed without affecting the detection effect, this paper selects one frame every five frames. In order to better carry out experimental comparison and analysis, this experiment detects 1000 vehicles in the video.

A. DETECTION RESULT

The focus of traffic incident detection is to detect the vehicle state. The vehicle state detection needs to track and judge the vehicle's speed. In case of dangerous speeding, illegal parking, and traffic accidents, these traffic events can be timely detected and continuously alarmed through speed analysis. In Fig 8-(a) and Fig 8-(b), the video data shows a two-lane expressway. Through the speed detection of the passenger car, when the speed exceeds the local specified Speed $>$120km/h, the algorithm can continuously alarm the traffic event and maintain the tracking state within the monitoring range. In Fig 8-(c), the tracked vehicle appears low speed in this video, and the vehicle stops illegally in the opposite lane, which is a severe traffic event. In Fig 8-(d), a severe traffic accident occurred between multiple vehicles in the video. During the experiment, the algorithm detects that the vehicle's speed at the lower left is lower than the threshold according to the restricted mask area and sends out alarm information. However, the algorithm can not effectively distinguish the vehicle type information and can not timely determine that the truck in the video is also in the alarm state, which is also the place that needs to be improved in the follow-up.

B. ALGORITHM EVALUATION INDEX

The performance evaluation of traffic event detection algorithm usually needs some indicators and methods. Generally,

(a) Vehicle overspeed detection (b) Vehicle overspeed detection

(c) Roadside illegal parking detection (d) Accident detection

FIGURE 8. Deepsort algorithm framework flow.

the algorithm will be evaluated from different aspects such as event detection effectiveness, detection efficiency, detection corresponding time, etc. In this paper, the detection rate R_T, false alarm rate R_F, false alarm rate R_A, and detection time T_h are selected as the criteria to evaluate the advantages and disadvantages of the algorithm. The detection rate R_T represents the ratio of the number of vehicles detected by the algorithm to all vehicles in the video. The formula is as follows:

$$R_T = \frac{TP}{TP + FA + FN} \tag{9}$$

where: TP——Number of vehicles detected correctly;

FN——Number of vehicles not detected;

FA——Number of false detected vehicles.

The number of undetected vehicles plus the correct number of detected vehicles is the total number of vehicles in the video. The false-negative rate R_F is the ratio of the number of vehicles not detected in the video to the number of all vehicles appearing in the video.

$$R_F = \frac{FN}{TP + FA + FN} \tag{10}$$

False alarm rate refers to the target number of vehicles incorrectly detected and the number of all vehicles detected. The number of all vehicles includes but is not limited to all vehicles in the video. The formula expresses the false alarm rate:

$$R_A = \frac{FA}{TP + FA + FN} \tag{11}$$

The length of detection time affects the efficiency of response processing for traffic event detection in detecting traffic events. In this experiment, the detection time T_h is used as the time-consuming evaluation index of the evaluation algorithm.

C. COMPARATIVE ANALYSIS OF EXPERIMENTAL RESULTS

In order to illustrate the adaptability of this algorithm to the detection of high-speed vehicles at night, this paper uses the background subtraction method, traditional optical flow method, and YOLO-v3 algorithm to compare with this algorithm for a total of 1000 vehicles. Because the vehicle detection results of the background subtraction method and traditional optical flow method are not ideal, in the comparison of tracking algorithms, the YOLO-v3+ Deepsort tracking algorithm is selected to compare with the method in this paper.

TABLE 1. Average detection time of four algorithms.

Detection algorithm	*Background subtraction method*	*Tradition al optical flow method*	*Yolo-v3 algorithm*	*Ppropose d algorith*
Computing time /ms	*16.6*	*33.7*	*46.4*	*55.6*

TABLE 2. Comparison of results of four algorithms.

Detection algorithm	*Background subtraction method*	*Traditional optical flow method*	*Yolo-v3 algorithm*	*Proposed algorithm*
TP	*117*	*329*	*823*	*931*
FN	*828*	*522*	*71*	*20*
FA	*55*	*149*	*106*	*49*
R_T	*11.7%*	*32.9%*	*82.3%*	*93.1%*
R_F	*82.8%*	*52.2%*	*7.1%*	*2%*
R_A	*5.5%*	*14.9%*	*10.6%*	*4.9%*

1) THE DETECTION ALGORITHM IS TIME-CONSUMING

In the vehicle detection method, for the background difference method and the traditional optical flow method, the video image is preprocessed to increase the brightness and contrast of the image, and then the vehicle detection is carried out. Table 1 shows the average time-consuming of single-frame images of the four detection methods.

Compared with the first two algorithms, the implementation of the algorithm in this paper calculates the dense optical flow combined with the deep learning network, so it takes a long time. At the same time, the detection time depends on different GPU running speeds. In this paper, the video collected by the camera is 20 frames per second. Therefore, when the detection is accurate, the algorithm in this paper can meet the real-time requirements of follow-up tracking.

2) DETECTION ACCURACY

The detection accuracy of the algorithm is an essential index of experimental analysis. The analysis result evaluates the four algorithms according to the evaluation criteria in Chapter 5.2, where TP + FA + FN = 1000. The detection rate R_T, false alarm rate R_F and false alarm rate R_A of the four algorithms are shown in table 2.

D. COMPARISON OF RESULTS

Combined with the detection results in table 2 and Fig 9, the detection rate of the algorithm in this paper is significantly higher than that based on the background difference method and traditional optical flow method, and the false alarm rate and false alarm rate are reduced. Compared with Yolo-v3

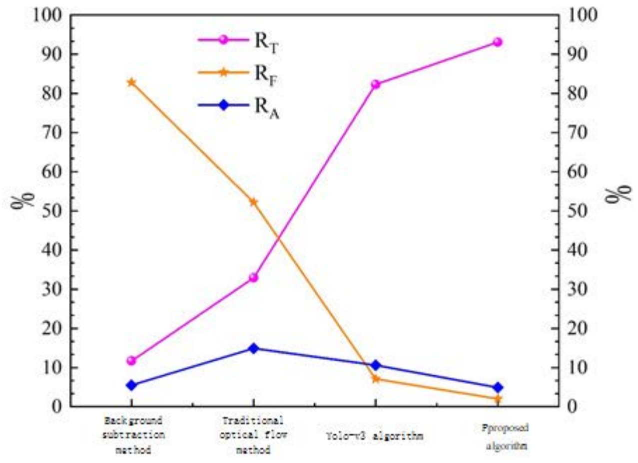

FIGURE 9. Comparison of results of four methods.

FIGURE 10. Frame 310 detection result.

FIGURE 11. Frame 17 detection result.

FIGURE 12. Frame 772 detection result.

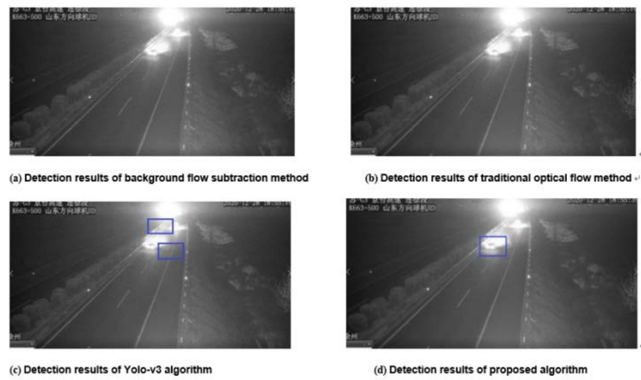

FIGURE 13. Comparison diagram of test results.

algorithm, the detection rate increases by 10%-11%, the false alarm rate decreases by 5%-6%, and the false alarm rate decreases by 5%-6%. Generally speaking, the light flow is blocked by each other due to the close parallel of vehicles, and the intense light interference of opposite lamps will lead to a higher false detection rate and missed detection rate than the available video. However, the accident video proves that the detection of this algorithm has a better detection effect than other similar detection algorithms in the event of traffic events.

Firstly, the representative images of frame 17, frame 310, and frame 772 are selected for result analysis. The detection result of frame 17 is shown in Fig 10. When the vehicle target in the image is driving into the ROI area, the Mask-SpyNet algorithm can extract the dense optical flow diagram of the front and rear lamps of the vehicle in real-time and then binarise the optical flow. After establishing the connected area, the vehicle target is successfully detected. The detection result of frame 310 is shown in Fig 11. The detection results show that when the opposite vehicle is close, the extraction of dense optical flow can effectively eliminate the interference of the opposite optical flow and distinguish it by colour, which will not affect the optical flow calculation of the target vehicle in the ROI area. The 772nd frame detection result is shown in Fig 12. When driving two target vehicles in the same direction, with the increase of dense optical flow, when the boundary threshold of dense optical flow is exceeded, it will be determined as two-vehicle targets with solid practicability.

Secondly, according to the detection results and other methods, the detection accuracy is compared and analyzed. Under the condition that no strong light interferes with the normal driving of the vehicle, both the background difference method and the traditional optical flow method can complete the detection. Faced with the interference of the strong light of the opposite lights, the background difference method and the traditional optical flow method cannot complete the detection of the target vehicle, as shown in Figure 13-(a), (b). The YOLO-v3 algorithm based on deep learning, through the training of data sets, can perform vehicle target recognition in a few cases, but the effect is not ideal. When the light of the opposite vehicle produces strong light interference, the algorithm cannot detect the target vehicle stably. As shown in Figure 13-(c). This paper extracts and calculates the traffic light flow, which is the largest characteristic factor when the vehicle is driving at night. It has strong robustness to other light disturbances in the current environment, and can effectively complete the detection of the target vehicle, as shown in Figure 13-(d). Show.

As shown in Figure 14, it is a comparison diagram of two tracking display results. Fig 14-(a) shows the tracking results with Yolo-v3 as the detection algorithm. Although the YOLO-v3 algorithm has good adaptability in vehicle detection, when combined with the deep sort tracking algorithm, there is a significant error in extracting the pixel displacement in the frame by frame image per unit time in the face of vehicles running at night. Therefore, for the dark environment, speed calculation has high accuracy. In Fig 14-(b), according to the previous analysis, the dense optical flow algorithm

(a) Yolo-v3 + Deepsort tracking results (b) Proposed algorithm tracking results

FIGURE 14. Comparison of two tracking display results.

FIGURE 15. Multi objective yolo-v3 + deepsort algorithm.

FIGURE 16. Single target detection of proposed algorithm.

FIGURE 17. Multi target detection of proposed algorithm.

based on Mask-SpyNet combined with Deepsort algorithm can analyse the optical flow of the lamp, calculate the size of the optical flow, and effectively calculate the displacement distance of pixels under the dense optical flow, to have high accuracy in speed detection.

Accuracy in parallel. There will inevitably be similar vehicles driving on the Expressway, especially in the same direction of multiple lanes. In the parallel state, the detection and tracking of target vehicles can be effectively distinguished, and accurate recognition has become the key of many multi-target tracking algorithms. As shown in Fig 15, in Fig 15, the YOLO-V3 + Deepsort algorithm does not have good adaptability at night. Due to the poor definition at night and the light interference of adjacent vehicles and opposite vehicles, the algorithm is more difficult to track experiments. This paper's algorithm is suitable for single target detection, as shown in Fig 16. Moreover, by using the optical flow extraction of light, the algorithm in this paper can judge whether it is the optical flow formed by multi-target vehicles at the same time according to the optical domain of dense optical flow when vehicles are parallel, to determine the multi-target vehicle and display the speed in combination with the optical flow size between consecutive frames, as shown in Fig 17.

VII. CONCLUSION

The traditional vehicle detection algorithm is suitable for good light conditions in the daytime and a relatively uncomplicated background environment. It can not play its efficient and straightforward characteristics in the night vehicle detection environment. The commonly used deep learning algorithm can detect and identify vehicles in the night environment through the training of data sets. However, it can not effectively detect and identify the common situation of night driving, such as opposite lamp interference, which affects the follow-up speed tracking and traffic event detection. This study uses the dense optical flow vehicle detection algorithm based on Mask-SpyNet, which improves vehicle detection accuracy at night and eliminates the light interference of opposite lamps. The Deepsort algorithm can detect traffic events at night well and has broad application prospects in the current low intelligent expressway environment.

Although this study relies on the monitoring equipment widely used in Expressway, with a high detection rate, low cost, and simple installation, there are still deficiencies in hardware facilities and detection technology. In the next step, the application of this study can be improved based on updating the expressway hardware equipment, such as the choice of the best photography angle of the target vehicle, adapting to the identification of continuous vehicle following under high density and the detection of traffic events in bad weather at night.

REFERENCES

[1] Li, "Vehicle detection on urban roads at night based on frame difference method," *Manuf. Automat.*, vol. 33, no. 4, pp. 212–214, 2011.

[2] Wang, "A new moving vehicle detection and tracking algorithm based on the combination of inter-frame difference and optical flow technology," *Comput. Appl. Softw.*, vol. 29, no. 5, pp. 117–120, 2012.

[3] G. Bo, H. Hao, R. Yan, and L. Xinyu, "Moving target detection and tracking in complex background," *Int. J. Smart Home*, vol. 9, no. 9, pp. 95–102, Sep. 2015.

[4] N. A. Prasath, G. Sivakumar, and N. M. Kumaresan, "Vehicle speed measurement and number plate detection using real time embedded system," *Netw. Complex Syst.*, vol. 5, no. 3, pp. 21–29, 2015.

[5] R. O'Malley, E. Jones, and M. Glavin, "Vision-based detection and tracking of vehicles to the rear with perspective correction in low-light conditions," *IET Intell. Transp. Syst.*, vol. 5, no. 1, pp. 1–10, Mar. 2011.

[6] Y.-C. Kuo and H.-W. Chen, "Vision-based vehicle detection in the nighttime," in *Proc. Int. Symp. Comput., Commun., Control Autom. (CA)*, May 2010, pp. 361–364.

[7] Du, "Night vehicle recognition method based on monocular vision," *Comput. Eng. Appl.*, vol. 50, no. 17, pp. 160–163, 2014.

[8] H. Juehui, "Improved optical flow method for vehicle recognition and tracking," *Sci. Technol. Eng.*, vol. 23, pp. 5814–5817, Oct. 2010.

[9] O. M. Sincan, V. B. Ajabshir, H. Y. Keles, and S. Tosun, "Moving object detection by a mounted moving camera," in *Proc. Int. Conf. Comput. as a Tool (EUROCON)*, Sep. 2015, pp. 1–6.

[10] J. Choi, "Realtime on-road vehicle detection with optical flows and Haar-like feature detectors," Tech. Rep., 2012.

[11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.

[12] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis.* Springer, 2016, pp. 21–37.
[13] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 580–587.
[14] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1440–1448.
[15] S. Ren *et al.*, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2015, pp. 91–99.
[16] J. Dai *et al.*, "R-FCN: Object detection via region-based fully convolutional networks," in *Proc. Adv. Neural Inf. Process. Syst.*, pp. 379–387.
[17] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 936–944.
[18] A. Dosovitskiy, P. Fischer, E. Ilg, P. Hausser, C. Hazirbas, V. Golkov, P. V. D. Smagt, D. Cremers, and T. Brox, "FlowNet: Learning optical flow with convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 2758–2766.
[19] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox, "FlowNet 2.0: Evolution of optical flow estimation with deep networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1647–1655.
[20] D. Sun, X. Yang, M.-Y. Liu, and J. Kautz, "PWC-Net: CNNs for optical flow using pyramid, warping, and cost volume," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8934–8943.
[21] J. Hur and S. Roth, "Iterative residual refinement for joint optical flow and occlusion estimation," 2019, *arXiv:1904.05290*.
[22] P. Liu, M. Lyu, I. King, and J. Xu, "SelFlow: Self-supervised learning of optical flow," 2019, *arXiv:1904.09117*.
[23] Q. Chen and W. Wang, "Multi-modal neural network for traffic event detection," in *Proc. IEEE 2nd Int. Conf. Electron. Commun. Eng. (ICECE)*, Dec. 2019, pp. 26–30.
[24] J. A. Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *IEEE Trans. Image Process.*, vol. 9, no. 5, pp. 889–896, May 2000, doi: 10.1109/83.841534.
[25] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1956–1963.
[26] X. Mo, C. Sun, D. Li, S. Huang, and T. Hu, "Research on the method of determining highway truck load limit based on image processing," *IEEE Access*, vol. 8, pp. 205477–205486, 2020.
[27] L. Jingyu and W. Yanjie, "Research on target tracking algorithms based on subspace," *Liquid Crystal Display Display*, vol. 29, no. 4, pp. 617–622, 2014.

XIANGLUN MO received the B.S. degree in traffic engineering from the Kunming University of Science and Technology, Kunming, Yunnan, in 2009, and the Ph.D. degree in traffic and transportation engineering from Jilin University, Changchun, Jilin, in 2014. Since 2014, he has been a Lecturer with the Transportation Engineering Department, China University of Mining and Technology. His research interests include analyzing the traffic systems, traffic information engineering, control, detection methods, and equipment, and open-pit mining systems.

CHUANPENG SUN received the B.S. degree in transportation from the China University of Mining and Technology, Xuzhou, Jiangsu, China, in 2019, where he is currently pursuing the master's degree, with a focus on the development of intelligent transportation and urban transportation planning and management.

CHENYU ZHANG received the bachelor's degree in transportation from the China University of Mining and Technology, in 2018, where she is currently pursuing the master's degree, with a focus on the development of construction robot and improvement of algorithms for tourist robot.

JINPENG TIAN received the B.S. degree in transportation engineering from the China University of Mining and Technology, Xuzhou, Jiangsu, China, in 2020, where he is currently pursuing the master's degree, with a focus on the application of intelligent transportation and small target detection on expressway pavement.

ZHUSHUAI SHAO received the B.S. degree in vehicle engineering from the Nanjing Institute of Technology, Nanjing, Jiangsu, in 2021. He is currently pursuing the master's degree with the Transportation Engineering Department, China University of Mining and Technology. His research interests include intelligent transportation technology and route optimization in interweaving areas.