Juan Alberto Antonio., et. al. International Journal of Engineering Research and Applications
www.ijera.com
ISSN: 2248-9622, Vol. 13, Issue 11, November 2023, pp 227-235

RESEARCH ARTICLE OPEN ACCESS

Video dataset analysis of pedestrians for their detection using the MDBS frame difference technique

Juan Alberto Antonio Velázquez*, Teresa Plata Hernández, Leopoldo Gil Antonio*, Adriana, Reyes Nava *, Merced Leodegario Urbina Díaz**

** Tecnológico Nacional de México/Tecnológico de Estudios Superiores de Jocotitlán, carretera Toluca-Atlacomulco km. 44.8, Ejido San Juan y San Agustín, Jocotitlán Edo. De México*
Email: juan.antonio@tesjo.edu.mx, teresa.plata@tesjo.edu.mx, leopoldo.gil@tesjo.edu.mx, adriana.reyes@tesjo.edu.mx
Centro Universitario UAEMex Atlacomulco, Universidad Autónoma del Estado de México – UAEMex Autopista Toluca-Atlacomulco km 60, C.P. 50400, Atlacomulco, Estado de México

Date of Submission: 12-11-2023 Date of acceptance: 28-11-2023

ABSTRACT

Nowadays video surveillance systems geared towards pedestrian detection has attracted a great deal of interest, leading to an increase in research on pedestrian datasets over the past decades. However, most existing datasets focus on images that are not specifically optimized for motion detection. To address this limitation, in this work is proposed a dataset that includes video recordings of pedestrians inside university facilities of the Autonomous University of the State of Mexico. It contains features that impede pedestrian motion detection. Here is considered the fact that pedestrian's transit through closed and open spaces, added to natural or artificial lighting, the type of background, occlusion and even the type of clothing. Therefore, pedestrian detection of these proposed scenes were performed with a technique called Motion Detection with Background Subtraction (MDBS), which shows promising results.

Keywords – Dataset, pedestrian's detection, frame difference, motion detection.

I. INTRODUCTION

Surveillance systems are an attractive field in computer vision and traffic monitoring systems [1]. Currently, there are several surveillance camera systems installed in our surroundings [2], But not all of them can be monitored continuously because human operators are needed to monitor events and not withdraw their attention from the screens, but this is difficult, as there are studies that indicate that a human operator tends to get tired after a certain time and with this situation it is not possible to detect events that seem to be important. In video surveillance systems, human detection is a major problem for the development of video surveillance technology[3]–[5]. The people detection process provides useful information that can be used to classify people's activities, such as walking, running, jumping, waiting, among others. Pedestrian detection continues to be a problem studied by different researchers, despite the fact that people move in real environments, such as parks, schools, shopping malls, lobbies, airports, borders, etc., important information is generated to be analyzed. Added to it, the type and color of clothing can make the detection problematic because of the lack of contrast. Also, lighting changes and distractors such as background clutter can affect detection. In the development of this task, the aim is to identify complications such as disturbances in the appearance of people in the sequence, obtaining false positives or false negatives in the detection of people, as well as the partial or total conclusion of people. That is the reason this work would focus on providing images of pedestrians walking in scenes with variations in background, type of clothing, etc., these images are intended to be stored in a dataset, which will be available on a website so that it could be used by researchers in the detection of pedestrians.

Nowadays, researches related to pedestrian detection in video recording systems obtained from CCTV cameras are being carried out. The state of the art studies the characteristics of the detection of people in different places, with different variations in lighting environments, size, shape and position, those factors prevent the detection of people walking in these scenes [6]–[8].

In order to perform pedestrian detection it is necessary to have a collection of images or datasets of images containing people wandering and which were videotaped in different parts of the world, it is worth mentioning that these datasets were not made for motion detection and another important point is that they were recorded in other places abroad, where there are certain restrictions to download these videos or pedestrians do not have the aforementioned characteristics. Some of these datasets are:

1. Penn-Fudan Database for Pedestrian Detection (https://www.cis.upenn.edu/~jshi/ped_html/), [9]
2. Caltech Pedestrian Detection Benchmark (http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/), [10]
3. INRIA Pedestrian Test Dataset, [11]
4. ETH Pedestrian Dataset, [12]
5. TUD-Brussels Pedestrian Dataset, [13]
6. Daimler Pedestrian Dataset, [14]

II. MATERIALS AND METHODS

The proposed technique is called Motion Detection with Background Subtraction (*MDBS*), which detects the region of change (*RoI, Region of Interest*), where the pedestrian can be located; therefore, the following methodology is proposed:

1. Convert the first frame of the video sequence to grayscale and it is defined as the background frame (F).
2. Reads the next frame in the video sequence (*I*) and converts it to grayscale.
3. Subtracts the background *F* from the image *I*, $B(x, y) = I(x,y) - F(x, y)$.
4. Binarize B, considering a threshold Th, whose value is calculated experimentally based on the state of the art, with a sample a few *n i*mages of x×y pixels.
5. Erosion B, using a circular structured element E of radius R = 5, considering that the silhouette of the human body is curvilinear, and a circular structure helps to round the contour.
6. Remove residues in *B* using the 4-neighbor technique, understanding residue as the set of pixels isolated from the silhouette of the person.
7. Fill holes in *B*, understanding that a hole is a set of pixels of *B* with value *0* inside the outline of the silhouette of the person.
8. Calculate the equation of the regression line, to delimit the cluster where the movement was detected.
9. Calculate the principal components considering x representing the n points with coordinates *(x,y)* corresponding to the pixels in B whose value is 1.

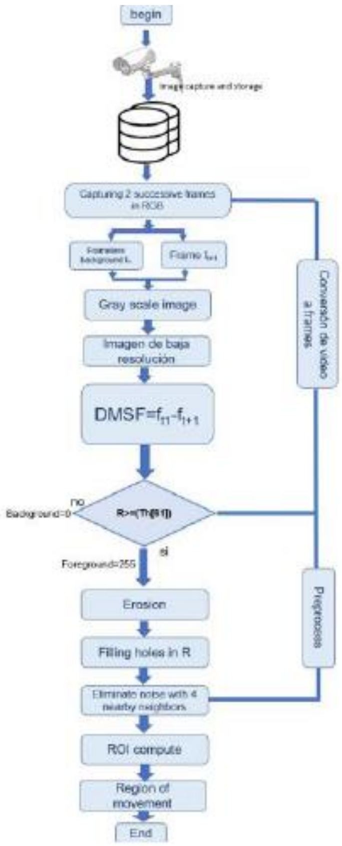

Fig. 1 Diagram showing the steps considered to create the boundary box that delimits the frame to a pedestrian with the motion detection technique.

10. Delimit the region of motion. For this, calculate the extreme points: $P1(x_1, y_1)$, $P2(x_2,y_2)$, $P3(x_3, y_3)$ and $P4(x_4,y_4)$ of the straight lines L_1 and L_2, whose direction vectors are the eigenvectors v_1 and v_2. These orthogonal lines intersect at the midpoint $O(x_i,y_i)$ of the pixels in *B* whose value is 1. Therefore, as shown in the Fig. 2, the points: $E1(x_4,y_1)$, $E2(x_2,y_1)$, $E3(x_2,y_3)$ and $E4(x_4, y_3)$ delimit the region of motion. In this case, the points forming the line L1 can also be calculated using the regression line.

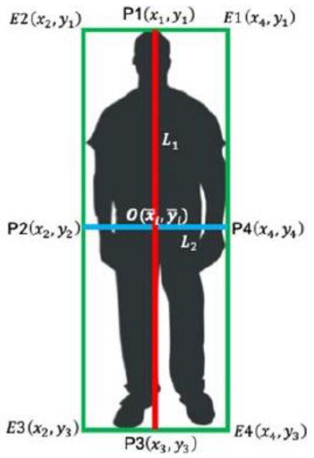

Fig. 2 Estimation of the rectangle or region of interest obtained by applying the proposed MDBS algorithm for motion detection.

Below is the graphical representation of the stages to segment and subsequently find the region of interest (ROI) which is delimited by the calculated rectangle.

III. PROPOSED FRAMES FOR DATASET CREATION

As mentioned previously, the purpose of this work is to present a collection of images (dataset), which can be useful to researchers working in motion detection.

To that purpose the frames shown contains pedestrians videotaped in the facilities of the Engineering Faculty of Autonomous University of the State of México (UAEMex), where specific features were searched to be considered in the study of motion detection, these videos can be obtained by sending an e-mail request to juan.antonio@tesjo.edu.mx.

In the next frames pedestrians can be seen walking in a forum of the Faculty of Engineering of the UAEMEX, the scenario of this forum was selected as a closed place with controlled artificial light. The pedestrians that were videotaped in these frames have a specific outfit (black clothes).

Fig. 3 Frame showing one pedestrian walking in the unique direction in the forum of the engineering faculty of UAEMex.

It is expected that they do not cause problems with the background, so that the proposed algorithm or algorithms do not have any problem in the detection of the possible pedestrian. For example, in Fig. 3 is showing a frame of the background without any pedestrian appearing in the scene, this frame shows that there is a white background immersed in a wooden environment that can cause some noise for detection as is the brightness of the floor and the brown color of the wood that might contrast with the pedestrian or pedestrians that are walking.

Fig. 4 Frame showing 2 pedestrians of different heights walking to the right in the forum of the engineering faculty of UAEMEX

In Fig. 4, you can see the selection of a frame with a single videotaped pedestrian, which should not cause any problem when detected because it does not cause occlusion. In Fig. 5, it can be seen that two pedestrians of different heights are walking in one direction to the right of the frame. However, in Fig. 6 and Fig. 7, two pedestrians can be seen walking in different directions generating occlusion at a specific moment of the crossing between the two pedestrians. This occlusion is useful for researchers because they can detect either one or the other pedestrian despite the occlusion that can be detected.

Fig. 5 Frame showing 2 pedestrians walking in different directions in the forum of the engineering faculty of UAEMex

Fig. 6 Frame where 2 pedestrians are observed walking in different directions, but at one point they are occluded.

It can also be seen that in Fig. 8, 3 pedestrians are shown entering the scene. This type of scene generates more noise at the moment of detecting the pedestrian due to the generation of occlusion between 3 pedestrians.

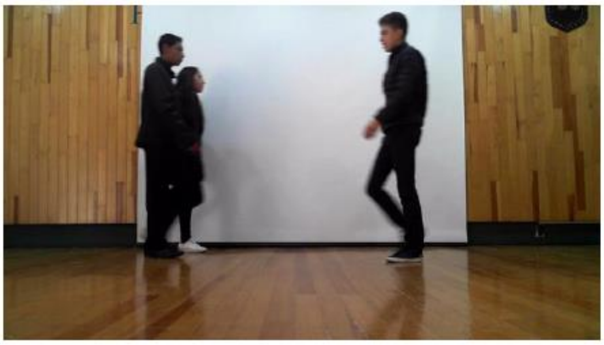

Fig. 7 Scene showing 3 pedestrians walking at the same time generating a problem known as occlusion.

Likewise shown below are the frames that were captured in the main forum of the faculty of engineering of the UAEMex. The forum has a wooden background, in addition to a polished wooden floor that reflects the shadow of people walking. These factors made it difficult to detect pedestrians in real time in that sequence. In contrast to the previous sequences, where a white background was placed to help avoiding contrast and thus easily detect pedestrians. This time the wooden background does not have a white screen (See Fig. 9), so the algorithm that detects possible pedestrians would find it difficult to locate these people. The pedestrians that were chosen for these scenes were dressed in white shirts and black pants, expecting satisfactory results at the time of detection. In Fig. 10 a pedestrian wearing the described clothing walks from the left to the right side of the screen.

Fig. 8 Background frame of a wooden forum where pedestrians will pass to detect

Fig. 9 Frame showing the video recording of a single pedestrian passing on a wooden background in a UAEMEX forum.

Fig. 10 Frame showing the video recording of 2 pedestrians passing on a wooden background in a UAEMEX forum.

ig. 11 shows a peculiar feature in the scene where there are two pedestrians walking in opposite directions. It can be noticed that when walking at a fast speed these pedestrians generate a blurred vision in the frame. That would be a disadvantage when detecting the pedestrians, in addition to the other disadvantages such as the pedestrians' clothes and the background. These factors make detection difficult with the proposed algorithm.

In order to verify the operation of the motion detection technique, another area of the faculty of engineering of the UAEMex was selected. It is a corridor of the graduate building, which has 2 types of lighting: natural and artificial lighting. Pedestrians were videotaped walking in a place that has a window to which a white background was purposely placed. This environment establishes a contrast between the pedestrian and the background. It is worth mentioning that the pedestrians that were videotaped in these scenes have a random clothing in order to see the behavior of the algorithm and that sometimes there was no contrast between the pedestrian and the background because of the clothing. In Fig.s 12 and 13, are shown pedestrians whose clothing does not contrast with the background because the windows were intentionally hidden with white wallpaper.

Fig. 12 In this frame a pedestrian was videotaped wearing a type of clothing that does not contrast with the background.

From these videotaped frames in the chosen environment, it can be concluded that the expected results tend to be more accurate when using the technique to detect pedestrians, because the background and the chosen environment is conducive to be able to enclose in a bounding box the detected target. Just in this scene the only difficulty is the floor where the pedestrians walk, as it reflects the silhouette of the pedestrian, and this would be considered a noise factor when applying a frame difference technique.

Fig. 13 A videotaped pedestrian contrasts with the background, due to his dark clothing.

IV. RESULTS

Here are some results obtained when using the MDBS algorithm. The best scene was chosen because it does not present difficulties with the current algorithm. Unlike the previous scenes, the one with a white background presents a better segmentation and a better detection is seen where the pedestrian is recognized by a rectangle known as region of interest (ROI) which was obtained thanks to the proposed algorithm. This rectangle is adapted to the shape of the supposed pedestrian detected. As mentioned before, some algorithms can be used for motion detection or can also be used some other techniques that researchers in the area of computer vision intend to analyze.

The results of applying the frame difference technique called MDBS are presented below, showing the performance of this algorithm in each of the scenes that were performed in the different enclosures explained previously.

Fig. 14 shows a pedestrian walking in one direction on the 3rd floor of the postgraduate building of the Autonomous University of the State of México; as it can be seen in the Fig., the result of the segmentation with the frame difference technique shows a silhouette of a pedestrian detected despite some problems such as the reflection of the floor.

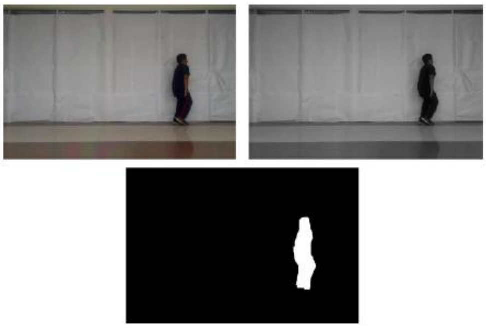

Fig. 14 This Fig. shows a pedestrian walking in the corridor of the 3rd floor of the postgraduate building, its original RGB frame and a frame in grayscale format are attached.

The Fig. 15 shows the result of two pedestrians walking in different directions. It can be noticed that by the time they are walking the two pedestrians are occluded when passing a pedestrian in front of another. A difficulty encountered by the DFSM algorithm is the floor which reflects the silhouette of each pedestrian, which makes detection difficult, as mentioned above.

Fig. 15 In this frame two pedestrians can be seen walking in a corridor on the 3rd floor of the graduate building in which you can notice their original image in RGB, grayscale and its binarized segmented result obtained by the DMFS technique, at the good result.

However, in Fig. 16, three pedestrians are seen walking in different directions, despite the occlusion and the floor brightness issue.

Fig. 16 It can be seen the segmented result of three pedestrians walking in the 3rd floor corridor of the graduate building, the original RGB, grayscale and binarized segmented images are shown.

Here are the results obtained in the main auditorium, in the auditorium, which has wooden walls and stave. There are some problems such as wood, which generates background noise. When pedestrians pass the background generates distortion with the motion detection algorithm, that is why in some of the scenes the passage of pedestrians is tested with a background containing a white screen that would help the detection without causing background problems. Fig. 17 and 18 show the results of a pedestrian walking alone in the scene, as well as pedestrians walking in different directions causing them to face each other at time t-1 and so occlude each other.

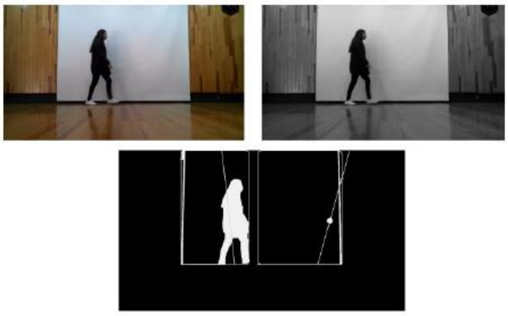

Fig. 17 This image shows a pedestrian walking from left to right in the auditorium with wooden stave and wooden walls, a white screen was placed to obtain the segmentation which has noise when applying the technique.

Fig. 18 This image shows two pedestrians walking in opposite directions in the auditorium with wooden stave and wooden walls, a white screen was placed in order to obtain a good segmentation.

The segmentation results obtained in the same wooden auditorium are presented below. However, the white screen was removed to verify the effects when applying the segmentation technique. The polished wood produces a brightness and reflection of pedestrians when walking and thus generates noise, therefore unfavorable results are expected.

Fig. 19 In this segmentation result that was applied to 2 pedestrians walking in opposite direction, it can be noticed that a good segmentation was not obtained due to the wooden background.

In addition to the normal segmentation obtained with the MDBS algorithm proposed in this research, a bounding box marking is performed which is calculated and plotted to enclose the pedestrian. This with the utility of being used as ground truth to be calculated with the predicted result, with the help of a metric such as Intersection Over Union (IoU) that can be verified with other detection algorithms. In Fig. 20 and Fig. 21 are the results obtained from a pedestrian walking in the third-floor corridor of the faculty of engineering building.

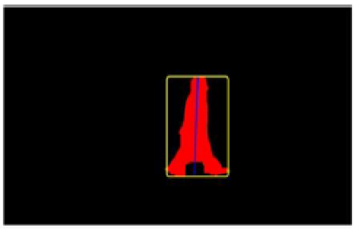

Fig. 20 shows the result of the bounding box obtained from the image of a pedestrian walking along

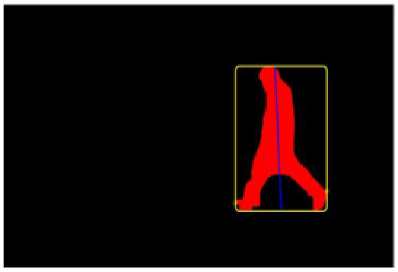

Fig. 21 It shows the result of the bounding box obtained from the image of a other pedestrian walking along the third floor corridor of the engineering faculty building.

In addition to the detection of a single pedestrian, detection results also show the region of interest (ROI) by bounding box, of two pedestrians walking in the third-floor corridor of the graduate building of the faculty of engineering of the UAEMEX.

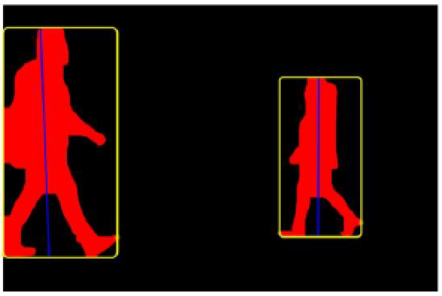

Fig. 22 It shows the result of the bounding box obtained from the image of a two pedestrian walking along the third floor corridor of the engineering faculty building.

Regarding the detection in the wooden auditorium in the stage where a white screen was placed so that there would be less noise at the time of detection, the Fig. shows a single pedestrian that at the time of detection shows noise when displaying its bounding box.

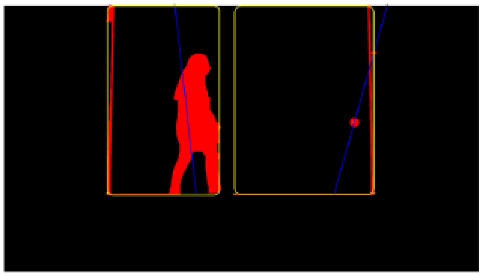

Fig. 23 The result of the detection of a single pedestrian walking in the auditorium with stave and wooden walls is shown, as can it be seen, noise is generated due to background problems.

But even though the auditorium with wooden stave and wooden walls has a screen, the detection results show some noise problems. Despite the efforts to perform pedestrian detection by removing the white screen in order to verify the results, unpromising results were obtained as we could not segment the pedestrians walking in the hallway of this auditorium because the wood reflects the silhouette of pedestrians, in addition to the brightness and reflection of the light generated by this auditorium, and therefore this algorithm is not efficient in this type of problem, as can be seen in Fig. 24 and 25.

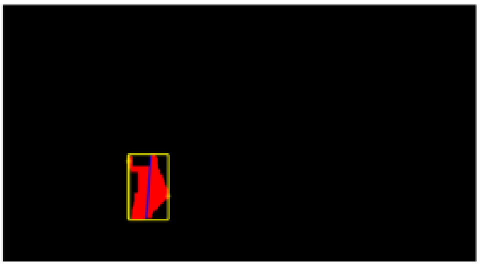

Fig. 24 In this result it can be noticed that the segmentation of a single pedestrian walking in the wooden auditorium is not satisfactory

Fig. 25 In this result it can be noticed that the segmentation of two pedestrians walking in the wooden auditorium is not satisfactory due to different background noise factors.

V. CONCLUSION

This paper presented the study of an algorithm called MDBS (Motion Detection with Background Subtraction), in which pedestrians walking in different parts of a dataset were videotaped to perform motion detection in each of the frames belonging to a set of images of the dataset. For this purpose, a set of datasets recorded at the Faculty of Engineering of the Autonomous University of the State of Mexico is proposed. These scenes have their own characteristics to be used by motion detection techniques, although it is open to be used by other techniques that can or want to detect pedestrians in the mentioned scenes. On the other hand, some results obtained by a proposed technique titled DMSF are shown, this technique has good results, as long as the pedestrians in the scenes have problems with the background. The development of this work is not finished. It is expected to use other motion detection techniques to analyze and compare their results.

A conclusion section should be included in which the advantages, limitations, and possible applications of the work are clearly stated. Although a conclusion may review the main points of the paper, do not repeat the abstract as a conclusion. A conclusion may explain the significance of the work or suggest applications and extensions.

Acknowledgements

We are grateful to the Tecnológico de Estudios Superiores de Jocotitlán for the support provided for the development of this article, as well as to the Faculty of Engineering of the UAEMex for giving us permission to videotape pedestrians in their facilities.

REFERENCES

[1]. A. Jordão and W. R. Schwartz, "The Good, The Fast and The Better Pedestrian Detector," Univ. Fed. Minas Gerais - Dep. Ciência da Comput., vol. 1, pp. 1–51, 2016, [Online]. Available: http://gibis.unifesp.br/sibgrapi16/eproceedings/wtd/19.pdf.

[2]. A. T. Angonese and P. F. Ferreira Rosa, "Multiple people detection and identification system integrated with a dynamic simultaneous localization and mapping system for an autonomous mobile robotic platform," ICMT 2017 - 6th Int. Conf. Mil. Technol., pp. 779–786, 2017, doi: 10.1109/MILTECHS.2017.7988861.

[3]. A. Dominguez-Sanchez, M. Cazorla, and S. Orts-Escolano, "Pedestrian Movement Direction Recognition Using Convolutional Neural Networks," IEEE Trans. Intell. Transp. Syst., 2017, doi: 10.1109/TITS.2017.2726140.

[4]. A. Pentland and B. Horowitz, "Recovery of Non-Rigid Motion and Structure," IEEE Trans. Pattern Anal. Mach. Intell., 1991, doi: 10.1109/34.85661.

[5]. N. Vandenbroucke, L. Macaire, C. Vieren, and J. G. Postaire, "Contribution of a color classification to soccer players tracking with snakes," 1997, doi: 10.1109/icsmc.1997.633237.

[6]. J. A. Antonio and M. Romero, "Detección de peatones con variaciones de forma al caminar con Modelos de Forma Activa," Cienc. ergo sum, vol. 27, no. 3, p. e100, 2020, doi: 10.30878/ces.v27n3a10.

[7]. S. Das Choudhury and T. Tjahjadi, "Gait recognition based on shape and motion analysis of silhouette contours," Comput. Vis. Image Underst., 2013, doi: 10.1016/j.cviu.2013.08.003.

[8]. C. Jang and K. Jung, "Human pose estimation using Active Shape Models," vol. 2, no. 8, pp. 530–534, 2008.

[9]. F. Xia, J. Zhu, P. Wang, and A. L. Yuille, "Pose-guided human parsing by an AND/OR graph using pose-context features," 30th AAAI Conf. Artif. Intell. AAAI 2016, vol. 30, pp. 3632–3640, 2016, [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/10460.

[10]. S. B. Zang Shanshan, Benenson Rodrigo, "CityPersons: A Diverse Dataset for Pedestrian Detection," CVPR 2017 open access, pp. 3213–3222, 2017, [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Zhang_CityPersons_A_Diverse_CVPR_2017_paper.html.

[11]. P. Dollár, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," IEEE Trans. Pattern Anal. Mach. Intell., 2012, doi: 10.1109/TPAMI.2011.155.

[12]. W. Ouyang, X. Zeng, and X. Wang, "Single-pedestrian detection aided by two-pedestrian detection," IEEE Trans. Pattern Anal. Mach. Intell., 2015, doi: 10.1109/TPAMI.2014.2377734.

[13]. X. Zeng, Xingyu; Ouyang, Wanli; Wang, "Multi-stage Contextual Deep Learning for Pedestrian Detection," Proc. IEEE Int. Conf. Comput. Vis., pp. 121–128, 2013, [Online]. Available: https://openaccess.thecvf.com/content_iccv_2013/html/Zeng_Multi-stage_Contextual_Deep_2013_ICCV_paper.html.

[14]. W. Ouyang and X. Wang, "A discriminative deep model for pedestrian detection with occlusion handling," CVPR'12, 2012, doi: 10.1109/CVPR.2012.6248062.