IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 22 July 2024, accepted 19 August 2024, date of publication 22 August 2024, date of current version 2 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3447716

RESEARCH ARTICLE

An Efficient Low Complex-Functional Link Artificial Neural Network-Based Framework for Uneven Light Image Thresholding

TAPASWINI PATTNAIK[1,2], (Member, IEEE), PRIYADARSHI KANUNGO[3], (Senior Member, IEEE), PRABODH KUMAR SAHOO[4], (Member, IEEE), TEJASWINI KAR[5], (Senior Member, IEEE), PRINCE JAIN[4], MOHAMED S. SOLIMAN[6], (Senior Member, IEEE), AND MOHAMMAD TARIQUL ISLAM[7], (Senior Member, IEEE)

[1]Department of Electronics and Telecommunication Engineering, Biju Patnaik University of Technology, Rourkela, Odisha 769015, India
[2]Department of Electronics and Communication Engineering, C. V. Raman Global University, Janla, Bhubaneswar, Odisha 752054, India
[3]Department of Electronics and Telecommunication Engineering, Driems University, Tangi, Cuttack, Odisha 754022, India
[4]Department of Mechatronics Engineering, Parul Institute of Technology, Parul University, Vadodara, Gujarat 391760, India
[5]School of Electronics Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha 751024, India
[6]Department of Electrical Engineering, College of Engineering, Taif University, Taif 21944, Saudi Arabia
[7]Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, UKM, Bangi, Selangor 43600, Malaysia

Corresponding authors: Prabodh Kumar Sahoo (sahooprabodhkumar@gmail.com) and Mohammad Tariqul Islam (tariqul@ukm.edu.my)

This research was funded by the Universiti Kebangsaan Malaysia Research Grant through the Dana Padanan Kolaborasi (DPK) under the grant number DPK-2022-006. Also, the research was funded by Taif University, Saudi Arabia, Project No. (TU-DSPP-2024-11).

ABSTRACT The most popular technique for converting two-class images into binary images is thresholding. However, thresholding methods tend to perform poorly when dealing with images affected by uneven lighting. To address this issue, local thresholding techniques are commonly used. While pixel-based local thresholding methods can achieve high accuracy, they are computationally complex. Window-based local thresholding presents challenges in selecting the initial window and determining the criterion function for dividing the image into smaller versions. In this study, a novel technique is proposed to improve the effectiveness of binarizing images with uneven lighting. The proposed method is based on a low-complexity functional neural network model (LC-FLANN) to estimate an image's illumination surface. The effectiveness of the proposed technique has been evaluated using five widely used uneven lighting image binarization techniques and various uneven light image variations. The results show that the proposed approach outperforms other alternatives in both qualitative and quantitative metrics. It achieved an average F-Measure score of 0.97, a Jaccard Index (JI) score of 0.95, and a Percentage of Misclassification Error (PME) 1.42%, demonstrating superior overall performance.

INDEX TERMS Image segmentation, adaptive thresholding, functional link artificial neural network, binarization.

LIST OF SYMBOLS AND ABBREVIATIONS

FLANN	Functional Link Artificial Neural Network.
JI	Jaccard Index.
ABC	Artificial Bee Colony.
IBAA	Image Binarization using Adaptive Averages.
GT	Ground Truth.
SVM	Support Vector Machine.
TSP	Training Sample Points.
G_v	Vertical Gradient.
$T_s(i,j)$	Threshold Surface.
LMS	Least Mean Square.

The associate editor coordinating the review of this manuscript and approving it for publication was Mu-Yen Chen.

$g(max)$	Maximum Intensity of Input Image.
h_i	Probability of Occurrence of Gray Level.
$\sigma^2_{Bi}(t)$	Between class variance.
k_0	Constant.
TP	True Positive.
LC	Low Complex.
PME	Percentage of Misclassification Error.
LIM	Lorentz Information Measure.
VO	Visual Odometry.
EM	Expectation Maximization.
PGA	Parallel Genetic Algorithm.
IBAW	Image Binarization using Adaptive Window.
P_r	Precision.
I_s	Illumination Surface.
I_N	Normalization Surface.
I_0	Observation Matrix.
$I_B(x, y)$	Binarized Image.
$\sigma^2_{wi}(t)$	Within Class Variance.
M_h	Horizontal Mask.
FP	False Positive.
$F1$	F-measure.
GMM	Gaussian Mixture Model.
ML	Machine Learning.
EB	Edge-Based.
T_{opt}	Optimal Threshold.
TST	Two-Step Thresholding.
G_h	Horizontal Gradient.
R_e	Recall.
Io	Original Image.
ε	Error.
δ	Learning Rate.
e	Entropy.
K	Sample Points.
M_v	Vertical Mask.
FN	False Negative.

I. INTRODUCTION

In any computer vision application, image segmentation plays a crucial role. The success of image segmentation determines how well objects can be detected, recognized, and tracked. Among the various image segmentation techniques, thresholding stands out as a simple and effective method for real-time applications. This technique involves dividing the image into two categories—background and object—making it ideal for two-class image binarization. The thresholding process can be grouped into global and local approaches. Otsu's method [1] is a popular algorithm for global thresholding in image binarization, aiming to maximize the variance between classes and minimize within-class variation. Meanwhile, Cai et al. [2] enhanced the Otsu approach by iteratively calculating the threshold using the image's flat distribution. Additionally, Sha et al. [3] presented a 2D Otsu thresholding approach based on the combined histogram of average and median characteristics to improve the accuracy of global thresholding. However, in the presence of uneven lighting, the bimodality of the gray level distribution vanishes, leading to a drastic decrease in the performance of global thresholding. To address this, Ma and Gheng [4] developed a quick and enhanced 2D Otsu's method based on the integral image and adaptive genetic algorithm to binarize unequal light and low contrast images. It's important to note that this method is effective for low-contrast picture binarization but may not work as well for other variations of uneven light image binarization. Consequently, local thresholding approaches have been developed to tackle the issue of uneven lighting.

Local thresholding techniques can be categorized into pixel-based and window-based approaches. Xing et al. [5] introduced an image partitioning method based on 2D Otsu for dealing with noisy and unevenly lit images. Niblack [6] determined a threshold for each window using the mean and standard deviation, but this method increases computational complexity with the size of the window and image. Bradley et al. [7] proposed a local binarization technique that reduces computational complexity by using an integral image approach, which is not affected by window size, but it doubles the memory requirement due to integral images. Zhao et al. suggested the Gaussian Mixture Model (GMM), which calculates each pixel's threshold within its neighbourhood and binary codes each pixel by comparing it to the local thresholding of a window-based area. It was noted that window-based local thresholding has less computational complexity than pixel-based methods. The primary objective of thresholding, regardless of being image partition-based or window-based, is to identify an area with both foreground and background distributions and minimal lighting effects. PAN [9] proposed a local image binarization technique using the ABC algorithm to divide the grayscale image into blood vessels and background by determining a suitable threshold.

To detect bimodality in each sub-image, Huang et al. [10] proposed using a criterion function called Lorentz information measure (LIM). The pyramid method is employed to divide the image into smaller sub-images. Next, Otsu's thresholding technique is applied to the sub-images that meet the LIM requirements. All the thresholded sub-images are then combined to create the final binarized image. The size of the chosen sub-image determines the effectiveness of this method for binarizing images with uneven lighting.

To determine the threshold, Bogiazis and Papadopoulos [11] developed a method for local thresholding based on fuzzy inclusions and entropy criteria. The window size required to binarize an image varies depending on the image. Proper bimodal criteria is crucial for selecting the relevant sub-image from any given image.

Kanungo et al. [12], [13] have devised a window-based methodology to enhance accuracy through adaptive window selection using window merging and window expansion approaches. The window selection criteria are based on feature entropy and entropy, and a Parallel Genetic Algorithm (PGA) is used to binarize each window. Additionally, the window expansion method has been employed to improve

precision. However, computational complexity is a significant issue with both window merging and window expansion based on local thresholding, and the accuracy of the window merger's determination is important.

To address this, Pattnaik et al. proposed employing the bimodal criterion of GMM-based [14] and normalized mean difference-based approach [15] to adaptively select windows for partitioning pictures and reduce complexity.

Various adaptive methods have been developed to create a threshold surface for identifying lighting patterns and binarizing each pixel based on this surface. Yanowitz et al. [16] utilized edge information and intensities to establish a threshold surface, but the computational complexity of this approach is significant. To reduce this complexity, Blayvas et al. [17] created a threshold surface using a multi-resolution technique. Meanwhile, Yazid and Arof [18] used gradients to develop an inverse threshold surface, which is adjusted for binarizing images with uneven lighting. Cai and Miklavcic [19] proposed precise background surface estimates by considering global and local edge points to address intensity in-homogeneities. Recent studies have compared machine learning techniques for enhancing metamaterial absorbers in terahertz applications and predicting dielectric properties of MoS2 nanofiller-reinforced epoxy composites [20], [21], highlighting their effectiveness in improving material performance. Sahoo et al. [22] enhanced the VGG-19 network for precise moving object detection in complex video scenes. Other works have used machine learning to predict metamaterial microwave absorption and develop THz metamaterial absorbers for biomedical applications [23], [24], [25], showing the potential of these methods in material science and biomedical sensing. He and Schomaker [26] proposed a novel method using SVM to binarize uneven-lighting images. Furthermore, Dehuri et al. [27] developed a deep learning-based method to enhance non-uniformity in images by training the network. Some of the applications of image segmentation techniques in challenging areas include the following:

Cheng et al. [34] developed a Holistic Prototype Activation (HPA) network to address issues with incorrect segmentation borders caused by network architecture or training. The HPA Network acts as a network without training, featuring a Cross-Referenced Decoder (CRD) for multilayer feature aggregation and a Prototype Activation Module (PAM) for creating activation modules for feature mapping. This method is used for video object segmentation, weak-level, and zero-shot tasks.

Lang et al. [35] introduced Few-shot Segmentation (FSS), a technique designed to identify objects from annotated samples, even if they belong to unseen classes. However, FSS has limitations in dealing with incomplete objects and ambiguous boundaries. To overcome these challenges, a new framework called Divide-and-Conquer Proxies (DCP) has been developed for video object segmentation. This framework divides the segmentation mask into regions with different properties and uses effective mask average pooling.

Lang et al. [36] presented a theoretically simple and easy-to-implement methodology to address a number of enduring problems in few-shot segmentation (FSS). The Progressive Parsing Module and the Commonality Distillation Module, two novel components introduced into the PCNet architecture, have greatly improved segmentation performance. Comprehensive tests using industry-standard benchmarks showed that PCNet outperforms earlier FSS techniques both quantitatively and qualitatively. The results demonstrate that in few-shot segmentation, PCNet sets a new state-of-the-art. However, the research does not specifically address the performance of PCNet under non-uniform lighting circumstances, which remains a challenging aspect without direct trials and data analysis.

The Base and Meta (BAM) architecture for Few-Shot Segmentation (FSS) was developed by Lang et al. [37] to address biases in data or models that are inherent and challenging to mitigate. They introduced an auxiliary branch to serve as a base learner to enhance the traditional FSS framework. The purpose of this auxiliary branch is to identify items belonging to basic classes and does not require segmentation. At the same time, the main FSS framework acts as the meta-learner, focusing on learning to segment new classes. To achieve accurate segmentation predictions, the base learner and meta-learner collaborate to produce coarse segmentation results that are effectively combined. Adjustment factors are used to measure the differences between the support (base class) and query (novel class) image pairings, considering both stylistic and appearance variations, to enhance the robustness of the meta-learner. By considering scene differences and improving segmentation performance, this approach aims to enhance the model's ensemble predictions. The paper does not specifically address the performance of the BAM framework under non-uniform illumination conditions. Due to its dual-branch approach and flexibility in handling variations, the novel design of the BAM framework suggests that it may be somewhat resilient to non-uniform light conditions. However, its actual performance in such scenarios needs to be demonstrated through focused testing. To ensure reliable operation, additional testing and potential adjustments may be necessary if managing uneven lighting is a critical requirement.

The main challenges of local thresholding approaches are: (i) the initial choice of window size, (ii) the partitioning approach, (iii) computational complexity, and (iv) the use of an appropriate bimodal criteria function. Similarly, for machine learning and deep learning, there is a requirement for very large training image sets. Hence, there is an opportunity for enhancements in uneven light thresholding methods. Building upon previous approaches [20], [24], [25], [26], a new technique has been developed to address uneven lighting issues. In aiming to achieve thresholding of uneven light images, this work makes a three-fold contribution:

1. A method is devised to generate valid training sample points (TSP) from input images with uneven lighting. 2. An efficient and less complex Functional Link Artificial Neural Network (LC-FLANN) model is proposed to estimate the illumination surface. 3. Normalization of uneven light through inverse modelling is applied to global thresholding. Therefore, the newly added value to the proposed methods lies in the extraction of more valid sample points for training the FLANN model to estimate illumination surfaces, as well as the normalization of the estimated image for further binarization.

II. RELATED WORK

A. LIGHT WEIGHT OBJECT DETECTION MODEL BY USING YOLOv8-LITE

Yang et al. [32] introduced a lightweight object detection model known as YOLOv8-Lite, which is based on the YOLOv8 framework. This model has been improved using the FastDet structure, TFPN pyramid structure, and CBAM attention mechanism. Traditional object detection models often face challenges such as large parameter sizes and high computational resource consumption, limiting their use on edge devices.

YOLOv8-Lite is designed for efficient object detection through simplified design and optimized computation. The framework incorporates the FastDet structure with simplified network complexity. The TFPN (Top-Down Feature Pyramid Network) pyramid structure addresses information loss, thus improving object detection accuracy. The TFPN structure combines feature maps from different levels to reduce information loss. The CBAM (Convolution Block Attention Module) is introduced to further enhance the network. This mechanism adaptively adjusts attention within feature maps, enabling the network to more accurately focus on important target areas. This CBAM can be represented as:

The channel attention map C is calculated as in Equation (1) the sigmoid activation of the average channel-wise feature response.

$$C = \sigma\left(\frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{w} ReLU\left(W_c.X_{i,j}\right)\right) \tag{1}$$

where C is the channel attention map. H and W are the height and width of the feature map. $x_{i,j}$ is the feature map at the position (i, j). W_c represents the weights used for channel-wise convolution. σ denotes the sigmoid activation function.

The spatial attention map S is computed in Equation (2) capturing the spatial dependencies across the feature map.

$$S = \sigma\left(\frac{1}{C}\sum_{c=1}^{C} ReLU\left(w_s.X\ :, :, c\right)\right) \tag{2}$$

where S is the spatial attention map. C is the no of channels in the feature map. $x_{:,:,c}$ cth channel of the feature map. w_s is the weight used for spatial-wise convolution. The output feature map Y is obtained by Equation (3) combining the channel and spatial attention maps with the original feature map X, enhancing the model's ability to focus on relevant regions.

$$Y = X \odot (C + S) \tag{3}$$

FASTDet structure is designed to improve the accuracy of object detection. Mathematically it can be represented as in Equation (4)

$$Y = WX + b \tag{4}$$

where Y is the feature map, W is the weight matrix, X is the input feature map, b is the bias vector.

The FastDet structure is proposed to improve the speed and accuracy of the object detection model. The FastDet structure mathematically expressed as in Equation (5)

$$y = Wx + b \tag{5}$$

where w is the weight matrix, b is the bias vector, and x is the input feature map and y is the output feature map expressed as in Equation (6)

$$y = Shuffle(x) \tag{6}$$

the channel shuffling operation denoted as Shuffle(.). Then dividing the feature map into two by splitting operation denoted as Equation (7)

$$y1, y2 = Split(x) \tag{7}$$

Then the output feature map after convolution in Equation (8) and polling is expressed as in Equation (9)

$$y = Conv(x) \tag{8}$$

$$y = Poolling(x) \tag{9}$$

this FastDet structure improve the efficiency and performance of the model.

B. LIGHTWEIGHT TWO-STAGE DETECTION NETWORK FOR REAL TIME OBJECT DETECTION

In his work, Wang [33] introduced two detection frameworks: RPN (Rapid Region Proposal Network) and Refinement Detection Network (RefinerNet). RPN is responsible for generating high-quality candidate regions, while RefinerNet analyzes these regions in detail to enhance detection accuracy. RPN uses hybrid data to generate high-quality regions, producing a set of bounding boxes, each with a score indicating the likelihood of containing an object. This process can be expressed as...

$$Regions = Softmax(DenseConv(Flatten(V))) \tag{10}$$

In Equation (10) V refers to voxelized data This network utilizes a series of convolution layers. These layers are process voxelized point cloud data. This voxelization divides continuous point-to-grids with aggregating information in each grid based on mean, maximum values, or a combination of features.

After the candidate regions are obtained RefinerNet is Refiner detection network is used for the localization

and classification of targets by deep convolutional neural network. This process mostly enhanced object edges in target regions expressed as in Equation (10)

$$F' = ReLU(N(Conv(Regions, F))) \tag{11}$$

$$result = PointNet(F'_j, F') \tag{12}$$

In Equation (11) F' is the collection of features extracted from the first stage including spatial and depth features to describe the contents of candidate regions. These features help refine the detection and classification of objects within the specified areas with high accuracy.

C. BINARIZATION OF IMAGES WITH CHANGING LIGHTING

In particular, image binarization with adaptive window (IBAW) is required for image binarization with adaptive illumination [28]. and adaptive averages (IBAA). A set of decision criteria derived from the local properties forms the basis of both algorithms. The visual Odometry (VO) approach is used to iteratively generate an effective window size across each pixel until the decision map remains unchanged.

The Bradley algorithm [7] is used in IBAA to achieve the binarization of each pixel throughout the window that is generated using the fusion image obtained from the selection criteria. In contrast, the fusion image is not utilized in IBAW to calculate the window size; rather, the Bradley technique [7] is used to approximate the binarized image.

Algorithm 1: VO Algorithm for Window size Computation:

For every pixel, there is a unique window $W(m, n)$ with an area of $(2W(m, n)+1)^2$. The maximum window size $W(m, n)$ in each pixel is determined by maximizing the window's area to ensure that the edges of the decision map, which provides information about areas that are lighter or darker in the original image, do not exceed the maximum edge number T and the maximum value of window size w_{max}.

The area $A(W(m, n))$ is located by binary search, and the integral image of the border matrix (ED) is searched maximize the function. In Equation (13), the border matrix is provided.

$$ED[m, n] = \begin{cases} 1 & If \;\; q(m, n) \neq q(m-1, n) \\ & or \\ & q(m, n) \neq q(m, n-1) \\ 0 & Otherwise \end{cases} \tag{13}$$

Using I_{ED}, a binary search is carried out for the integral image I_{ED} for each coordinate (m, n) to find the ideal window size.

Algorithm 2: IBAA for image Binarization:

The Molina algorithm [29] was enhanced with an adaptive window for every pixel in the IBAA algorithm. The IBAA method uses the VO algorithm 1 and takes into account the fused image I_F to determine the ideal window size, $W(m, n)$, for each pixel (m, n). The original image $I(m, n)$ and the inversely illuminated image $I_L(m, n)$ are combined to generate the fused image I_F. The following formula is used to get the inversely illuminated image:

Step 1: Determine the Image I's Luminance L using the morphological kernel C, or $L = I.C$.

Step 2: entails computing the reflectance (R) at the pixel level, which is the ratio of the matching lighting intensity ($L(m, n)$) at that pixel to the image intensity ($I(m, n)$).

Step 3: Determine the Inverse Luminance by deducting the Luminance value of each pixel from 255, or $LI(m, n) = 255 - L(m, n)$.

*Step 4:*In order to acquire the image with inverse lighting, Assign the result to $I_L(m, n)$, that is, $I_L(m, n) = R(m, n) \times LI(m, n)$, by multiplying the reflectance image $R(m, n)$ by the light intensity image $LI(m, n)$. *Step 5:* The difference image DI is computed as follows: $DI = I - I_L$.

Step 6: The Integral Image H_{DI} of DI is to be computed.

Step 7: $S(H_{DI}, W(m, n), m, n)$, is the sum of values within a square window in which $W(m, n)$ denotes the window size. Maximizing $S(H_{DI}, W(m, n), m, n)$ yields the decision criteria $p^{(it)}$. $I_F(m, n) = I(m, n)p^{(it)}(m, n) + I_L(m, n)p^{(it)}(m, n)$ is the fused image that we finally compute.

Step 8: Using the fused image I_F and a window size W via algorithm-1, we use the Bradley method [7] to produce a binarized image.

By applying the Bradley method [7], the IBAW methodology compares a quality function to a threshold value to produce a binarized image. An adaptive window is used as the basis for the quality function calculation.

Algorithm 3: IBAW for Image Binarization:

To convert images into binary format, the Bradley [7] IBAW algorithm is employed. To do this, apply algorithm 1 to compute a window over each pixel $W(m, n)$. The Decision map was derived from the difference of the quality function $DF(m, n)$ rather than the fused image $I_F(m, n)$, in contrast to IBAA. The quality function, which is stated as follows, establishes the distance between two modes m_1 and m_2 from which pixels are placed.

Step 1: The absolute difference between m_1 and $I_1(m, n)$ defines the function $F_1(m, n)$.

Step 2: The absolute difference between m_2 and $I_1(m, n)$ defines the function $F_2(m, n)$.

Step 3: The difference between the quality functions is calculated as $F1[m, n] - F2[m, n] = DF[m, n]$ to find $DF(m, n)$.

Step 4: Determine the integral image H_{DF} that is a part of DF.

Step 5: A square window's total values can be found using the formula $S(H_{DF}, W(m, n), m, n)$, where $W(m, n)$ indicates the window's size. The decision criterion is obtained by adding the integral image of the difference in quality function $S(H_{DF}, W(m, n), m, n)$ over the window $W(m, n)$. The binarized image based on this choice criterion is obtained using the Bradley technique [7] and the matching window based on the VO algorithm1.

D. FUZZY INCLUSION AND ENTROPY BASED IMAGE BINARIZATION

A method to determine the threshold for every pixel in a picture was presented by Bogiatzis et al. [11]. By dividing 255 by all the $x \times y$ nearby pixel values, they produced a fuzzy set M. The fuzzy inclusion between the M and the three images: the gray picture (G), the black image (B), and the white image (W). W, B, and G have sizes of $m \times n$. All of the elements in W are 1, all of the elements in B are 0, and all of the elements in G are 0.5. It is thought that W is a universal set. Equation (14) defines $S(A, B)$ as the fuzzy inclusion function between two sets A and B.

$$S(A,B) = \begin{cases} \dfrac{\sum_{w \epsilon W} I_z(m_A(w), m_B(w)) m_A(w)}{\sum_{w \epsilon W} m_A(w)} & if\ A \neq \phi \\ 1 & if\ A = \phi \end{cases} \tag{14}$$

where the membership functions are denoted by $m_A(w)$ and $m_B(w)$. Fuzzy inclusion indicators $s_1 = S(W, M)$, $s_2 = S(M, B)$, $s_3 = S(G, M)$ and $s_4 = S(M, G)$ provide information. about the brightness, darkness, and the window's grayness.

Any set M has an entropy e that is determined by Equation (15).

$$e = 2 \frac{\sum_{w \epsilon W} min(1 - m_M(w)) m_M(w)}{\sum_{w \epsilon W} max(1 - m_M(w)) m_M(w)} \tag{15}$$

Based on the values of s_1 and s_2, eight criteria are defined. Using these criteria, the set M can be classified into one of eight groups. Using linear regression analysis, a function r, i.e., $r = r(s_1, s_2)$, is defined. Equation (16). defines a fuzzy symmetric triangular number in the end.

$$\tilde{T} = (r - k, r, r + k), where\ k = |s_3 - s_4| \tag{16}$$

Eq. (17) is used to determine the T_1 and T_2 given the parameters e, k, and r.

$$T_1 = k(e - 1) + r\ and\ T_2 = k(1 - e) + r \tag{17}$$

if $s_1 \leq s_2$ then the threshold $T = T_1$, else $T = T_2$ for the $p(x, y)$ pixel.

E. GAUSSIAN MIXTURE MODEL BASED BINARIZATION

Zhao et al. [8] proposed binarization of images with non-uniform light and noise using a Gaussian Mixture Model (GMM). By utilizing the image mean as a threshold to divide pixels into two classes, the method divides the gray level distribution into two groups, modeling it as a two-class GMM. The optimal parameter vector $\Theta = [w_1, w_2, \mu_1, \mu_2, \sigma_1, \sigma_2]$ is then estimated using the EM algorithm. The image is binarized to suppress impulse and Gaussian noise, using the posterior probability of the pixel, which is evaluated with spatial neighborhood information h_{jk} (as in Equation (18)).

$$h_{jk} = mean\left[median\left\{ \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\left(\frac{\bar{x_j} - \mu_k}{\sigma_k}\right)^2} \right\} \right] \tag{18}$$

where $\bar{x_j} \epsilon N(x_j)$ and $N(x_j)$ is the set of neighborhood pixels across a window that is 3×3 or 5×5, with the center pixel x_j. Using Equation (19), the pixel's posterior probability x_j corresponding to the k^{th} class is calculated.

$$q(k \mid x_j, \hat{\Theta}) = \frac{w_k h_{jk}}{\sum_{i=1}^{k} w_i h_{ji}} \tag{19}$$

where $k = 1 or 2$, and the mixture coefficient with constraints is denoted by w_k. With $\sum_{k=1}^{2} w_k = 1$, $0 < w_k < 1$ and 3×3 and 5×5 is a window of size that contains spatial neighborhood information h_{ik}. Using Equation (20), the binarized image $C(m, n)$ is produced.

$$C(m,n) = \begin{cases} 0 & if\ (q(k = 1 \mid x_j, \hat{\Theta}) > q(k = 2 \mid x_j, \hat{\Theta}) \\ 255 & Otherwise \end{cases} \tag{20}$$

F. GRADIENT-BASED THRESHOLDING SURFACE FOR IMAGE BINARIZATION

Yazid et al. [18] developed a method for binarizing images with non-uniform lighting. Every pixel in the image is subjected to vertical, horizontal, diagonal, and anti-diagonal Prewitt operators as part of the method's gradient-based thresholding surface approach. The edge or border information of the picture is determined by the largest absolute response for each pixel among these four operators.

Strong edge and weak edge pixels are the two categories for these edge or boundary pixels. The non-border pixels are those with weak edges. Equation (21) is used to generate the thresholding surface based on these edge pixels.

$$T_s(i,j) = \frac{\sum_{m=1}^{8} P_m(i,j)}{8} \tag{21}$$

In image processing, binarization involves separating an image into two distinct classes of pixels - black and white. The thresholding surface $T_s(i, j)$ is used to accomplish this task. Additionally, there are eight primary surfaces $P_m(i, j)$, in which $m = 1 to 8$ created through eight passes in different directions: $LRTB$, $TBLR$, $RLBT$, $BTRL$, $RLTB$, $TBRL$, $LRBT$, and $BTLR$. Here, L, R, T, and B refer to left, right, top, and bottom, respectively.

A constant value k_o is added to or removed from the threshold surface $T_s(i, j)$ intensities to modify the inverse threshold surface's position concerning the original image $I_o(i, j)$ for image binarization. The threshold surface may

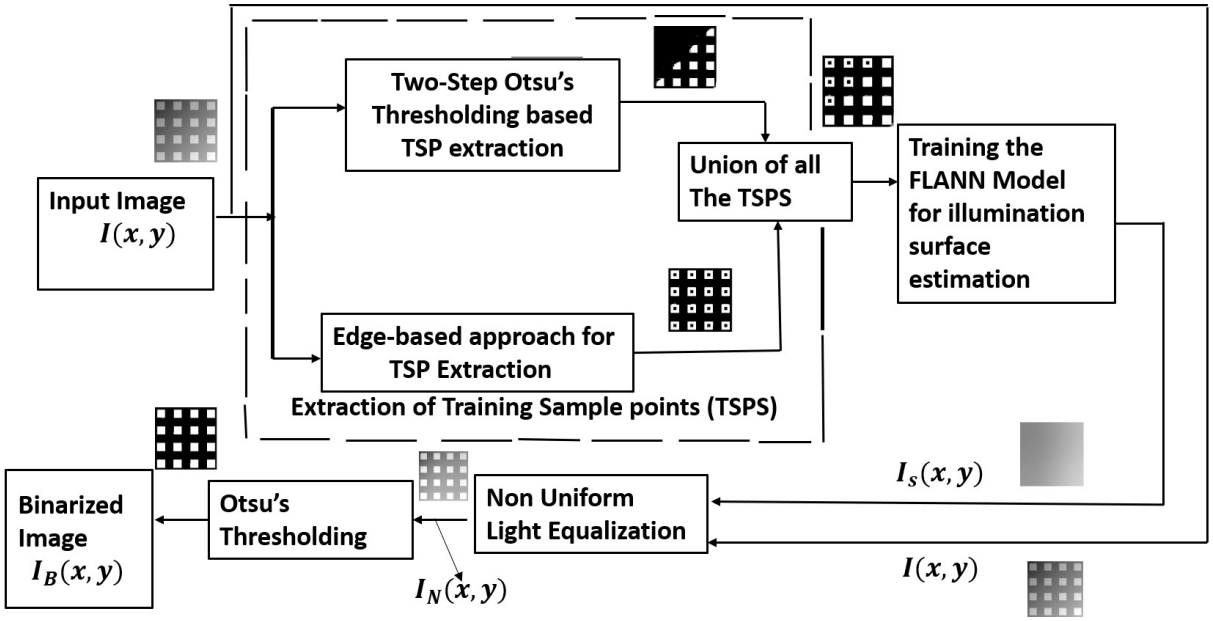

FIGURE 1. Block schematic of the suggested method for thresholding images with nonuniform lighting.

move up or down as a result. Equation (22) is used to find the value of k_o.

$$k_o = \underset{k}{arg\,min}\left[\sum_{borderpixel(i,j))}\left|(I_o(i,j) - Ts(i,j)) + k\right|; \times k\epsilon\left\{-255..255\right\}\right] \quad (22)$$

Equation (22) defines the value of k, denoted as k_o, which represents the optimal value of k, which is obtained when the adjusted thresholding surface $T_s(i,j) + k$ and the pixels of the original input image $I_o(i,j)$ are added together, the sum of the absolute differences between them is at its minimum. To distinguish between backgrounds and objects, the threshold surface assigns a distinct threshold value to each pixel. Next, we apply Eq. (22) to binarize the original image Io.

$$g(i,j) = \begin{cases} 0 & if\ I_o(i,j) > Ts(i,j) + k_o \\ 1 & if\ I_o(i,j) < Ts(i,j) + k_o \end{cases} \quad (23)$$

III. MATERIALS AND METHODS

Due to the overlapping of the gray level distributions of object and background, it is difficult to binarize any uneven light images using global thresholding. Further, the constraint of local thresholding approaches are (i)initial choice of window size, (ii) Computational complexity, (iii) optimal choice of bimodal criteria function, and (iv) Optimal partitioning approaches, for any uneven lighting image binarization. To improve the performance of uneven light image binarization, the FLANN Model-based estimation of illumination surface neutralizes nonuniformity and increases the global thresholding's effectiveness. The proposed approach consists of three steps, which are outlined below (i) the extraction of training sample points(TSP) that are valid; (ii) the Development of a FLANN model for estimation of the illumination surface and (iii) Illumination normalization for global thresholding.

Compared to multi-layer neural networks, the Functional Link Artificial Neural Network (FLANN) has a simpler design and requires less computing power. FLANN is suited for real-time applications because of its straightforward architecture, making it simpler to develop and using less processing power. Functional expansions of the input features, such as polynomials, trigonometric functions, or other non-linear transformations, are used by FLANN to increase the representative capability of a single-layer network. This enhances the network's capacity to tackle challenging issues by enabling FLANN to represent non-linear relationships in the data without additional layers. For applications that require quick training and adaptation, FLANN's speedier training process resulting from the lack of several hidden layers is a benefit. The utilization of straightforward algorithms for weight modification and the reduction of parameters contribute to the efficiency of the learning process. Functional expansions allow FLANN to reduce the number of weights needed to achieve improved generalization, which lowers the possibility of overfitting—especially on smaller datasets.

The network is better able to identify the underlying patterns in the data thanks to the functional expansion, which functions as a type of regularization. FLANN is a flexible technique that can be used in many domains because

it can be used for different kinds of data and challenges. It may readily adjust to various problem specifications by selecting the relevant functional extensions. Because FLANN can detect non-linear trends, it is used to predict stock prices, exchange rates, and other financial time series data. It uses trend analysis on previous data to forecast atmospheric conditions and weather patterns. FLANN is used to extract and categorize relevant features for tasks including object detection, video analysis, and picture recognition.

FLANN is a superior model than the other network models due to its lower complexity. Unlike conventional methods, FLANN can accurately estimate the illumination surface by expanding features to a higher-dimensional space, which improves input discrimination. Its single-layer structure results in reduced weight updating time compared to convolutional or deep learning network models. The normalized coordinates of the sample points are used to train the FLANN model. Based on the discrepancy between the sample point's actual and anticipated gray values, the weights are changed. All of the image's coordinates' lighting is calculated after the FLANN has been trained. The non-uniform lighting is neutralized once the illumination is estimated, enabling the use of global thresholding for the separation of objects from backgrounds. Consequently, the FLANN model is preferred for estimating the illumination surface more accurately and efficiently, with less time and complexity than other models.

In this work, the non-uniform lighting surface of an image is estimated using the FLANN model. Global thresholding is further used to neutralize the non-uniform illumination to binaries the image. On the other hand, either object points or background points must be extracted to estimate the non-uniform illumination surface. Extracting valid sample points from an image with non-uniform lighting, either from the background or the foreground region, is an extremely difficult process. To extract the legitimate taring sample points, a two-stage method i.e. the dual thresholding technique and the edge-based approach is devised in this study.

The block diagram for the proposed uneven lighting image thresholding is shown in Fig. 1. To help estimate the illumination surface, the proposed approach in Fig. 1 extracts valid training sample points (TSP) using a multi-step process. These sample points correspond to either the two-class image's background or object pixels. In the first stage, a two-step thresholding is utilized to produce more valid sample points. The two-stage Otsu's thresholding method extracts the object sample points more successfully in non-uniform lighting situations. One threshold may not be sufficient to properly partition the image's object region in scenarios with different textures or illumination. In the second step, the edge-based method is employed to extract the legitimate sample points from the object regions. The object's boundary edges are extracted using the edge-based technique. The distribution of the original edge point pixels is subjected to a global threshold in order to extract various object sample points from the object boundary. The dual threshold and edge-based approaches' valid sample points work together to aid in the collection of sample points throughout the object region.

A. EXTRACTION OF APPROPRIATE TSP

To accurately estimate the illumination surface, valid training sample points (TSP) are required. The images used for binarization are two-class images. As a result, it is crucial to choose pixels only either the background area or the object to serve as the TSP. The proposed FLANN model uses the object (foreground) region points as the TSP for training. Considering the quantity of reliable training sample points, a two-step method has been developed. Two-step thresholding is used in the first stage, and an edge-based technique is used in the second stage to extract valid training sample points.

1) TWO-STEP THRESHOLDING (TST) METHODOLOGY

During the first step of the TST process, binarization of the uneven light gray image $I(x, y)$ is obtained using Otsu's thresholding. This method is based on the assumption that the image contains a brighter object than the background. The global threshold value, is represented as T_1. Next, every pixel having a gray value higher than T_1 are assembled into a set S.

$$S = \{I(x, y)|I(x, y) \geq T_1\} \tag{24}$$

Once all the elements in S have been taken into account, a histogram h_g is produced, with a gray level range of $T_1 \leq g \leq 255$. The threshold T_2 is then assessed by applying Otsu's thresholding criterion to h_g. The pixels in the image $I(x, y)$ that have a value greater than T_2 are leveled as the first set of TSP, which is S_{0_1}.

Fig. 2(a) represents the original images with uneven lighting, and Fig. 2(b), is the histogram of Fig. 2(a) with the evaluated thresholds T_1 and T_2. Fig. 2(c) shows the binarized image of Fig. 2(a) with the threshold T_1. However, it is observed that not all white pixels belong to the object only. Most of the background pixels are also detected as object pixels.

The binary image with threshold T_2 is shown in Fig. 2(d). Here, all the white pixels belong to only the object class, which means that this thresholding method has successfully identified the object in the image.

The valid TSP in the 1st step are considered as set S_{0_1} as follows

$$S_{0_1} = [\{x, y, I(x, y)\} | I(x, y) \geq T_2] \tag{25}$$

To reduce the impact of outliers in the training sample points, a dual threshold is utilized. However, when observing Figure 2(c), it is evident that the white area (TSP) - specifically S_{0_1} - is mostly concentrated towards the object's right side. This can cause the illumination surface to be estimated incorrectly. An edge-based technique in its second stage is designed to tackle this problem. This approach is used to collect extra observation points made just from the object region's edge, which results in a more accurate estimation of the illumination surface.

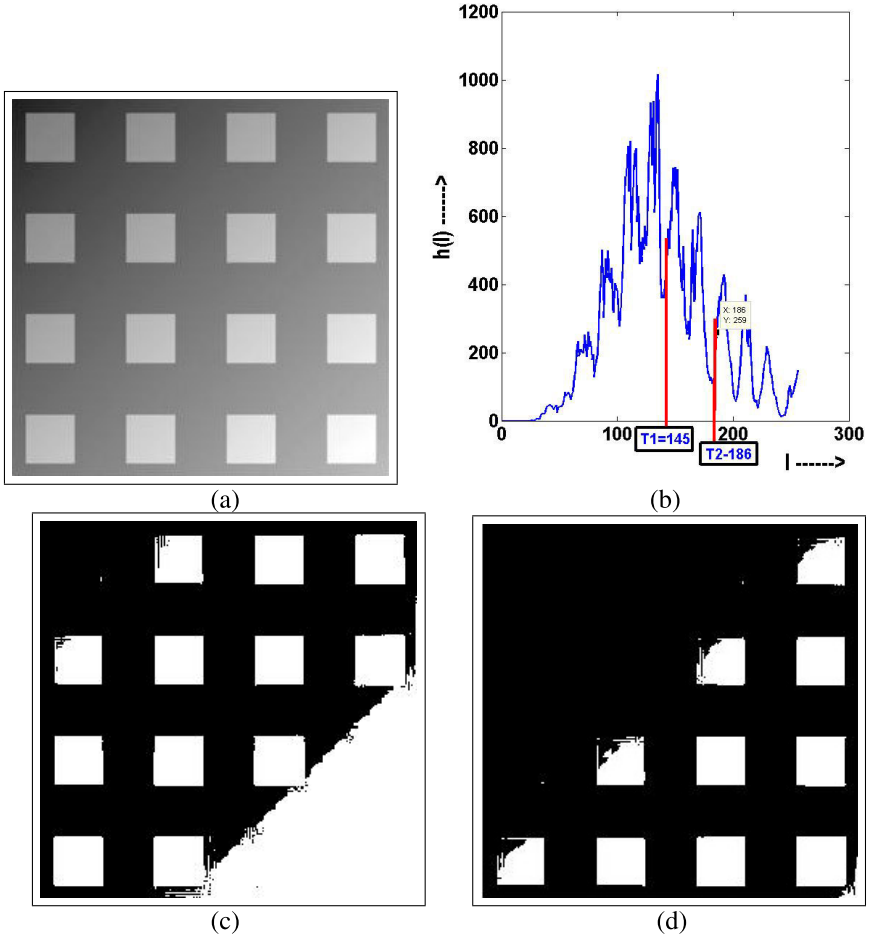

FIGURE 2. a)Original non-uniform lighting block image b) Histogram of Block image with thresholds T_1 and T_2 (c) Binarize Block image with T_1 (d) Binarize Block image with T_2.

2) EDGE BASED (EB)APPROACH

The first step in detecting edges of the gray image I is to extract edge points by applying the gradient-based operator i.e Sobel edge operator. The Sobel operator calculates an image's intensity function's gradient approximately to detect edges in an image. Two 3×3 masks, M_h and M_v, are used in Equation (26) to identify the horizontal and vertical derivative approximations. After convolving these masks with the input image, two gradient images, G_h and G_v, are obtained., as shown in Equation (27).

$$M_h = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad M_v = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{26}$$

$$G_h = M_h * I \text{ and } G_v = M_v * I \tag{27}$$

The magnitude of the gradient approximation is calculated using the formula $I_F(x, y) = \sqrt{G_h^2(x, y) + G_v^2(x, y)}$. This value is used to obtain the edge image I_E by identifying those points where the gradient of the image is greater than a certain threshold T. In other words, I_E is obtained as $max(I_M, T)$, where T is considered to be 100 based on the simulation results.

$$I_E(x, y) = \begin{cases} 1\ (edgepoint) & if\ I_M(x, y) \geqslant T \\ 0\ (Nonedgepoint) & Otherwise \end{cases} \tag{28}$$

In Figure 3(a), the Sobel edge detector is used to detect edges that cross the borders of an object. Since not every edge point can be considered an accurate sample point, these edge points indicate the gap between an object and the background. Every edge pixel in the object region is given a 5×5 window, It is employed to retrieve just valid sample points. when the window has more than five non-zero values, or white pixels, it is considered that the object's area and the background areas are almost equal. Therefore, all pixel values in the window at the same position on the input image are binarized applying global thresholding by Otsu. Based on the optimal threshold T_0, the pixel values in the window greater than T_0 are regarded as accurate sample points and constitute a part of the set S_{0_2}.

$$S_{0_2} = \{x, y, I(x, y) | (I(x, y) \in w) \geq T_0\} \tag{29}$$

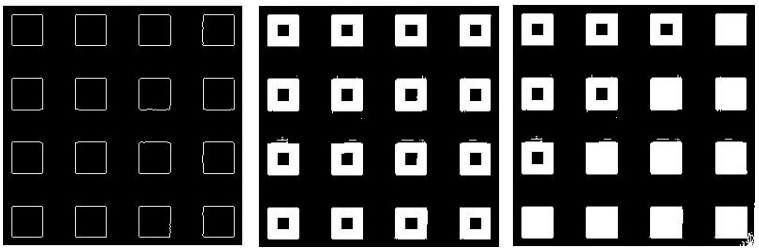

FIGURE 3. a) Edge image (b) Edge based observation points (c) The set $S_0 = S_{0_1} \cup S_{0_2}$ both segmented image considering histogram and edge.

The process of generating all the sample points in set S_{0_2} involves applying the same procedure to all the edge pixels. The TSP (Training Sample Point) of S_{0_2} are highlighted in white color in Fig. 3(b). The final TSP set S_0 is defined as the union of S_{0_1} and S_{0_2} as shown in Fig. 3(c). It can be seen from Fig. 3(c), that the number of training sample points has grown dramatically and now encompasses nearly the whole object region. Fig. 4(a), (b), and (c) present the three-dimensional plot of the TSP obtained by combining two-step thresholding and an edge-based approach. As can be seen from Fig. 4(c), the sample points belong to the object region only and are well-distributed across the object area in the image. This means that the estimation of the illumination surface of the image will be highly accurate.

B. PROPOSED LC-FLANN BASED FRAMEWORK FOR ILLUMINATION SURFACE ESTIMATION

The proposed FLANN model is shown in Fig. 5. The objective is to train the FLANN model using the Section A recognized valid TSP is used to estimate the illumination surface. FLANN model, increases the dimensionality to get the space and discriminate the non linear input. Also in FLANN, due to its single layer, the number of weights and training time decreases. In the proposed FLANN Model, the input is expanded by trigonometric expansion, which produce more compact representation by using the Fourier basis function. In this expansion, sine and cosine functions are used to decrease the computation complexity. The inputs to the FLANN model are the coordinates of each sample point belongs to set S. The input data to FLANN is two dimensional pattern $S_i = [x_i\ y_i]^T$, where x_i, y_i are the normalized value of coordinate of sample point $S_i \in S$. The intensity of the object point is considered as desired data output 'd_i', which is the normalized intensity value of the sample point $S_i \in S$. The coordinate x_i, and y_i and the intensity d_i of sample points are normalized values. A simple FLANN model is constructed with a pattern of two features, considering the i^{th} sample point, the input vector $X(i)$ is defined as $X(i) = [x_1(i)\, x_2(i)]^T$. This i^{th} input sample is mapped to high dimensional space by functional expansion like trigonometric expansion as follows:

$$\phi_i = [x_i, sin\pi x_i, cos\pi x_i, y_i, sin\pi y_i, cos\pi y_i, x_i y_i] \quad (30)$$

where i, is the i^{th} input vector is functionally expanded. The sum of weighted components of the expanded input is represented as $\varphi(i,k) = \phi_i^T.w$ where w is the weight vector $w = [w_1, w_2, ..w_M]$. Here M is the total no of expanded term. $\varphi(i,k)$ is represented as

$$\varphi(i,k) = w_1 x_i + w_2 sin\pi * x_i + w_3 cos\pi * x_i + w_4 y_i + w_5 sin\pi * y_i + w_6 cos\pi * y_i + w_7 x_i * y_i \quad (31)$$

This input is passed through a sigmoid activation function to produce the model output with a probability of 0 and 1. To improve the linearity, sigmoid activation function is used at the output of the model as follows.

$$\hat{d}(i,k) = sig(\varphi(i,k)) = \frac{1}{1+e^{-\varphi(i,k)}} \quad (32)$$

where i is the i^{th} training sample point and k is the k^{th} iteration. The error term of the model, $e(i,k)$, is defined as

$$e(i,k) = d_i - \hat{d}(i,k) \quad (33)$$

$\hat{d}(i,k)$ is the estimated intensity value of i^{th} sample points at the output of the activation function, at the k^{th} epoch.

The weights are updated by LMS algorithm as in (34) after each epoch in the training phase as follows:

$$w_m(k+1) = w_m(k) + \Delta w_m(k); \quad (34)$$

where, $m = 1, 2, 3, 4, 5, 6, and\ 7$. The change of weight at k^{th} iteration $\Delta w_m(k)$ is calculated as given in (35)

$$\Delta w_m(k) = 2\mu\phi_m(i)\delta(i,k) \quad (35)$$

where, $\phi_m(i)$ is the m^{th} expanded value with i^{th} input pattern and $\delta(i,k)$ is the learning rate defined as

$$\delta(i,k) = \left[\frac{1-\hat{d}^2(i,k)}{2}\right] e(i,k) \quad (36)$$

Training the model with all training sample points once is one iteration. The objective is to train the model with the same training sample points iteratively to minimize the error e. Once the error does not change, the model is trained with optimal weights. After the convergence of the weights, the illumination intensity of all pixels of the uneven light image is estimated using the trained LC-FLANN

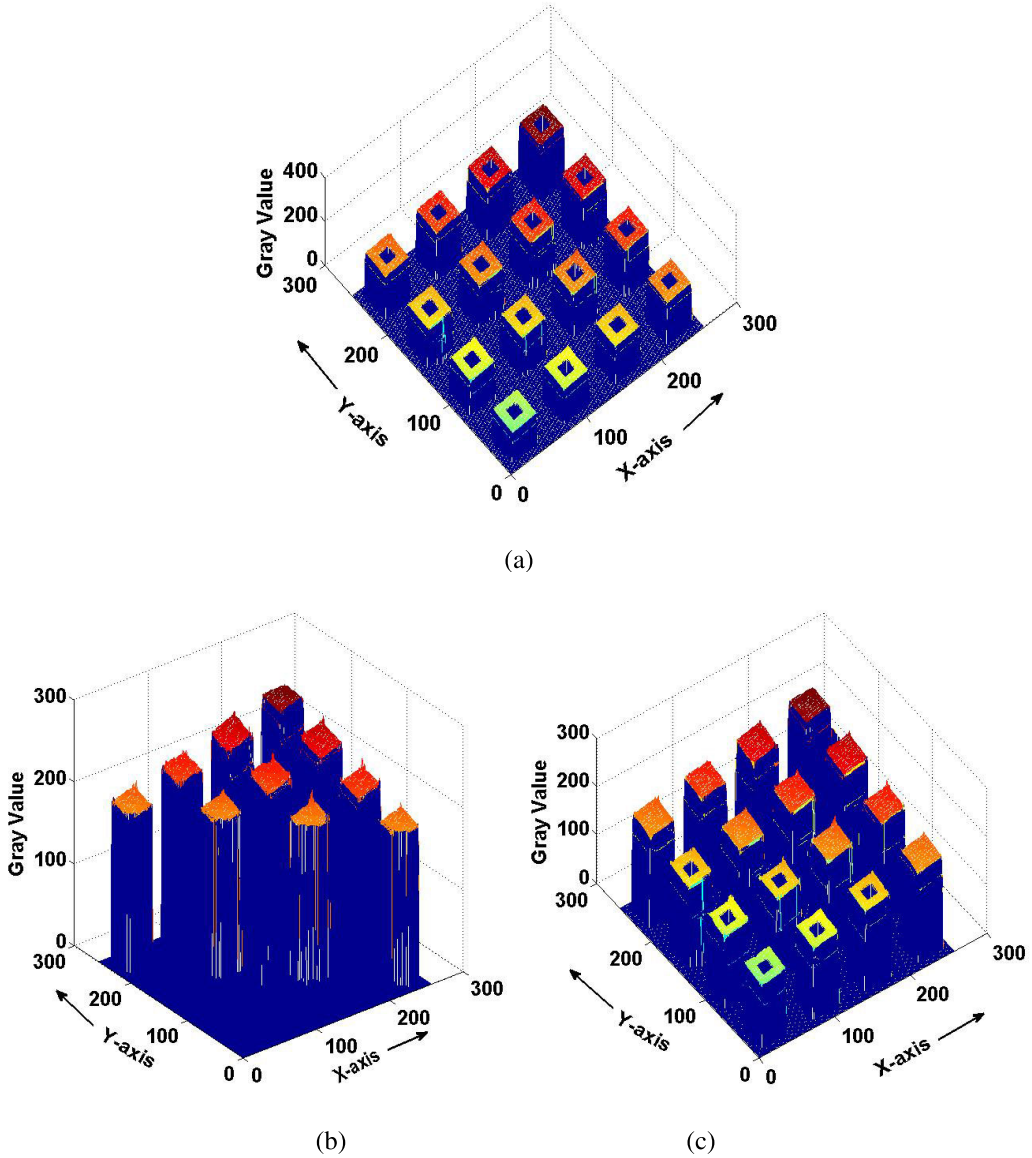

FIGURE 4. a) 3D image detected observation points with Edge b) 3D image with dual threshold Edge (c) 3D image of the combination of edge based and dual threshold based approach.

model. This estimated illumination surface of Fig. 2(a) is placed in Fig. 6(a). The 3D plot of Fig. 6(a) is shown in Fig. 6(b)

C. ILLUMINATION NORMALIZATION & THRESHOLDING

The image I with uneven lighting is normalized using the estimated illumination surface I_s in (37).

$$I_N(x, y) = \frac{I(x, y) \times g_{s(max)}}{I_s(x, y) \times g_{(max)}} \times 255 \tag{37}$$

In Equation (37), $g_{(max)}$ represents the maximum intensity value of the input gray image $I(x, y)$, and $g_{s(max)}$ represents the maximum intensity value of the estimated illumination surface image $I_s(x, y)$. Fig. 7(a) displays the illumination normalized image I_N of I, while Fig. 7(b) shows the 3D illumination surface of I_N. The normalized image I_N appears to have a nearly uniform illumination variation, as shown in Fig. 7(a) and 7(b).

Binarization is often a critical step in various image processing tasks, especially when the goal is to distinguish objects from the background. Depending on the specific requirements of the task and the characteristics of the images, other methods might be used as alternatives. However, in many cases, binarization provides a simple and effective way to achieve the desired outcome. To convert the image to a binary format, Otsu's thresholding method is applied on I_N. Fig. 7(b) displays the result of Otsu's global thresholding approach on the binarized image from Fig. 7(a).

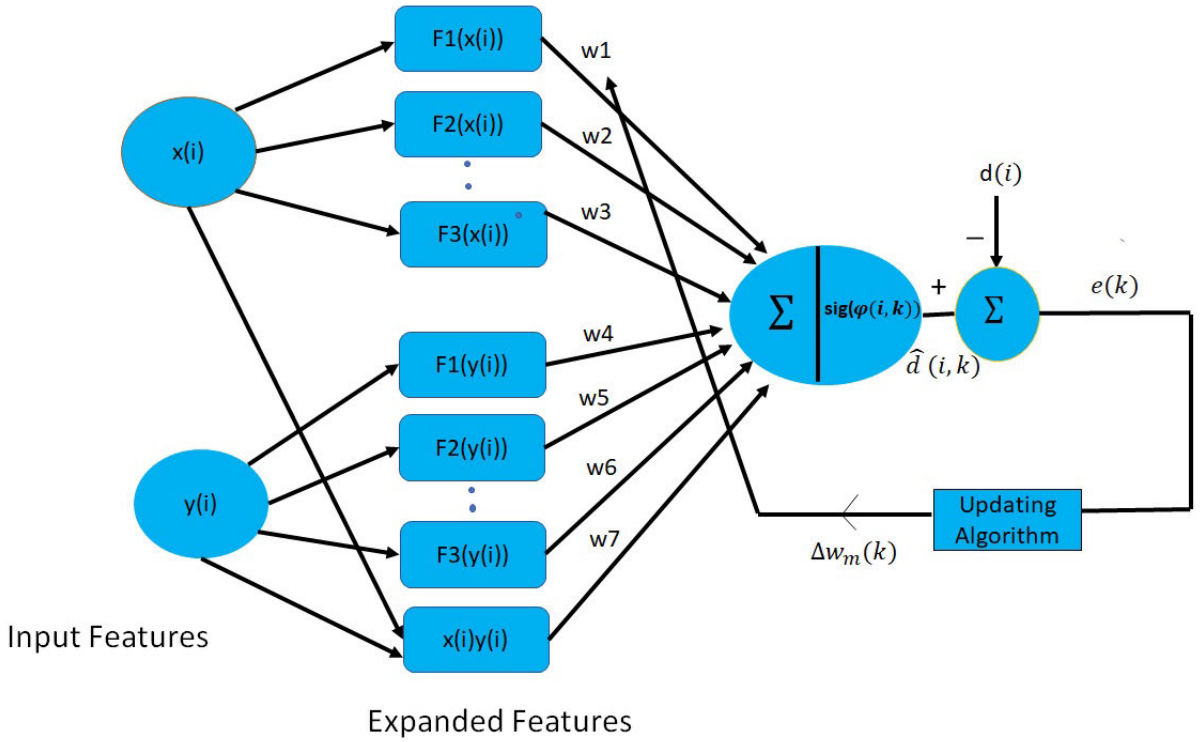

FIGURE 5. Proposed LC-FLANN framework for illumination surface estimation.

In Otsu's thresholding, the optimal threshold T_{opt} is obtained by minimizing within-class variance σ^2_{wi} and maximizing between-class variance σ^2_{Bi}. The σ^2_{wi} and σ^2_{Bi} at a threshold t are defined as

$$\sigma^2_{wi} = w_B(t)\sigma^2_B(t) + w_F(t)\sigma^2_F(t) \tag{38}$$

where, $w_B(t)$ and $w_F(t)$ represent weight of background and foreground represented as $w_B(t) = \sum_{i=1}^{t} p_i$, and $w_F(t) = \sum_{i=t+1}^{L} p_i$, where p_i is the probability distribution of gray level $i = [1, 2, \ldots L]$ with, $L = 255$. Similarly the $\sigma^2_{Bi}(t)$ is represented as

$$\sigma^2_{Bi}(t) = w_B(t)w_F(t)(\mu_B(t) - \mu_F(t))^2 \tag{39}$$

where $\mu_B(t)$ and $\mu_F(t)$ are the mean value of background and foreground pixels based on threshold t represented as $\mu_B = \sum_{i=1}^{t} iP_i$ and $\mu_F = \sum_{i=t+1}^{L} iP_i$. The variance of background and foreground pixels are represented as. $\sigma_B(t) = \sum_{i=1}^{T}(i-\mu_B(t))^2 \times p_i$ and $\sigma_F(t) = \sum_{i=T+1}^{L}(i-\mu_F(t))^2 \times p_i$. The parameters $w_B(t)$ and $w_F(t)$, $\mu_B(t)$ and $\mu_F(t)$ are the function of the threshold t. Therefore the optimal threshold T_{opt} is evaluated as follows:

$$T_{opt} = \max_t \left(\frac{\sigma^2_{B_i}(t)}{\sigma^2_{w_i}(t)} \right) \tag{40}$$

Finally the binarized image $I_B(x, y)$ is obtained as

$$I_B(x, y) = \begin{cases} 1 & if\ (I_N(x, y) > T_{opt}) \\ 0 & otherwise \end{cases} \tag{41}$$

Some sample uneven light images with their estimated 2D illumination surface image and corresponding illumination normalized images generated using the proposed approach are placed in Fig. 8 to Fig.11. The avg. convergence plot of the proposed Illumination surface estimation is shown in Fig. 12. The convergence plot is obtained by considering the average convergence plot of Mean square error (MSE) for all sample images, by taking 1000 iterations for each image. It is observed that this plot is converged near about 600 iterations.

IV. EXPERIMENTS AND EVALUATION

To validate the performance of a suggested method, five different approaches are used. These include Niblack [6], Bogiazis et al. [11], Yazid et al. [18], Cai and Miklavcic [19], and Zhao et al. [8]. These approaches are simulated on a computer with an Intel Core i5, 4MB L2 cache, 8 GB RAM, and a 2.4 GHz processor. 21 test images with varying degrees of uneven lighting are used to confirm each of these techniques. Ten of the images are sourced from the Berkeley database [31] and Weizmann [30], while eleven images are collected from various online sources. The original test images are presented in Fig. 13 and 14 along with the ground truth and binarized results. Fig. 13 and 14's first and second columns display the original test images and corresponding ground truth (GT) images. Similarly, the binarized images in columns three through eight are correspondingly produced using the suggested method and the landmark methods mentioned above.

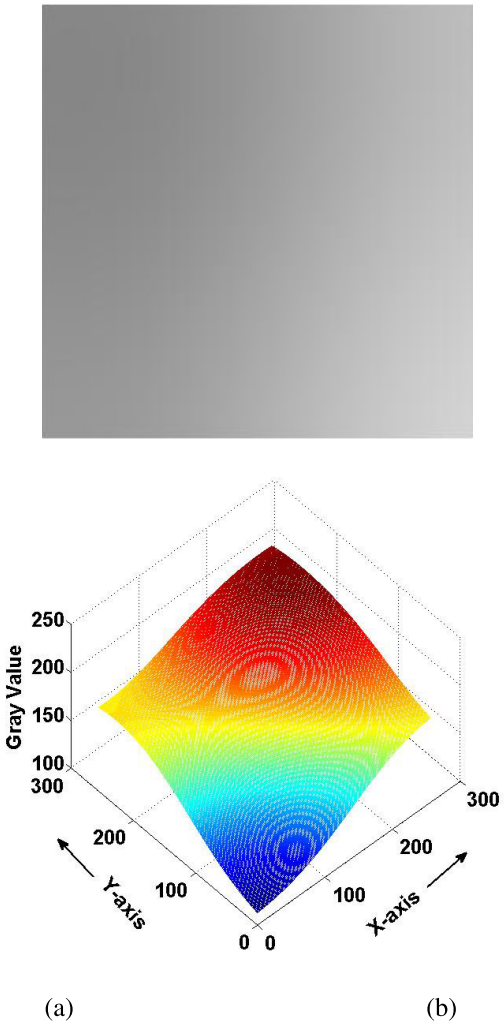

FIGURE 6. a) Estimated 2D illumination surface image I_S b) Estimated 3D Illumination plane.

In Fig. 13, the objects and background are blended in images that were processed using the Niblack and Bogiatzis techniques, leaving only the outer edges visible. The highlighted areas of the Yazid method's objects are improperly segmented, except the L-image and Synth-1 images. However, pics like the coin, Synth1, Crane, Synth-4, thread, and hand reveal backdrop that is incorrectly categorized as an object in the Cai technique. According to the Zhao approach, the majority of the object pixels in images like rice, and synth-1, are accurately segmented in the suggested approach.

Both quantitative and qualitative metrics are used to evaluate the performance. Based on the visual performance, the qualitative performance was assessed. On the other hand, Recall(R_e), Precision(P_r), F-measure(F_1), Jaccard Index(JI), and the percentage of misclassification error (PME) are taken into consideration while evaluating the quantitative measures. Images of the ground reality are crucial for assessing quantitative metrics. The databases contain the

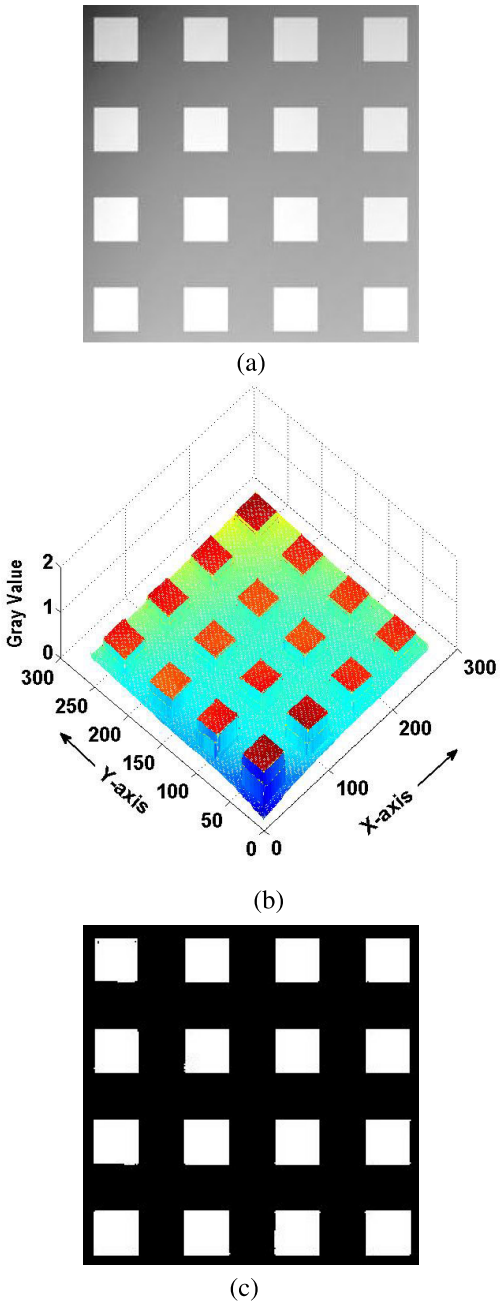

FIGURE 7. a) Illumination normalized image I_N b) 3D Intensity balanced surface c) Binarized Image.

Berkeley and Weizmann ground truths. Nevertheless, the real-world test images on the internet are created manually.

The Precision measure (P_r) measure is defined as

$$(P_r) = \frac{T_P}{T_P + F_P}, \tag{42}$$

where T_p is the true positive and F_p is the false positive.

The Recall measure R_e is defined as

$$(R_e) = \frac{T_P}{T_P + F_N} \tag{43}$$

where, F_N is the false Negative.

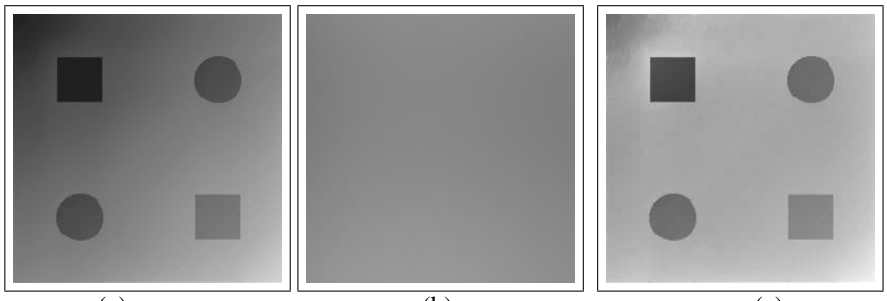

FIGURE 8. a) Input Synth-4 image b) Estimated 2D illumination surface image I_S (c) Illumination normalized image I_N.

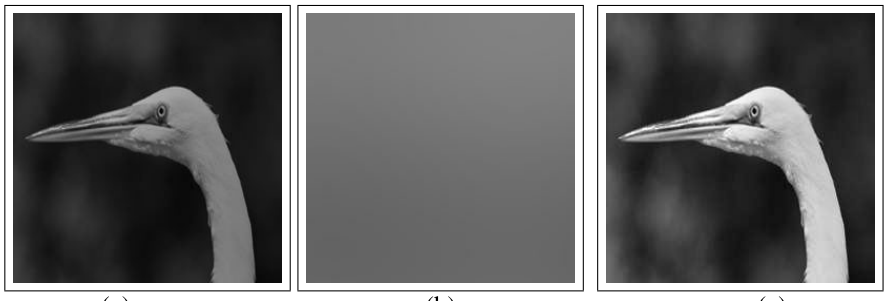

FIGURE 9. a) Input crane image b) Estimated 2D illumination surface image I_S (c) Illumination normalized image I_N.

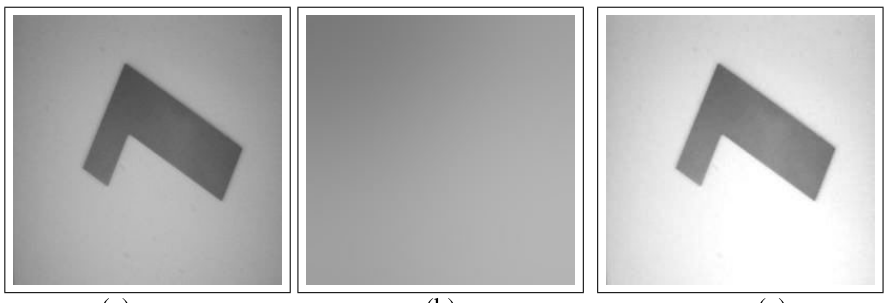

FIGURE 10. a) Input Limage image b) Estimated 2D illumination surface image I_S (c) Illumination normalized image I_N.

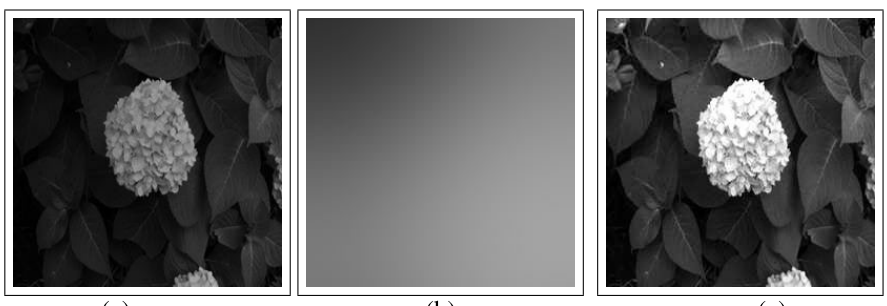

FIGURE 11. a) Input flower image b) Estimated 2D illumination surface image I_S (c) Illumination normalized image I_N.

The combination of the P_r and R_e measures is known as the F-Measure ($F1$).

$$F - Measure(F1) = 2 \times \frac{P_r \times R_e}{P_r + R_e} \tag{44}$$

The JI measure is defined as

$$Jaccard\ Index(JI) = \frac{SI \bigcap G_T}{SI \bigcup G_T} \tag{45}$$

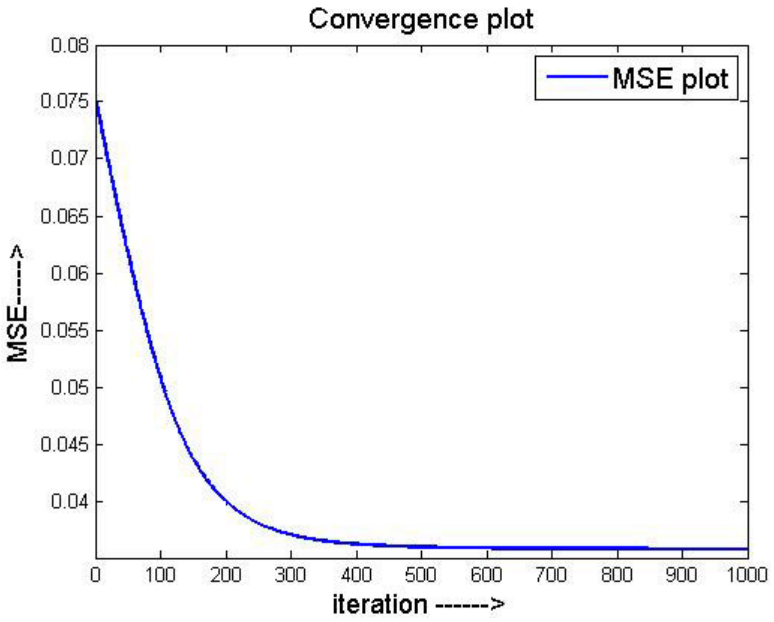

FIGURE 12. convergence plot of MSE.

TABLE 1. Visual perspective based qualitative measure.

Different method	Binarized images visually identical to the appropriate ground truth in Fig. 13	Total No	Binarized images visually identical to the appropriate ground truth in Fig. 14	Total No
Niblack [6]	NIL		Fingerprint	01
Bogiazis [11]	NIL		Fingerprint	01
Yazid [18]	Synthetic1, L image, Synthetic-4	03	Bandage,Plane,Tower	03
Cai [19]	rice, L image, synthetic2, synthetic3, thread	05	Light, Ninety eight,Bandage, Plane, Tower, Fingerprint Boat, 3096	08
Zhao [8]	L image, Synthetic2, Synthetic-3, Crane	04	Pins1, Light, Plane, Tower, Boat	05
Proposed	coin,rice,synthetic1,Limage, crane,pimage,Block2 Synthetic4, thread	**09**	Light, Bandage,Plane, Flower,Tower,Boat ,3096	**07**

where the corresponding ground truth image is G_T, and the binarized image is SI. The (PME) measure defined as

$$PME = \left[1 - \frac{(F_s \cap F_g) \cup (B_s \cap B_g)}{(F_g + B_g)}\right] \times 100 \qquad (46)$$

where F_g and B_g denote the set of foreground and background pixels in the ground truth image and F_s and B_s represent the set of foreground and background pixels in the binarized image respectively. The qualitative performance is evaluated as in Table 1, which shows the degree of visual perspective similarity between the binarized image and the ground truth test images in Fig. 13 and 14. It is observed from table 1 that the two methods Niblack et al. [6] and Bogiazis et al. [11] are satisfactory on one image like a fingerprint. However the performance of Yazid et al. [18] and Cai and Miklavcic [19] is satisfactory on 40% of test images in Fig. 13 and 60% in Fig. 14. The performance of Zhao et al. [8] is 40% in both the set of test images in Fig. 13 and Fig. 14. However, the proposed approach outperformed in most of the test images in Fig. 13 and Fig. 14. As a result, the suggested strategy is the only one among all of them that can binarize any variation of uneven light image. Fig. 15 displays the average P_r, R_e, F_1, and JI score for each of these five approaches using the suggested FLANN-based methodology. It is noted that the suggested FLANN technique performs best out of all the methodologies taken into consideration for the validation and is consistent in every measure.

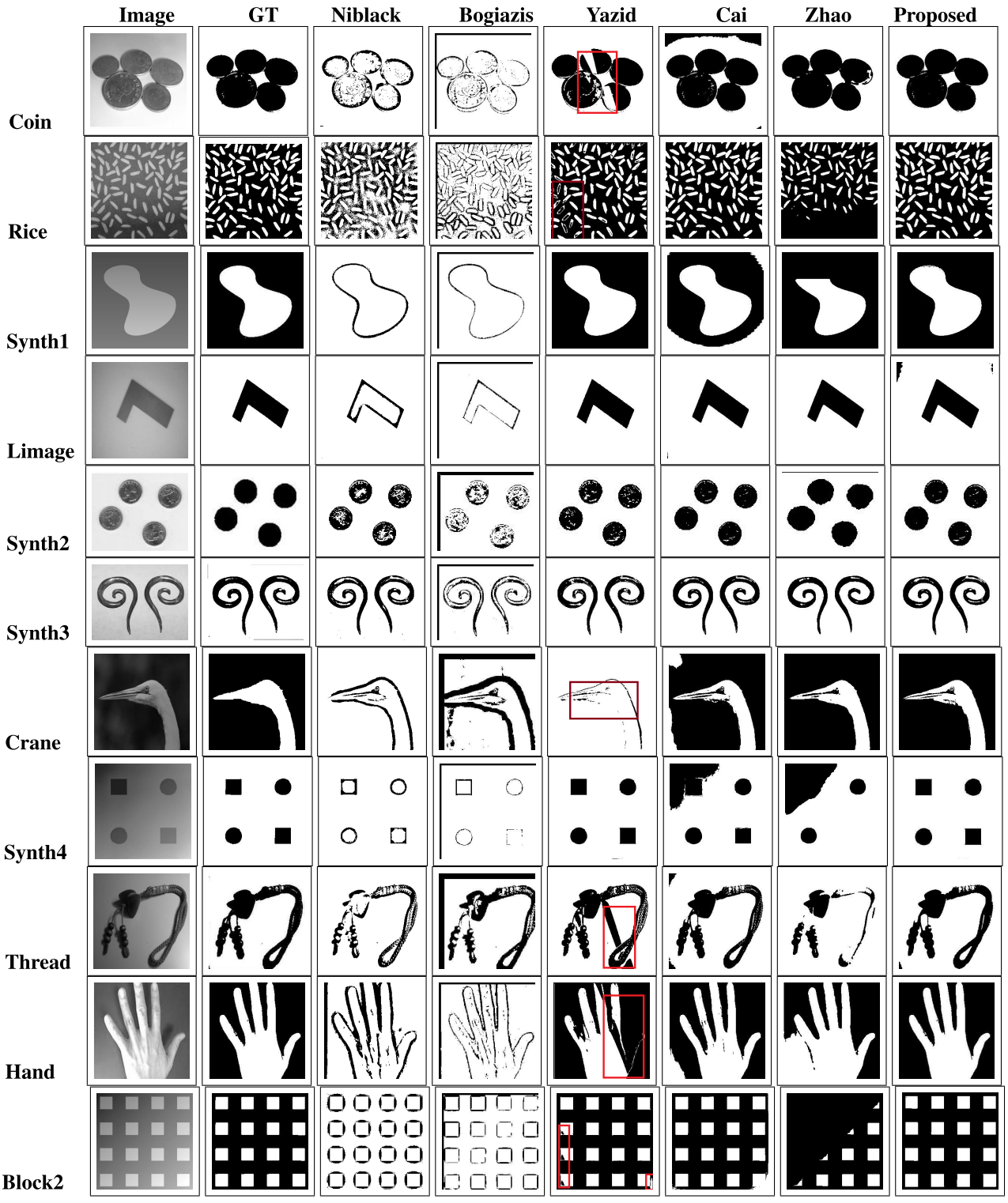

FIGURE 13. Segmentation results of Internet images: Col 1: Original uneven light images; Col 2: GT images; Col 3: to Col 8: are the binarized images of Niblack [6], Bogiatzis [11], Yazid [18], Cai and Miklavcic [19], Zhao [8], and Proposed Method respectively.

To further verify the effectiveness of each of these approaches, quantitative metrics such as P_r, R_e, F_1, JI, PME, and time complexity are employed. The P_r, R_e, F_1, JI, and PME measurements for each of these approaches are displayed, correspondingly, on all 21 test pictures in Tables 2, 3, 4, 5, and 6. These methods on all 21 test images.

Table 2 shows that for the majority of the images in the suggested technique with an average, the P_r value is high.

TABLE 2. Performance measure precision (P_r).

Image	Niblack [6]	Bogiazis [11]	Yazid [18]	Cai [19]	Zhao [8]	Proposed Method
Coin	0.67	0.6	0.9	0.98	0.4	**0.99**
Rice	0.46	0.26	0.96	0.96	**0.99**	**0.95**
Synth1	0.3	0.32	0.98	0.72	**0.99**	0.98
Limage	0.9	0.87	0.99	1	0.99	**1**
Synth2	0.92	0.8	0.96	0.97	0.99	**0.99**
Synth3	0.97	0.87	0.97	0.98	0.97	**0.98**
Crane	0.18	0.19	0.18	0.81	**1**	0.99
Synth4	0.93	0.90	0.99	0.99	0.96	**0.99**
Thread	0.86	0.80	0.94	0.95	0.85	**0.96**
Hand	0.46	0.41	0.96	0.92	0.74	**0.97**
Block2	0.32	0.22	0.98	0.97	**0.99**	0.98
Pins1	0.95	0.91	0.97	0.97	0.99	**0.98**
Light	0.7	0.6	0.87	0.99	0.98	**0.99**
Nightyeight	0.75	0.62	0.9	0.94	0.94	**0.95**
Bandage	0.74	0.72	0.98	0.98	0.98	**0.98**
Plane	0.98	0.96	0.99	0.98	0.98	**0.99**
Flower	0.11	0.11	0.10	0.55	0.90	**0.99**
Tower	0.86	0.74	0.99	0.99	0.99	**0.99**
Finger	0.88	0.65	0.8	0.98	0.93	**0.99**
Boat	0.13	0.13	0.98	0.96	0.99	**0.99**
3096	0.98	0.96	0.99	0.99	0.99	**0.99**
AVG	0.66	0.60	0.87	0.93	0.93	**0.98**

TABLE 3. Performance measure recall (R_e).

Image	Niblack [6]	Bogiazis [11]	Yazid [18]	Cai [19]	Zhao [8]	Proposed Method
Coin	0.99	0.91	**0.99**	0.88	0.21	0.98
Rice	**0.99**	0.77	0.88	0.98	0.56	0.98
Synth1	0.96	**0.99**	0.96	0.96	0.8	0.96
Limage	0.99	0.94	0.99	0.99	**0.99**	0.98
Synth2	**0.99**	0.88	0.98	0.98	**0.93**	0.92
Synth3	0.99	0.91	0.99	0.99	0.99	**0.99**
Crane	0.92	0.74	**0.96**	0.83	0.86	0.88
Synth4	0.99	0.95	**0.99**	0.86	0.82	**0.99**
Thread	0.99	0.74	**0.91**	0.97	**0.99**	0.98
Hand	0.92	0.89	**0.81**	0.94	**0.99**	0.95
Block2	0.99	0.93	0.97	0.98	0.36	**0.99**
Pins1	0.99	0.95	0.99	0.99	0.99	**0.99**
Light	0.99	0.84	0.97	0.99	**0.99**	0.98
Nightyeight	0.97	0.85	0.98	0.98	0.98	**0.99**
Bandage	0.99	0.93	0.99	0.99	0.98	**0.99**
Plane	0.99	0.94	0.99	0.99	0.99	**0.99**
Flower	0.85	0.80	**0.86**	**0.94**	0.92	0.93
Tower	0.96	0.91	0.99	0.99	0.99	**0.99**
Finger	0.99	0.74	0.98	**0.99**	0.95	0.98
Boat	**0.94**	0.87	0.83	0.82	0.88	0.85
3096	0.99	0.96	0.99	0.99	0.94	**0.99**
AVG	**0.97**	0.87	0.95	0.95	0.86	0.96

The value of P_r score of 0.98. The Yazid technique [18] has the third-highest average P_r, at 0.87, while Cai and Miklavcic [19] and Zhao [8] have the second-highest average, at 0.93. The high value of P_r suggests that there are more

TABLE 4. Performace measure F-Measure (F_1).

Image	Niblack [6]	Bogiazis [11]	Yazid [18]	Cai [19]	Zhao [8]	Proposed Method
Coin	0.8	0.73	0.94	0.93	0.28	**0.99**
Rice	0.63	0.39	0.92	0.97	0.72	**0.97**
Synth1	0.46	0.48	0.97	0.82	0.89	**0.97**
Limage	0.94	0.9	0.99	0.99	0.99	**0.99**
Synth2	0.95	0.83	0.97	**0.98**	0.96	0.96
Synth3	0.98	0.89	0.98	0.99	0.98	**0.99**
Crane	0.30	0.30	0.30	0.82	0.92	**0.93**
synth4	0.96	0.92	0.99	0.92	0.88	**0.99**
thread	0.92	0.77	0.93	0.96	0.91	**0.97**
Hand	0.62	0.56	0.88	0.93	0.84	**0.96**
Block2	0.49	0.44	0.97	0.98	0.53	**0.99**
Pins1	0.97	0.93	0.98	0.98	0.99	**0.99**
Light	0.82	0.7	0.92	**0.99**	0.98	0.98
Nightyeight	0.84	0.72	0.94	0.96	0.96	**0.97**
Bandage	0.85	0.81	0.98	0.98	0.98	**0.98**
Plane	0.99	0.95	0.99	0.99	0.99	**0.99**
Flower	0.20	0.19	0.19	0.69	0.91	**0.95**
Tower	0.87	0.82	0.99	0.99	0.99	**0.99**
Finger	0.93	0.69	0.88	0.98	0.94	**0.98**
Boat	0.24	0.22	0.9	0.88	**0.93**	0.91
3096	0.99	0.96	0.99	0.99	0.96	**0.99**
AVG	0.75	0.67	0.88	0.93	0.88	**0.97**

TABLE 5. Performance measure jaccard index (*JI*).

Image	Niblack [6]	Bogiazis [11]	Yazid [18]	Cai [19]	Zhao [8]	Proposed Method
Coin	0.67	0.57	0.9	0.87	0.16	**0.98**
Rice	0.44	0.24	0.85	**0.95**	0.56	0.94
Synth1	0.3	0.31	**0.95**	0.7	0.8	**0.95**
Limage	0.9	0.83	0.99	**0.99**	0.99	0.98
Synth2	0.91	0.72	0.95	**0.96**	0.93	0.92
Synth3	0.97	0.81	0.97	0.98	0.96	**0.98**
Crane	0.18	0.18	0.17	0.69	0.86	**0.88**
synth4	0.93	0.86	0.99	0.86	0.79	**0.98**
thread	0.85	0.62	0.86	0.93	0.85	**0.95**
Hand	0.44	0.39	0.79	0.87	0.73	**0.93**
Block2	0.32	0.23	0.95	0.96	0.36	**0.98**
Pins1	0.94	0.87	0.96	0.97	0.98	**0.98**
Light	0.7	0.54	0.85	**0.99**	0.97	0.97
Nightyeight	0.73	0.56	0.89	0.94	0.93	**0.94**
Bandage	0.74	0.69	0.97	0.97	0.97	**0.97**
Plane	0.98	0.91	**0.99**	0.98	0.98	0.98
Flower	0.11	0.11	0.10	0.53	0.83	**0.92**
Tower	0.78	0.69	0.99	0.99	0.99	**0.99**
Finger	0.88	0.53	0.79	0.97	0.89	**0.97**
Boat	0.13	0.12	0.82	0.79	**0.88**	0.84
3096	0.98	0.93	0.98	0.99	0.93	**0.99**
AVG	0.66	0.55	0.84	0.89	0.82	**0.95**

true positives and fewer false positives. The Niblack approach has the highest average when taking into account the R_e metric of 0.97 as shown in Table 3. The proposed technique has the second highest 0.96, whereas the Yazid and Cai

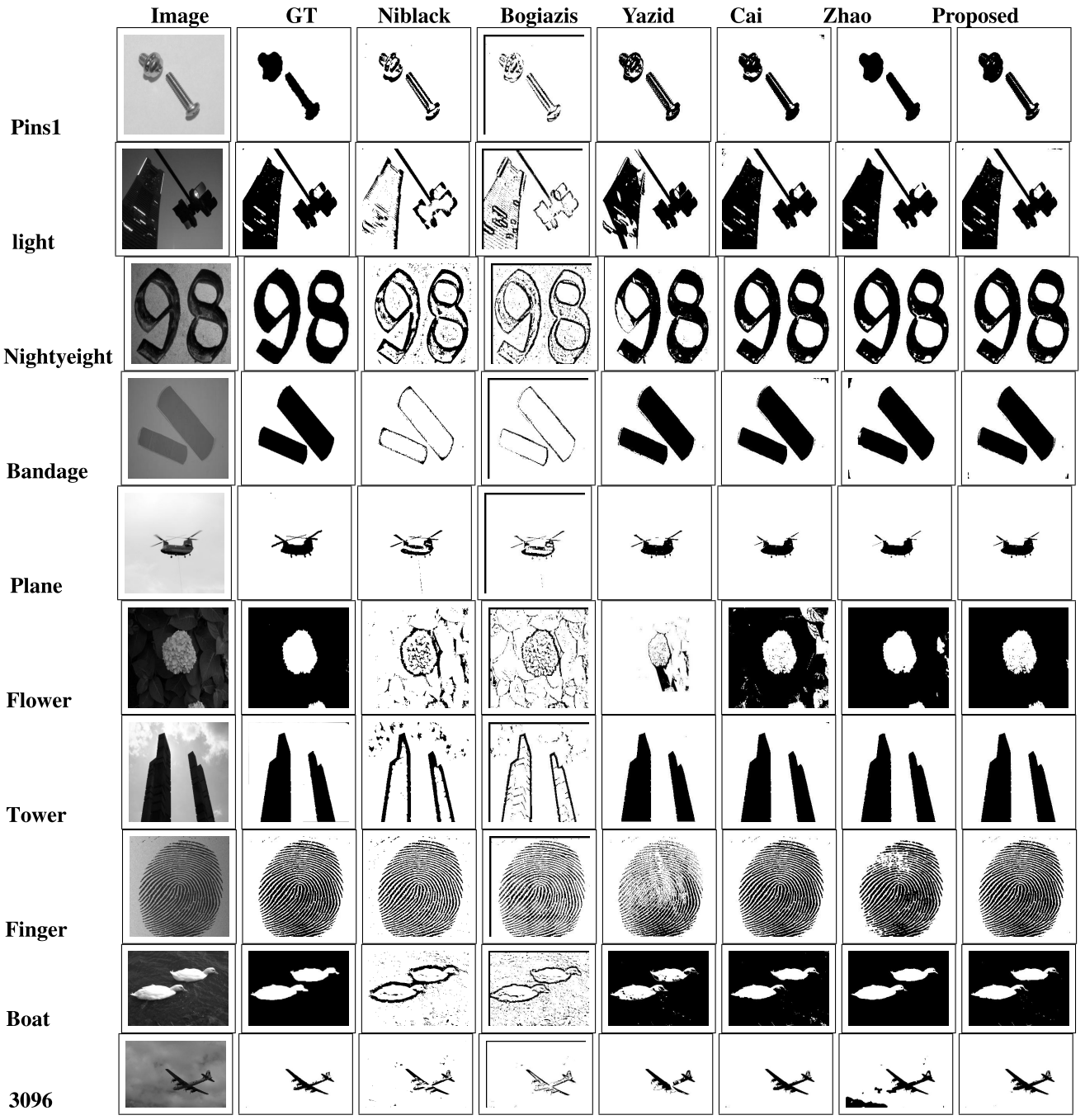

FIGURE 14. Segmentation results of dataset images: Col 1: Original uneven light images; Col 2: GT images; The binarized images of Niblack [6], Bogiatzis [11], Yazid [18], Cai and Miklavcic [19], Zhao [8], and the Proposed Method are the outputs, corresponding to Cols 3 through Col 8.

method has the highest score of 0.95. Consequently, when comparing the average R_e, the amount of false negatives in the suggested method is nearly negligible. It has been noted that the suggested approach has the lowest false positive and false negative values and the highest true positive value.

Likewise, Table 4 presents the $F1$ measure, which is a combination of R_e and P_r. The average for the suggested method is F_1 is 0.97. The average for Cai is F_1 is 0.93, whereas the average for Yazid and Zhao is F_1 is 0.88. The similarity metric between the segmented image and its ground truth is called the Jaccard Index (JI). Table 5 demonstrates that, when compared to other methods, the suggested technique has the highest JI measure of 0.95. Fig. 15 displays the average P_r, R_e, F_1, and JI score for each of these five approaches using the suggested FLANN-based methodology. It is noted that the proposed FLANN technique performs best out of all the methodologies taken

TABLE 6. Performance measure (*PME*).

Image	Niblack [6]	Bogiazis [11]	Yazid [18]	Cai [19]	Zhao [8]	Proposed Method
Coin	29.2	40.59	6.56	7.39	2.24	**0.9**
Rice	30.25	61.5	3.88	**1.3**	11.5	1.5
Synth1	66.96	63.8	1.45	12.27	5.87	**1.4**
Limage	10.28	16.3	0.46	0.34	**0.12**	1.3
Synth2	12.77	25.69	3.72	2.41	5.03	**2.3**
Synth3	7.07	17.24	2.1	1.34	2.68	**1.3**
Crane	76.4	62.29	80.9	6.60	2.4	**2.0**
synth4	6.28	13.4	**0.3**	12.69	18.62	1.0
Thread	12.28	32.2	10.26	5.0	13.09	**2.3**
Hand	45.1	55.1	8.6	5.3	14.08	**2.4**
Block2	59.3	63.11	1.19	2.41	17.15	**0.37**
Pins1	4.9	11.7	2.9	2.04	**1.15**	1.4
Light	26.6	33.8	9.9	**0.48**	1.41	0.56
Nightyeight	24.87	37.8	6.71	3.47	3.69	**3.4**
Bandage	23.68	30.24	1.4	**1.2**	2	1.4
Plane	1.8	7.9	0.96	1	1.13	**1.01**
Flower	80.7	78.5	86.6	9.9	2.15	**0.95**
Tower	18.7	28.4	0.5	0.40	0.62	**0.39**
Finger	18.2	40.03	15.3	1.58	23.01	**1.3**
Boat	80.06	76.08	2.2	2.73	1.47	1.9
3096	3.2	6.53	1.5	**0.57**	5.89	0.78
AVG	30.4	38.2	11.7	3.83	6.44	**1.42**

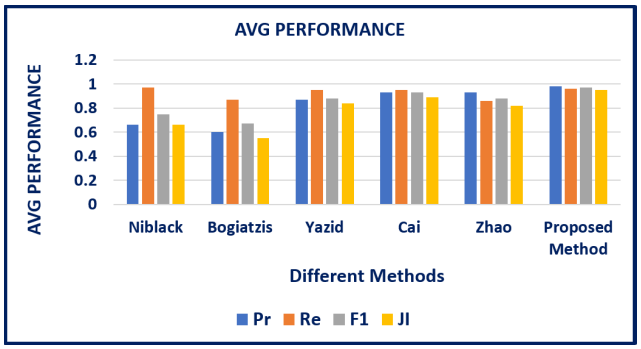

FIGURE 15. Average performace of P_r, R_e, $F1$, and JI.

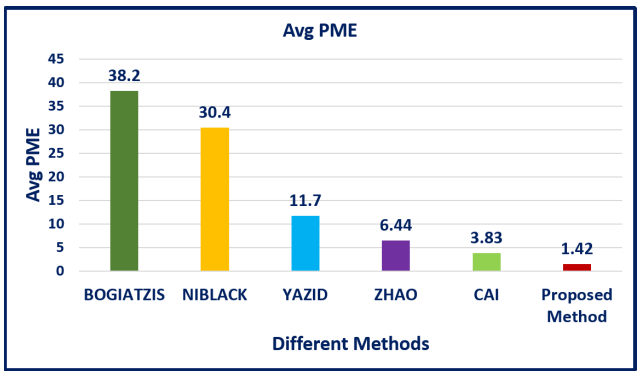

FIGURE 16. AVG percentage of misclassification error *PME*.

into consideration for the validation and is consistent in every measure.

TABLE 7. Average performance for different approach.

Different Approach	Pr	Re	F1	JI	PME
Niblack [6]	0.66	**0.97**	0.75	0.66	30.4
Bogiatzis [11]	0.60	0.87	0.67	0.55	38.2
Yazid [18]	0.87	0.95	0.88	0.84	11.7
Cai [19]	0.93	0.95	0.93	0.89	3.83
Zhao [8]	0.93	0.86	0.88	0.82	6.44
Proposed	**0.98**	0.96	**0.97**	**0.95**	**1.42**

Comparably, in Table 6 the PME measure shows suggested approach has the least amount of misclassification (*PME*), of 1.42%. However, *PME* of 4% and 6%, respectively, are found in the Cai and Zhao et al. approaches. Fig. 16 displays the average *PME* of the five approaches as well as the suggested FLANN method. When evaluating both qualitative and quantitative metrics, the suggested method performed better than all state-of-the-art techniques on every variation of example images with non-uniform illumination. It is observed from Table 7 shows that the proposed method has the highest average *JI* measure, of 0.95. The methods of Cai et al. [19] and Yazid et al. [18] rank second and third, respectively, with average *JI* scores of 0.89 and 0.84. Similarly, average F1-measure for suggested method is of 0.97, and Cai and Miklavcic [19] and Zhao et al. [8] is having the second and third highest of 0.93 and 0.88. For all types of uneven images, the proposed approach performs better

TABLE 8. Computation time in seconds.

Image	Niblack [6]	Bogiazis [11]	Yazid [18]	Cai [19]	Zhao [8]	Proposed Method
Coin	**0.09**	30.83	4.34	0.36	3.37	24.0
Rice	**0.13**	30.45	3.74	0.39	3.65	27.0
Synth1	**0.08**	30	3.3	**0.35**	3.67	12.09
Limage	0.68	30.63	3.16	**0.33**	3.69	27.0
Synth2	**0.11**	7.9	1.21	0.29	1.33	18.34
Synth3	**0.1**	23.9	2.39	0.36	3.41	31.0
Crane	0.17	55.38	1.39	0.18	5.34	**4.3**
synth4	0.08	29.87	1.46	0.29	3.65	**5.86**
thread	0.68	36.29	1.50	0.79	6.86	**13.71**
Hand	0.08	31.4	1.40	0.20	3.75	**6.7**
Block2	0.08	30.02	3.32	0.47	4.02	**2.44**
Pins1	0.92	31.6	3.33	**0.33**	3.62	25.0
Light	**0.17**	30.67	3.19	0.37	3.78	25.3
Nightyeight	**0.17**	31.05	2.84	0.35	4.14	28.0
Bandage	**0.23**	30.22	3.2	0.36	3.62	25.0
Plane	**0.17**	29.9	2.71	0.41	3.44	21.0
Flower	0.04	36.13	1.41	0.21	3.96	**4.3**
Tower	**0.15**	30.35	3.11	0.43	3.82	31.0
Finger	**0.16**	30.95	3.57	0.40	3.78	21.0
Boat	**0.2**	31.26	2.97	0.39	5.45	11.7
3096	**0.13**	75.7	6.02	0.47	9.18	31.0
AVG	**0.22**	33.07	2.83	0.37	4.23	18.8

TABLE 9. Ablation experiment results on database images and sample images.

Modules	Database Images						Sample Images					
	Pr	Re	F	JI	PME	TIME	Pr	Re	F	JI	PME	TIME
TST	0.95	0.95	0.96	0.94	2.0	6.99	0.90	0.88	0.78	0.79	11.20	6.56
EB	0.94	0.90	0.91	0.85	6.5	3.78	0.85	0.88	0.85	0.76	12.17	5.31
Combined (TST+EB))	0.98	0.96	0.97	0.95	1.3	20.0	0.98	0.96	0.97	0.95	1.52	17.6

than the other approaches when taking into account both qualitative and quantitative metrics. Table 8 computes the CPU time in order to assess the methods' complexity further. The average calculation time for the suggested approach is 18.8 seconds, the second-lowest average computation time is 0.37 seconds for the Cai et al. method, and the lowest average computation time is 0.22 seconds for the Niblack [6] method. On the other hand, the JI and F_1 measures in the Niblack technique are located in the 5^{th} position. In terms of F_1 and JI metrics, the Cai and Miklavcic [19] is ranked second. The suggested approach, however, performs best in terms of averages for P_r, R_e, F_1, JI, and PME.

In Table 9, a study was conducted to compare the specific impacts of different modules and their combined performance on various sample images through an ablation experimental approach. The comparison involved two modules: TST (Two Stage Thresholding) or dual threshold method, and EB (Edge Based) method, along with the performance of the combined Dual threshold and Edge based method. The evaluation included precision (Pr), recall (Re), F-measure (F1), Jaccard Index (JI), percentage of misclassification error (PME), and time complexity for database images such as Berkeley and Weizmann database images, as well as sample internet images.

In the table, the performance values for Precision (Pr), Recall (Re), F-measure (F), and Jaccard Index (JI) are of 0.98, 0.96, 0.97, and 0.95, respectively, for both the database and sample images. Similarly, the percentage of misclassification error(PME) of 1.3% for the database images and 1.52% for the sample internet images. Overall, the combined effect of the two modules results in an average PME is of 1.42%, which is better than the individual impacts of the TST and EB modules. Subsequently, the combined effect of two modules TST and EB-based method for extraction of valid sample point generation increases the time complexity compared to the individual impact. However, with the time complexities, the percentage of misclassification for those images are in the range of 1%-2%.

In some of the failure cases of the proposed method like: for some images such as synth-2, crane, thread, hand, and ninety-eight, the percentage of misclassification is in the range of 2%-3% as in Table. 6 quantitatively. However, qualitatively also these images are not accurate w.r.t to ground truth as discussed in Table 1. Also from the ablation study of the experiment of database and sample images of Table 9 due to the combined effect of two modules TST and EB-based methods, the time complexity for database and sample images is in the range of 15-20sec.

V. CONCLUSION

In this study, A quick and effective method for images with variable illumination conditions, local/adaptive thresholding computes a distinct threshold for each pixel based on its local vicinity. Nonetheless, the following are the main difficulties with local/adaptive thresholding methods: (i) It is important to choose the ideal window or sub-image size. A window that is too big might not be able to catch local differences, while a window that is too narrow will be more susceptible to noise. (ii) Because local/adaptive thresholding relies on computations based on pixels or sub-images, it is very sophisticated and memory-intensive. (iii) The presence of both background and object regions in every window or sub-image is necessary to prevent significant misclassification errors during local thresholding. (iv) Compared to global approaches, implementing local thresholding algorithms might be more complicated, presenting difficulties for developers and needing more sophisticated programming abilities. (v)The specific environmental characteristics of the image, such as differences in lighting and texture, can have a significant impact on how effective local thresholding parameters are. The goal is to normalize uneven illumination situations in order to improve the accuracy of global thresholding. By utilizing the FLANN model to normalize the illumination surface, the suggested approach lessens complexity and boosts efficiency in order to overcome these difficulties. Acquiring legitimate training sample points (TSP) is essential for training models in an efficient manner. Better training and more accurate segmentation results are guaranteed by the suggested approach, which improves the extraction process. The suggested approach addresses these issues and enhances image binarization accuracy and dependability in a range of lighting scenarios.

The suggested method is found to have the lowest $F1$ score of 0.97 and lowest PME on average, 1.42%, average JI score of 0.95, and an average time for computation of 18.8 seconds. To shorten the computation time at the sample point creation, parallel processing and FLANN functional expansion part parallelization are recommended.

ACKNOWLEDGMENT

The authors would like to acknowledge to the Universiti Kebangsaan Malaysia Research Grant through the Dana Padanan Kolaborasi (DPK) under the grant number DPK-2022-006. Also, extend their appreciation to Taif University, Saudi Arabia, for supporting this work through project number (TU-DSPP-2024-11).

REFERENCES

[1] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
[2] H. Cai, Z. Yang, X. Cao, W. Xia, and X. Xu, "A new iterative triclass thresholding technique in image segmentation," *IEEE Trans. Image Process.*, vol. 23, no. 3, pp. 1038–1046, Mar. 2014.
[3] C. Sha, J. Hou, and H. Cui, "A robust 2D Otsu's thresholding method in image segmentation," *J. Vis. Commun. Image Represent.*, vol. 41, pp. 339–351, Nov. 2016.
[4] J. Ma and X. Cheng, "Fast segmentation algorithm of PCB image using 2D OTSU improved by adaptive genetic algorithm and integral image," *J. Real-Time Image Process.*, vol. 20, pp. 596–605, Feb. 2023.
[5] J. Xing, P. Yang, and L. Qingge, "Robust 2D Otsu's algorithm for uneven illumination image segmentation," *Comput. Intell. Neurosci.*, vol. 2020, no. 1, 2020, Art. no. 5047976, doi: 10.1155/2020/5047976.
[6] W. Niblack, *An Introduction to Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1986.
[7] D. Bradley and G. Roth, "Adaptive thresholding using the integral image," *J. Graph. Tools*, vol. 12, no. 2, pp. 13–21, 2007.
[8] L. Zhao, S. Zheng, W. Yang, H. Wei, and X. Huang, "An image thresholding approach based on Gaussian mixture model," *Pattern Anal. Appl.*, vol. 22, pp. 75–88, Jan. 2019.
[9] X. Pan, Q. Zhang, and H. Pan, "Improved artificial bee colony algorithm and its application to fundus retinal blood vessel image binarization," *IEEE Access*, vol. 8, pp. 123726–123734, 2020.
[10] Q. Huang, W. Gao, and W. Cai, "Thresholding technique with adaptive window selection for uneven lighting image," *Pattern Recognit. Lett.*, vol. 28, pp. 801–808, May 2005.
[11] A. C. Bogiatzis and B. K. Papadopoulos, "Local thresholding of degraded or unevenly illuminated documents using fuzzy inclusion and entropy measures," *Evolving Syst.*, vol. 10, no. 4, pp. 593–619, Dec. 2019.
[12] P. Kanungo, P. K. Nanda, and A. Ghosh, "Parallel genetic algorithm based adaptive thresholding for image segmentation under uneven lighting conditions," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, Oct. 2010, pp. 1904–1911.
[13] P. Kanungo, P. K. Nanda, and A. Ghosh, "Detection of Earth surface cracks using parallel genetic algorithm based thresholding," in *Proc. Int. Conf. Adv. Comput., Control, Telecommun. Technol.*, Dec. 2009, pp. 4–11.
[14] T. Pattnaik and P. Kanungo, "GMM based adaptive thresholding for uneven lighting image binarization," *J. Signal Process. Syst.*, vol. 93, pp. 1253–1270, Nov. 2021.
[15] T. Pattnaik and P. Kanungo, "Adaptive window selection for non-uniform lighting image thresholding," *ELCVIA: Electron. Lett. Comput. Vis. Image Anal.*, vol. 20, no. 1, pp. 42–54, 2021.
[16] S. D. Yanowitz and A. M. Bruckstein, "A new method for image segmentation," *Comput. Vis., Graph., Image Process.*, vol. 46, no. 1, pp. 82–95, 1989.
[17] I. Blayvas, A. Bruckstein, and R. Kimmel, "Efficient computation of adaptive threshold surfaces for image binarization," *Pattern Recognit.*, vol. 39, no. 1, pp. 89–101, Jan. 2006.
[18] H. Yazid and H. Arof, "Gradient based adaptive thresholding," *J. Vis. Commun. Image Represent.*, vol. 24, no. 7, pp. 926–936, 2013.
[19] J. Cai and S. Miklavcic, "Surface fitting for individual image thresholding and beyond," *IET Image Process.*, vol. 7, no. 6, pp. 596–605, 2013.
[20] P. Jain, M. T. Islam, and A. S. Alshammari, "Comparative analysis of machine learning techniques for metamaterial absorber performance in terahertz applications," *Alexandria Engineering J.*, vol. 103, pp. 51–59, Sep. 2024.
[21] A. D. Watpade, S. Thakor, P. Jain, P. P. Mohapatra, C. R. Vaja, A. Joshi, D. V. Shah, and M. Tariqul Islam, "Comparative analysis of machine learning models for predicting dielectric properties in MoS_2 nanofiller-reinforced epoxy composites," *Ain Shams Eng. J.*, vol. 15, no. 6, Jun. 2024, Art. no. 102754.

[22] P. K. Sahoo, M. K. Panda, U. Panigrahi, G. Panda, P. Jain, M. S. Islam, and M. T. Islam, "An improved VGG-19 network induced enhanced feature pooling for precise moving object detection in complex video scenes," *IEEE Access*, vol. 12, pp. 45847–45864, 2024.
[23] P. Jain, H. Chhabra, U. Chauhan, K. Prakash, P. Samant, D. K. Singh, M. S. Soliman, and M. T. Islam, "Machine learning techniques for predicting metamaterial microwave absorption performance: A comparison," *IEEE Access*, vol. 11, pp. 128774–128783, 2023.
[24] N. Misran, M. T. Islam, M. Y. Ismail, and S. H. Yusop, "Analisis pencirian parameter ketebalan dan kebertelusan substrat bagi elemen cincin segiempat sepusat bersela antena tatasusun pantulan, *Jurnal Kejuruteraan*, vol. 23, p. 1115, Nov. 2011.
[25] P. Jain, H. Chhabra, U. Chauhan, K. Prakash, A. Gupta, M. S. Soliman, M. S. Islam, and M. T. Islam, "Machine learning assisted hepta band THz metamaterial absorber for biomedical applications," *Sci. Rep.*, vol. 13, no. 1, p. 1792, Jan. 2023.
[26] S. He and L. Schomaker, "DeepOtsu: Document enhancement and binarization using iterative deep learning," *Pattern Recognit.*, vol. 91, pp. 379–390, Jul. 2019.
[27] S. Dehuri, R. Roy, S.-B. Cho, and A. Ghosh, "An improved swarm optimized functional link artificial neural network (ISO-FLANN) for classification," *J. Syst. Softw.*, vol. 85, no. 6, pp. 1333–1345, Jun. 2012.
[28] F. Calderon, A. Garnica-Carrillo, and C. Reyes-Zuñiga, "Binarization of images with variable lighting using adaptive windows," *Signal, Image Video Process.*, vol. 16, no. 7, pp. 1905–1912, Oct. 2022.
[29] E. Molina, J. Diaz, H. Silva, and E. Chvez, "Algoritmos de binarizacion robusta de imágenes con iluminación no uniforme," *Revista Iberoamericana de Automática e Informática Ind.*, vol. 15, no. 3, pp. 252–261, 2018.
[30] Alpert. *Weizmann Segmentation Evaluation Database*. Accessed: Oct. 30, 2027. [Online]. Available: http://www.wisdom.weizmann.ac.il/~vision/Seg-Evaluation-DB
[31] Martin. *Berkeley Segmentation Dataset*. Accessed: Oct. 30, 2017. [Online]. Available: http://www2.eecs.berkeley.edu/Research/projects/CS/vision/bsds/
[32] M. Yang and X. Fan, "YOLOv8-Lite: A lightweight object detection model for real-time autonomous driving systems," *IECE Trans. Emerg. Topics Artif. Intell.*, vol. 1, no. 1, pp. 1–16, 2024.
[33] S. Wang, "Real-time object detection using a lightweight two-stage detection network with efficient data representation," *IECE Trans. Emerg. Topics Artif. Intell.*, vol. 1, no. 1, pp. 17–30, 2024.
[34] G. Cheng, C. Lang, and J. Han, "Holistic prototype activation for few-shot segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 4, pp. 4650–4666, Apr. 2023.
[35] C. Lang, G. Cheng, B. Tu, and J. Han, "Few-shot segmentation via divide-and-conquer proxies," *Int. J. Comput. Vis.*, vol. 132, no. 1, pp. 261–283, Jan. 2024.
[36] C. Lang, J. Wang, G. Cheng, B. Tu, and J. Han, "Progressive parsing and commonality distillation for few-shot remote sensing segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5613610.
[37] C. Lang, G. Cheng, B. Tu, C. Li, and J. Han, "Base and meta: A new perspective on few-shot segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 9, pp. 10669–10686, Sep. 2023.

TAPASWINI PATTNAIK (Member, IEEE) received the B.Tech. degree in electronics and telecommunication engineering and the M.Tech. degree in electronics and communication engineering from Biju Patnaik University (BPUT), Rourkela, Odisha, in 2004 and 2010, respectively, where she is currently pursuing the Ph.D. degree in image processing. Currently, she is an Assistant Professor with the Department of Electronics and Communication Engineering, C. V. Raman Global University. She has published many research papers in refereed international conferences and journals. Her research interests include image processing and machine learning and deep learning.

PRIYADARSHI KANUNGO (Senior Member, IEEE) received the degree in electrical engineering from the Institute of Engineers, Calcutta, in 1997, the M.Tech. degree in electronics system and communication from REC, Rourkela, Orissa, in 2001, and the Ph.D. degree in engineering from the National Institute of Technology, Rourkela, in 2010. He is currently holds the position of a Professor with the Department of Electronics and Telecommunication Engineering, Driems University, Cuttack, Odisha, India. With over 22 years of experience in both teaching and research, his primary areas of expertise encompass signal processing, image analysis, computer vision, parallel genetic algorithms, evolutionary computation and bioinformatics, machine learning, MIMO systems, soft computing, and robotics. He has contributed to over 50 scientific publications in esteemed journals and international conferences.

PRABODH KUMAR SAHOO (Member, IEEE) received the M.E. degree from Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, India, in 2005, and the Ph.D. degree from the Centurion University of Technology and Management, Odisha, India, in 2019. He is currently an Associate Professor with the Department of Mechatronics, Parul University, Vadodara, Gujarat, India. He has six peer reviewed journals, five international conference articles, and two patents to his credit. His research interests include image processing, computer vision, and cyber-physical systems.

TEJASWINI KAR (Senior Member, IEEE) received the B.Tech. degree in electronics and telecommunication engineering from BPUT, in 2003, and the M.Tech. degree in communication system engineering and the Ph.D. degree in electronics and telecommunication engineering from KIIT Deemed to be University, Bhubaneswar, India, in 2008 and 2018, respectively. She has more than 18 years of teaching experience. She is currently an Assistant Professor with the School of Electronics Engineering, KIIT Deemed to be University. She has published many research papers in refereed international conferences and journals. Her current research interests include image and video processing and machine learning and deep learning.

PRINCE JAIN received the Doctor of Philosophy (Ph.D.) degree from Punjab Engineering College (Deemed to be University), Chandigarh, India. He is currently an Assistant Professor (Research Cadre) with the Mechatronics Engineering Department, Parul Institute of Technology, Parul University, Vadodara, India. He received the Visvesvaraya Ph.D. Scheme Fellowship to complete the Doctor of Philosophy (Ph.D.) Dissertation. He is the author or co-author of about 28 research journal articles, 20 conference papers, and a few book chapters on various topics related to machine learning, artificial intelligence and metamaterials. He has contributed as a peer reviewer for prestigious publishers, including IEEE, Springer Nature, Elsevier, IOPscience, Wiley, PIER, Emerald, and PLOS. His research interests include machine learning, artificial intelligence, optimization techniques, metamaterial absorbers/antennas at RF, THz and visible frequencies, material science, nanotechnology, and biomedical signal processing. He serving as an Academic Editor for *Journal of Electrical and Computer Engineering* (Wiley), *Discover Applied Sciences* (Springer), and *PLOS One* journal. He is serving as a Topical Advisory Panel Member for *Micromachines* and *Materials* (MDPI).

MOHAMED S. SOLIMAN (Senior Member, IEEE) received the Ph.D. degree in communications engineering from the Graduate School of Engineering, Osaka University, Japan. He is currently an Associate Professor with the Department of Electrical Engineering, College of Engineering, Taif University, Saudi Arabia, where he was granted many research projects from the Deanship of Scientific Researches. He is the author or co-author of more than 140 papers indexed in Scopus and WOS databases and five book chapters. He has been recognized as a Distinguished Researcher (one of the 100 most active researchers) from the Deanship of Scientific Research, Taif University, in 2023, where he was rewarded the Distinguished Researchers Supporting Project and Scientific Publication Reward for 2023 from the Deanship of Scientific Research. His research interests include wireless communications, phased and timed array signal processing, UWB antennas, MIMO antennas, dielectric resonant antennas, optimization techniques in antenna design, antenna measurement techniques, metamaterial structures, biosensors, RF energy harvesting systems, and numerical methods in electromagnetics. He is a Senior Member of the IEEE-MTT/AP Society. Also, he serves as a Reviewer for many scientific journals [i.e., PIER—PhotonIcs and Electromagnetics Research Symposium, *International Journal of RF and Microwave Computer-Aided Engineering* (Wiley Hindawi Partnership), and *Wireless Personal Communications* (Springer)]; and a TPC member for many international conferences.

MOHAMMAD TARIQUL ISLAM (Senior Member, IEEE) is currently a Professor with the Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia (UKM), and a Visiting Professor with Kyushu Institute of Technology, Japan. He is the author or co-author of about 600 research journal articles, nearly 250 conference articles, and a few book chapters on various topics related to antennas, metamaterials, and microwave imaging with 25 inventory patents filed. Thus far, his publications have been cited 14,000 times and his H-index is 53 (Source: Scopus). His Google scholar citation is 22500 and H-index is 62. He was a recipient of more than 40 research grants from Malaysian Ministry of Science, Technology and Innovation, Ministry of Education, UKM research grant, international research grants from Japan, Saudi Arabia, and Kuwait. His research interests include communication antenna design, metamaterial, satellite antennas, and microwave imaging. He served as an Executive Committee Member for IEEE AP/MTT/EMC Malaysia Chapter, from 2019 to 2020, the Chartered Professional Engineer (C.Eng.), a fellow of IET, U.K., and a Senior Member of IEICE, Japan. He received several international gold medal awards, a Best Invention in Telecommunication Award for his Research and Innovation and Best Researcher Awards from UKM. He was a recipient of 2018, 2019, and 2020 IEEE AP/MTT/EMC Malaysia Chapter, Excellent Award. He also won the Best Innovation Award and the Best Researcher Award by UKM, in different years. He was a recipient of Publication Award from Malaysian Space Agency, in several years. He has supervised about 50 Ph.D. theses, 30 M.Sc. theses, and has mentored more than ten postdoctoral's and visiting scholars. He has developed the Antenna Measurement Laboratory which includes antenna design and measurement facility till 40 GHz. He was an Associate Editor of *IET Electronics Letter*. He also serves as a Guest Editor for *Sensors* journal and *Nanomaterials* and an Associate Editor for IEEE ACCESS.