Deep Learning Based Shadow Detection in Images

Naman Bansal, Akashdeep and Naveen Aggarwal

Abstract Various computer vision applications required Shadow detection and removal for example object tracking and recognition, scene interpretation, and video supervision. Since shadows have similar characteristics as that of the objects shadow pixels can be classified as part of object. This may cause problems such as merging or loss of object, alternation, and misinterpretation of object shape. To deal with this problem, we represent a deep learning based framework to automatically detect shadows in images. Our method learns many significant features automatically using supervised approach in Convolutional Neural Networks (ConvNets). The approach also makes use of drop out to improve results. The proposed methodology is tested on SBU dataset and results are promising.

Keywords Convolutional neural network · Shadow detection · Feature extraction

1 Introduction

In general, a shadow is formed as soon as a light source is blocked by an opaque item. In other words, when an opaque object is positioned between background surface and a light source, it results in variation of illumination in that area, as it does not allow the light to access the adjoining regions of foreground object. Change of illumination is less significant at the outer boundaries as compared to the center region of shadows. Hence, shadow can be classified as: penumbra and umbra region. Umbra can be defined as darker regions in the shadow in which direct light is completely obstructed whereas penumbra are the lighter regions of shadow.

N. Bansal (✉) · Akashdeep · N. Aggarwal
UIET, Panjab University, Chandigarh, India
e-mail: namanbansal07@gmail.com

Akashdeep
e-mail: akashdeep@pu.ac.in

N. Aggarwal
e-mail: navagg@gmail.com

C. R. Krishna et al. (eds.), *Proceedings of 2nd International Conference on Communication, Computing and Networking*, Lecture Notes in Networks and Systems 46, https://doi.org/10.1007/978-981-13-1217-5_37

In various computer vision applications, object detection is an essential phase. A precise evaluation of object shape and size is required so that object tracking or recognition can be performed. The problem gets more complex when shadows are also part of video or image. Because shadow shows similar characteristics as that of the object therefore the shadow region is classified as part of object. This leads to reduction in performance of various computer vision applications for example segmentation, image recognition, and object tracking as shadows cause object merging or occlusion.

Although shadows have many disadvantages, they can be helpful to find the information such as object shape, size and orientation as well as direction and intensity of light source. Shadow detection is required to overcome misclassification problem and to extract useful features.

This paper presents a deep CNN-based approach for detecting shadows in images. CNN automatically extracts features from the input image and uses them to detect shadows. The proposed CNN architecture consists of six layers organized as two pair of convolutional and ReLU layers, followed by a dropout layer and at end a convolutional layer [Conv-ReLU-Conv-ReLU-Dropout-Conv]. It is trained over 2700 images of size 256 * 256.

2 Related Work

Several shadow detection techniques have been presented in the literature. Prati et al. [1] conducted a study on moving shadow detection techniques in which classification was done on the basis of algorithm-based taxonomy. Sanin et al. [2] observed that choosing features has more impact in shadow detection than the selection of any algorithm.

On the basis of [1, 2] shadow detection techniques are divided into major categories of: Spectrum-based and Geometry-based models (Fig. 1).

Geometry-based methods detect shadows by taking advantage of geometric information such as camera location, background scene, and location of light source.

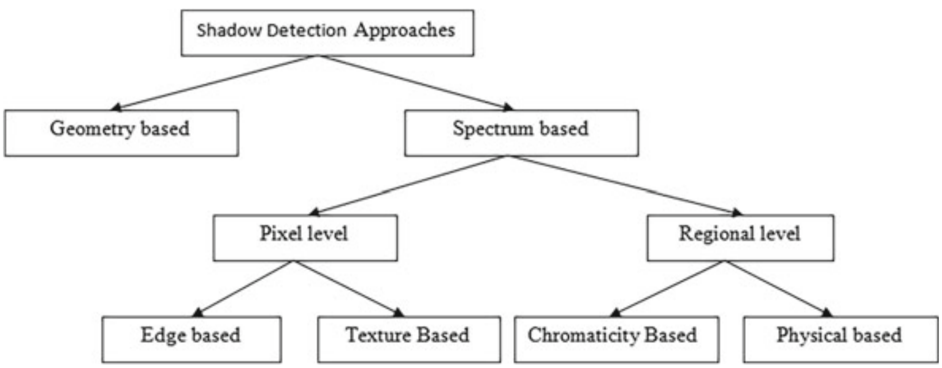

Fig. 1 Classification of shadow detection methods

A three-stage algorithm presented by Chen et al. [3] is used in shadow detection of pedestrians posing vertically. First, to detect shadow features Support Vector Machine was trained and applied on foreground camouflage. Then, the foreground camouflage is partitioned into shadow sub-regions and humans by a linear classifier. A 2D Gaussian filter was used in the final stage to reconstruct the shadow area with the help of background region. A coarse to fine method for shadow removal is presented by Hsieh et al. [4]. A moment-based technique was used at the coarse stage, for estimating rough boundaries among moving object and the shadows. By Gaussian shadow modeling approximation of shadow region is further computed at the fine stage. Algorithms using **Chromaticity based Methods** attempt to use appropriate color spaces to depict the difference in pixel value and appearance when shadow appears. Cucchiara et al. [5] used Hue-Saturation-Value (HSV) color space in shadow recognition. Shadows were detected by calculating rate of change in HSV component of the frame and background referred. Chen et al. [6] proposed YUV color space method to obtain luminance information and to keep chrominance component intact for shadow detection. These techniques are easy to execute but are prone to noise and do not work with strong shadows. **Physical-based Methods** were proposed to model the particular appearance of shadow points according to reflections and illumination. Physical-based color features were used to detect shadows by Huang et al. [7] in outdoor and indoor environments. When object have chromaticity similar to background, physical methods does not give good results. **Edge-based Methods** are very useful in detecting shadows as edges do not vary under changing illumination. Panicker et al. [8] extracted foreground and background mask using Sobel operator. Then these two edge maps were correlated to preserve internal edges of the object. Finally, vertical and horizontal operations were applied to reconstruct object shape. For detecting shadows in indoor sequences, Xu et al. [9] proposed a static edge correlation method. In this method Change Detection Mask (CDM) was generated and canny edge detection was applied to detect shadow regions. **Texture based Methods** usually apply two steps: (i) selection of shadow points with the help of weak shadow detector and then (ii) categorizing these candidates as object or shadow on the basis of texture correlation. The technique presented by Sanin et al. [10] makes use of color features to recognize shadow regions and then applied gradient texture correlation to discriminate them from object. Zhang et al. [11] presented a different algorithm for detection of shadow established on ratio edge, providing that ratio edge is illumination unvaried. As texture based methods have to compute different neighborhood evaluations for every pixel, they are usually slow.

In literature along with the handcrafted features which are used to extract shadows, many approaches based on automatic feature extraction have significant weightage. Shen et al. [12] presented an effective learning based algorithm for detecting shadows in particular image. Local structure of shadow edges was extracted using Convolutional Neural Network (CNN), to increase local consistency in pixel labels. Network architecture consists of seven layers with multiple filters of size 5 * 5. The shadow and bright measures were calculated from the detected shadow edges. Khan et al. [13] proposed a framework which learns the features automatically with multiple convolutional deep neural networks. The architecture consists of a seven-layer network which has alternate convolutional and sub-sampling layers. The given method extracts elements alongside of the object boundaries and at the super-pixel level. Features were learned using a window which is focused at interest points in both the cases. The calculated posteriors were then given to a Conditional Random Field (CRF) model for generating efficient shadow outlines for shadow detection. Author attained highest accuracy of 93.16% on UIUC dataset. Khan et al. [14] presented an algorithm for automatically detecting and removing shadows. The algorithm automatically extracted relevant features using convolutional neural networks in a supervised manner. The extracted features were given to a CRF model for generating efficient mask of shadow. A Bayesian formulation was proposed by using these shadow masks, for correctly extracting shadow matte and removing the shadows. Proposed framework gives good result for both umbra and penumbra regions.

The above approaches require user assistance, or assumptions regarding scene properties or object surface. To overcome these problems, we present a deep learning architecture for shadow detection independent of user assistance or any similar assumptions.

3 Proposed Scheme

For detection of shadows in images we propose a CNN-based framework for automatic detection of required features. The CNN architecture used consists of alternating Convolutional and ReLU layers (Fig. 2) followed by a dropout layer. The CNN layer will help to extract detailed features from raw images which are then fed to Relu layer for activation map. Different filters are used in convolutional layer, which are then convolved with input feature map. The convo layer uses filters of size 5 * 5 for extracting features. Number and dimensions of filters have been set after experimentation. The sequence of convo-relu is repeated twice and its output is given to drop out layer. Dropout layer has been added to improve accuracy within obtained results.

The proposed technique uses 32 filters of size 5 * 5 at each conv layer. ReLU layer provides an activation function, proposed framework uses $f(x) = \text{Max}(0, x)$ activation. The reason for using Max Activation function is to suppress impartial results. Dropout layer randomly ignores nodes to prevent interdependencies between different nodes. This layered architecture allows CNN to learn features at multilevel hierarchy. A dropout of 0.2 has been used in the proposed framework. The final layer of network is convolutional layer in which 1 filter of dimensions 5 * 5 * 32 is used to output the required shadow in image. During training process, framework uses stochastic gradient descent to automatically learn features in supervised manner and uses back-propagation for gradient computation. Once the network is trained, it takes 256 * 256 grayscale image as an input and processes it to provide shadow in that image.

Input: Image of size $[256 \times 256 \times 1]$ is provided to proposed architecture

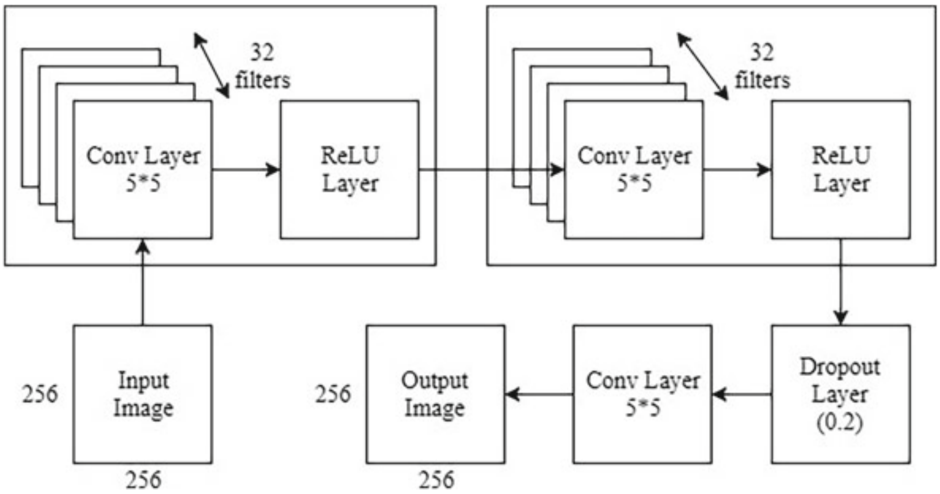

Fig. 2 Proposed CNN architecture for shadow detection

Table 1 Details of each layer in proposed CNN architecture

Layer type	Input size	Filter size	No. of filters	Stride	Padding value
conv	256 * 256 * 1	5 * 5 * 1	32	1	2
Relu	256 * 256 * 32	–	–	–	–
conv	256 * 256 * 32	5 * 5 * 32	32	1	2
Relu	256 * 256 * 32	–	–	–	–
dropout	256 * 256 * 32	–	–	–	–
conv	256 * 256 * 32	5 * 5 * 32	1	1	2

Conv layer: It will calculate a dot product between filter weights and region of input image. Output of this layer is [256 * 256 * 32] feature map.

ReLU layer: It will apply an activation function, max (0, x) thresholding at zero. This will not change the size of feature map.

Dropout layer: It will randomly select the nodes to be dropout with a probability of 0.2.

In this architecture, CNN is trained over 2700 images for batch size 1 with 50 epochs and learning rate of 0.01 and. Table 1 provides details of each layer used in the architecture.

4 Results and Discussion

4.1 Dataset

SBU Shadow Dataset: The dataset consists of 4089 images along with their ground truth. Only 3600/4089 images are used. For training purpose 2700 images are used and 900 images are used for the purpose of testing. The dataset contains variable size RGB images which are resized to 256 * 256 grayscale images.

4.2 Results

To evaluate the superiority and effectiveness of the proposed method, Qualitative results obtained by proposed framework are shown on three different types of images in Figs. 3, 4 and 5. First column of Figs. 3, 4 and 5 consists of input image having shadows, second column consists of the expected shadow results, and column third, shows the obtained results after the implementation of proposed CNN framework. The results show that our proposed framework detects shadows successfully. Results are equally good on satellite images (Fig. 3), outdoor scenes (Fig. 4) and images having shadows of humans (Fig. 5). As can be seen in Fig. 3c the obtained results are very close to the expected results showing high accuracy in shadow detection by proposed framework. The achieved results as shown in Fig. 4, depicts the presence of some white spots along with the shadow. This is due to the similar characteristics of the region and shadow. Some shadow information is lost in Fig. 5, but result are parallel to the expected results of the image consisting shadow.

(a) InputImage (b) Expected Result (c) Achieved Results

Fig. 3 Results of proposed approach on image#10 of SBU dataset

(a) Input Image (b) Expected Result (c) Achieved Results

Fig. 4 Results of proposed approach on image#17 of SBU dataset

(a) Input Image (b) Expected Result (c) Achieved Results

Fig. 5 Results of proposed approach on image#35 of SBU dataset

5 Conclusion

This research paper has proposed a deep learning based architecture for solving the problem of shadow detection in images. An automatic feature learning CNN approach is presented to extract the most significant features from a single image. CNN architecture consists of multiple hidden layers along with input layer and an output layer. Hidden layers are either convolutional, ReLU or dropout layer. The approach uses a 6-layer CNN framework [Conv-ReLU-Conv-ReLU-Dropout-Conv] on single images for detection of shadows. 32 filters of size 5 * 5 are used in each convolutional layer and a dropout of 0.2 is applied in the framework. Qualitative evaluation of results obtained through proposed method shows that it performed well in various types of single images such as in outdoor and aerial images. The approach is required to be tested quantitatively for further approval. The proposed approach can be further improved by developing more deep CNN architectures for precise results. The approach can be further tested on videos and other benchmark datasets for further testing.

References

1. A. Prati, I. Mikic, M.M. Trivedi, R. Cucchiara, Detecting moving shadows: algorithms and evaluation. IEEE Trans. Pattern Anal. Mach. Intell. **25**(7), 918–923 (2003)
2. A. Sanin, C. Sanderson, B.C. Lovell, Shadow detection: a survey and comparative evaluation of recent methods. Pattern Recogn. **45**(4), 1684–1695 (2012)
3. C.-C. Chen, J.K. Aggarwal, Human shadow removal with unknown light source, in *IEEE 2010 20th International Conference on Pattern Recognition (ICPR)* (2010), pp. 2407–2410
4. J.-W. Hsieh, W.-F. Hu, C.-J. Chang, Y.-S. Chen, Shadow elimination for effective moving object detection by Gaussian shadow modeling. Image Vis. Comput. **21**(6) (2003), 505–516
5. R. Cucchiara, C. Grana, M. Piccardi, A. Prati, Detecting moving objects, ghosts, and shadows in video streams. IEEE Trans. Pattern Anal. Mach. Intell. **25**(10), 1337–1342 (2003)
6. C.-T. Chen, C.-Y. Su, W.-C. Kao, An enhanced segmentation on vision-based shadow removal for vehicle detection, in *IEEE 2010 International Conference on Green Circuits and Systems (ICGCS)* (2010), pp. 679–682
7. J.-B. Huang, C.-S. Chen, Moving cast shadow detection using physics-based features, in *IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009* (2009), pp. 2310–2317
8. J. Panicker, Vinitha, M. Wilscy, Detection of moving cast shadows using edge information, in *IEEE 2010 The 2nd International Conference on Computer and Automation Engineering (ICCAE)*, vol. 5, pp. 817–821 (2010)
9. D. Xu, X. Li, Z. Liu, Y. Yuan, Cast shadow detection in video segmentation. Pattern Recogn. Lett. **26**(1), 91–99 (2005)
10. A. Sanin, C. Sanderson, B.C. Lovell, Improved shadow removal for robust person tracking in surveillance scenarios, in *IEEE 2010 20th International Conference on Pattern Recognition (ICPR)* (2010), pp. 141–144
11. W. Zhang, X.Z. Fang, X.K. Yang, Q.MJ. Wu, Moving cast shadows detection using ratio edge. IEEE Trans. Multimed. **9**(6), 1202–1214 (2007)
12. L. Shen, T.W. Chua, K. Leman, Shadow optimization from structured deep edge detection, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 2067–2074
13. S.H. Khan, M. Bennamoun, F. Sohel, R. Togneri, Automatic feature learning for robust shadow detection, in *2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2014), pp. 1939–1946
14. S.H. Khan, M. Bennamoun, F. Sohel, R. Togneri, Automatic shadow detection and removal from a single image. IEEE Trans. Pattern Anal. Mach. Intell. **38**(3), 431–446 (2016)