Expert Systems With Applications 209 (2022) 118272

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

A feature level image fusion for Night-Vision context enhancement using Arithmetic optimization algorithm based image segmentation

Simrandeep Singh [a,g], Harbinder Singh [b], Nitin Mittal [c], Harbinder Singh [d], Abdelazim G. Hussien [e,h], Filip Sroubek [f]

[a] Department of Electronics & Communication Engineering, UCRD Chandigarh University, Gharuan, Punjab, India
[b] Department of Electronics & Communication Engineering, Chandigarh Engineering College, Landran, Punjab, India
[c] Department of Skill Faculty of Science and Technology, Shri Vishwakarma Skill University
[d] Department of Electronics & Communication Engineering, Chandigarh University, Gharuan, Punjab, India
[e] Faculty of Science, Fayoum University, Fayoum 63514, Egypt
[f] Institute of Information Theory and Automation, The Czech Academy of Sciences, 18208 Prague 8, Czech Republic
[g] Department of Computer Science and Engineering, IIT Ropar, Rupnagar 140001, India
[h] Department of Computer and Information Science, Linköping University, SE-581 83 Linköping, Sweden

ARTICLE INFO

Keywords:
Infrared (IR)
visible (Vis) image
Image Fusion
AOA
Image segmentation
WLS

ABSTRACT

Images are fused to produce a composite image by combining key characteristics of the source images in image fusion. It makes the fused image better for human vision and machine vision. A novel procedure of Infrared (IR) and Visible **(Vis)** image fusion is proposed in this manuscript. The main challenges of feature level image fusion are that it will introduce artifacts and noise in the fused image. To preserve the meaningful information without adding artifacts from the source input images, weight map computed from Arithmetic optimization algorithm (AOA) is used for the image fusion process. In this manuscript, feature level fusion is performed after refining the weight maps using a weighted least square optimization (WLS) technique. Through this, the derived salient object details are merged into the visual image without introducing distortion. To affirm the validity of the proposed methodology simulation results are carried for twenty-one image data sets. It is concluded from the qualitative and quantitative experimental analysis that the proposed method works well for most of the image data sets and shows better performance than certain traditional existing models.

1. Introduction

Owing to weak night-time illumination conditions, visible images are mostly fused with the accompanying infrared (IR) pictures to improve the background of cinematic sequences. A hybrid approach using local Laplacian filter for edge preserving and enhancement along with segmentation-based weight map fusion approach has been proposed. It enhances the night vision context of infrared and visible image fusion and also keeps edges intact without adding any artifacts.

In recent times, the detectability of military targets has reduced substantially with defence system (Meher, Agrawal, Panda, Dora, & Abraham, 2022; James & Kavitha, 2014). The military specifications under those circumstances have stimulated the improvement of multi-mode image fusion technology, which typically uses IR and visible images to obtain complementary information. The IR image tracks heat energy emitted from objects in the scene and can be used to discover objectives because it has a hot contrast, whilst the visible picture has considerably more high-frequency background information, which is important if the target positions and circumstances are to be accurately identified.

The improved image processing system performance has led to the development of several image fusion techniques which combine information gathered through the various sensors (Jiang, Jin, Hou, Lee, & Yao, 2018; Pan, Shi, & Xu, 2017; Piella, 2003). The goal of image fusion is to combine the proportion able attributes of the images acquired with the best visual effects in a fused image. This fused image generates details which cannot be obtained by separately analysing various images. Other cameras such as infrared cameras, which capture images in different wavelengths, are often favoured along with a digital charge-coupled device (CCD) camera (Li & Yang, 2008). Using different

E-mail addresses: simrandeepsingh.ece@cumail.in (S. Singh), harbinder.ece@cgc.edu.in (H. Singh), nitinmittal.me@cumail.in (N. Mittal), harbindersingh.ece@cumail.in (H. Singh), abdelazim.hussien@liu.se (A.G. Hussien), sroubekf@utia.cas.cz (F. Sroubek).

https://doi.org/10.1016/j.eswa.2022.118272
Received 25 September 2021; Received in revised form 21 June 2022; Accepted 22 July 2022
Available online 26 July 2022

imaging sensors to capture images for the same location helps to get enhanced outputs once they are fused (Singh, Singh, Gehlot, kaur, & Gagandeep, 2022). Multi-modal image fusion is a combination of data acquired from different sensors (Singh et al., 2013, 2020), so that the information finally generated has lesser uncertainties and more information as compared to the individual performance of each sensor (Li & Wang, 2022; Liu & Wang, 2015; Xu, Ma, Jiang, Guo, & Ling, 2020; Zhang, Xu, Tian, Jiang, & Ma, 2021). In order to perform segmentation, OTSU as an objective function to perform segmentation is discussed in the forthcoming section 2. Section 3 conferred fusion methodology AOA algorithm and proposed modified algorithm AOA. Result analysis along with performance metric are discussed section 4 and section 5 presents conclusion.

2. Related work

It is an important preliminary task which plays a crucial role in the field of medicine and other computer vision applications. It is a hot research topic gaining rapid pace in recent years. In region based fusion image thresholding provides significant support (Singh, Mittal, & Singh, 2020; 2021), (Kaur & Singh, 2017). Some of the approaches to image segmentation are addressed in subsequent sections.

2.1. Image segmentation

In general, monochrome image segmentation algorithms are based on one of two fundamental characteristics of the intensity values: discontinuity and similarity. Partitioning a picture in the first group is based on the sudden gray level shifts, where discrete point, line and edge detection are discovered in the pictures. Thresholding and region-based algorithms fall into the second category. Thresholding technique based on histogram analysis is found to be the simplest in segregating the object regions and the background (Singh et al., 2022). The drawback of this technique is that it may result in an improper segmentation, if the valley points of the histogram are not sharp. Each category of segmentation has got its own advantages and disadvantages.

Segmentation of fused image finds its applications in medical field, pattern recognition, machine object recognition, feature extraction etc., Segmentation technique in medical field is used for segmenting the tumor part from the non-cancerous part. Fuzzy techniques are better suited for the analysis of such complex natural systems and have been realized in various application domains. All the major ideas involved in fuzzy set theory, fuzzy logic and fuzzy systems are found in and are effectively utilized in imaging analysis. These ideas on fuzzy techniques lend a helping hand in dealing with the inherent imprecision in gray values present in images and fuzzy statistics has proved its superiority under such circumstances. Fuzzy C-means clustering, which is the widely used method for image segmentation is also instrumental in developing most of the other analytic fuzzy clustering approaches. In monochrome image segmentation techniques, information is obtained only from gray values. Since, color plays a vital role in visual experience for human being and also because of the reason that color conveys more information, it is considered to be of paramount importance in image analysis. This led to the development of many fuzzy based color image segmentation algorithms in recent years which help in achieving a more meaningful and robust segmentation performance. However, color image processing involves special challenges, since they are multidimensional.

If bilevel thresholding is carried out image may be segregated into two parts using single optimal threshold value (Th). Image will be classified into two classes b_1 and b_2 as mentioned in the following equation for an image $c(m,n)$ having size (x * y).

$$a(m,n) = \left\{ \begin{array}{l} b_1 \mathit{if} c(m,n) > \mathit{Th} \\ b_2 \ \mathit{if} c(m,n) < \mathit{Th} \end{array} \right\} \tag{1}$$

The main objective behind thresholding is to differentiate background and objects, which possess different range of pixel level intensities. The results produced using bi-level thresholding may convert gray level image into the binary image. Global thresholding and local thresholding are two commonly used methods to perform thresholding. A single optimal threshold value is chosen in global thresholding to differentiate foreground and background details in global thresholding method. Any how the global thresholding method performs well only in the case of image having bi-modal histogram, otherwise if wrong threshold value is chosen it may lead to calamitous. Remedy, for this case is to use multilevel thresholding for segmentation. It includes multiple thresholding values for different regions and lead to formation of different classes as mentioned in the following equation (Bohat & Arya, 2019). Different classes $[b_1, b_2, b_3, b_i, \cdots\cdots..b_n]$ represents multilevel thresholding and following equation mentions n different classes for different thresholding values.

$$\begin{array}{l} b_1 \leftarrow C_1 \text{ if } 0 < c(x,y) < Th_1 \\ b_2 \leftarrow C_2 \text{ if } Th_1 < c(x,y) < Th_2 \\ b_3 \leftarrow C_3 \text{ if } Th_2 < c(x,y) < Th_3 \\ b_i \leftarrow C_i \text{ if } Th_i < c(x,y) < Th_{i+1} \\ b_n \leftarrow C_n \text{ if } Th_n < c(x,y) < Th_n \end{array} \tag{2}$$

2.2. Otsu

Otsu's method is derived from the name of inventor Nobuyuki. It is an nonparametric and unsupervised method to perform thresholding automatically based on clustering (Otsu et al., 1979). It works on the principle of interclass and intra-class variance to perform segmentation. Distinct classes must contain maximum interclass variance and minimum intra-class variance. Within class or intra-class variance σ_W and between class or interclass variance is calculated as per the below mentioned expressions.

$$\sigma_W = W_b \sigma_b^2 + W_f \sigma_f^2 \tag{3}$$

$$\sigma_B = \sigma_0 + \sigma_1 \tag{4}$$

$$= W_b(\mu_b - \mu_T)^2 + W_f(\mu_f - \mu_T)^2, = W_b W_f(\mu_b - \mu_f)^2 \tag{5}$$

where $\mu_T = W_b\mu_b + W_f\mu_f$, σ_W is with in class variance, σ_B is between class variance of two classes, σ_b represents background variance, σ_f represents foreground variance, W_b is weight of background, W_f is foreground weight, μ_b & μ_f are mean of background and foreground respectively and μ_T represents total mean. Otsu's thresholding method is commonly required for multilevel thresholding, which involve complex calculations. Multilevel thresholding Otsu's method for n different classes may be defined by the Eq. (6).

$$P_i = \frac{f_i}{\sum_{i=0}^{L-1} f_i} \tag{6}$$

P_i represents probability gray level occurring, f_i is frequency of given i^{th} level, and L denotes total gray levels. $\sigma_B^2(th)$ is provided in the Eq. (7).

$$\sigma_B^2(th) = \sum_{k=0}^{K} W_k(\mu_k - \mu_T)^2 \tag{7}$$

where $\mu_T = \sum_{i=0}^{L-1} \frac{iP_1}{W_{n-1}}$, For bi-level thresholding fitness function using Otsu method is given by $f_{OTSU}(th)$, whereas for multilevel thresholding is represented by $f_{OTSU}(TH)$ having n different classes in Eq. (8) and (9) respectively. Maximum value this function would correspond to optimum threshold value.

$$f_{OTSU}(th) = \varnothing_o = \max\left(\sigma_B^2(th)\right), \qquad 0 \le th \le L-1 \tag{8}$$

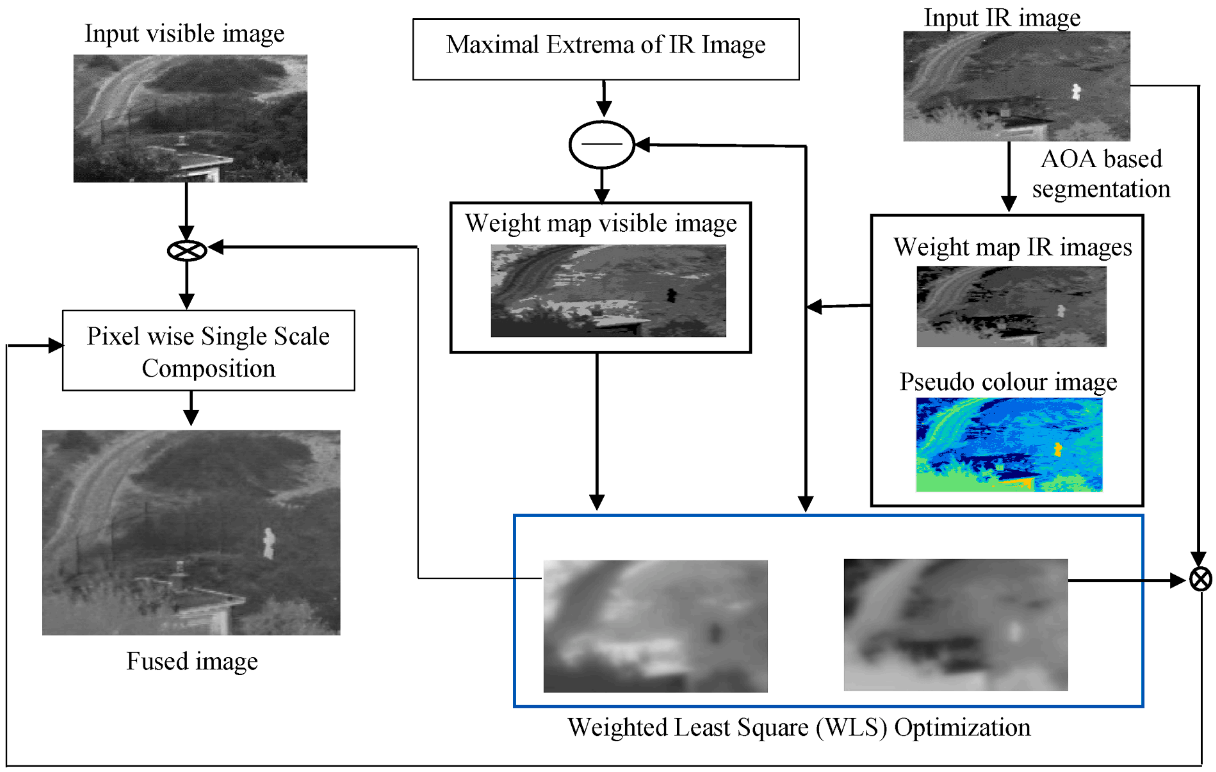

Fig. 1. Proposed IR and Visible Image fusion framework for Night Vision Context Enhancement.

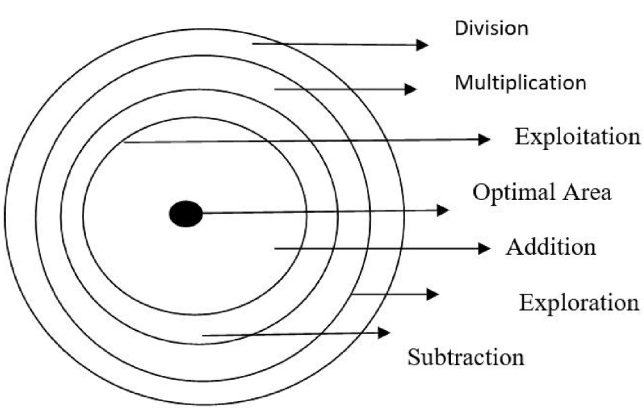

Fig. 2. Shows arithmetic operators according to superiority.

$$f_{OTSU}(TH) = \varnothing_o = \max\left(\sigma_B^2(TH)\right), \qquad 0 \le th \le L-1, \quad i = 1, 2, 3\cdots\cdots\cdots, n \tag{9}$$

where$TH = [th_1, th_2, th_3, \cdots..th_{n-1}]$,

Object tracking in IR photos can be achieved simply as it displays unimodal peaks for objects, but in complicated contexts of equivalent intensity values, it becomes a difficult process. In such cases global threshold value doesn't work, adaptive threshold value-based segmentation is required. Multilevel segmentation of the IR image is illustrated in Fig. 2 including five thresholding values and IR segmented images with pseudo colors.

3. IR and visible image fusion methodology

In this section, a novel segmentation based fusion methodology is proposed. The objective behind the proposed technique is to fuse the Infrared $I_{IR}(x,y)$ and Visible $I_{Vis}(x,y)$ images without introducing any artefacts and keeping meaningful information intact. The detailed block diagram of proposed manuscript with appropriated images is shown in Fig. 1. Each step involved in this algorithm is described in depth in forthcoming sub-sections. In which, obtain the segmented image $I_S(x,y)$ using metaheuristic-based segmentation. A metaheuristic algorithm AOA is proposed in this technique is used to obtain weight maps. The optimal infrared weight W_{IR}^1 is extracted using AOA and visible image W_{Vis}^1 weights are derived from the IR weight W_{IR}^1. Then the resultant weight maps are refined, using weighted least square (WLS) (Song et al., 2016) algorithm and normalization. The final fused resultant image is obtained by fusing the IR I_{IR} and Vis I_{Vis} source image using corresponding visible and IR weight maps using pixel-wise single scale composition.

Image segmentation is a significant task and pre-processing step in computer vision and image processing. It is the simplest, fastest, and most effective image segmentation techniques capable of discriminating against objects from the background through a set of pixel-level thresholds. In certain examples of image processing, it is required to segregate the foreground object from gray-level pixels of background (Kamel & Zhao, 1993). It possesses a variety of applications in different fields like in medical (Liu et al., 2018), IR images, artificial intelligence, surveillance, remote sensing for specific target recognition, medical imaging, etc. (Singh, Mittal, & Singh, 2021; 2022). The basic principle behind thresholding is to calculate the optimum threshold value to differentiate the target from the background (Mousavirad & Ebrahimpour-komleh, 2019; Shi, Yang, Hospedales, & Xiang, 2017; Singh, Hrisheekesha, & Cristobal, 2019).

3.1. Arithmetic optimization algorithm (AOA)

The paper's aim is to present a methodology for image segmentation using AOA (Abualigah, Diabat, Mirjalili, Abd Elaziz, & Gandomi, 2021). Due to its simple and easy application, AOA has been used to address many real-world optimization issues such as extending the lifetime of the radio frequency identification (RFID) network, photovoltaic systems, multi-level digital image segmentation threshold, and range-based wireless node localization. Although randomization and static swarm behavior have a great worldwide search capability for AOA, its local search capabilities are limited and results in local optima capturing. At the original phases, iteration level hybridization method guarantees exploration capability and exploitation capability at the subsequent phases and ensures an enhanced accuracy of the global optimum. Furthermore, details and description of this algorithm is discussed as follows:

AOA is new *meta*-heuristic method that uses common mathematical operations such as Division (D), Addition (A), Multiplication (M), Subtraction (S) as shown in Fig. 1, which is applied and modeled to execute optimization in a wide variety of search fields (Abualigah et al., 2021). Commonly, population-based algorithms (PBA) launch their improvement processes by randomly selecting a number of candidate strategies. This defined solution is enhanced incrementally by a set of optimization standards and analyzed sequentially by a particular objective function; and that's the basis of an optimization techniques. Although PBA are stochastically trying to find some efficient strategy to optimization problems, a single run solution is not guaranteed. However, the chance of an optimum global solution to the problem is improved by a large set of possible solutions and optimization simulations (Singh et al., 2021). Considering the variations among *meta*-heuristic methods in PBA approaches, the optimization process comprises of two cycles: exploitation vs exploration (Wang, 2022). The previous examples for extensive coverage are of search fields by means of search agents of method to bypass local solutions. Above is the increase in the performance of solutions achieved during exploration process.

Arithmetic is a key component of mathematics and is most important components of modern math's, together with analysis, geometry and algebra. Arithmetic operators (AO) are traditionally used for study of numbers. These basic math's functions are used for optimization for finding ideal element particularly with selected solutions. Optimization challenges have appeared in all mathematical fields, such as engineering, economics and computer science to organizational analysis and technology, and the advancement of optimization techniques has drawn the attention of mathematics from time to time. The key motivation of the new AOA is use of AO to solve problems. The behavior of AO and their effect mostly on existing algorithms, arrangement of AO and their superiority is shown in Fig. 2. AOA is then proposed on the basis of a statistical model.

3.1.1. *Initial stage*

The method of optimization starts with selected sets denoted by A as in Eq.10. The ideal set in every iteration is created randomly and is taken as optimum solution.

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & .. & .. & a_{1,j} & a_{1,1} & a_{1,n} \\ a_{2,1} & a_{2,2} & .. & .. & a_{2,j} & .. & a_{2,n} \\ a_{3,1} & a_{3,2} & .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. & .. & .. \\ a_{N-1,1} & .. & .. & .. & a_{N-1,j} & .. & a_{N-1,n} \\ a_{N,1} & .. & .. & .. & a_{N,j} & a_{N,n-1} & a_{N,n} \end{bmatrix} \tag{10}$$

Exploitation/Exploration should be carefully chosen at start of AOA. Coefficient of math optimizer accelerated (MOA) is defined in Eq.11.

$$MOA(C_{iter}) = Min + C_{iter}x\left(\frac{Max - Min}{M_{iter}}\right) \tag{11}$$

where, $MOA(C_{iter}) = i^{th}$ iteration function value, $M_{iter} =$ Max. no. of iteration, $Max\&Min =$ Accelerated Function of Max. and Min. Values, $C_{iter} =$ current iteration (within 1 and M_{iter}).

The exploratory nature of AOA is discussed, as per the AO, mathematical calculations whether using Division (D) or Multiplication (M) operator have obtained high distribution values or decisions that contribute to an exploration search method. However, as opposed to other operators, these D and M operators never easily reach the objective due to high distribution of S and A operators. AOA exploration operators exploit the search field arbitrarily through many regions and seek a better alternative dependent on two key search techniques M and D search techniques as shown in Eq.12.

$$a_{ij}(C_{iter}+1) = \begin{cases} besta_j \div (MOP \ \div \varepsilon) \times \left(\left(UB_j - LB_j\right) \times \mu + LB_j\right), r_2 < 0.5 \\ besta_j \times MOP \times \left(\left(UB_j - LB_j\right) \times \mu + LB_j\right), \ otherwise \end{cases} \tag{12}$$

where, $a_i(C_{iter}+1) = i^{th}$ solution of next iteration, $a_{i,j}(C_{iter}+1) = j^{th}$ position in current iteration, $\mu =$ control parameter ≤ 0.5, $LB_j\&UB_j =$ Lower & Upper bound limit, $\varepsilon =$ smallest integer no., $besta_j = j^{th}$ position of optimum solution till now,

$$MOP(C_{iter}) = 1 - \frac{C_{iter}^{\frac{1}{\alpha}}}{M_{iter}^{\frac{1}{\alpha}}} \tag{13}$$

where, Math Optimizer Probability $(MOP) =$ coefficient, $MOP(C_{iter}) = i^{th}$ iteration function value, $C_{iter} =$ current iteration, $M_{iter} =$ Max. iterations ≤ 5.

The exploitation nature of AOA is discussed, as per AO mathematical formulas whether using addition (A) or subtraction (S) as they provided high-density results. AOA exploitation operators exploit the search field deeply through many regions and seek a better alternative dependent on two key search techniques A and S search techniques as shown in Eq. (14) and algorithm of AOA is shown below in algorithm 1,

$$a_{i,j}(C_{iter}+1) = \begin{cases} besta_j - MOP \times \left(\left(UB_j - LB_j\right) \times \mu + LB_j\right), r_3 < 0.5 \\ besta_j + MOP \times \left(\left(UB_j - LB_j\right) \times \mu + LB_j\right), otherwise \end{cases} \tag{14}$$

3.2. *Segmentation based fusion*

In this section, a novel segmentation based fusion methodology is proposed. The objective behind the proposed technique is to fuse the infrared (I_{IR}) and visible (I_V) images without introducing any artefacts. The block diagram of the proposed approach is shown in Fig. 1. Each step involved in this algorithm is described in depth in forthcoming subsections. In the first step, the segmented image (I_S) is computed based on metaheuristic-based segmentation. For weight map computation of IR image, a modified metaheuristic algorithm AOA is proposed.

Further the weight map function for Vis (W_V^1) is derived from the IR weight (W_{IR}^1) using Eq. (16). The resultant weight maps are refined based on WLS optimization (Song et al., 2016). The refined weighted maps for IR image (W_{IR}^2) and Vis image (W_V^2) are calculated using Eq. (17). The resultant fused image is obtained by fusing the IR (I_{IR}) and Vis (I_V) source image using corresponding visible and IR weight maps. Weighted average based pixel-wise fusion is done to produce final fused image.

3.3. *Weight map computation*

In the proposed method, values calculated using AOA based thresholding are used to obtain segmented image which acts as initial weight map. Let the IR image is operated by AOA to produce segmented image, that is to say $I_s(x,y) = AOA(IR(x,y))$. The weight map for IR and Vis image is given by the following equations.

$$W_{IR}^1 = I_s(x,y) \tag{15}$$

$$W_V^1 = \max\{I_{IR}\} - W_{IR}^1(x,y) \tag{16}$$

where $W_{IR}^1(x,y)$ is weight map function of IR image, $W_V^1(x,y)$ is a weight map function computed for the Vis image, and $I_s(x,y)$ is a segmented image.

The weight maps computes in Eq. (15) and Eq. (16) are hard and

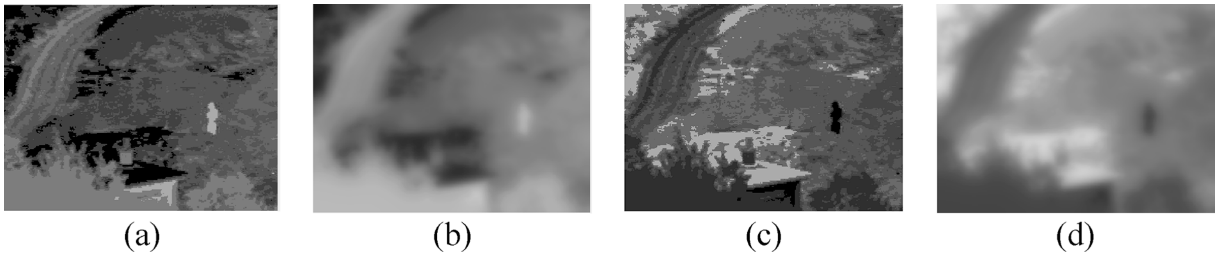
Fig. 3. (a) Initial weight map of IR (b) Refined IR weight map (c) Initial weight map of Vis (d) Refined Vis weight map.

noisy and are not suitable for single scale-weighted average fusion of input images. If these weight map functions are multiplied by the input picture directly, artefacts will appear in the fused image. To eliminate the artefacts in fused image due to fusion process, the weight maps are to be refined using WLS (Song et al., 2016). Nonetheless, attributes of IR and Vis photos are very dissimilar, Vis photos mostly include fine-scale structural data, whereas IR typically includes coarse-scale structures or other dissenting IR specifics and noise. Direct fusion may consider more insignificant information or noise from the IR image and less fine-scale information from the Vis photos. In order to avoid such circumstances, prior to fusion weights are refined using the WLS optimization scheme. In the proposed fusion algorithm, WLS filter is utilized to refine the weighted maps of IR and Vis images that is illustrated in Fig. 3. Many researchers have used WLS filter to develop edge-preserving multi-scale decomposition of image (Song et al., 2016).

The WLS is an edge preserving filter and smoothing filter, it helps to maintain the edges by finding the best possible balance between blurring and sharpening. It has the greatest potential to progressively sharpen the picture, while preserving the spatial information consistent. WLS-based weight map refinement is used in our fusion approach to improve these noisy weight maps. The edge-preserving operator based on WLS might be seen as a compromise between two opposing aims. In particular, given an input weight $W^1_{(IR)p}$, an output weight map $W^2_{(IR)p}$ having maximum similarity with $W^1_{(IR)p}$ is required, also it should be as smooth as possible everywhere except around the edges and minimizing the following equation would give refined weights.

$$W^2_{IR} = \sum_p \left(\left(W^2_{(IR)p} - W^1_{(IR)p} \right)^2 + \lambda_1 \left(g_{x,p,\alpha_1,\epsilon_1} \left(W^1_{IR} \right) \left(\frac{\partial W^2_{IR}}{\partial x} \right)^2_p + g_{y,p,\alpha_1,\epsilon_1} \left(W^1_{IR} \right) \left(\frac{\partial W^2_{IR}}{\partial x} \right)^2_p \right) \right) \tag{17}$$

$$W^2_{V} = \sum_p \left(\left(W^2_{(V)p} - W^1_{(V)p} \right)^2 + \lambda_2 \left(g_{x,p,\alpha_2,\epsilon_1} \left(W^1_{V} \right) \left(\frac{\partial W^2_{V}}{\partial x} \right)^2_p + g_{y,p,\alpha_2,\epsilon_1} \left(W^1_{V} \right) \left(\frac{\partial W^2_{V}}{\partial x} \right)^2_p \right) \right) \tag{18}$$

The spatial coordinate of a pixel is denoted by subscript p. And λ_1 and λ_2 are regularization constraint used to balance sharp and smooth parameters to preserve edge details, a higher value will carry out more smoothening operation controlled by gradients (g) computed across input wight maps. In the proposed fusion approach, λ_1 and λ_2 are empirically set to 1.5 and 0.7, respectively. Where, $\alpha_{1,}$and$\alpha_{2,}$ (are set empirically to 0.8 and 0.4, respectively) that determines the sensitivity to the gradient values of input weight maps to be refined. The smaller value of α_2 helps to preserve finer details from input Vis image. ϵ_1 is a small constant (typically 0.0001) that prevents division by zero in areas where input weight is constant. $W_{x,p}$ and $W_{y,p}$ are the horizontal and vertical smoothness weights and $\left(W^2_{(IR)p} - W^1_{(IR)p} \right)^2$ ensures maximal similarity between input and output.

3.4. Fusion strategy

In this section, final image fusion is computed by considering source input images and their weight maps. After obtaining the weight maps for IR image (W^2_{IR}) and visible image (W^2_V), the pixel-wise single-scale weighted average composition is performed that is given in Eq. (19). The pixel wise weighted average fusion provides robustness to the fusion algorithm, as it minimizes the information loss.

$$I_F(x,y) = W^2_{IR}(x,y) \times I_{IR}(x,y) + W^2_V(x,y) \times I_{Vis}(x,y) \tag{19}$$

The final results of the proposed algorithm are analyzed with seven existing techniques and they are executed based on the codes available in the public domain. The techniques used to carry out comparative analysis are fusion method using cross bilateral filter (CBF) (Shreyamsha Kumar, 2015) and discrete cosine harmonic wavelet transform (DCHWT) proposed by Shreyamsha Kumar (Shreyamsha Kumar, 2013), JSR model with saliency detection fusion method (JSRSD) offered by Liu et al. (Liu, Qi, & Ding, 2017), the gradient transfer fusion method (GTF) given by Zhang et al. (Zhang, Fu, Li, & Zou, 2013), convolutional sparse representation(LP-GAN) and deep convolutional neural network based method (CNN) proposed by Liu et al. (Liu, Chen, Ward, & Wang, 2016) and latent low-rank representation fusion method (LATLRR) suggested by (Li & Wu, 2018).

The names of these different images used in the algorithm are addressed in Fig. 4. Different IR and VI image sets are used with the suggested methodology. The presented technique is applied on different picture datasets and is qualitatively and quantitatively analysed. Twenty-one picture pairings datasets are chosen from TNO data set ('TNO Image Fusion Dataset'). The names of chosen datasets are (UN Camp, Traffic, Steamboat, Building, Tree, Home post, Airplane, Bench, Bunker, Heather, Helicopter, Kaptein, Light hut, Lake Man in Doorway, Jeep, Road Car, Sand path, Soldier, and Trench image couples) are subjected to qualitative analysis owing to space restrictions. However, all 21 picture pairings are analysed quantitatively.

4. Experimental results

The final fusion results of the proposed approach are compared with seven existing state-of-the-art fusion techniques. CCD cameras with low-light sensitivity or a standard sensitivity can capture the input visible images. Some image details in both visible and infrared imagery may have to be boosted to maximise their visibility. In addition, the infrared sensor frequently uses the mid-wave and long-wave spectral bands to better identify details from objects in dark and obstructed areas. We can finally improve night vision by incorporating IR spectral information into visible images. Four fusion performance measures were used in the quantitative analysis, and the MATLAB scripts for each are freely available for the research purposes. The techniques used to carry out comparative analysis are fusion method using cross bilateral filter (CBF) and discrete cosine harmonic wavelet transform (DCH) proposed by Shreyamsha Kumar, JSR model with saliency detection fusion method (JSR) proposed by Liu et al. (Liu et al., 2017; Zhang et al., 2013), and the gradient transfer fusion method (GTF) proposed by Zhang et al. (Q.

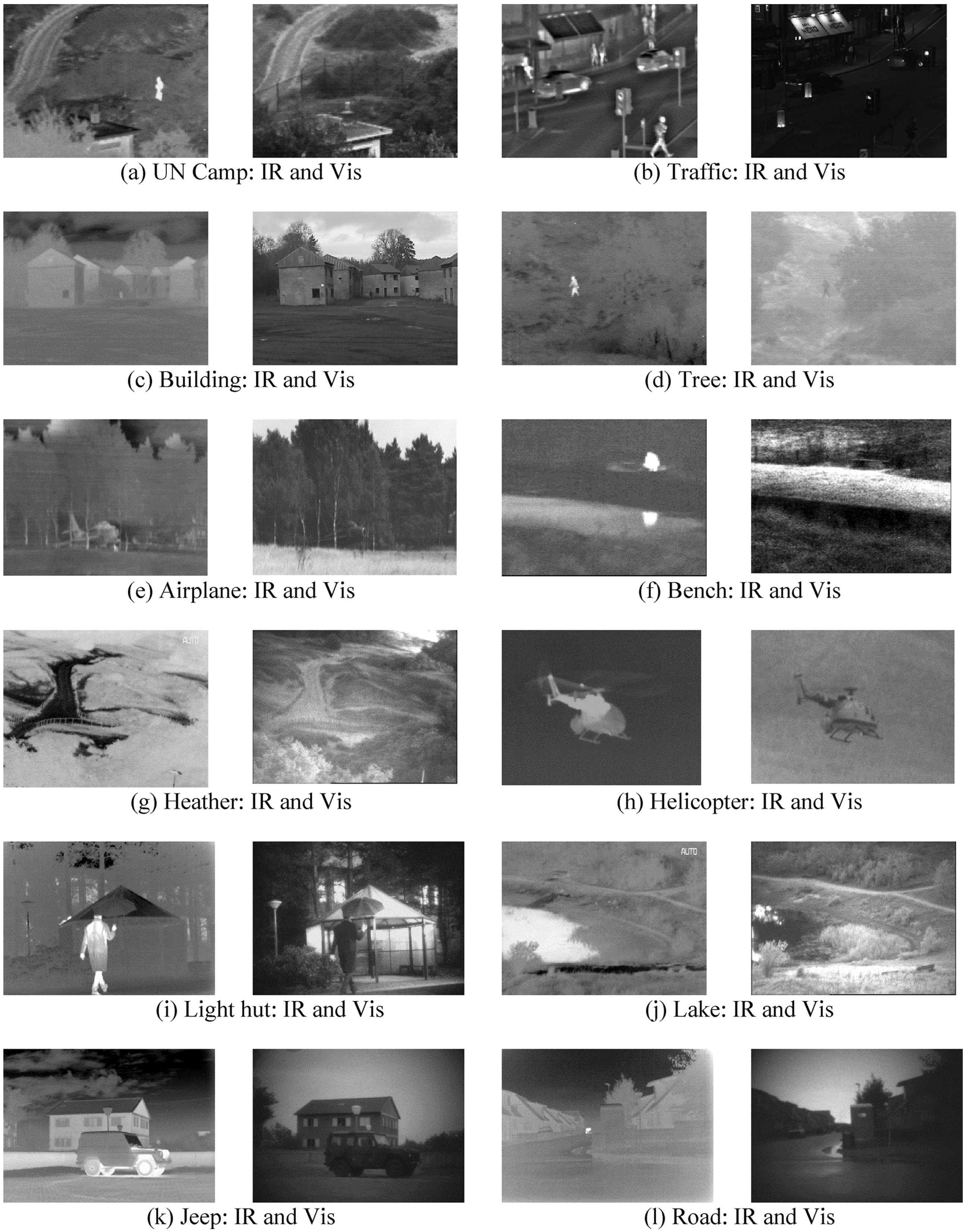

Fig. 4. Data set of IR and Visible images used for context enhancement.

Zhang et al., 2013). Furthermore, deep convolutional neural network based method (CNN) (Liu, Chen, Peng, & Wang, 2017), laplacian pyramid and generative adversarial network (LP-GAN) (Wang, Ke, Wu, Liu, & Zeng, 2021) and latent low-rank representation fusion method (LLR) (Li & Wu, 2018) based fusion methods are also considered for qualitative and quantitative analysis.

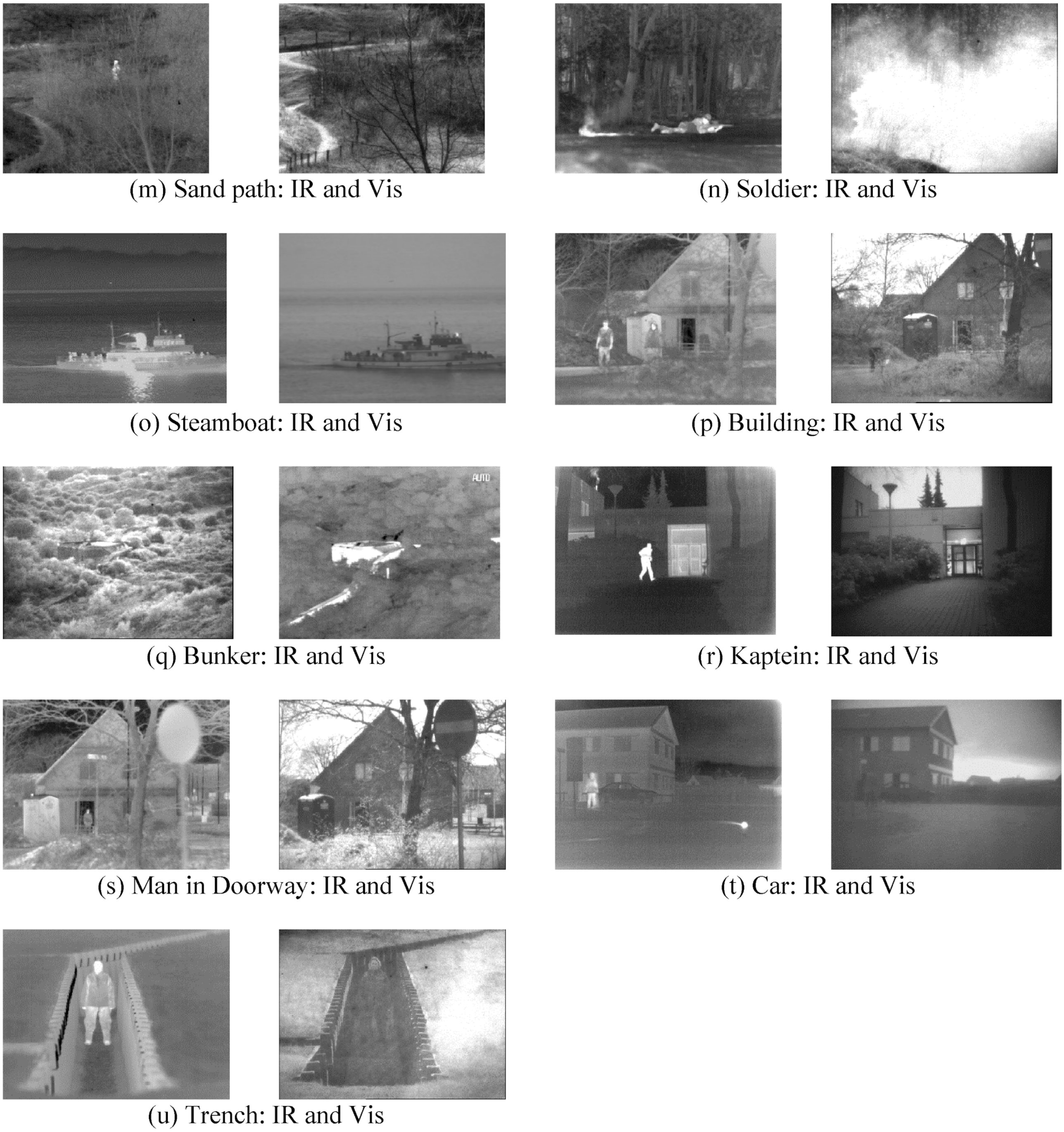

Fig. 4. (*continued*).

4.1. Qualitative evaluation of fusion results

Because of the limited amount of space available, six image datasets used for qualitative evaluation were selected from the TNO data collection (Alexander, 2014). The visual quality comparison of UN Camp, Sand path, tree, Traffic, Steamboat and Kaptein datasets are presented in Figs. 5–9. In all cases, figures (a), (b), and (c) show IR images, VI images, and a suggested output, respectively. And (d)-(j) are the fusion results of DCH, JSR, GTF, LP-GAN, CNN, LLR, and CBF methods respectively. Fences, the room's roof, and a pedestrian can all be seen enhanced in the fused image shown in Fig. 5(c), which gives us a comprehensive detail of the battleground region. Aside from that, the key IR target is properly transferred to the final fused image to pinpoint the precise location of the individual being tracked.

In Fig. 6 and Fig. 7, Fig. 6(c-j) and Fig. 7(c-j) depict the fusion results of compared methods and results proposed by the proposed method for "*Sand path*" and "*Tree*" image data sets, respectively. Fig. 6(a) and (b) show an infrared imagery of a person standing behind trees and near to a fence, as well as a visible image of a sandy path, trees, and fences. Due to the fact that the person is more prominent in the fused images and the landscape structure is more accurately portrayed when looking at the fusion results from other state-of-the-art approaches, the suggested approach is clearly superior to the others. We can see that the IR target and scenery in Fig. 7(c) is clearer than those results depicted in Fig. 7(d-j). The huge brightness difference between the visible and infrared source images causes the DCH, GTF and CBF techniques to generate artificial details, as can be seen in the Fig. 7(d) (f) and (j). Overall, the suggested AOA-based context enhancement fusion approach is capable

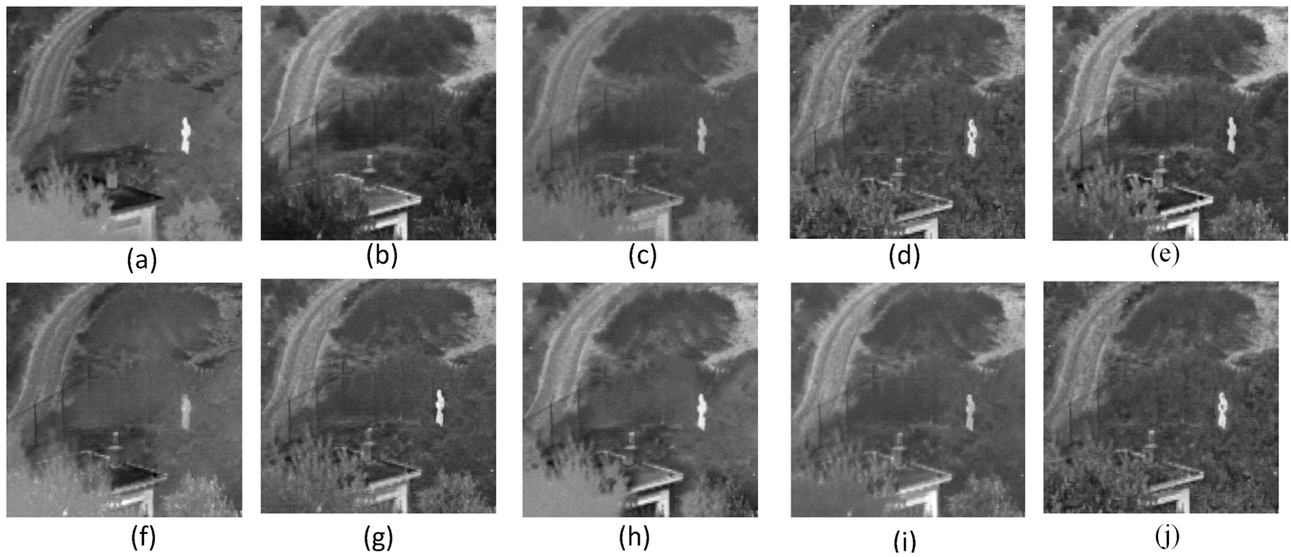

Fig. 5. Comparison of results "UN C*amp*": (a) IR image (b) VIS image (c) Proposed (d) DCH (e) JSR (f) GTF (g) LP-GAN (h) CNN (i) LLR and (j) CBF.

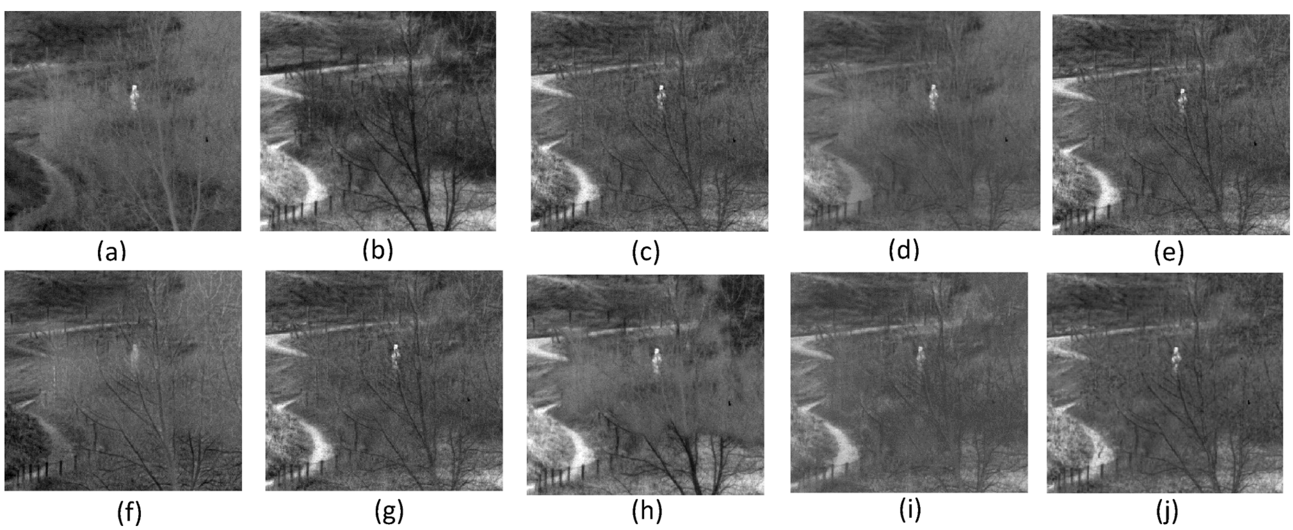

Fig. 6. Comparison of results "*Sand path*": (a) IR image, (b) VIS image (c) Proposed (d) DCH (e) JSR (f) GTF (g) LP-GAN (h) CNN (i) LLR and (j) CBF.

of producing improved night vision fusion outcomes.

In another example, traffic lights and automobiles are clearly visible in Fig. 8(a) but the advertising board is obscured (Fig. 8(b)). Cars and traffic signals, on the other hand, are virtually invisible in the visible imagery, but the advertising board is very apparent. We can notice from the fusion results of JSR (Fig. 8(e)), GTF (Fig. 8(f)), LP-GAN (Fig. 8(g)), CNN (Fig. 8(h)), and LLR (Fig. 8(i)) that they yield less details from VIS image of "*Traffic*". The fusion results of DCH (Fig. 8(d)), and CBF (Fig. 8 (j)) are depicting some visual artifacts. As compared to other fusion methods, the proposed approach is able to transfer details from VIS and IR images into the output fused image (Fig. 8(c)). The fused image looks natural and does a good job of preserving the brightness of the people. Moreover, the text "NERO" displayed on the light board are accurately preserved in the fused image.

Fig. 9 shows pairs of input imagery of "Steamboat" dataset and fusion results for visual examination. In the "Steamboat" image series (see Fig. 9(a, b)), distinct parts of the boat may be identified by their thermal profiles, while the complementing features are visible in the VIS picture. Image fusion makes it easier to follow things and spot various boat operations. The most difficult part is transforming IR spectral information into a VIS imagery that can be used to improve night vision. The results produced by DCH and CBF shown in Fig. 9(d) and (j) produce artifacts due to the intensity map difference between VIS and IR images. As it can be seen from the fusion results shown in Fig. 9(c), the proposed method preserves background and IR target details accurately with lesser artifacts. IR spectrum information as well as apparent picture features in low-light lit zones are readily evident with these context enhancement approach. The fusion results for "Kaptein" image datasets are shown in Fig. 10. We can notice from fusion results shown in Fig. 10(c-j) that IR target (such as human) appear to be less prominent in other fusion methods. Furthermore, we can see that the proposed AOA-based fusion approach preserves backdrop landscape features while reducing distortion.

To summarise, the fusion results created by suggested algorithm and alternative approaches are described as follows: artifacts can be found in CBF and DCH, and their saliency isn't perfect for image fusion. The fusion results obtained from JSR, GTF, LP-GAN, and CNN include many ringing artefacts, as well as the detailed information that is not readily apparent. LLR's output doesn't provide much in the way of specifics on important characteristics of IR target. The suggested method, on the other hand, creates fusion results that preserve greater detail while still preserving saliency features of IR targets. The suggested fusion

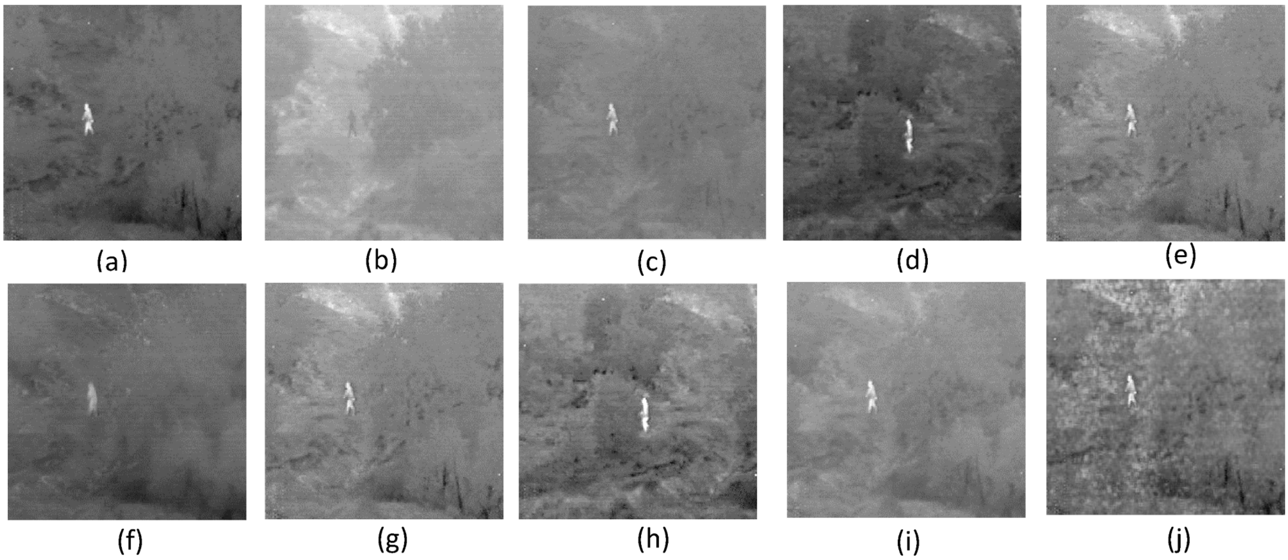

Fig. 7. Comparison of results "*tree*": (a) IR image, (b) VIS image (c) Proposed (d) DCH (e) JSR (f) GTF (g) LP-GAN (h) CNN (i) LLR and (j) CBF.

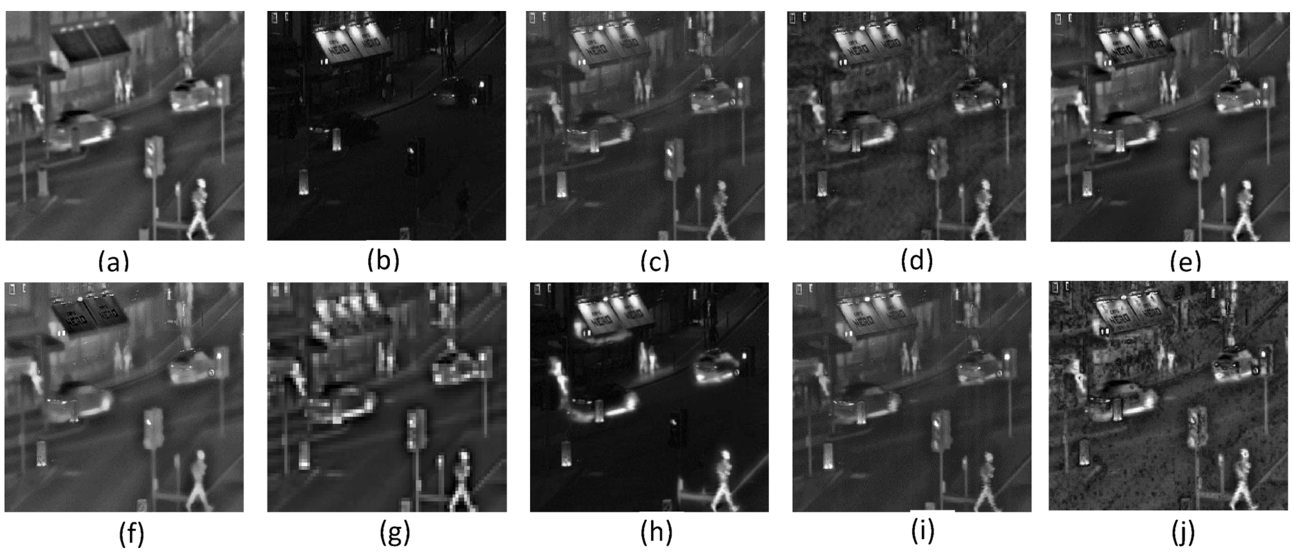

Fig. 8. Comparison of results "*Traffic*": (a) IR image, (b) VIS image (c) Proposed (d) DCH (e) JSR (f) GTF (g) LP-GAN (h) CNN (i) LLR and (j) CBF.

technique produces fused images that are more human-perceivable and has increased efficiency in the qualitative assessment.

4.2. Objective fusion metrics

Performance evaluation of a fusion method is a challenging task in the absence of reference fused image. In literature, many fusion metrics has been proposed. For better assessment, we choose four objective fusion metrics. Subjective analysis alone is not sufficient to authenticate the results; more quantitative review is necessary to evaluate the performance of proposed algorithm (Singh et al., 2020). In order to perform further analysis of fusion result, four widely used quantitative image fusion metrics edge based similarity index ($Q^{AB/F}$), sum of correlation difference (SCD), structural similarity index (SSIM), and artifact measure ($N^{AB/F}$) have been chosen. $Q^{AB/F}$ provides the most important feature of fused image i.e. edge preserving details and SCD calculates the quality by considering the source images and their impact on the fused image. SSIM quantifies image quality degradation caused by various processing and the $N^{AB/F}$ provides the quantitative information related to noise or artefact added to fused image. Above mentioned quality parameters provide an in-depth analytic report of fusion performance by quantifying: important feature and their impact, total fusion performance, fusion loss and fusion artifacts (artificial information created). Based on these fusion quality metrics, quantitative analysis of pairs of twenty-one.

4.2.1. Edge based similarity index ($Q^{AB/F}$)

It gives the edge preservation detail in the final image from the source input images and given by Eq. (20) higher value of this metric is desired for good results and it ranges between (0–1) (Singh et al., 2021b).

$$Q^{AB/F} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} Q^{A/F}(i,j)g_A(i,j) + Q^{B/F}(i,j)g_B(i,j)}{\sum_{i=1}^{M}\sum_{j=1}^{N} g_A(i,j) + g_B(i,j)} \tag{20}$$

where A, B are source and F is fused imagery and $Q^{A/F}(i,j)$ and $Q^{B/F}(i,j)$ are given as

$$Q^{A/F}(i,j) = Q_g^{A/F}(i,j)Q_\alpha^{A/F}(i,j)$$

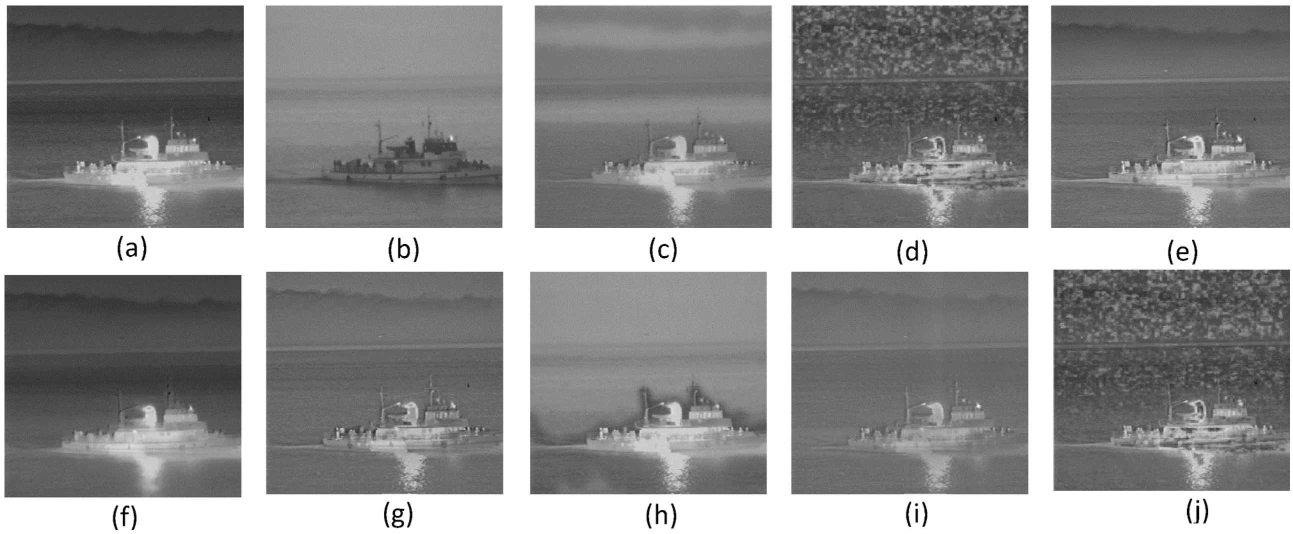

Fig. 9. Comparison of results "*Steamboat*": (a) IR image (b) VIS image (c) Proposed (d) DCH (e) JSR (f) GTF (g) LP-GAN (h) CNN (i) LLR and (j) CBF.

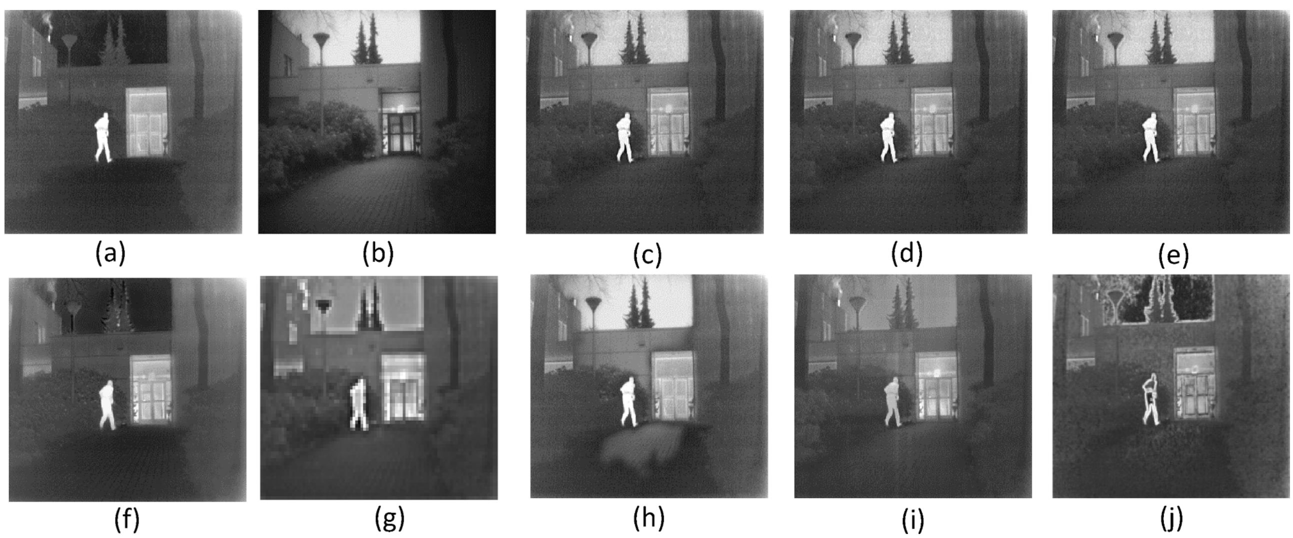

Fig. 10. Comparison of results "*Kaptein*": (a) IR image, (b) VIS image (c) Proposed (d) DCH (e) JSR (f) GTF (g) LP-GAN (h) CNN (i) LLR and (j) CBF.

$$Q^{B/F}(i,j) = Q_g^{B/F}(i,j)Q_\alpha^{B/F}(i,j) \tag{21}$$

where $Q_g^{A/F}$ and $Q_g^{B/F}$ denotes strength of edge, $Q_\alpha^{A/F}$ and $Q_\alpha^{B/F}$ are the parameters representing preservation of orientation values, and $g_A(i,j)$ and $g_B(i,j)$ are weight values for images *A* and *B* respectively for pixel position *(i,j)*.

4.2.2. Sum of Correlation Difference (SCD)

It demonstrates the extent of meaningful data being transmitted from input sources to the final result (Li & Wu, 2018; Meher, Agrawal, Panda, & Abraham, 2019; Qiu, Wang, Zhang, & Xia, 2017). Higher value of *SCD* denotes better fusion results and is given by following equation.

$$SCD = r(D_1, A) + r(D_2, B) \tag{22}$$

where $D_1 = F\text{-}B$ and $D_2 = F\text{-}A$, *F* is fused image, *A* and *B* are input images and *r(.)* function calculates the correlation between,*A* and D_1, and *B* and D_2.

4.2.3. Structural Similarity Index (SSIM)

It is the most consistent technique to find similarity between two images as compared to other existing techniques like peak signal to noise ratio (PSNR) and mean square error (MSE). SSIM measures the extent of degradation of fused image as compared to input image (Bhandari & Kumar, 2019; Krishnamoorthy & Soman, 2010; Liang, Jia, Xing, Ma, & Peng, 2019; Sreeja & Hariharan, 2018). It is mostly preferred in the presence of ground truth but a modified version of *SSIM* is used as given in Eq. (25). The higher value of *SSIM* denotes better fusion results.

$$SSIM(A,F) = \frac{[2\mu_A\mu_F + C_1](2\sigma_{AF} + C_2)}{[\mu_A^2 + \mu_F^2 + C_1](\sigma_A^2 + \sigma_F^2 + C_2)} \tag{23}$$

$$SSIM(B,F) = \frac{[2\mu_B\mu_F + C_1](2\sigma_{BF} + C_2)}{[\mu_B^2 + \mu_F^2 + C_1](\sigma_B^2 + \sigma_F^2 + C_2)} \tag{24}$$

where μ_A, μ_B and μ_F are mean intensities, σ_A, σ_B and σ_{BF} are the standard deviation of two input images *A* and *B*, and fused image *C*, respectively. σ_{AF} and σ_{BF} are the square root of covariance of two input and fused images respectively and C_1 and C_2 are constants. Modified *SSIM* for two source images in the absence of ground truth is calculated by taking average of two values as given by.

Table 1
Quantitative analysis based on $N^{AB/F}$. Input image datasets courtesy of Alexander Toet.

Image name	Method name							
	CBF	DCH	JSR	GTF	LP-GAN	CNN	LLR	Proposed
UN Camp	0.23167	0.05118	0.34153	0.07027	0.01445	0.02976	0.0172	0.00031
Traffic	0.487	0.2184	0.19889	0.11237	0.01201	0.05862	0.02922	0.00119
Steamboat	0.54477	0.28619	0.38627	0.07335	0.1224	0.01563	0.03738	0.0056
Building	0.45288	0.19093	0.42353	0.143	0.02235	0.03609	0.02083	0.125
Tree	0.43257	0.07415	0.49804	0.08501	0.01091	0.01134	0.02234	0.00486
Home post	0.23932	0.10268	0.36509	0.06399	0.02396	0.02096	0.00873	0.0243
Airplane	0.41779	0.1506	0.5222	0.05249	0.01614	0.02886	0.03166	0.00267
Bench	0.15233	0.05781	0.21536	0.12329	0.03504	0.03424	0.01116	0.0019
Bunker	0.11741	0.05342	0.30761	0.09009	0.02471	0.0077	0.00485	0.00367
Heather	0.2009	0.09272	0.34271	0.10404	0.02113	0.02796	0.00588	0.00541
Helicopter	0.47632	0.1034	0.32941	0.07322	0.01122	0.01685	0.03266	0.0014
Kaptein	0.25544	0.0726	0.32502	0.03647	0.01645	0.01355	0.01399	0.00198
Light hut	0.36066	0.09951	0.2822	0.06577	0.01988	0.05044	0.0156	0.0271
Lake	0.18971	0.08579	0.40261	0.09124	0.02137	0.01751	0.00796	0.0257
Man in Doorway	0.21509	0.09529	0.35013	0.0752	0.02307	0.02381	0.00971	0.00086
Jeep	0.52783	0.18409	0.26888	0.08291	0.02036	0.03734	0.02138	0.0057
Road	0.52887	0.19714	0.3372	0.03276	0.01661	0.01785	0.01364	0.0196
Car	0.26649	0.06879	0.55732	0.03287	0.01599	0.01155	0.00715	0.00702
Sandpath	0.12582	0.01886	0.27302	0.0416	0.01479	0.00997	0.00682	0.00267
Soldier	0.25892	0.24507	0.16541	0.09293	0.02674	0.02271	0.00913	0.0148
Trench	0.18091	0.13342	0.38546	0.12682	0.02846	0.01798	0.00783	0.00045
Average	0.31727	0.12295	0.34657	0.07951	0.02058	0.02432	0.01596	0.013438

Table 2
Quantitative analysis based on $\boldsymbol{SSIM(a)}$.

Image name	Method name							
	CBF	DCH	JSR	GTF	LP-GAN	CNN	LLR	Proposed
UN Camp	0.62376	0.74834	0.52715	0.69181	0.75054	0.70536	0.76182	0.8267
Traffic	0.49861	0.64468	0.62353	0.61109	0.67574	0.64394	0.67029	0.6758
Steamboat	0.5828	0.82592	0.63503	0.82392	0.87605	0.82589	0.87728	0.9156
Building	0.58315	0.77224	0.59028	0.76132	0.81299	0.764	0.81653	0.8512
Tree	0.59632	0.80619	0.46767	0.73386	0.81483	0.76013	0.83372	0.8468
Home post	0.61724	0.73449	0.50549	0.73213	0.76272	0.71112	0.76892	0.7567
Airplane	0.67194	0.83192	0.5813	0.83434	0.8658	0.82501	0.86276	0.9127
Bench	0.5236	0.57614	0.45458	0.50273	0.56229	0.55761	0.60854	0.6435
Bunker	0.61793	0.65663	0.4725	0.62984	0.66793	0.62314	0.68909	0.6703
Heather	0.63041	0.70979	0.44712	0.67283	0.72958	0.6714	0.74291	0.7408
Helicopter	0.66486	0.83815	0.70365	0.80499	0.85922	0.80705	0.85903	0.8625
Kaptein	0.64975	0.75453	0.58333	0.70077	0.76211	0.71975	0.77222	0.78344
Light hut	0.53699	0.70607	0.56569	0.65676	0.72027	0.66818	0.73233	0.7315
Lake	0.69888	0.75918	0.52336	0.74249	0.78159	0.7315	0.79005	0.8125
Man in Doorway	0.59021	0.6883	0.48033	0.67176	0.71394	0.66418	0.72292	0.7731
Jeep	0.45747	0.68872	0.57664	0.66056	0.72278	0.68157	0.72532	0.7568
Road	0.50982	0.72735	0.60078	0.69419	0.77148	0.71527	0.77284	0.8403
Car	0.68824	0.81124	0.43761	0.77948	0.81941	0.7754	0.8342	0.8215
Sandpath	0.63683	0.7127	0.49591	0.63685	0.68763	0.64363	0.71667	0.8268
Soldier	0.53005	0.62304	0.57412	0.62966	0.70404	0.65703	0.71191	0.7373
Trench	0.68207	0.74211	0.52062	0.73201	0.78033	0.7277	0.79277	0.8627
Average	0.59957	0.73132	0.54127	0.70016	0.75435	0.70852	0.76486	0.792788

$$SSIM(a) = \frac{SSIM(F,A) + SSIM(F,B)}{2} \tag{25}$$

4.2.4. *Artifact Measure* ($N^{AB/F}$)

Fusion artifacts reflect visual details that the fusion cycle adds into the fused picture. Fusion objects are essentially false information which directly detracts from the usefulness of the fused picture and can have serious implications for fusion applications (Bavirisetti & Dhuli, 2016; Petrović & Xydeas, 2004; Petrović & Xydeas, 2003). It gives the volume of noise or artifacts added in the fused image during the image fusion process and its minimum value is preferred. It is preferred for comparison purposes because it gives in-depth evaluation of an image fusion method.

The values of each measured quality index for the 21 image data sets are given in the tables (see Tables 1–4). For each calculated quality metric the best value is shown in bold. From Table 1, we can notice that proposed method is performing well for 16 out of 21 image datasets in term of $N^{AB/F}$. As shown in Table 2, proposed method exhibits highest value of *SSIM* for 16 out of 21 image datasets. For *SCD* fusion metric, proposed method yields larger values for 17 image data sets. As can be noticed from Table 4, the proposed method obtains the largest value of $Q^{AB/F}$ for 18 out of 21 image datasets. Table 5 displays the average values of each quality metric for the proposed technique and current state-of-the-art methods for the 21 data sets and Table 6 depicts comparison of execution time. The proposed approach shows best average values for all datasets in minimum time. This indicates that the fused image generated by the proposed technique integrates more significant and accurate information from the input source data. Through this analysis, one can deduce that the current image fusion process yields superior efficiency over other methods.

Table 3
Quantitative analysis based on ***SCD***.

Image name	Method name							
	CBF	DCH	JSR	GTF	LP-GAN	CNN	LLR	Proposed
UN Camp	1.30442	1.44541	1.4111	0.96895	1.51833	1.33851	1.73308	1.7539
Traffic	1.27191	1.43676	1.70293	1.10568	1.55366	1.31964	1.73867	1.7465
Steamboat	1.5629	1.85025	1.74666	1.13871	1.91793	1.69729	1.91612	1.9238
Building	1.4797	1.60591	1.58464	0.97451	1.71889	0.91136	1.74929	1.7592
Tree	1.27176	1.48353	0.74887	0.70368	1.26624	0.9776	1.4225	1.4992
Home post	1.39012	1.71128	1.77756	1.05256	1.78225	1.54743	1.79913	1.7852
Airplane	1.25217	1.36439	1.41748	0.55475	1.38379	1.52066	1.44336	1.5589
Bench	1.62827	1.81492	1.79545	1.08473	1.69096	1.71422	1.78461	1.5858
Bunker	1.36803	1.56726	1.62841	1.15649	1.51395	1.51574	1.56165	1.6325
Heather	1.39251	1.6286	1.46388	1.19997	1.65355	0.96053	1.60757	1.6867
Helicopter	1.45458	1.71228	1.7184	1.38069	1.70129	1.58412	1.82698	1.8367
Kaptein	1.24703	1.5279	1.67259	0.79876	1.56028	1.63543	1.6587	1.6468
Light hut	1.45965	1.65352	1.54558	0.92405	1.62665	1.52093	1.67302	1.6728
Lake	1.36132	1.60896	1.70845	1.17275	1.68071	1.66239	1.72449	1.7832
Man in Doorway	1.45324	1.68888	1.61929	1.11126	1.71831	1.44563	1.69547	1.7367
Jeep	1.606	1.77072	1.64616	1.05059	1.829	1.71916	1.84254	1.8591
Road	1.73554	1.89848	1.81639	1.01399	1.91695	1.85046	1.92459	1.9958
Car	1.56524	1.7909	1.79824	0.97778	1.88302	1.7979	1.89505	1.9058
Sandpath	1.6343	1.69318	1.53392	1.0281	1.65549	1.43181	1.71426	1.7488
Soldier	0.53538	0.94824	1.34397	0.73046	1.47726	1.43624	1.51764	1.5358
Trench	1.20809	1.60723	1.73614	0.97403	1.6484	1.46113	1.61818	1.7429
Average	1.38963	1.60993	1.59124	1.00488	1.65223	1.4806	1.70699	1.73314

Table 4
Quantitative analysis based on..$Q^{AB/F}$

Image name	Method name							
	CBF	DCH	JSR	GTF	LP-GAN	CNN	LLR	Proposed
UN Camp	0.392	0.39101	0.30318	0.40514	0.48138	0.38339	0.42174	0.5596
Traffic	0.45063	0.47003	0.5899	0.37349	0.5472	0.33248	0.50307	0.9158
Steamboat	0.39473	0.43656	0.30916	0.28487	0.5112	0.28548	0.45012	0.7125
Building	0.54724	0.60936	0.3571	0.5382	0.64838	0.42222	0.3921	0.7185
Tree	0.32964	0.41983	0.28216	0.31927	0.51201	0.23977	0.43114	0.5658
Home post	0.47643	0.48	0.28404	0.38912	0.53647	0.27618	0.39209	0.5854
Airplane	0.43608	0.47029	0.33704	0.37148	0.52131	0.40783	0.45296	0.5861
Bench	0.63068	0.63058	0.38013	0.57943	0.57995	0.42865	0.47565	0.5767
Bunker	0.59259	0.57579	0.30221	0.48723	0.57303	0.24415	0.39526	0.6694
Heather	0.40948	0.42805	0.19807	0.46613	0.47647	0.18651	0.3324	0.4852
Helicopter	0.35362	0.44056	0.34079	0.51853	0.56364	0.30004	0.48154	0.5498
Kaptein	0.38801	0.40492	0.3394	0.31391	0.51513	0.25551	0.38622	0.5697
Light hut	0.43149	0.43398	0.38772	0.39907	0.52989	0.3491	0.40088	0.6528
Lake	0.53923	0.53952	0.2751	0.48131	0.55189	0.21755	0.38518	0.5591
Man in Doorway	0.50734	0.50494	0.29007	0.47549	0.5474	0.32607	0.39336	0.5598
Jeep	0.30252	0.33871	0.36911	0.31636	0.45615	0.2652	0.37892	0.5814
Road	0.3284	0.39212	0.39525	0.22889	0.52269	0.23365	0.42163	0.6146
Car	0.36369	0.40214	0.21891	0.29312	0.41424	0.24205	0.41306	0.41173
Sandpath	0.35552	0.35457	0.26653	0.3999	0.47539	0.26409	0.35011	0.5762
Soldier	0.40809	0.45498	0.26842	0.41948	0.57381	0.15541	0.38706	0.5847
Trench	0.59442	0.60639	0.28463	0.55742	0.61756	0.23039	0.42367	0.6658
Average	0.43961	0.46592	0.32281	0.41037	0.53985	0.28789	0.41277	0.60521

Table 5
The average comparison for all four fusion quality metrics.

Metric name	Method name							
	CBF	DCH	JSR	GTF	LP-GAN	CNN	LLR	Proposed
$N^{AB/F}$	0.31727	0.12295	0.34657	0.07951	0.02058	0.02432	0.01596	0.013438
SSIM	0.59957	0.73132	0.54127	0.70016	0.75435	0.70852	0.76486	0.79279
$Q_{AB/F}$	0.43961	0.46592	0.32281	0.41037	0.53985	0.28789	0.41277	0.599348
SCD	1.38963	1.60993	1.59124	1.00488	1.65223	1.4806	1.70699	0.60521

5. Discussion and conclusion

A novel feature level image fusion algorithm using segmentation based on AOA is proposed in this manuscript. Firstly, using AOA based segmentation, the IR image is segmented into different groups. These weights map functions are refined using WLS optimization. The resulting segmented image is used for measuring weight map function used in the process of fusion. Finally, the fused image is reconstructed using pixel-wise weighted average fusion. The efficiency of the approach proposed is measured using both qualitative and quantitative methods. Comparison of the proposed algorithm's experimental results show better performance in terms of different performance metrics which

Table 6
Comparison of execution time (in seconds).

Image name	Method name							
	CBF	DCH	JSR	GTF	LP-GAN	CNN	LLR	Proposed
UN Camp (360 × 270)	10.11	3.45	1.648	1.313	1.59	0.591	25.22	**0.587**
Traffic (632 × 496)	31.38	6.55	2.219	2.001	2.73	0.853	238.08	**0.791**
Steamboat (510 × 505)	25.32	6.82	1.901	1.658	2.13	0.725	179.49	**0.617**

include *SSIM*, *SCD*, $Q^{AB/F}$, and $N^{AB/F}$. The proposed fusion results show complete preservation of an object and background details without introducing any undesired artifacts and thus, the proposed technique is suitable for the fusion of IR and VIS image data sets. Findings from experiments show that using AOA weight maps to compute weight maps lead to significant increases in fusion performance. For example, military surveillance and medical imaging would be benefited from this technique. It is our future vision to implement a method to produce threshold values with more precise parameter selection, which will enhance segmentation in visible and IR images.

CRediT authorship contribution statement

Simrandeep Singh: Conceptualization, Methodology, Software. **Harbinder Singh:** Data curation, Writing – original draft. **Nitin Mittal:** Investigation, Visualization. **Abdelazim G. Hussien:** Writing – review & editing. **Filip Sroubek:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

Abualigah, L., Diabat, A., Mirjalili, S., Abd Elaziz, M., & Gandomi, A. H. (2021). The arithmetic optimization algorithm. *Computer Methods in Applied Mechanics and Engineering, 376*, Article 113609. https://doi.org/10.1016/j.cma.2020.113609

Alexander, T. (2014). *TNO Image Fusion Dataset*. Retrieved from https://figshare.com/articles/dataset/TNO_Image_Fusion_Dataset/1008029/1.

Bavirisetti, D. P., & Dhuli, R. (2016). Fusion of infrared and visible sensor images based on anisotropic diffusion and Karhunen-Loeve transform. *IEEE Sensors Journal, 16*(1), 203–209. https://doi.org/10.1109/JSEN.2015.2478655

Bhandari, A. K., & Kumar, I. V. (2019). A context sensitive energy thresholding based 3D Otsu function for image segmentation using human learning optimization. *Applied Soft Computing Journal, 82*, Article 105570. https://doi.org/10.1016/j.asoc.2019.105570

Bohat, V. K., & Arya, K. V. (2019). A new heuristic for multilevel thresholding of images. *Expert Systems with Applications, 117*, 176–203. https://doi.org/10.1016/j.eswa.2018.08.045

James, A. P., & Kavitha, A. P. (2014). Mean-variance blind noise estimation for CT images. *Advances in Intelligent Systems and Computing, 264*, 235–243. https://doi.org/10.1007/978-3-319-04960-1_21

Jiang, Q., Jin, X., Hou, J., Lee, S. J., & Yao, S. (2018). Multi-sensor image fusion based on interval type-2 fuzzy sets and regional features in nonsubsampled shearlet transform domain. *IEEE Sensors Journal, 18*(6), 2494–2505. https://doi.org/10.1109/JSEN.2018.2791642

Kamel, M., & Zhao, A. (1993). Extraction of binary character/graphics images from grayscale document images. *CVGIP: Graphical Models and Image Processing, 55*, 203–217. https://doi.org/10.1006/cgip.1993.1015

Kaur, R., & Singh, S. (2017). An artificial neural network based approach to calculate BER in CDMA for multiuser detection using MEM. *Proceedings on 2016 2nd International Conference on Next Generation Computing Technologies, NGCT 2016*, (October), 450–455. 10.1109/NGCT.2016.7877458.

Krishnamoorthy, S., & Soman, K. P. (2010). Implementation and comparative study of image fusion algorithms. *International Journal of Computer Applications, 9*(2), 25–35. https://doi.org/10.5120/1357-1832

Li, H., X.-J. Wu. (2018). Infrared and visible image fusion using Latent Low-Rank Representation. *ArXiv Preprint ArXiv:1804.08992*. Retrieved from http://arxiv.org/abs/1804.08992.

Li, J., & Wang, Q. (2022). Multi-modal bioelectrical signal fusion analysis based on different acquisition devices and scene settings: Overview, challenges, and novel orientation. *Information Fusion, 79*, 229–247. https://doi.org/10.1016/j.inffus.2021.10.018

Li, S., & Yang, B. (2008). Multifocus image fusion by combining curvelet and wavelet transform. *Pattern Recognition Letters, 29*(9), 1295–1301. https://doi.org/10.1016/J.PATREC.2008.02.002

Liang, H., Jia, H., Xing, Z., Ma, J., & Peng, X. (2019). Modified grasshopper algorithm-based multilevel thresholding for color image segmentation. *IEEE Access, 7*, 11258–11295. https://doi.org/10.1109/ACCESS.2019.2891673

Liu, C. H., Qi, Y., & Ding, W. R. (2017). Infrared and visible image fusion method based on saliency detection in sparse domain. *Infrared Physics and Technology, 83*, 94–102. https://doi.org/10.1016/j.infrared.2017.04.018

Liu, Y., Chen, X., Peng, H., & Wang, Z. (2017). Multi-focus image fusion with a deep convolutional neural network. *Information Fusion, 36*, 191–207. https://doi.org/10.1016/j.inffus.2016.12.001

Liu, Y., Chen, X., Wang, Z., Wang, Z. J., Ward, R. K., & Wang, X. (2018). Deep learning for pixel-level image fusion: Recent advances and future prospects. *Information Fusion, 42*, 158–173. https://doi.org/10.1016/J.INFFUS.2017.10.007

Liu, Y., Chen, X., Ward, R. K., & Wang, J. (2016). Image fusion with convolutional sparse representation. *IEEE Signal Processing Letters, 23*(12), 1882–1886. https://doi.org/10.1109/LSP.2016.2618776

Liu, Y., & Wang, Z. (2015). Simultaneous image fusion and denoising with adaptive sparse representation. *IET Image Processing, 9*(5), 347–357. https://doi.org/10.1049/iet-ipr.2014.0311

Meher, B., Agrawal, S., Panda, R., & Abraham, A. (2019). A survey on region based image fusion methods. *Information Fusion*. https://doi.org/10.1016/j.inffus.2018.07.010

Meher, B., Agrawal, S., Panda, R., Dora, L., & Abraham, A. (2022). Visible and infrared image fusion using an efficient adaptive transition region extraction technique. *Engineering Science and Technology, an International Journal, 29*, Article 101037. https://doi.org/10.1016/j.jestch.2021.06.017

Mousavirad, S. J., & Ebrahimpour-komleh, H. (2019). Human mental search-based multilevel thresholding for image segmentation. *Applied Soft Computing Journal*. https://doi.org/10.1016/j.asoc.2019.04.002

Otsu, N., Smith, P. L., Reid, D. B., Environment, C., Palo, L., Alto, P., & Smith, P. L. (1979). Otsu_1979_otsu_method. *IEEE Transactions on Systems, Man, and Cybernetics, C*(1), 62–66.

Pan, B., Shi, Z., & Xu, X. (2017). Hierarchical guidance filtering-based ensemble classification for hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing, 55*(7), 4177–4189. https://doi.org/10.1109/TGRS.2017.2689805

Petrović, V. S., & Xydeas, C. S. (2004). Gradient-based multiresolution image fusion. *IEEE Transactions on Image Processing, 13*(2), 228–237. https://doi.org/10.1109/TIP.2004.823821

Petrović, V. S., & Xydeas, C. S. (2003). Sensor noise effects on signal-level image fusion performance. *Information Fusion, 4*(3), 167–183. https://doi.org/10.1016/S1566-2535(03)00035-6

Piella, G. (2003). A general framework for multiresolution image fusion: From pixels to regions. *Information Fusion, 4*(4), 259–280. https://doi.org/10.1016/S1566-2535(03)00046-0

Qiu, C., Wang, Y., Zhang, H., & Xia, S. (2017). Image fusion of CT and MR with sparse representation in NSST domain. *Computational and Mathematical Methods in Medicine, 2017*. https://doi.org/10.1155/2017/9308745

Shi, Z., Yang, Y., Hospedales, T. M., & Xiang, T. (2017). Weakly-supervised image annotation and segmentation with objects and attributes. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(12), 2525–2538. https://doi.org/10.1109/TPAMI.2016.2645157

Shreyamsha Kumar, B. K. (2013). Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform. *Signal, Image and Video Processing, 7*(6), 1125–1143. https://doi.org/10.1007/s11760-012-0361-x

Shreyamsha Kumar, B. K. (2015). Image fusion based on pixel significance using cross bilateral filter. *Signal, Image and Video Processing, 9*(5), 1193–1204. https://doi.org/10.1007/s11760-013-0556-9

Singh, H., Hrisheekesha, P. N., & Cristobal, G. (2019). Infrared and visible image fusion based on nonparametric segmentation. *International Journal of Innovative Technology and Exploring Engineering, 8*(9 Special Issue), 29–35. 10.35940/ijitee.I1005.0789S19.

Singh, S., Mittal, N., & Singh, H. (2020). A multilevel thresholding algorithm using LebTLBO for image segmentation. *Neural Computing and Applications, 32*(21), 16681–16706. https://doi.org/10.1007/s00521-020-04989-2

Singh, S., Mittal, N., & Singh, H. (2021a). A multilevel thresholding algorithm using HDAFA for image segmentation. *Soft Computing, 25*. https://doi.org/10.1007/s00500-021-05956-2

Singh, S., Mittal, N., & Singh, H. (2021b). Review of various image fusion algorithms and image fusion performance metric. *Archives of Computational Methods in Engineering, 28*(5), 3645–3659. https://doi.org/10.1007/s11831-020-09518-x

Singh, S., Mittal, N., & Singh, H. (2022). A feature level image fusion for IR and visible image using mNMRA based segmentation. *Neural Computing and Applications, 7*. https://doi.org/10.1007/s00521-022-06900-7

Singh, S., Singh, H., Kaur, A. J., & Gagandeep. (2022). IR and visible image fusion using DWT and bilateral filter. *Microsystem Technologies, 7*. https://doi.org/10.1007/s00542-022-05315-7

Singh, S., et al. (2020). Classification of various image fusion algorithms and their performance evaluation metrics. *Computational Intelligence for Machine Learning and Healthcare Informatics* (pp. 179–198). De Gruyter.

Singh, S., et al. (2022). Nature and Biologically Inspired Image Segmentation Techniques. *Archives of Computational Methods in Engineering, 29*, 1415–1442. https://doi.org/10.1007/s11831-021-09619-1

Singh, S., et al. (2013). Multi-resolution representation of multifocus image fusion using Gaussian and Laplacian pyramids. *Int J Adv Res Comput Sci Softw Eng, 3*(11), 1639–1642.

Singh, S., et al. (2020). Multifocus image fusion based on multiresolution pyramid and bilateral filter. *IETE Journal of Research*, 1–12. https://doi.org/10.1080/03772063.2019.1711205

Song, Y., Wu, W., Liu, Z., Yang, X., Liu, K., & Lu, W. (2016). An adaptive pansharpening method by using weighted least squares filter. *IEEE Geoscience and Remote Sensing Letters, 13*(1), 18–22. https://doi.org/10.1109/LGRS.2015.2492569

Sreeja, P., & Hariharan, S. (2018). An improved feature based image fusion technique for enhancement of liver lesions. *Biocybernetics and Biomedical Engineering, 38*(3), 611–623. https://doi.org/10.1016/j.bbe.2018.03.004

Wang, J., Ke, C., Wu, M., Liu, M., & Zeng, C. (2021). Infrared and visible image fusion based on laplacian pyramid and generative adversarial network. *KSII Transactions on Internet and Information Systems, 15*(5), 1761–1777. https://doi.org/10.3837/TIIS.2021.05.010

Wang, S. (2022). Enhanced Remora Optimization Algorithm for Solving Constrained Engineering Optimization Problems. *Mathematics, 10*, 1696. https://doi.org/10.3390/math10101696

Xu, H., Ma, J., Jiang, J., Guo, X., & Ling, H. (2020). U2Fusion: A unified unsupervised image fusion network. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 44*(1), 502–518. https://doi.org/10.1109/tpami.2020.3012548

Zhang, H., Xu, H., Tian, X., Jiang, J., & Ma, J. (2021). Image fusion meets deep learning: A survey and perspective. *Information Fusion, 76*, 323–336. https://doi.org/10.1016/J.INFFUS.2021.06.008

Zhang, Q., Fu, Y., Li, H., & Zou, J. (2013). Dictionary learning method for joint sparse representation-based image fusion. *Optical Engineering, 52*(5), Article 057006. https://doi.org/10.1117/1.oe.52.5.057006