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

# Performance Analysis of Object Recognition using Wavelet Transform and Machine Learning Approaches

**Meghna Kanaujia[1], Prof. Suresh. S. Gawande[2]**
M. Tech. Scholar, Department of Electronics and Communication, RKDF College of Engineering, Bhabha University, Bhopal[1]
Guide, Department of Electronics and Communication, RKDF College of Engineering, Bhabha University, Bhopal[2]

***Abstract:*** Because of article discovery's cozy relationship with video examination and picture understanding, it has drawn in much exploration consideration lately. Conventional item discovery techniques are based on carefully assembled elements and shallow teachable models. Their presentation effectively deteriorates by building complex groups which consolidate different low-level picture highlights with significant level setting from object identifiers and scene classifiers. With the quick advancement in profound learning, all the more integral assets, which can learn semantic, undeniable level, further highlights, are acquainted with address the issues existing in customary designs. In this paper, an approach for object recognition based on wavelet transform is presented. This approach decomposes the input image into sub-bands by using the multiresolution analysis, Discrete Wavelet Transform (DWT). As each sub-band in the decomposed image contains useful information about the object, the mean of each sub-band is considered as features. This approach is tested on Object Image database. All the objects are considered for the classification based on one machine learning approach and calculate MSE and PSNR.

***Index Terms*** – MSE, PSNR, Machine Learning, Discrete Wavelet Transform.

## I. INTRODUCTION

Over the past years domains like image analysis and video analysis has gained a wide scope of applications. CV and AI are two main technologies dominating technical society. Technologies try to depict the biology of human. Human vision is the sense through which a perception of outer 3D world is perceived. Human Intelligence is trained over years to distinguish and process scene captured by eyes. These intuitions acts as a crux to budding new technologies. Rich resource is now accelerating researchers to excavate more details form the images. These developments are due to stateof the-art methods like CNN. Applications from Google, Facebook, Microsoft, and Snapchat are all results of tremendous improvement in Computer vision and Deep learning [1].

During time, the vision-based technology has transformed from just a sensing modality to intelligent computing systems which can understand the real world. Computer vision applications like vehicle navigation, surveillance and autonomous robot navigation find Object detection and tracking as important challenges. For tracking vehicles and other real word objects, video surveillance is a dynamic environment. In this paper, efficient algorithm is designed for object detection and tracking for video Surveillance in complex environment. Object detection and tracking goes hand in hand for computer vision applications. Object detection is identifying object or locating the instance of interest in-group of suspected frames [2]. Object tracking is identifying trajectory or path; object takes in the concurrent frames. Image obtained from dataset is, collection of frames. Basic block diagram of object detection and tracking is shown in Fig. 1. Data set is divided into two parts. 80 % of images in dataset are used for training and 20 % for testing. Image is considered to find objects in it by using algorithms CNN and YOLOv3. A bounding box is formed across object with Intersection over union (IoU) > 0.5. Detected bounding box is sent as references for neural networks aiding them to perform Tracking. Bounded box is tracked in concurrent frames using Multi Object Tracking (MOT). Importance of this research work is used to estimate traffic density in traffic junctions, in autonomous vehicles to detect various kinds of objects with varying illumination, smart city development and intelligent transport systems [3].

Intermittence: Division of the advanced picture predicated because of abrupt changes on power. For instance, edge location, point identification and line recognition.

Similitude: Dividing the computerized picture into district dependent on predicated on set of predefined rules. For instance, thresholding, locale developing, area parting and blending. The picture division strategies are classified as [4]:

An edge is characterized as limits of items or unexpected change in a picture which isn't in a ceaseless structure that assists with distinguishing and recognize the articles in a given picture [5]. The primary point behind edge discovery [6, 7] strategy is to recognize and find the focuses in an advanced picture at which force of the picture changes. Among different strategy of different edge recognition procedure Canny [8] administrator gives preferred yield over Sobel [9], Prewitt [10], and Laplacian technique. In Thresholding strategy grayscale worth of the picture is looked at with predefined worth of the limit. Assuming grayscale worth of the information pixel is huge, yield worth of that pixel becomes 1 or, more than likely 0.

Picture procurement digitizes the picture caught by camera. Picture improvement is the method involved with controlling a picture so the outcomes are more appropriate for explicit applications. Picture rebuilding works on an appearance of a picture

which watches out for probabilities model of picture corruption Morphological cycles are the apparatuses of removing picture parts that are helpful in the portrayal and show of a picture.

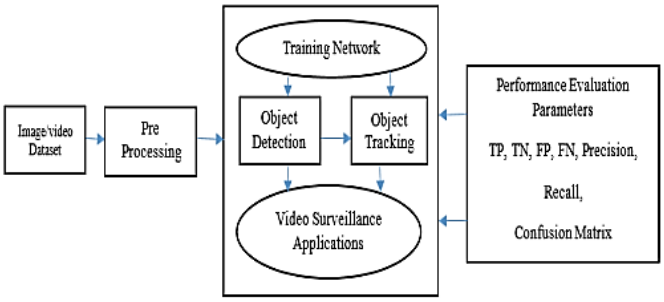


Fig. 1: Block Diagram of Object Detection and Tracking

Image segmentation is the most difficult ask in digital image processing which separates objects from the background. Representation makes the decision whether to represent data as boundary or as a complete region. Recognition is the process that assigns label to an object based on information provided by its descriptor.

## II. PROPOSED ALGORITHM

Let the phantom spatial hyperspectral information block be signified by I ∈ RM×N×D, where I is the first information, M is the width, N is the stature, and D is the quantity of otherworldly groups/profundity. Each HSI pixel in I contains D ghastly measures and structures a one-hot mark vector Y = (y1, y2,..., yC) ∈ R1×1×C, where C addresses the land-cover classifications. In any case, the hyperspectral pixels display the blended land-cover classes, presenting the high intraclass fluctuation and interclass likeness into I. It is of extraordinary test for any model to handle this issue. To eliminate the unearthly repetition first, the customary head segment investigation (PCA) is applied over the first HSI information (I) along ghastly groups. The PCA lessens the quantity of phantom groups from D to B while keeping up with similar spatial measurements (i.e., width M and stature N). We have decreased just ghostly groups with the end goal that it saves the spatial data which is vital for perceiving any article. We address the PCA diminished information 3D shape by X ∈ RM×N×B , where X is the adjusted contribution after PCA, M is the width, N is the stature, and B is the quantity of unearthly groups after PCA.

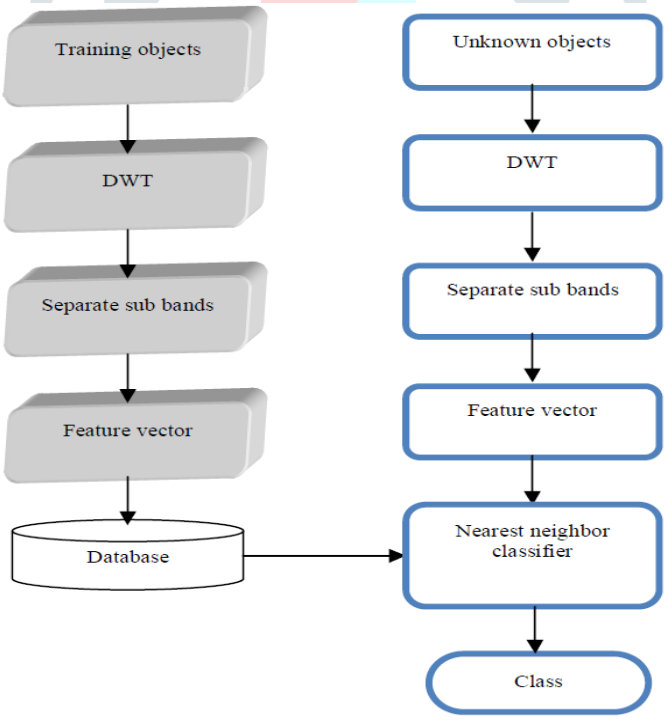


Fig. 2: Proposed model that integrates 3-D and 2-D convolutions for HSI classification

The absolute number of created 3-D-patches (n) from X is given by (M − S + 1) × (N − S + 1). Consequently, the 3-D-fix at area (α, β), signified by Pα,β, covers the width from α − (S − 1)/2 to α + (S − 1)/2, range from β − (S − 1)/2 to β + (S − 1)/2, and all B ghastly groups of PCA diminished information solid shape X. In 2-D-CNN, the information are convolved with 2-D portions. The convolution occurs by figuring the amount of the spot item between input information and piece. The part is strided over the info information to cover full spatial measurement.

### Classification Approach

The word ML and AI are driving the Figures by a huge degree. This consistent ascent in the notoriety of AI is a direct result of its rising use in our everyday lives. It is these days being utilized in different gadgets and machines just as devices. Be that as it may, everyone is still are careful about it. Along these lines, to get rid of such fantasies, let us view the short history of ML. ML can be characterized as a cycle of contributing information to the PC frameworks such that the PC will gain proficiency with the capacity

to measure and play out the action in the future without being unequivocally modified or being taken care of with comparable or additional information. In case PCs are furnished with the capacity to think, they become more intelligent and along these lines simpler to utilize. Their usefulness will increment by an enormous degree, and they become a necessary resource for mankind. ML can be utilized in practically every one of the fields of epistemology. The present moment, it is being utilized in regions, for example, cheminformatics, computational life structures, gaming, versatile sites, regular language preparing, robot development and motion, clinical finding, arrangement mining, conduct examination, etymology, interpretation, misrepresentation identification, and so forth The rundown goes on. It is inspected that the acoustics signals have been gained from the rotating machines and they have been utilized with the wavelets for the helpful determination of the outcomes.

**DT:-**

A DT is a choice help instrument that utilizes a tree-like model of choices and their potential results, including chance occasion results, asset expenses, and utility. It is one method for showing a calculation that just holds back restrictive control explanations. DT are ordinarily utilized in tasks research, explicitly in choice examination, to assist with recognizing a technique probably going to arrive at an objective, but at the same time are a well known device in ML.

**GB:-**

GB calculation is one of the most remarkable calculations in the field of AI. As we realize that the blunders in AI calculations are extensively characterized into two classifications for example Inclination Error and Variance Error. As inclination supporting is one of the helping calculations limiting predisposition mistake of the model is utilized.

**SVM:-**

In ML, SVM are directed learning models with related learning calculations that examine information for grouping and relapse examination.

To isolate the two classes of data of interest, there are numerous conceivable hyperplanes that could be picked. Our goal is to find a plane that has the greatest edge, for example the greatest distance between data of interest of the two classes. Boosting the edge distance gives some support so future information focuses can be grouped with more certainty.

**Discrete Wavelet Transform**

Multiresolution analysis (MRA) is a characteristic feature of sub-band and it is used for better spectral representation of the signal. In MRA, the signal is decomposed for more than one DWT level known as multilevel DWT. It means the low-pass output of first DWT level is further decomposed in a similar manner in order to get the second level of DWT decomposition and the process is repeated for higher DWT levels [7]. Some algorithms have been suggested for computation of multilevel DWT. One of the most important algorithm are pyramid algorithm (PA), this algorithm are proposed Mallet for parallel computation of multilevel DWT. PA for 1-D DWT is given by

$$Y_l^{\,j}(n) = \sum_{i=0}^{k-1} h(i) Y_l^{\,j-1}(2n-i) \overset{\|O}{\underset{\|}{C}} \tag{1}$$

$$Y_h^{\,j}(n) = \sum_{i=0}^{k-1} g(i) Y_h^{\,j-1}(2n-i) \tag{2}$$

Where $Y_l^{\,j}(n)$ is the n[th] low-pass sub band component of the j[th] DWT level and $Y_h^{\,j}(n)$ is the n[th] high-pass sub band component of the j[th] DWT level. Two-dimensional signal, such as images, are analyzed using the 2-D DWT as shown in Figure 1. Currently 2-D DWT is applied in many image processing applications such as image compression and reconstruction [8]. The 2-D DWT is a mathematical technique that disintegrates an input image in the multi-resolution frequency space [9]. The 2-D DWT disintegrate an enter photo into four sub bands called low-low (LL), low-excessive (LH), high-low (HL) and high-excessive (HH) sub band.

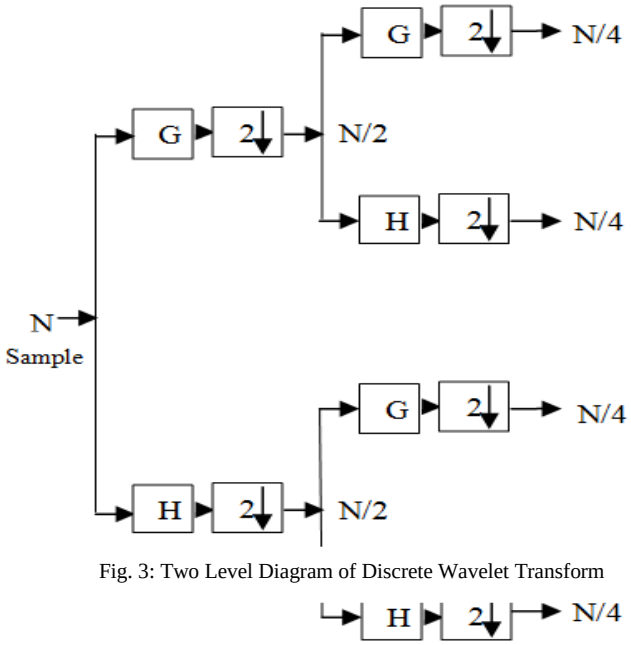


Fig. 3: Two Level Diagram of Discrete Wavelet Transform

## III. SIMULATION RESULTS

As shown in table 1 the error, smoothness, uniformity and processing time are obtained from the proposed image segmentation using fuzzy canny method algorithm.

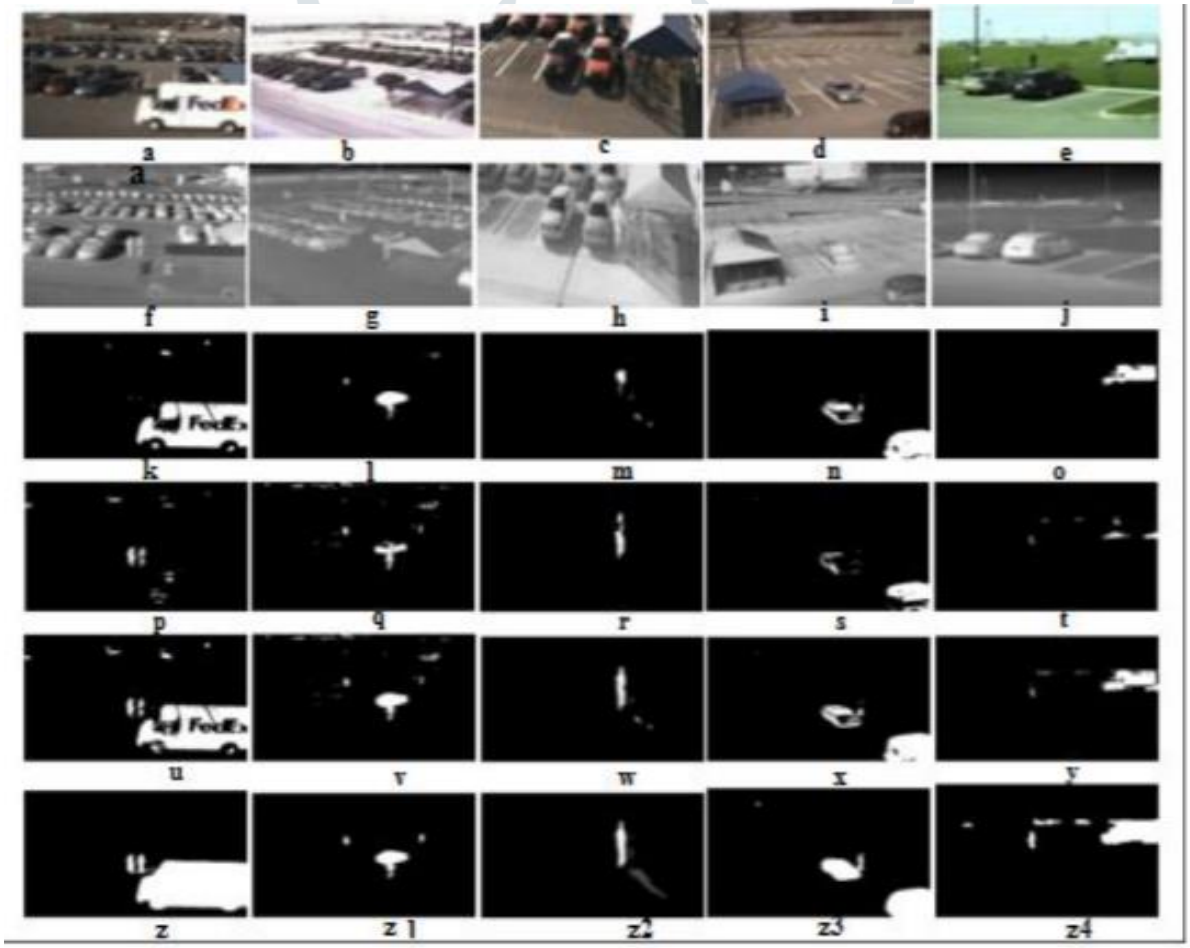

Fig. 4: Results using Proposed Method on different INO Datasets

Fig. 5: Pre-processing of thermal images using Gamma correction

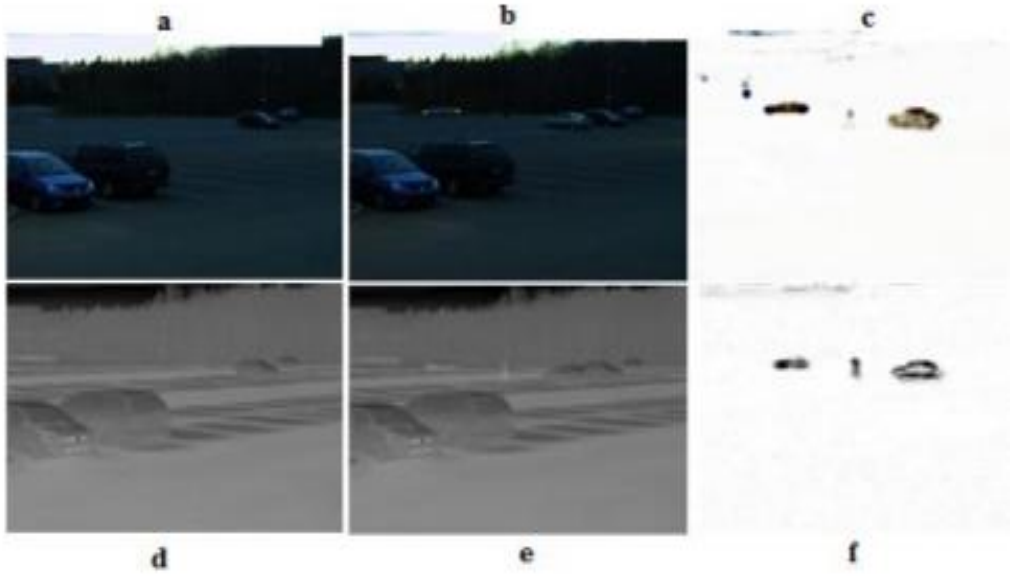


Fig. 6: MOD using Structural Similarity information. Row 1: a) Background frame, b) Foreground frame and c) SSIM map of this two images with background frame as reference frame in visible domain. Row 2: d) Background frame, e) Foreground frame, f) SSIM map of these two images with background frame as reference frame in Thermal domain.

Image captured in VSI and TIR domain are pre-processed for image enhancement. The pre-processing methods: Gamma correction is used to improve contrast. Due to Contrast improvement, foreground objects are clearly distinguished from the background.

Table 1: Results of simple background & simple foreground Image

| Image | NAE | MSE | PSNR (dB) |
|---|---|---|---|
| Airplane Image | 0.3128 | 2.3493 | 50.4759 |
| Eagle Image | 0.3906 | 3.3464 | 48.9396 |
| Bird Image | 0.2543 | 1.7663 | 51.7149 |

Table 2: Results of Textured foreground & simple foreground Image

| Image | NAE | MSE | PSNR (dB) |
|---|---|---|---|
| Island Image | 4.300 | 22.514 | 40.660 |
| Iceberg Image | 0.4068 | 3.3977 | 48.873 |
| Duck Image | 1.241 | 9.680 | 44.326 |

## IV. CONCLUSION

In this paper another division procedure is survey utilizing morphological activities. An automated approach for object recognition based on DWT and one nearest neighbor classifier is presented. The proposed approach uses all sub-band energy of DWT as feature vector to represent the COIL database objects. The whole COIL database is separated into six training sets based on object rotations. The experimental results show that the training objects of 10 degree rotation gives the better classification accuracy when compared with the other degrees.

## REFERENCES


[1] Jaehoon Lee, Myungjin Cho and Min-Chul Lee, “3D Visualization of Objects in Heavy Scattering Media by using Wavelet Peplography”, IEEE Access 2022.

[2] J. Lee, M.-C. Lee, and M. Cho, ``Three-dimensional photon counting imaging with enhanced visual quality," *J. Inf. Commun. Converg. Eng.*, vol. 19, no. 3, pp. 180_187, 2021.

[3] J. Lee, M. Cho, and M.-C. Lee, ``Three-dimensional photon counting optical encryption with enhanced visual quality and security level," *IEEE Access*, vol. 9, pp. 128862_128869, 2021.

[4] J. Lee, M. Kurosaki, M. Cho, and M.-C. Lee, ``Noise reduction for photon counting imaging using discrete wavelet transform," *J. Inf. Commun. Converg. Eng.*, vol. 19, no. 4, pp. 276_283, 2021.

[5] M. Hupfel, A. Y. Kobitski, W. Zhang, and G. U. Nienhaus, ``Wavelet- based background and noise subtraction for _uorescence microscopy images," *Biomed. Opt. Exp.*, vol. 12, no. 2, pp. 969_980, 2021.

[6] Sunit Vaidya, Naisha Shah, Niti Shah and Prof. Radha Shankarmani, “Real-Time Object Detection for Visually Challenged People”, International Conference on Intelligent Computing and Control Systems (ICICCS 2020).

[7] Ganbayar Batchuluun, Jin Kyu Kang, Dat Tien Nguyen, Tuyen Danh Pham, Muhammad Arsalan And Kang Ryoung Park, “Deep Learning-based Thermal Image Reconstruction and Object Detection”, IEEE Access, 2020.

[8] G. Batchuluun, Y. W. Lee, D. T. Nguyen, T. D. Pham, and K. R. Park, "Thermal Image Reconstruction Using Deep Learning", IEEE Access, vol. 8, pp. 126839-126858, 2020.

[9] G. Batchuluun, H. S. Yoon, D. T. Nguyen, T. D. Pham, and K. R. Park, “A Study on the Elimination of Thermal Reflections,” IEEE Access, vol. 7, pp. 174597–174611, 2019.

[10] D. Bianli, R. Xiaokang, and R. Jie, “CNN-based Image Super-Resolution and Deblurring,” in Proc. Association for Computing Machinery, Wuhan, China, 29–31 Oct. 2019, pp. 70–74.

[11] X. Zhang, F. Wang, H. Dong, and Y. Guo, "A Deep Encoder-Decoder Networks for Joint Deblurring and Super-Resolution," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, Calgary, Canada, 15–20 Apr. 2018, pp. 1448–1452.

[12] F. Albluwi, V.A. Krylov, and R. Dahyot, “Image Deblurring and Super-Resolution Using Deep Convolutional Networks,” in Proc. IEEE 28th International Workshop on Machine Learning for Signal Processing, Aalborg, Denmark, 17–20 Sept. 2018, pp.1–6.

[13] B. Oswald-Tranta, “Temperature Reconstruction of Infrared Images with Motion Deblurring,” J. Sens. Sens. Syst., vol. 7, pp. 13–20, 2018.

[14] H. Choi and J. Jeong, ``Despeckling images using a preprocessing _filter and discrete wavelet transform-based noise reduction techniques," *IEEE Sensors J.*, vol. 18, no. 8, pp. 3131_3139, Apr. 2018.

[15] Y. Matsushita, H. Kawasaki, S. Ono, and K. Ikeuchi, "Simultaneous Deblur and Super-Resolution Technique for Video Sequence Captured by Hand-Held Video Camera," in Proc. IEEE International Conference on Image Processing, Paris, France, 27–30 Oct. 2014, pp. 4562–4566.

[16] H. Linyang, L. Gang, and L. Jinghong, “Joint Motion Deblurring and Superresolution from Single Blurry Image,” Math. Probl. Eng., vol. 2015, pp. 1–10, 2015.