JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

OBJECT DETECTION AND FEATURES EXTRACTION IN VIDEO FRAMES USING DEEP LEARNING

[1]Mrs.M.Angel Shalini,[2]Ms.V.Vivitha

[1]Assisant Professor,MCA,(Ph.D),[2]Final MCA Student

[1]PG & Research Department Of Computer Applications(MCA),

[1]Hindusthan College of Arts and Science (Autonomous), Coimbatore, India

Abstract: Due to the importance of real-time applications such as video surveillance systems, person identification, automated driver assistance, and automotive safety, pedestrian detection is a rapidly growing research area in computer vision. Because pedestrians are vulnerable in heavy traffic, especially in urban areas, pedestrian detection is critical for road safety. In various scenarios for pedestrian detection, the existing techniques of Deformable Part Model, extended deep model, RealBoost method, and Deep Neural Networks have been used. This study presents a method for detecting objects and extracting features from static video imagery that uses color/gray-scale frames captured by common digital cameras or images readily available from external sources. A deep leaning technique is used to segment objects. The supporting hardware is a dedicated P4 PC-based image processing environment, while the intrinsic software was simulated and validated on the MATLAB 2013a platform. We develop a deep leaning algorithm to detach the moving object from the background in this study. The deep leaning algorithm is very efficient and robust, as demonstrated by our experiments. This method could be used in a variety of fields, including vehicle and pedestrian traffic flow measurement, athletic and dancing performance evaluation, public, private, and military security, and so on.

Keywords: Deep Learning algorithm, pedestrian detection, classification, video processing, MATLAB platform.

I. Introduction

To provide significant improvements in regulating road traffic in urban scenes, image processing is used to analyse instances of traffic flow in the form of images or videos. sivanantham kalimuthu and colleagues

(2021). Image processing can be used to visualise traffic flows, vehicle speed measurements, multiple-point vehicle counts, and vehicle categorization in the field of road traffic monitoring. Image processing has a wide range of applications in autonomous vehicle guidance, such as estimating the vehicle's relative location and detecting obstacles.

Object detection, as defined by Seungwon Lee et al., is the process of identifying instances of real-world objects in an image (2015). Object detection is critical in a variety of computer vision tasks, including video surveillance, person recognition, and behavioural analysis. Human detection accuracy in visual surveillance systems is critical for a variety of applications. Pedestrian detection is a classic example of object detection that has a wide range of applications, including surveillance, autonomous vehicle systems, and so on. Detecting pedestrians in road scenes is becoming increasingly important as the number of pedestrian fatalities continues to rise. Pedestrian detection usually necessitates the examination of multiple resolutions of input images. This paves the way for wavelets to be used in the detection of pedestrians. Wavelets are used to recognise objects and extract features. Wavelets isolate the two singularities caused by edges in noisy data. As a result, pedestrian detection performance improves. Deep learning neural networks with more than three or four hidden layers are referred to as deep models. Sivanantham kalimuthu uses deep neural networks to solve localization issues and achieve high pedestrian detection accuracy (2021).

Several occlusions are present in real-world traffic conditions, making it difficult to detect pedestrians from input images. When there are other objects in the image, accurate pedestrian identification becomes critical for road traffic regulation. The process of detecting pedestrians is divided into two categories: Offline detection process• Real-time detection process• The data input is in the form of video or images acquired from a separate input device in the offline detection process. The image frames from the video are extracted directly through input devices during the real-time detection process. After that, there is a pre-processing step for 3 noise removal and resizing. In order to detect pedestrians, significant features must be extracted from the background or other objects. In the case of real-time pedestrian detection, David Vazquez et al. have identified a number of challenges that should be considered in order to improve detection speed (2014). Furthermore, the increased number of vehicles complicates the management of traffic issues in large cities. To manage traffic flows in various situations, various traffic control procedures are used. These methods, on the other hand, necessitate manual interpretation and are difficult to implement. As a result, automatic pedestrian detection systems in vehicles and traffic signals are required to avoid collisions and regulate traffic.

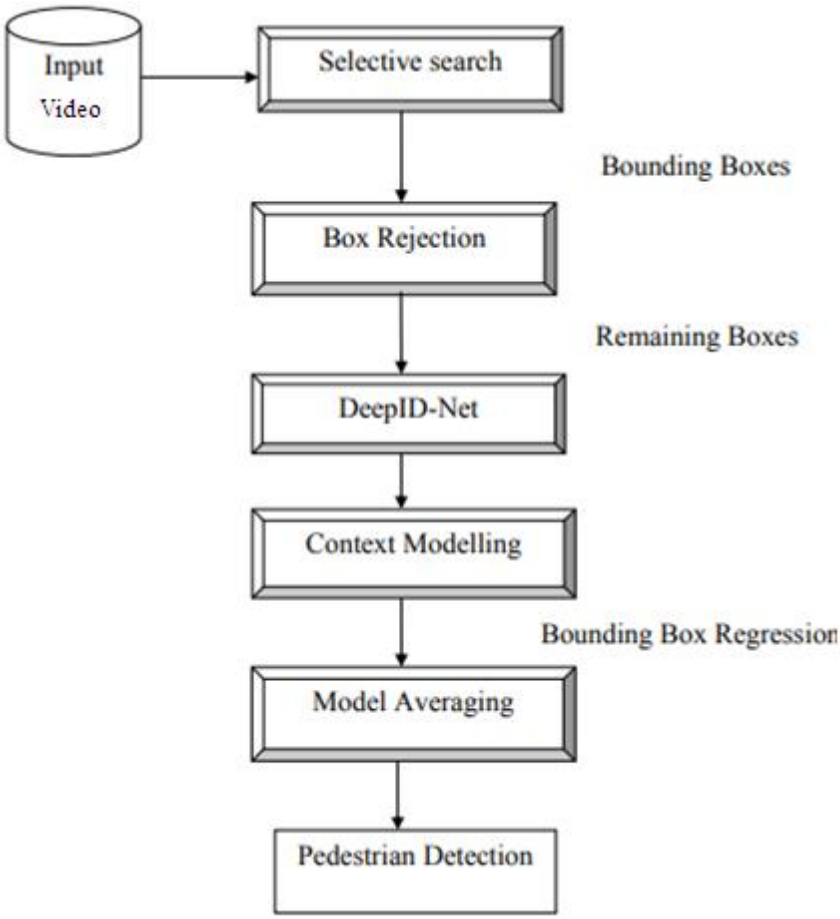

Figure 1 Flow Diagram of Pedestrian Detection

The remainder of this paper is organised in the following manner. The second section gives a quick overview of in-cloud approach control conservancy. Section III delves into the methodology of the proposed approach control techniques. Section IV displays the experimental results for evaluating and comparing performance. The proposed system's conclusion is found in Section V.

II. Literature review

Image processing techniques are increasingly being used in real-time object detection applications as computer technology advances. Object detection is an important and difficult task in computer vision. Pedestrian detection is used in a variety of applications, including car safety systems, video surveillance, robotics, and self-driving vehicles. Pedestrian detection has been studied by a number of researchers. Windowing, feature extraction, and classification are all performed by pedestrian detection techniques. The road traffic control system in urban areas is ineffective due to flaws in existing pedestrian detection techniques. The following sections summarise the existing literature on pedestrian detection techniques.

Zhihui Wang et al. (2014) presented an accurate pedestrian detection system based on machine learning techniques such as the cascade AdaBoost detector and random vector functional-link net. The cascade AdaBoost detector and the random vector functional-link net were trained offline. Sivanantham Kalimuthu et al. (2022) proposed a new method for calculating the pedestrian count based on multisource video data. To determine the number of pedestrians from single-source video, the partial least squares regression (PLSR)

model was introduced. On the basis of visible light video, the temporal feature of daytime or nighttime was recognised. With pre-set equivalent confidence levels, pedestrian count detection results from visible light and infrared video data were obtained. Ying Guo et al. (2019) developed a principal component analysis algorithm to reduce the dimension of extracted multidimensional gait features. To distinguish the phone-carrying pose for pedestrian detection, the extracted features were random forest modelled. Yang Zhang et al. (2016) presented an improved and efficient pedestrian detection method for auto driver assistance systems. To improve the discriminating power and minimise the illumination effects, an improved Accumulate Binary Haar (ABH) feature extraction algorithm was introduced. The improved Deep Belief Network (DBN) classification algorithm was used to develop a pedestrian classification method. The information on pedestrian features was adapted to Bernoulli distribution for recognition using a Restricted Bolzmann Machine (RBM) with T distribution function visible layer nodes. To address the class-imbalance issues, a cost-sensitive Support Vector Machine (SVM) 27 classifier was used. Song Tang et al. (2019) used a classifier regression model in the source domain, where each sample has its own proprietary classifier. For each candidate window in the image, a pedestrian classifier was predicted. The source samples were not kept when a new adaptive detector was introduced. To improve the reconstruction and classification performance, a new dimensionality reduction method for classifier vector was introduced. Rajkumar Soundrapandiyan and Chandra Mouli (2018) proposed an adaptive pedestrian detection method based on human visual mechanisms and support vector machines. To reduce background noise and improve SNR, a mean and Laplacian of Gaussian (LoG) filter was used. To reduce noise, a morphological process was applied to the filtered image. Using the local thresholding segmentation process, pedestrians and non-pedestrians were joined. Masoud Afrakhteh and Park Miryong (2017) developed an image categorization method to reduce the false positive rate. As a base detector, an aggregate channel features method was investigated and recommended. With night images, a pre-trained pedestrian detector was used, and with daytime images, a daytime detector was used.

Despite occlusion and prior knowledge of objects, a new framework identified and tracked moving objects. Sukanyathara & Alphonsa Kuriakose's robust threshold decision algorithm through multi-background model was used in the segmentation step (2014). Multi-target tracking was combined with multi-background registration in an effective formulation. Victor Vaquero et al. (2015) introduced a computationally low-cost and robust Detection and Tracking Moving Objects (DATMO) algorithm for autonomous guided vehicles and autonomous trucks for cargo transportation. The algorithm tracked moving objects in port terminals after identifying them. The DATMO system was put to the test on two different platforms. The first was designed to detect moving obstacles and included tracking and detection filtering. The second was designed to keep the targets alive when there was no detection. During the object detection process, however, the number of incorrect identifications did not go down.

lll. System design

Softsign Gaussian Recurrent Deep Neural Network (SGRDNN) is a proposed research project that aims to identify pedestrians in images with fewer time measures. The Convolutional Neural Network (CNN) is a feed forward artificial neural network that processes only the current input state. The Recurrent Neural Network (RNN) differs from a feedforward neural network in that it has feedback connections that feed back previous states as input to the current state of the network. Recurrent Neural Network, also known as Recurrent Deep Neural Network, is used to detect pedestrians within a deep learning framework (RDNN). The proposed technique uses the integration of deep convolutional and recurrent neural network learning to create a directed cycle by making connections between neurons. The output of the hidden layer in the proposed technique is obtained not only based on the current inputs, but also on the output of the previous neuron state for extracting relevant features. The SGRDNN technique uses a standard multilayer perceptron structure, which consists of connections between hidden units that are time delayed. The information about previous outputs is well learned thanks to these connections. This aids in the identification of significant features by estimating temporal correlations between features in the image that are far apart. In comparison to other neural network techniques for recognising pedestrians in images, the proposed SGRDNN technique effectively handles large image datasets with higher accuracy and less time consumption. Without having to manually extract features from the image, the features from the image are directly analysed. As a result, the amount of time it takes to detect pedestrians is cut in half. The proposed technique is distinguished by the fact that 73 relevant features are not pre-trained. The proposed technique discovers the features of pedestrian objects when the network is trained on a set of images. For detecting pedestrians in images, the proposed SGRDNN technique uses two essential functions: Gaussian activation function and softsign activation function. The relevant features are extracted by estimating the connection between features on images using the Gaussian activation function hidden layer. The Softsign Activation Function matches extracted relevant features with pre-stored templates to classify objects in images as pedestrian or background. The flow diagram of the Softsign Gaussian Recurrent Deep Neural Network for pedestrian detection is shown in Figure 2. As shown in Figure 2, the proposed technique implements the pedestrian detection process using different layers such as input, hidden, and output. The process begins with the input layer receiving the number of images. The input layer sends the received images to the hidden layer, which extracts the features. Using the Gaussian activation function in the hidden layer, only significant (i.e. relevant) features are extracted during the feature extraction process. The determination of the connection between features on images is used to extract features. This aids in the detection of pedestrians in images by using the Gaussian activation function to learn images repeatedly. The extracted features are then transmitted to the output layer. The softsign activation function is used in the output layer to find pedestrians on input images using features extracted from the hidden layer. This is accomplished by comparing extracted features to previously stored templates.

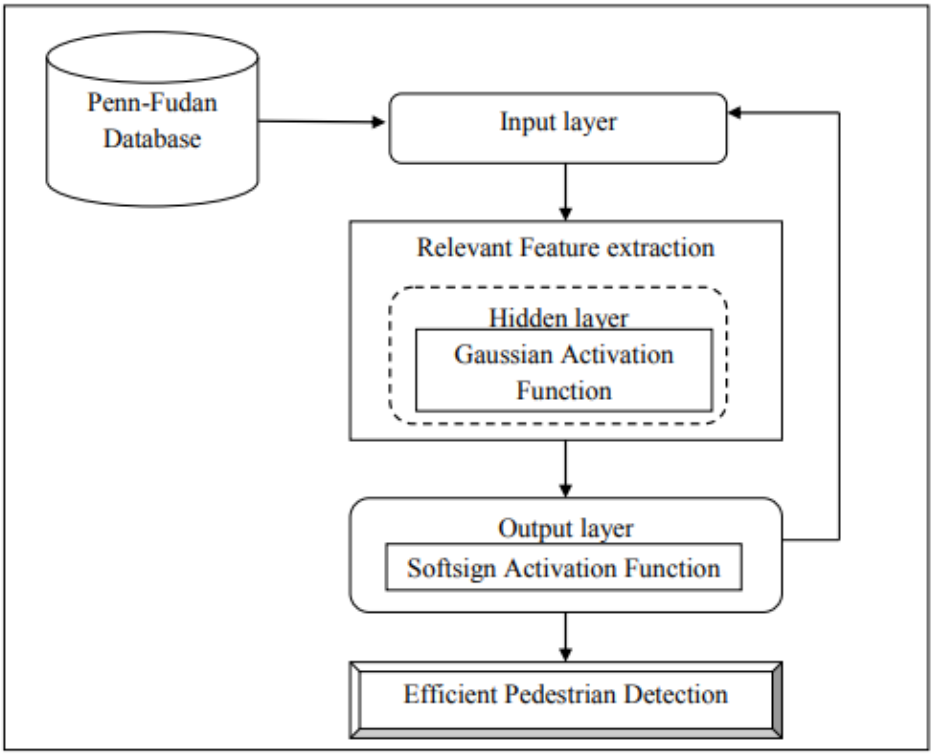

Figure 2 Flow Diagram of Recurrent Deep Neural Network technique

At each node of the neural network, the proposed SGRDNN technique starts with random weights. Each input image is sent from the input layer to the hidden layer first. Second, the hidden layer uses recurrent behaviour and a Gaussian activation function to learn the various features on images in depth. The relationship between the features is determined in the hidden layer using the Gaussian activation function. As a result, the most important features for performing pedestrian detection are extracted in an efficient manner. When the hidden layer result is 1, the extracted 84 feature is said to be important for identifying pedestrians on images. In order to perform pedestrian detection, the extracted feature is said to be irrelevant if it is not used. As a result, the hidden layer transmits the image's extracted relevant features to the output layer. The output layer then compares the extracted features to pre-stored templates in order to recognise pedestrians on images using the softsign activation function. The above steps are repeated until the network error is low enough to accurately recognise pedestrians on images in a short amount of time. As a result, the proposed SGRDNN technique improves road traffic control performance by recognising pedestrians on images in less time.

III. Result and discussion

The proposed method is implemented in MATLAB with the help of a database. Figure 3 shows some of the sample images used in the experiments, which were taken from a video database. The first module uses segmentation to find possible detections. Everything in this module is our own work, with the exception of the edge detection function (for which we used the MATLAB version), such as the structure, image dilation, splitting component, non-maximum suppression, image resizing, and marking. The second module is HoG feature extraction. We wrote our own version of HoG feature extractor, but it only gave us 70% accuracy in

the experiment, so we decided to use the MATLAB built-in function instead. The classification module is the final module in which we experiment with different neuron network structures and window sizes to train the network. We use It's used to classify images and output the target image once the most appropriate one has been found. We created the NN from scratch, with only one function. The classifier is our system's final module. We are eager to use neural networks now that we have recognised their potential. However, due to the dataset's size limit, we must also limit the size of our network. We chose a simple neural network with one linear layer for our project. The gui design for pedestrian detection is shown in Figure 3. Figures 4 and 5 show how to extract the input video to the corresponding frame separation. This frame separation is 16 frames per second.

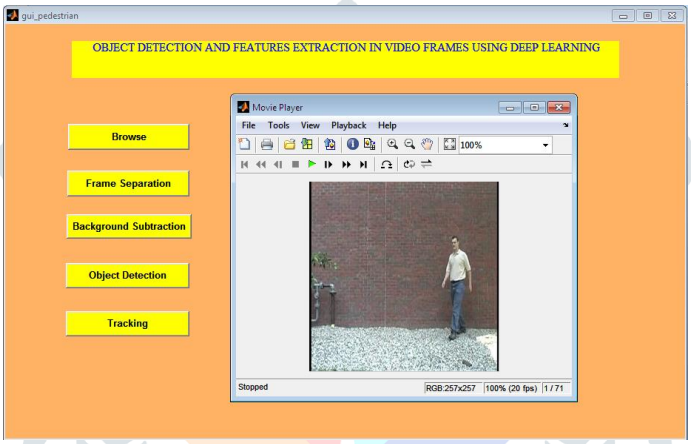

Figure 3 input video data selection

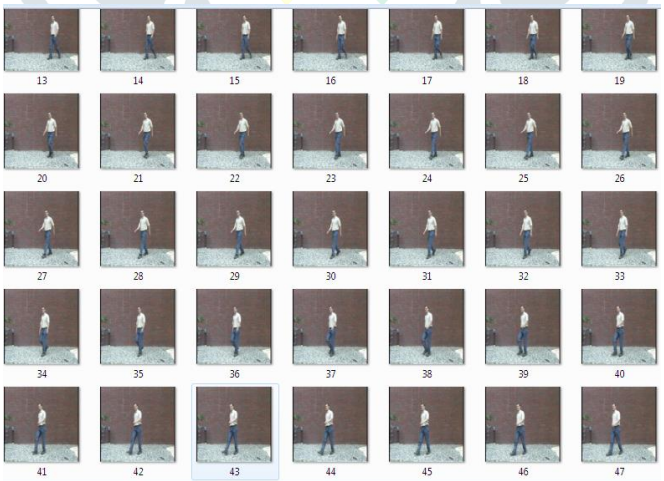

Figure 4 frame separation

The detection object navigation in Figure 5 is explained in this section, which proposes different window sizes in the hopes of finding the one that best captures the common appearance of pedestrians. We get the HoG features of the processed image for each window size after resizing the training images. Then, using the HoG features as inputs, we build a neuron network that outputs the label 2 if it's a pedestrian and 1

otherwise. We can get different lowest cross-entropy losses for different window sizes, so we need to set different thresholds for stopping the training process. In general, a larger window size yields a larger HoG feature vector and lower cross-entropy loss, but due to the small number of training data we have, a larger window size is more likely to result in an overfitted model. For different window sizes, see the results below.

Figure 5 object navigation

Figure 6 depicts the final navigation and object identification, as well as how object tracking is accomplished. The first experiment we conducted was to keep using our example image. We did not scale the image in this experiment so that the detection is on a fixed scale. We only use a window size of 74*34 for detection. We tested our system using that image, as shown above, and the resulting image is as follows, with the rectangle denoting a detected pedestrian.

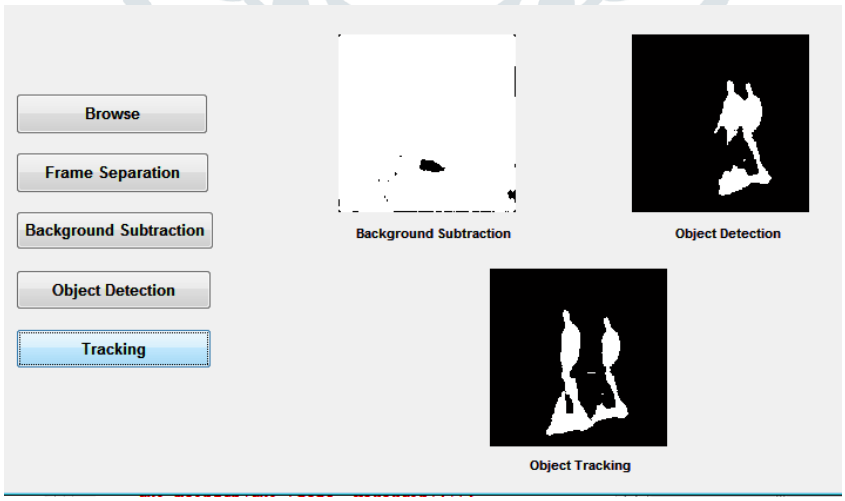

Figure 6 pedestrian detection

We introduced image scaling in the second set of experiments. We scale the image to have a height of 1, 4/3, 5/3, 2, 7/3, 8/3, 3 times the window height, and use detect pedestrian for all of these images. This way, we can almost find all pedestrian in the image, even if their sizes are different in the original image. We can see in the following diagrams that two of them are capable of holding all of the pedestrians.

IV. Conclusion

The goal of this study is to regulate road traffic by detecting pedestrians in input images of traffic scenes. Pedestrian detection is successfully carried out with the implementation of proposed techniques - Softsign Gaussian Recurrent Deep Neural Network (SGRDNN) technique, as shown in previous chapters. Pedestrians are detected more accurately, with less time and memory requirements, using the proposed techniques. The proposed Softsign Gaussian Recurrent Deep Neural Network (SGRDNN) technique is a Recurrent Neural Network (RNN) for pedestrian detection that is implemented 154 with a deep learning framework. The SGRDNN process begins with input images from the database provided to the input layer. The images are sent from the input layer to the hidden layer, where the significant (i.e. relevant) features are extracted by using a Gaussian activation function to estimate the connections between features on the images. The extracted features are sent to the output layer, which uses the softsign activation function to match them to pre-stored templates. As a result, pedestrians in images can be accurately identified in less time and with less memory.

References

1. Kalimuthu, S., Naït-Abdesselam, F., & Jaishankar, B. (2021). Multimedia Data Protection Using Hybridized Crystal Payload Algorithm With Chicken Swarm Optimization. In *Multidisciplinary Approach to Modern Digital Steganography* (pp. 235-257). IGI Global.
2. Kalimuthu, S. (2021). Sentiment Analysis on Social Media for Emotional Prediction During COVID-19 Pandemic Using Efficient Machine Learning Approach. *Computational Intelligence and Healthcare Informatics*, 215.
3. Sivanantham, K., Kalaiarasi, I., & Leena, B. (2022). Brain Tumor Classification Using Hybrid Artificial Neural Network with Chicken Swarm Optimization Algorithm in Digital Image Processing Application. *Advance Concepts of Image Processing and Pattern Recognition: Effective Solution for Global Challenges*, 91.
4. Seungwon Lee, Nahyun Kim, KyungwonJeong, Kyungju Park & Joonki Paik 2015, 'Moving Object Detection Using Unstable Camera for Video Surveillance Systems', Optik - International Journal for Light and Electron Optics, Elsevier, vol. 126, issue no. 20, pp. 2436 - 2441.
5. David Vazquez, Antonio M. Lopez, Javier Marın, Daniel Ponsa & David Geronimo 2014, 'Virtual and Real World Adaptation for Pedestrian Detection', IEEE Transactions On Pattern Analysis And Machine Intelligence, vol. 36, issue. 4,pp. 797-809.
6. Zhihui Wang, Sook Yoon, Shan Juan Xie, Yu Lu & Dong Sun Park 2014, 'A High Accuracy Pedestrian Detection System Combining a Cascade AdaBoost Detector and Random Vector Functional-Link Net', The Scientific World Journal, Hindawi Publishing Corporation, vol. 2014, pp. 1-7.

7. Ying Guo, Qinghua Liu, Xianlei Ji, Shengli Wang, Mingyang Feng & Yuxi Sun 2019, 'Multimode Pedestrian Dead Reckoning Gait Detection Algorithm Based on Identification of Pedestrian Phone Carrying Position', Mobile Information Systems, Hindawi Publishing Corporation, vol. 2019, pp. 1-14.
8. Yang Zhang, Dong-Rong Xin & Yi-Hu Wu, 'Pedestrian detection for traffic safety based on Accumulate Binary Haar features and improved deep belief network algorithm', Journal of Transportation Planning and Technology, vol. 39, issue. 8, pp. 791-800.
9. Song Tang, Mao Ye, Pei Xu & Xudong Li 2019, 'Adaptive pedestrian detection by predicting classifier', Neural Computing and Applications, Springer, vol. 31, issue 4, pp. 1189–1200
10. Rajkumar Soundrapandiyan & Chandra Mouli, PVSSR 2018, 'An Approach to Adaptive Pedestrian Detection and Classification in Infrared Images Based on Human Visual Mechanism and Support Vector Machine', Arabian Journal for Science and Engineering, Springer, vol. 43, issue.8, pp. 3951–3963
11. Masoud Afrakhteh & Park Miryong 2017, 'Pedestrian Detection with Minimal False Positives per Color-Thermal Image', Arabian Journal for Science and Engineering, Springer, vol. 42, issue. 8, pp. 3207–3219.
12. Sukanyathara, J & Alphonsa Kuriakose 2014, 'Detection and Tracking of Multiple Objects in Cluttered Backgrounds With Occlusion Handling', Computer Science & Information Technology (CS & IT), pp. 41–48.
13. Victor Vaquero, Ely Repiso, Alberto Sanfeliu, John Vissers & Maurice Kwakkernaat 2015, 'Low Cost, Robust and Real Time System for Detecting and Tracking Moving Objects to Automate Cargo Handling in Port Terminals', Robot 2015: Second Iberian Robotics Conference, pp. 491-502.