Face Detection in Blurred Surveillance Videos for Crime Investigation

K. Menaka¹ and B. Yogameena²
¹Associate Professor, Department of ECE, Mangayarkarasi College of Engineering, Madurai, India.
²Associate Professor, Department of ECE, Thiagarajar College of Engineering, Madurai, India.

ece.menaka@gmail.com, ymece@tce.edu

Abstract. Face detection itself acts as a significant problem in low-resolution surveillance video, primarily due to out-of-focus blur. Real-time and forensic analysis for Law enforcement desires superior face recognition system for person identification in surveillance applications. However, face detection influences face recognition in critical situations. If faces on the blacklist are not detected, recognition fails, which increases the burden and is a tremendous challenge for police authority. This motivates us to present the deblurring algorithm to improve the face detection rate and reduce the false-positive rate. Hence, this paper focuses on removing blur by applying a blind deconvolution algorithm, suppressing the ringing artifact in surveillance video by adopting Discrete Wavelet Transform (DWT). Finally, after this preliminary work, faces are detected by Yolo v2. The proposed framework works well in the sparsely crowded scenario and improves face detection rate compared to the Lucy-Richardson-Yolo v2 framework, which suffers from the problem mentioned earlier. Experiments and evaluations on the surveillance dataset show that the proposed blind deconvolution with ringing suppression-based solution outperforms the state-of-the-art methods in detecting both frontal and profile faces.

1. Introduction

Over the past decade, there have been many crimes frequently happening worldwide every day, mainly in areas like jewelry shops, ATMs, banks, and educational institutions. To curb such crimes occurring across the globe, there arises a need to evolve facial recognition technology. In such a scenario, face detection and recognition of criminals play a significant part in most surveillance applications such as biometrics, person-specific identification, and so on. However, numerous parameters degrade the quality of the video captured by the camera, and consequently, the criminal cannot be recognized. If faces in the blacklist could not be detected due to blur or other degradations, the criminal's recognition is a failure. Hence, this problem turns to be a great challenge and further increases the burden of police authority. There are various face detection challenges, such as blur, pose variation, uneven illumination, facial expression, and occlusion. While considering surveillance applications, blur acts as a significant issue in face detection in surveillance camera recordings due to the movement of the camera [1]. This motivates us to present the restoration or deblurring algorithm for blurred videos to improve the face detection rate and reduce the false-positive rate. This system is very invaluable for real-life applications such as person-specific identification during criminal investigation.
2. Related work

The factors affecting any surveillance video's efficiency and reliability are blur, noise, illumination, apart from other environmental factors. Though there are robust algorithms against illumination and poses, blur removal is less focused. It is evident from the literature that, despite the limited presence of blur, especially in surveillance videos, only some research has been done to recognize the face [1]. The proposed work is intended to deal with this type of scenarios where blur impacts a significant role in face detection. In particular, the work targets to detect and recognize the criminals, as found in the footage of a woman stabbed over 22 times by a stalker in Delhi, collected from CCTV footage. The footage obtained from Non-ATM1 of Oneindia News is cited as an example of an excessive blur, due to which the guilty could not be recognized authentically. In this case, even though the stabber faces the camera, his face could not be identified due to excess blur.

Blur cannot be removed unless one has prior knowledge about the type of blur present in the video. Various restoration-based deblurring algorithms, such as the Lucy-Richardson algorithm [2] and blind deconvolution [3]. Other linear filters such as wiener, constrained least square, inverse, and pseudo-inverse filters can also be used to restore the blurred image. Various motion deblurring algorithms such as gradient magnitudes with zero mean GMM, blind deconvolution are more appropriate for removing motion blur. Gaussian and cylindrical blur is the reasonable estimates of out-of-focus blur present in the camera.

The work reported in [4] has utilized deconvolution and super-resolution algorithm to handle the blurring effect in videos. However, its performance has deteriorated during the estimation of optical flow, layer separation, and deconvolution algorithm, and motion blur that varies spatially is more intricate to handle. Blind deconvolution is implemented in [3] at the final step for deblurring after preprocessing an image. However, in the presence of noise, it is challenging to estimate PSF accurately. Least Mean Square (LMS) optimization in [5] reduces the coding artifact. However, it cannot be applied for real-life applications as the critical quality metrics to categorize the sharpness and artifact level are not employed. Co-prime blurred pair (CBP) model [6] is implemented in a surveillance camera to improve data security and reduces the frame rate impose challenges sensitive to time for face or object recognition. There is a need to detect faces in surveillance videos, even though they are distorted by out-of-focus blur. The prevalent Viola-Jones algorithm fails to detect faces in surveillance videos due to severe lighting conditions and produces more false positives. From the literature review, it is inferred that traditional methods rely on feature extraction for face detection with various constraints. Nowadays, a deep learning methodology has attained enormous success in challenging face detection as a subclass of face classification, localization, and detection. CNN for face detection is significantly computational with numerous different sliding windows [7]. Softmax loss function discourages the intra-class firmness acts as the major disadvantage of region-based Faster R-CNN methods [8] in face detection. Face R-CNN, proposed in [9], has tackled the above problem by attempting the center loss function for the first time. However, Faster R-CNN performance is poor for low-resolution videos [10], could not be used for surveillance videos. More recently, Yolo has been the newly boomed single deep network, applied to achieve less training and testing time for face detection. However, low recall values than region-based methods and critical localization errors create complexity. Therefore, it is suggested that Yolo v2 [11] improve the accuracy of detection. Compared to Yolo, it generates improved recall values and identifies more [17] faces. The survey addresses the positives and negatives of traditional and a deep network of face detection algorithms. It is assumed that after removing the blur, no deep learning algorithm detects face [1] that mainly exists in the security applications surveillance video for personal identification. Hence, after deblurring, the proposed framework subsequently adopts Yolo v2 [11] for face detection.

2.1 Motivation
The crime rate has not decreased. However, there is a need to identify the persons involved in any criminal activity. The face information provides higher authentication. Consequently, the critical information, i.e., sharp features in the video sequences, are blurred. These features impact face detection and worsen the concentration of the video captured by a surveillance camera [11].

2.2 Problem Formulation
It is observed from the literature review that traditional methods do not deal effectively with this issue and generate False Positives (FP). Therefore, a thorough approach is required in surveillance videos to perform face detection, considering blur removal as a pre-processing phase after face detection. This work focuses on the detection of face and profile in the presence of blur in surveillance videos.

2.3 Objective
The primary aim of this work is to introduce an improved intelligent surveillance system for face detection even under blurred conditions for forensic and safety identification of individuals.

2.4 Key Contribution
The proposed face detection framework in surveillance video includes blur removal, blind deconvolution, and face detection using Yolo v2 with Darknet-19. The dataset collection of the authors of this work is a significant contribution to this framework. It helps to identify the person for crime detection. Another vital contribution is selecting appropriate deblurring algorithms suitable for pre-processing surveillance videos before face detection. After blur removal and ringing suppression, faces are detected using deep learning-based Yolo v2 with the Darknet-19 object detection algorithm. Hence, detecting faces after pre-processing the blur is the crucial contribution.

3. Proposed Methodology

3.1 Methodology
This work aims to develop an improved face detection system that helps to identify persons even under blurred conditions for surveillance applications. The proposed structure starts by transforming the video input into a frame sequence. After that, a pre-processing phase eliminates the out-of-focus blur in the surveillance video. The blind deconvolution is performed with the modeled Gaussian blur PSF to identify the sharp edges in the higher frequency band. Followed that, the deblurred video is smoothened using a Gaussian filter to remove noise. Further, the ringing artifacts are suppressed using an adaptive edge map of DWT in the post-processing step. Finally, the faces are detected using Yolo v2 with the Darknet-19 algorithm. The overview of the proposed framework for improved face detection is given in figure 1.

![Diagram of proposed method](image)

**Figure 1.** Overview of the proposed framework for improved face detection.
3.2 Estimation of Out-of-focus Blur PSF

Surveillance footages that have low resolution, emphasize the presence of blur and they are considered in this proposed work. First, the videos are converted into a frame sequence. The blurred sequence is generally of the form presented by

\[ g(x, y) = f(x, y) * h(x, y) + \eta(x, y) \]

(1)

Here \( f(x, y) \) is the desired un-blurred sequence, \( h(x, y) \) is the blur PSF, \( g(x, y) \) is the perceived blurred surveillance sequence, and \( \eta(x, y) \) is the additive noise due to camera error or quantization.

In general, out-of-focus blur exists in the monitoring video due to out-of-focus of the lens in the camera, as discussed earlier. The blur highly degrades the video quality and hence it has to be restored for further analysis. Here, the simple, blind deconvolution algorithm has been adapted for the deblurring process. However, the blurring kernel i.e the Point Spread Function (PSF) is unknown, and so it has to be estimated. Since the proposed work considers the out-of-focus blur, the PSF of out-of-focus blur has to be modelled. Blurring, in general, is the low pass filter. Similarly, the Gaussian filter functions as a low pass filter which attenuates the element of the higher frequency. The standard Gaussian function is considered to be the blur model with standard deviation, \( \delta \) as the blur parameter. Therefore, out-of-focus blur kernel is modelled as the Gaussian function and the PSF is,

\[ h(x, y) = e^{-\frac{x^2 + y^2}{2\delta^2}} \]

(2)

The equation (2) in the frequency domain can be rewritten as

\[ H(U, V) = \sqrt{2\pi\delta} e^{-\frac{1}{2}\left|\frac{U}{\delta}\right|} \]

(3)

Here \( U, V \) are frequency components. \( H(U, V) \) is the modelled Gaussian blur PSF for out-of-focus blur.

3.3 Blind Deconvolution

Blind deconvolution, given a blurred kernel, is the process of recovering an unknown image from the blurred sequence. In the frequency domain, ignoring the noise term equation (1) is expressed as

\[ G(U, V) = F(U, V)H(U, V) \]

(4)

The blurring process in equation (4) suppresses the higher frequency component, comprising sharp edge features that are required for face detection. In order to restore the sharp features, the blind deconvolution deblurring algorithm is utilized.

The deconvolution, given a blurred kernel, is recovering an unknown image from the blurred sequence. The blind deconvolution [12, 13] deblurring algorithm is therefore expressed as

\[ \hat{F}(U, V) = H(U, V)G(U, V) \]

(5)

The deblurring process in equation (5) deconvolves the blurred surveillance sequence \( G(U, V) \) with the modeled Gaussian blur PSF \( H(U, V) \). It retrieves the higher frequency component that consists of sharp edge features needed for face detection.

3.4 Ringing Suppression using DWT

After the deblurring process, DWT is used to suppress the ringing artifacts as a pre-processing step before they are fed into the Yolov2 face detector to improve the face detection rate. Blind deconvolution and other non-blind deblurring techniques such as Lucy Richardson deconvolution and Weiner and inverse filters produce ringing artifacts after deblurring. The adapted DWT finds the primary edge map [14] regions and is dilated to determine the location of the ringing artifacts. The
average DWT domain pixel variance is applied for measuring the degradation. Finally, the deblurred sequences are obtained, and they are fed into the Yolo v2 object detector for face detection.

3.5 Face Detection using Yolov2
The proposed framework adopts Yolo v2 with Darknet-19 for detecting faces in the surveillance after pre-processing using blind deconvolution. Yolo v2 consists of 19 convolutional layers and five max-pooling layers. It uses the Darknet-19 feature extractor as a base so that the speed and accuracy are improved.

4. Experimental Results and Discussion
The proposed approach for face detection has experimented with the following surveillance dataset: OWN, Chokepoint, ATM1, ATM2, ATM3, IISC Bangalore, Non-ATM 1 SC Face, ETH, ILIDS-VID, and PRID 2011. The results are evaluated using Matlab 2018a. Specifications of surveillance video datasets for the proposed work are shown in Table 1. The dataset contains frontal and profile faces with challenging conditions such as blur, pose variation, varying illumination, and shadow.

Table 1. Surveillance video datasets and their specifications.

| Dataset   | Identities | Total frames | Resolution | Challenging Conditions | View angle          |
|-----------|------------|--------------|------------|------------------------|---------------------|
| OWN       | Vary       | 11339        | 704x576    | Blur, Pose variation   | Front, Profile      |
| Choke     | 29         | 48 video     | 240X180    | Pose variation, Varying illumination | Front               |
| SC Face   | 130        | 1950         | 680x556    | Pose variation, Varying illumination | Front               |
| ETH       | 8,53,528   | 8580         | Vary       | Pose variation, poor resolution | Front               |
| ILIDS-VID PRID 2011 | 300   | 42495        | Vary       | Similar Clothing, Varying illumination | Vary               |
| ATM1      | 2          | 6950         | 352x288    | Sudden illumination, blur | Front               |
| ATM2      | 2          | 1935         | 640x360    | Blur, Camouflage       | Profile             |
| ATM3      | 2          | 1620         | 448X336    | Extreme lighting       | Front               |
| OWN1      | Vary       | 20,641       | 704x228    | Shadow, blur           | Front               |
| Non-ATM1  | Vary       | 175          | 1280X720   | Excess Blur            | Profile             |

Comparatively, sequences are restored using state-of-the-art deblurring algorithms such as Lucy-Richardson and blind deconvolution and Weiner filter and are qualitatively revealed in Table 2. Timestamp characters recorded in the ATM 1 dataset are visible at the perception level even at extreme lighting conditions, using the blind deconvolution in the restored sequence. It impacts the proposed approach in deblurring for improved face detection. Lucy Richardson algorithm produces ringing artifacts in the regions of significant object motion for the IISC dataset and the background of video for the OWN dataset, which is the common problem of this algorithm [15]. Ringing artifacts that are highly produced in OWN 1 dataset by Lucy Richardson are suppressed using the DWT after blind deconvolution algorithm. In the Chokepoint dataset, the blurring is increased even after applying the
Weiner filter, which is the principal disadvantage of this filter. These qualitative results demonstrate the merit of blind deconvolution with DWT and the demerits of the other two algorithms.

Further, it is evaluated quantitatively. Performance metric for deblurring is Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) [3]. It is observed that OWN1 and ATM1 are excessively blurred using Weiner with very low PSNR compared with Lucy and Blind deconvolution. Comparatively, Blind deconvolution produces better PSNR and MSE for all the tested datasets while suppressing the ringing artifacts with DWT. The quantitative results are enumerated in Table 3. After deblurring, Yolo v2 is applied for face detection. Faces bounded in red indicate the detected frontal faces, whereas faces bounded in green indicate the detected profile faces.

Table 2. Comparison of various deblurring algorithms on surveillance dataset for face detection.

| Dataset | Deblurring using |
|---------|-----------------|
|         | Weiner          | Lucy Richardson | Blind Deconvolution |
| OWN     | ![Image](OWN.png) | ![Image](OWN.png) | ![Image](OWN.png) |
| Chokepoint | ![Image](Chokepoint.png) | ![Image](Chokepoint.png) | ![Image](Chokepoint.png) |
| ATM1    | ![Image](ATM1.png) | ![Image](ATM1.png) | ![Image](ATM1.png) |
| OWN1   | ![Image](OWN1.png) | ![Image](OWN1.png) | ![Image](OWN1.png) |

It is observed that fewer faces are detected in OWN and Chokepoint dataset using Weiner filter due to noise variance, and FP occurred in ATM1, and face detection fails in the OWN1 dataset. More faces are detected in the OWN and Chokepoint dataset with the blind deconvolution algorithm, compared with Lucy Richardson algorithm [2] and Weiner. However, FP occurs in ATM1, and faces are not detected in the OWN1 dataset due to ringing artifacts after deblurring without ringing suppression is analyzed. It is qualitatively shown in figure 2.

Table 3. MSE and PSNR of various deblurring algorithms on a few surveillance datasets.

| Dataset | Weiner Filter | Lucy Richardson | Blind Deconvolution |
|---------|---------------|-----------------|---------------------|

6
Similarly, after performing deblurring using blind deconvolution, it is analyzed that ringing suppression is done by adopting DWT. As a result, more faces are detected in OWN and Chokepoint dataset. Notably, faces have also been detected in OWN1 and ATM1, which proves the significance of the proposed approach with the impact of ringing suppression using DWT. Its qualitative results are depicted in figure 3.

![Figure 2](image-a)

**(Figure 2.** Several faces detected using Yolov2 after deblurring using Blind deconvolution without Ringing Suppression (RS) for sample surveillance datasets a) OWN, b) Choke Point, c) ATM 1 and d) OWN 1.)

![Figure 3](image-b)

**(Figure 3.** A number of faces detected using Yolov2 after deblurring with Ringing Suppression (RS) for sample surveillance datasets a) OWN, b) Choke Point, c) ATM 1 and d) OWN 1.)

The precision (p), recall(r), accuracy, and Average Precision (AP) performance metrics are calculated.0

\[ AP = \frac{1}{11} \sum_{r} p_{max}(r) \]  

(6)

Here, AP [16] is the mean average precision at 11 equally spaced recall levels [0, 0.1, 0.2……1.0]. Further, the precision-recall values are interpolated by replacing them with the maximum precision \( p(r) \), measured at higher recall levels \( r \). The mAP is the average of all classes.

To validate the experimentation, mAP (%) is obtained for the proposed Yolo v2 (Y) based face detection and is compared with the state-of-the-art prevalent Viola-Jones (V) algorithm. Table 4 depicts the improvement in face detection rate, using Yolo v2, after deblurring by adopting blind deconvolution and DWT-based ringing suppression of the experimented datasets. These experimental results reveal that Yolo v2 with blind deconvolution provides a higher mAP even in challenging real-world conditions such as sudden illumination changes, low resolution in surveillance data, as compared to Yolo v2 with other deblurring algorithms such as Weiner and Lucy Richardson. Also, the proposed framework produces higher accuracy than the state-of-the-art Viola-Jones with deblurring.

|        | MSE  | PSNR | MSE  | PSNR | MSE  | PSNR |
|--------|------|------|------|------|------|------|
| OWN    | 14.85| 26.14| 14.78| 21.12| 14.53| 45.15|
| Choke  | 13.82| 24.20| 13.70| 26.19| 13.20| 38.82|
| ATM 1  | 16.67| 21.04| 16.69| 26.94| 16.58| 35.93|
| OWN1   | 15.40| 18.84| 15.05| 22.93| 14.70| 41.96|
Apart from this quantitative analysis, the ablation study has been carried out based on blur removal. In this work, face detection has also been experimented straight on the frames without performing blur removal. It is found that fewer faces are detected using Viola-Jones, as compared with Yolo v2. Also, Yolo v2 does not produce the result up to the expected target. Remarkably, it is perceived that if the blur removal procedure is neglected, more FP's (faces like object in the background and foreground) results occur using Yolo v2. In the ATM1 dataset, face like object in the background has been detected as a face. This kind of FP is certainly avoided by removing blur. This authenticates the impact of blur removal as a pre-processing step before face detection for processing any surveillance video.

### Table 4. Face detection rate using Yolo V2 (Y) before and after blur removal

| Dataset | Face detection rate before blur removal | Face detection rate after blur removal |
|---------|---------------------------------------|--------------------------------------|
|         | Viola (ACC)     | Yolov2 (mAP)     | Weiner Filter (Viola (ACC) | Yolov2 (mAP) | Lucy Richardson (Viola (ACC) | Yolov2 (mAP) | Blind Deconvolution (Viola (ACC) | Yolov2 (mAP) |
| OWN1    | 60 | 71 | 55.5 | 61.4 | 65 | 67 | 70.8 | 79 |
| Choke   | 55 | 67 | 58.8 | 63.2 | 73.7 | 75 | 85 | 87 |
| SCFace  | 58.2 | 65 | 60.3 | 68 | 62.3 | 65 | 65.2 | 81.2 |
| ETH     | 51 | 63 | 53 | 68 | 55 | 59 | 62 | 75 |
| PRID 2011 | 52 | 70 | 54.9 | 61 | 65 | 66.4 | 75.4 | 82 |
| ILIDS-VID | 60 | 74 | 62.4 | 67 | 64.6 | 68.1 | 70 | 84 |
| ATM1    | 50 | 62 | 58.5 | 64 | 61 | 64 | 70 | 83 |

The proposed framework ensures whether the faces are detected in surveillance scenarios. However, face detection fails in specific scenarios; using Yolo v2, further person identification for criminal investigation could not be achieved. Such failure cases, using Yolo v2 in ATM2, ATM3 due to extreme lighting and specifically in NonATM1 due to excess blur. Also, more failure cases are identified using prevalent Viola Jones in OWN1, ATM1, ATM2, ATM3, and Non-ATM1 datasets compared with Yolo v2.

### 5. Conclusion

The proposed framework adopts blind deconvolution with the modeled Gaussian blur to remove the blur in surveillance videos. The ringing artifact that is caused due to blur is suppressed using DWT. Subsequently, Yolo v2 with Darknet-19 is applied for face detection. The proposed solution has significantly increased the face detection rate for its scope of profile and frontal faces. It works out well with real-world surveillance applications. It can be extended to a person's Re-ID function for a criminal investigation where the crime goes unnoticed and identifies the traffic monitoring of the unhelmeted person to reduce the traffic police pressure. Further, the work can be extended to recognize the stabber's facial expression, involved in any stabbing or stealing events, specifically in finance-based sectors such as ATM, banks, and jewellery shops. In the future, based on the authors' proposals in this work, the research may further focus on a fully automated face recognition system, insensitive to illumination and other blur types with accurate PSF estimation. For a robust face detection and recognition system to create an intelligent surveillance system, specific criteria such as occlusion, pose variations, and small faces can also be considered.

### 6. References

[1] Pei Li; Patrick J and Flynn Loreto Prieto Domingo Mery, 2019 Face Recognition in Low Quality Images: A Survey, ACM Computation Survey, 1, no. 1, arXiv:1805.11519

[2] X. Deng; Y. Shen; M. Song; D. Tao; J. Bu; and C. Chen, 2012 Video-based non-uniform object motion blur estimation and deblurring, Neurocomputing, Elsevier, 86, pp. 170–178
[3] X, Zhu and P, Milanfar, 2013 Removing Atmospheric Turbulence via Space-Invariant Deconvolution, *IEEE Transactions On Pattern Analysis and Machine Intelligence*, IEEE, 35, no. 1, pp. 157-170.

[4] Y. Tai; W, Du; H, S, M, Brown and S, Lin, 2010 Correction of Spatially Varying Image and Video Motion Blur using a Hybrid Camera, *IEEE Transactions On Pattern Analysis and Machine Intelligence*, IEEE, 32, no. 6, pp. 1012 – 1028

[5] L. Shoa; 2008 Simultaneous Coding Artifact reduction and sharpness enhancement for Block based compressed images and videos, *In. Signal Processing: Image Communication*, Elsevier, 23, no. 6, pp. 463-470

[6] C, Thorpe; F, Li; Z, Li; Z, Yu; D, Saunders; and J, Yu, 2013 A Coprime Blur Scheme for Data Security in Video Surveillance, *IEEE Transactions on Pattern Analysis and Machine Intelligence*,IEEE, 35, no. 12

[7] A, Yicheng and Y, Chang, 2017 CNNs for Face Detection and Recognition, *in An2017CNNsFF*

[8] Guanglu Song; Yu Liu; Ming Jiang; Yujie Wang; Junjie Yan; Biao Leng, 2018 Beyond Trade-off: Accelerate FCN-based Face Detector with Higher Accuracy, *International Conference on Computer Vision and Pattern Recognition*, IEEE, pp.7756-7764

[9] R. Girshick, 2015 Fast R-CNN, *Proc. International Conference on Computer Vision*, IEEE, Santiago, Chile, pp. 1440-1448

[10] Ren, S; He, K; Girshick, R, et al, 2017 Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*,IEEE, 39, no. 6, pp. 1137-1149

[11] Redmon, J; and Farhadi, A, 2017 YOLO9000: Better, Faster, Stronger, *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, US*, IEEE, pp. 6517 – 6525

[12] Liu, T; Chen, Z; Liu, S; Zhang, Z; and Shu, J, 2016 Blind image restoration with sparse priori regularization for passive millimeter-wave images, *Journal of Visual Communication and Image Representation*, Elsevier, 40, pp. 58-66

[13] Nobuharu Nakajima, 1993 Blind deconvolution using the maximum likelihood estimation and the iterative algorithm, *Optics Communications*, Elsevier, 100, no. 1-4, pp. 59-66

[14] Joao P. Oliveira; Mario A. T. Figueiredo and Jose M. Bioucas-Dias, 2010 Parametric Blur Estimation for Blind Restoration of Natural Images, Linear Motion and Out-of-Focus, *IEEE Trans. Image Processing*, IEEE, 23, no. 1

[15] Ansari, R; Yadav, H and Jain, A, 2013 A Survey on Blurred Images with Restoration and Transformation Techniques, *International Journal of Computer Applications*, Inderscience, 68, no. 22, pp. 0975-8887

[16] Everingham, M; Van Gool, L and Williams, C.K.I, 2010 The PASCAL Visual Object lasses (VOC) Challenge, *International Journal of Computer Vision*, Springer, 88, no. 2, pp. 303-338

[17] DweepnaGarg, ParthGoel, and Sharnil Pandya, 2019 A Deep Learning Approach for Face Detection using YOLO, *IEEE Conference*, IEEE, Pune, India

**Acknowledgment**

This work has been supported under “Video Analytics Development System DST Project” sponsored by IISC Bangalore.