Australasian Physical & Engineering Sciences in Medicine (2019) 42:647–657
https://doi.org/10.1007/s13246-019-00753-6

TECHNICAL PAPER

Accurate segmentation of inflammatory and abnormal regions using medical thermal imagery

Kakali Das[1] · Mrinal Kanti Bhowmik[1] · Omkar Chowdhuary[1] · Debotosh Bhattacharjee[2] · Barin Kumar De[3]

Received: 1 June 2018 / Accepted: 30 March 2019 / Published online: 5 April 2019

Abstract

Methodologies reported in the existing literature for identification of a region of interest (ROI) in medical thermograms suffer from over- and under-extraction of the abnormal and/or inflammatory region, thereby causing inaccurate diagnoses of the spread of an abnormality. We overcome this limitation by exploiting the advantages of a logarithmic transformation. Our algorithm extends the conventional region growing segmentation technique with a modified similarity criteria and a stopping rule. In this method, the ROI is generated by taking common information from two independent regions produced by two different versions of a region-growing algorithm that use different parameters. An automatic multi-seed selection procedure prevents missed segmentations in the proposed approach. We validate our technique by experimentation on various thermal images of the inflammation of affected knees and abnormal breasts. The images were obtained from three databases, namely the Knee joint dataset, the DBT-TU-JU dataset, and the DMR-IR dataset. The superiority of the proposed technique is established by comparison to the performance of state-of-the-art competing methodologies. This study performed temperature emitted inflammatory area segmentation on thermal images of knees and breasts. The proposed segmentation method is of potential value in thermal image processing applications that require expediency and automation.

Keywords Hotspot detection · Inflammation · Region growing · Thermal imaging

Introduction

Thermal imaging is an emerging area in the field of medical diagnosis that involves recording of the skin temperature of the human body. The surface temperature distribution of the human body contains essential information regarding pathological activities and is sensitive to different body abnormalities. Thermal imaging has already proven its efficiency in the early prediction of breast abnormalities (occurring near the surface area of the breast, such as ductal carcinomas) [1–3], surface inflammation in patients suffering from different types of joint pain [4, 5], prediction of diabetic foot ulcers [6, 7], dermatological disorders, vascular disorders, ophthalmology, and surgery [8, 9]. Although proven to be useful, thermal medical images do have a few limitations, such as (1) thermal images only detect the surface temperature, (2) abnormalities such as benign breast tumours may be inaccurately identified in thermal images due to a lack of inflammation. The information in thermal images may differ based on the concerned diseases. For knee joint diseases, those that are responsible for inflammation, produce local heat by an increase in blood flow. In cases of breast pathologies too, the mechanism that generates heat is due to the increase in local blood flow. Cancerous lesions are more prone to increase that flow by inducing inflammation. Inflammation helps the spread of malignant lesions. The inflammation causes the production of heat.

✉ Kakali Das
kakalids54@gmail.com

Mrinal Kanti Bhowmik
mrinalkantibhowmik@tripurauniv.in

Omkar Chowdhuary
cse.omkar@gmail.com

Debotosh Bhattacharjee
debotosh@ieee.org

Barin Kumar De
barin_de@yahoo.com

[1] Computer Science and Engineering, Tripura University, Suryamaninagar, Tripura 799022, India

[2] Department of Computer Science and Engineering, Jadavpur University, Kolkata, West Bengal, India

[3] Department of Physics, Tripura University, Suryamaninagar, Tripura 799022, India

Although, the heat is mostly detected by thermal imaging when the lesions occur near the skin surface. However the mechanism, thermography can be used in most situations with the same accuracy. In both cases, early detection of the disease often becomes difficult due to the non-specific symptoms. Nevertheless, due to the onset of inflammation, the increased blood flow starts before the noticeable symptoms, by increasing the local temperature, however little it may be. Thermography can detect small thermal changes, helping in early detection of the targeted pathologies. The increase in the temperature of the affected region is referred to as the hot region (hotspot) in thermal images. In such conditions, the detection and analysis of the hotregions from thermograms are found to be very effective for an early diagnosis. The distant goal of this paper is to provide an alternative opinion for the clinicians (radiologists, rheumatologists, etc.), who may feel suspicious about the possible presence of a specific inflammation in human breasts and knees. This paper, though restricted to knee joints and breast tissues, paves the path for future use in the detection of inflammation or abnormalities in other areas of the body. Studies have shown the significance of thermal image segmentation for abnormality prediction in breast cancer [13] and knee joint pain related diseases [14], such as Rheumatoid Arthritis, Osteo Arthritis, etc. Therefore, in brief, abnormalities in breasts due to breast cancer and affected inflammatory knee joints due to arthritis produce hotspots in the corresponding thermal images [9, 15, 17]. The hotspot area is our region of interest (ROI), and may differ according to different disease conditions. The abnormality could also be identified by an analysis of the texture, shape and other parameters of the hotspot area. Analysis of the ROI can provide a better understanding of disease activity and can define the type of disease. In the implementation of automatic disease detection, recognition of the disease type and ROI activity segmentation will play an important role. Figure 1 represents some examples of thermal images with distinct hotspots caused by breast abnormalities and knee joint diseases. Segmenting the hotspot regions from the thermal images (shown in Fig. 1) is a very challenging problem due to the presence of smooth edges. The segmentation of thermal images must overcome this challenge. Several segmentation techniques exist for hotspot segmentation in thermal images, as shown in Table 1. Accordant with our observations from Table 1, we found that FCM, K-Means, and Otsu's thresholding based segmentation techniques were popularly employed. Furthermore, region growing (RG) is found to be very efficient in thermal image segmentation because of the natural choice of pixels with the highest intensity as initial seed points and it is also found to be stable to noise [10]. The proposed scheme uses the benefits of a logarithmic transformation along with the RG. The significant contributions and advantages of this article are summarized as follows:

1. Our proposed technique offers an effective hotspot segmentation mechanism of the ROI devoid of the associated problem of segmentation.
2. The proposed method introduces an automatic multi-seed point selection criterion for seed point initialization. This solves the missed segmentation problem, which arises due to the presence of disconnected regions in the hotspot.
3. It experiments on two breast thermal image datasets, DBT-TU-JU [15] and DMR-IR [16], and one newly created inflammatory knee joint dataset [17], with clinically validated ground truths for hotspot segmentation.
4. A comparative study, for proper validation and inference, is presented between the state-of-the-art segmentation methods used in medical thermal images and the proposed method.

The structure of the paper is: "Proposed Segmentation Method" section describes the proposed segmentation technique. In "Experimental results"and "Discussions" sections,

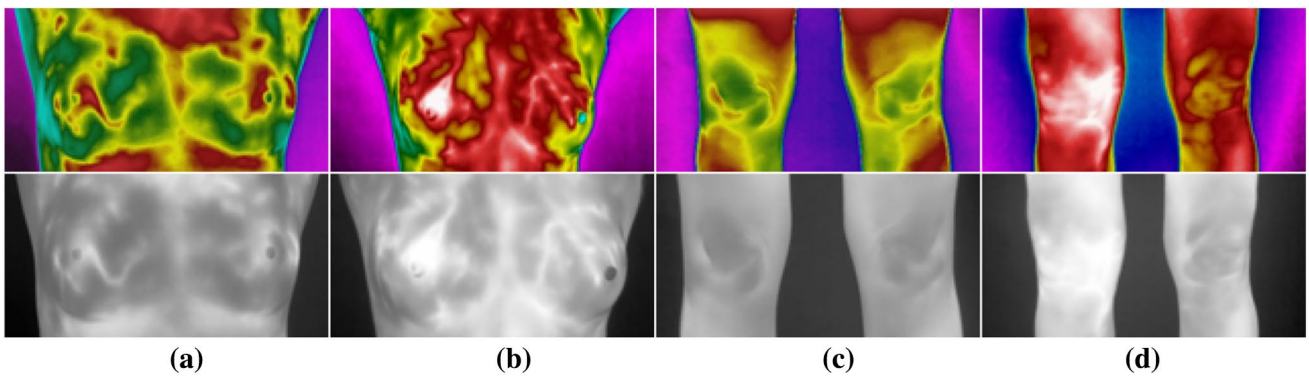

Fig. 1 Pseudo color images with their corresponding Gray Images of: **a** breast image of patient's with no abnormality. **b** Abnormality in the patient's right breast. **c** Image of knee of a patient with no arthritis. **d** Osteo-Arthritis in patient's left knee

Table 1 Thermal image segmentation in literature

References	Comments
Jadin and Taib [41]	A comparative study of methods such as Kittler & Illingworth's Minimum Error Thresholding, Thresholding based on Entropy of the image properties, and first order statistical properties, OTSU Thresholding, Fuzzy-C-Means Clustering (FCM), Graph Cut, Mean shift, SIFT was performed on thermal images of electric systems. All the methods other than the SIFT results in over-segmentation. SIFT also works for thermal images having repetitive patterns.
Etehadtavakol et al. [42]	Considered color segmentation rather than the gray image segmentation using Fuzzy-C-means and K-Means algorithm, using limited thermal breast images. They showed FCM to correctly acquire the ROI in every cases.
Snekhalatha et al. [43]	Segmentation was performed using K-means which provided better feature extraction from thermal images of hands.
Shahari et al. [44]	Showed the K-means algorithm is able to segment the hot region from the thermal images of breast.
Etehadtavakol et al. [45]	Compared Mean Shift and K-means with FCM and concluded FCM performs better. Experiments shown in a limited number of thermal breast images. Also showed the importance of fractal geometry in the analysis of the ROI.
Etehadtavakol et al. [46]	Used a Fuzzy-C-means algorithm for segmentation of ROI from Breast Thermograms. Used bispectrum features from the ROI for classification between malignant, non-malignant, benign and normal.

we compare and analyse the outcomes of the proposed segmentation with the state-of-the-art techniques. Finally, we conclude in "Conclusion" section.

Proposed segmentation method

Thermal images of a human body part with any abnormality creates a persistent hotspot on that image. An RG produces good accuracy in the segmentation of the hotspot from the thermal image. However, smooth edges in a thermal image leads to a decrease in accuracy of the hotspot segmentation procedure, as it may include an unwanted region around the edge or it may discard important areas around the edge. Therefore, performing an RG on the original image will not provide accurate segmentation results. We proposed an RG by modifying the stopping criteria and seed selection procedure. We employed the proposed RG on the original image as well as the log-transformed image. One low pass filter is used after a logarithmic transformation. An RG was performed on the filtered image. The output from this filtered image was not enough to produce an accurate segmentation. That is why we employed an RG on both the original image and filtered image. The flow of the proposed method is depicted in Fig. 2.

Filtering for noise removal

Images are subjected to noise and the removal of noise is considered a preprocessing step in image processing. Image segmentation is also followed by noise removal in the proposed work of different authors [20]. Medical thermal images are captured in a controlled environment. However, some variations in surface emissivity are generally present. To remove the effect of these variations, a logarithmic transformation is used along with an average filter (low pass filter). This procedure is equivalent to using the homomorphic filter [40]. When using the homomorphic filter in thermal images the slow spatial variations are removed from the thermal image. The output, considered as 'filtered image'

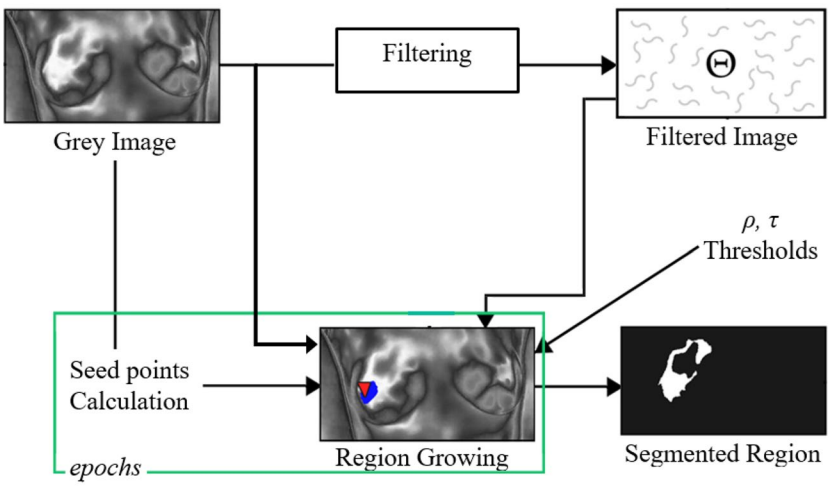

Fig. 2 Flow diagram of the proposed method

(denoted by Ω), is obtained after implementing this procedure step-by-step on the original image (denoted by I). An RG is performed on the 'filtered image,' Ω.

An RG with similar stopping criteria and seed selection procedure is performed on the 'filtered image.' For stopping criteria, $\tau = 0.02$ is used, which will be the same for all the datasets, as already discussed.

Extended RG algorithm with modified selection criteria

The Conventional Multi-seeded RG technique [18] (CMRG) requires single/multiple seed points and a threshold to start the segmentation [19]. By assimilating homogeneous neighbouring pixels, the segment grows iteratively. The resulting segmentation could highly depend on the initial chosen seed, threshold values and the criteria on which neighbouring pixels are examined. The selection of the homogeneity criteria in an RG depends not only on the problem under consideration but also on the type of image subjected to segmentation. There are three significant issues in the CMRG [19]. They are (1) Selection of seed points, (2) Similarity and connectivity criteria, (3) Termination of the segmentation process (Stopping rule). The similarity criteria and the stopping rule are discussed in a later section, and seed selection is described "Automatic multi seed selection" section.

Similarity criteria and stopping rule

CMRG uses a single threshold value for the similarity and stopping criteria. We introduce two thresholds ρ and τ, for the same purpose in our technique. Hence, justifiably we use the pixels with the highest intensity value as the seed points in every *epoch*. The *epochs* are stopped if no more seed points matching a selection criterion given in (1) are found. The ith seed point Z_i is taken from the set of seed points Z as the starting pixel for the RG. When considering the seed point Z_i with the coordinate (u, v) as a part of the region to be segmented, we define two pseudo-variables μ_k and δ_k assigning to them the value $\Omega_{u,v}$. Where Ω is defined as the image matrix and on this, the RG method will be performed. Each pixel in the 8-neighbourhood of the seed point is considered for comparison using (1), providing thereby the similarity criteria and stopping rule.

$$X = \begin{cases} True, \ if |\mu_s - \Omega_{i,j}| < \tau \\ \quad if \left|\delta_k - I_{i,j}\right| < \rho \\ False, \ otherwise \end{cases} \tag{1}$$

The pixels for which X becomes *'True'* as per (8), are labelled as the *"detected region"* and are included in the set Δ_i, where i is the number of iterations. The number of iterations depends on the number of seed points chosen in a particular *epoch*. The pseudo-variables μ_s and δ_k are updated using (2) when new pixels from the neighbourhood of any boundary pixel is/are identified in the *"detected region,"*

$$\mu_s = \frac{1}{n}\sum_{i}^{n} Mi \, and, \ \delta_k = \frac{1}{n}\sum_{i}^{n} Ni \tag{2}$$

where, M_n is the set of the corresponding pixel values of Ω of n number of pixel coordinates present in Δ_i and μ_s is the mean of these n energy values in M_n. Similarly, N_n is the set of the corresponding pixel values of I of pixel points present in Φ_i and δ_k is the mean of the n entropy values in N_n. The two conditions in (1) are followed in each iteration growing the detected region while updating the pseudo-variables μ_s and $\delta_k$in every iteration. When the iteration stops, the final set of pixel coordinates Δ_i is produced. Δ_i is the detected intermediate ROI grown from the ith seed point of Z. The same process is repeated for all the remaining seed points only if they do not belong to the already detected hotspot region Δ_j.

$$\Delta_j = \bigcup_{i}^{m} \Delta_i \forall j \tag{3}$$

The final region output Δ_j is computed using (3) which is the union of Δ_i's detected from all the m seed points where j is the epoch number. After each epoch, the detected ROI (Δ_{roi}) is updated using (4).

$$\Delta_{roi} = \Delta_{roi} + \Delta_j \forall j \tag{4}$$

The RG algorithm described above is multi-seeded. A multi-seeded RG is required for proper segmentation of a disconnected sub-region. Single seed segmentations capable of partitioning one region from the image as one seed grows in one region may lead to missed segmentation problems in the thermal image segmentation. Thermal images contain disconnected hotspot regions. So, a multi-seeded RG is required for thermal image segmentation for the hotspot. Seeds can be selected manually or automatically. A manual seed selection burdens the complexity of the segmentation algorithm and the precision of the algorithm is also compromised through the manual intervention. We proposed an automatic seed selection to overcome this problem, and it is described below.

Automatic multi seed selection

Thermal images may contain disconnected hotspot regions, with the pixel values of different ranges. Since in thermal images, the CMRG technique assigns the pixels with the highest grey value as seed points, the hotspot regions that have pixels smaller than the highest grey value are bound to remain undetected. In that event, there is a high probability

of missed segmentations. To solve this issue, we introduce criterion for seed point selection. The pixels with the highest grey value are selected as the initial seed points for the first *epoch*. After the *i*th *epoch,* seed points are again selected for the next *epoch* using two criteria. First, the seed points are the pixels with the highest grey values among all the pixels of the input image excluding the ones already included in the first epoch,i.e., the current Δ_{roi}. Second, the highest pixel (say P) must satisfy (5),

$$P > \min(\Delta roi) + \lceil \eta \times \{\max(\Delta roi) - \min(\Delta roi)\} \rceil \quad (5)$$

where min (Δ_{roi}) is the minimum grey value in Δ_{roi}, max(Δ_{roi}) is the maximum grey value in Δ_{roi}. In addition η is a real number whose value can lie in the range (0,1). The value of η can vary for different datasets. We obtained the best results using $\eta = 0.3$ for all the data sets (see section IV-D). At the beginning of each *epoch,* new seed points are calculated. The criterion (5) chooses a pixel as a seed point only if its grey value exceeds a certain integer (say k). Let us assume the intermediate segmented set of pixels Δ_{roi} in the jth *epoch* has the lowest grey value (say a) and the highest grey value (say b). Therefore, in the $(j+1)$th *epoch*, as per (5), k must lie in between a and b where, $k = a + \eta(b - a)$. This indicates the grey value of the seed point must be greater than k. We assume each disconnected hotspot region in the ROI must contain at least one pixel with a grey value greater than k. Hence, the missed segmentations can be reduced by using (5) but may not be eliminated.

Computation of common area (intersection) from detected regions

A logarithmic transformation is not suitable for applications where high-intensity values contain information. This transformation will cause the loss of some information regarding hotspot segmentation as that area of the thermal image contains the highest pixel values. Therefore, we choose seed values from the original image and the RG is performed on the filtered image with this seed.

An RG is also performed on the original image. The output of this RG is also considered in the final step of the proposed method. We individually perform the proposed RG technique (RG), on I and Ω using thresholds ρ, τ to obtain two separate regions, π, and λ respectively. There are two possibilities, these two obtained regions can be merged by a union or can be intersected. Union of these two regions produces an ROI that surely includes the abnormal region, but is unable to produce a clear boundary as is present in the ground truth. Whereas intersecting these two regions will produce an ROI that includes the actual boundary of the ROI. The aim of this proposed method is to obtain an actual regional boundary without over-extraction and under-extraction. Therefore, the two obtained regions are intersected to get the final output.

$$\Delta_{roi} = \pi \cap \lambda \quad (6)$$

The proposed method gives the final segmented region Δ_{roi} as the intersection of π and λ, given in (6). Since the intersection of the two independent sets of pixels is output as the final segmented ROI (Δ_{roi}), the problem of RG segmentation may be specifically avoided.

Experimental results

Database preparation

The proposed segmentation technique in our work was tested on three thermal medical image datasets: (1) the DBT-TU-JU breast dataset [15] (40 images), (2) the DMR-IR breast dataset [16], designed by Silva et al. (44 images), and (3) a new inflammatory knee joint dataset [17] (50 images), which was created and collected from the Physical Medicine and Rehabilitation (PMR) Department, Agartala Government Medical College (AGMC) and Regional Cancer Center (RCC), Govind Ballav Pant Hospital (GBP), Agartala, Tripura, India. The first and second dataset contain breast images with an abnormality such as ductal breast cancer, malignant, or benign tumour. The third dataset consists of thermal images of different inflammatory pain-related diseases of knee joints such as osteo arthritis, rheumatoid arthritis, reactive arthritis, and mono arthritis. The dataset images were manually trimmed to remove the temperature scales for the experiments. The dataset was designed by preserving all the protocols that are proposed in the recognized literature. Table 2 precisely contains all the protocols that were maintained to create the dataset.

There are some factors that adversely affect the thermogram acquisition [21, 29]. The thermogram acquisition factors for pain diagnosis can be categorized into four groups: infrared camera specifications, infrared camera positioning, environmental conditions and patient preparation [17]. Camera specification, camera positioning, and patient preparation are easy to handle manually compared to the environmental conditions. Control of the environmental conditions is not so much in human hands. Radiation of objects present in the environment surrounding the acquisition setup can degrade the quality of the infrared image. The infrared image captures the surface temperature of objects. Other objects, having radiation properties will produce artefacts on the temperature profile of the actual object during capturing. The radiation property refers to the infrared energy property of some objects, such as electric wires, pipes, outlets, etc. The best place to do the acquisition, to avoid these interferences, is in a dark room made of a black piece of cloth and wood

Table 2 Protocols maintained during acquisition

Type of protocols	Protocol
Camera specification	Camera name: FLIR T650sc
	Sensitivity: < 20mK @ 30 °C
	Spectral range: 7.5–14.0 µm
	Image resolution: 640 × 480 pixels
Patient preparation [25–28]	Undressed the part to be captured
	Stabilization for 15 min in at the aquisition room and should avoid contact of other parts.
	Patients should avoid smoking, nerve stimulation, acupunture, hot or cold presses, Physiotherapy or intense physical excercises etc. for 24 h of prior to thermography
	Any kind of jewelary not allowed at that time of imaging.
	Patient should avoid tight fitting clothes.
Camera positioning	Camera positioned at a distance of 2-m from the patient.
	The 90° alignment between the camera and the body part to be captures is also an important requirement

[22]. The room should also be protected from direct sunlight and air drifts [23]. The field of view (FOV) and required distance between object and camera is used to determine the size of the room. We also tried to control other factors such as room temperature and room humidity. The room temperature was fixed between 22 and 25 °C and the humidity was between 40 and 60% [24]. Other protocols that are maintained in the acquisition of infrared images regarding all the groups are mentioned in Table 2.

In previous work [15], they used colour thermograms for segmentation. Colour thermograms use pseudo colours for representation of and pseudo colours do not share the properties of the RGB colour palette. Therefore, segmentation based on the pseudo colours will not be able to provide accurate results. Therefore, the grey palette images are used in this work. The difference of results in using grey palette versus the colour palette are provided in Table 3. From Table 3 we find that the results of the FCM and K-means are better when the grey palette is used rather than the colour palette along with the I-indexing. Whereas the results of the other two methods, FO-DPSO and the Mean shift algorithm, decreased when we use grey palette images with I-indexing. In the same manner the performance of the RG also decreases with the grey palette image.

Clinical validation of the datasets

Clinical validation of medical images highlights their performance as well as their limitations. Hence, they are considered an important part of medical image processing. The DBT-TU-JU and DMR-IR datasets were already validated clinically in Refs. [15, 16], respectively. The newly created knee joint dataset was validated by examining the arthritis patients. Arthritis symptoms, as described by clinicians, are an acute pain in one joint or multiple joints with swelling and inflammation. Joint deformation, tenderness, and immobility are important symptoms of such a disease. There are two ways to validate arthritis:

Table 3 Quantitative difference of segmentation results between the color palette and Grey palette thermograms

	Color palette thermograms				Grey palette thermograms			
	JI	DS	RC	PRC	JI	DS	RC	PRC
DBT-TU-JU dataset								
FCM	0.36	0.51	0.82	0.40	0.45	0.58	0.46	0.92
K-means	0.32	0.47	0.86	0.36	0.36	0.47	0.36	0.93
RG	0.64	0.77	0.91	0.70	0.50	0.61	0.64	0.68
FO-DPSO	0.70	0.80	0.86	0.78	0.08	0.15	0.08	0.97
Meanshift	0.64	0.75	0.80	0.75	0.07	0.14	0.07	0.99
DMR-IR								
FCM	0.25	0.38	0.62	0.35	0.24	0.36	1	0.24
K-means	0.22	0.35	0.84	0.25	0.28	0.41	1	0.08
RG	0.54	0.67	0.75	0.66	0.08	0.13	0.08	0.83
FO-DPSO	0.61	0.74	0.67	0.92	0.08	0.14	1	0.08
Meanshift	0.56	0.71	0.80	0.73	0.03	0.06	1	0.03

(1) a subjective measure, i.e., joint count and patient assessment of the diseases's activity, and, (2) an objective measure, i.e., acute phase reactants such as ESR (Erythrocyte sedimentation rate), CRP (C-reactive protein) and radiographic imaging. The subjective measure is based on four criteria: tenderness, swelling, redness, and restriction of movement. These are based on the assessment of both the patients and the clinicians. These tuples are found positive for thermal images of patients having a hotspot, which shows an abnormality. Objective measures incorporate ESR and CRP measures. Patients with higher ESR and CRP from the normal range are diagnosed as having arthritis. The ESR and CRP reports of patients examined here validate the scope of thermal imaging in the diagnosis of arthritis.

Generation of ground truth

Ground truth creation for all the three datasets was necessary to verify the competence of the proposed technique in comparison to the state-of-the-art methods. An alternative solution is the creation of ground truths using manual segmentation [30]. A manual segmentation is considered the most reliable method for a segmentation to identify the shape and the structure of a particular clinical task. Generally, a doctor/clinician has more knowledge and experience with the disease, its spread, and its origin. Only they can appropriately recognize the affected area, and for that, ground truths should be validated by doctors with expertise. Ground truths for the knee joint dataset were created in two consecutive steps. First, seven thermal image experts/technicians independently traverse the hotspot region of the thermal images to create ground truth images using the GNU Image Manipulation Program (GIMP) [31] software. In the next step, a pixel wise voting policy is considered to find the final output ground truth image. In a pixel-wise voting policy, if a pixel is considered a foreground pixel by a minimum of four technicians, then it is counted as a foreground pixel. By classifying each pixel based on this method a final ground truth is generated. In the second step, the medical expert (doctor) analyses a final ground truth and finalizes them. A different policy is used for ground truth generation of the breast images (from DBT-TU-JU and DMR-IR). Four medical experts are employed for the creation of a ground truth using four different software packages: (a) Safexa segmentation tool (b) a Photoshop tool, (c) a GNU Image Manipulation Program Tool and (d) FLIR Research IR tool. After the generation of ground truths, the same pixel wise voting policy/majority voting rule is conducted for a final ground truth selection.

Performance measures

To analyse the performance of the proposed method compared with the state-of-the-art methods different quantitative measures have been used. These following measures are used for a performance analysis: (1) Jaccard index (JI) [32, 35, 38] (2) Dice similarity index (DS) [32, 39] (3) Recall (RC) [32], (4) Precision (PRC) [32], (5) Relative Volume Distance (RVD) [33] and (6) Accuracy [37, 38]. The Jaccard Similarity Index (JI) is an accepted measure for evaluating the efficiency of any segmentation method. A JI works well for larger and simple shaped objects [34]. The JI measures the similarity between the ground truth and the segmented image and varies between (0,1), where 1 (one) indicates a perfect overlap of the compared segmentation. A DS index quantifies the overlapping of the ground truth and segmented image. The value of the DS also varies between the range of (0,1) where 1 indicates a good segmentation. Recall (RC) & Precision (PRC) are popularly used together to identify the effectiveness of a segmentation technique. The higher the value of an RC and PRC, the better the performance of the technique. Between these two measures, if the value of one is higher than the other with a significant difference, less accuracy is applied. The Relative Volume Distance (RVD) identifies the volume size error that concerns the ground truth.

Thresholds and parameter selection

An optimal energy threshold ρ and entropy threshold τ were selected separately for both the datasets to test our technique. Both ρ and τ vary in the range of (0, 1). Based on trial and error, we choose ρ as 0.03 and τ as 0.02 for the inflammatory knee joint dataset. ρ as 0.02 and τ as 0.07 are used for the DBT-TU-JU breast dataset, whereas ρ as 0.03 and τ as 0.02 was used for the DMR-IR breast dataset. The plot in Fig. 3a illustrates that the best segmentation accuracy was attained using the selected thresholds for the knee data set. The performance of the algorithm does show dependency on the choice of the parameter η, selected between the range of (0,1). The plot in Fig. 3b illustrates that the best segmentation was achieved with η in the range 0.6 and 0.7 for all the datasets ($\eta = 0.66$ for the DBT-TU-JU dataset, $\eta = 0.7$ for the DMR-IR dataset, and, $\eta = 0.7$ for the knee joint dataset).

Discussions

We compared the proposed technique with other recent and/or widely accepted state-of-the-art segmentation methods which were successfully used for medical thermal image segmentation. Such techniques for this purpose include (1) Otsu's Thresholding, (2) K-means

segmentation, (3) Fuzzy-C-Means (FCM), (4) Mean-shift Algorithm, (5) RG (RG) [18], and (6) FO-DPSO: Fractional Order Darwinian Particle Swarm Optimization [36]. A comparison of the proposed method with the state-of-the-art methods, based on the average of JI, DS, RC, PRC, RVD, and accuracy for knee joint and breast datasets respectively are presented in Tables 4 and 5. The results imply the proposed method can serve as an efficient alternative technique for thermal medical image segmentation. The segmentation outputs of the proposed technique along with the state-of-the-art methodsare compared in Fig. 4.

From the experimental analysis, we have observed that segmenting the two breast datasets was more challenging than the knee joint dataset. Compared to the breast datasets, all the segmentation methods discussed in this paper

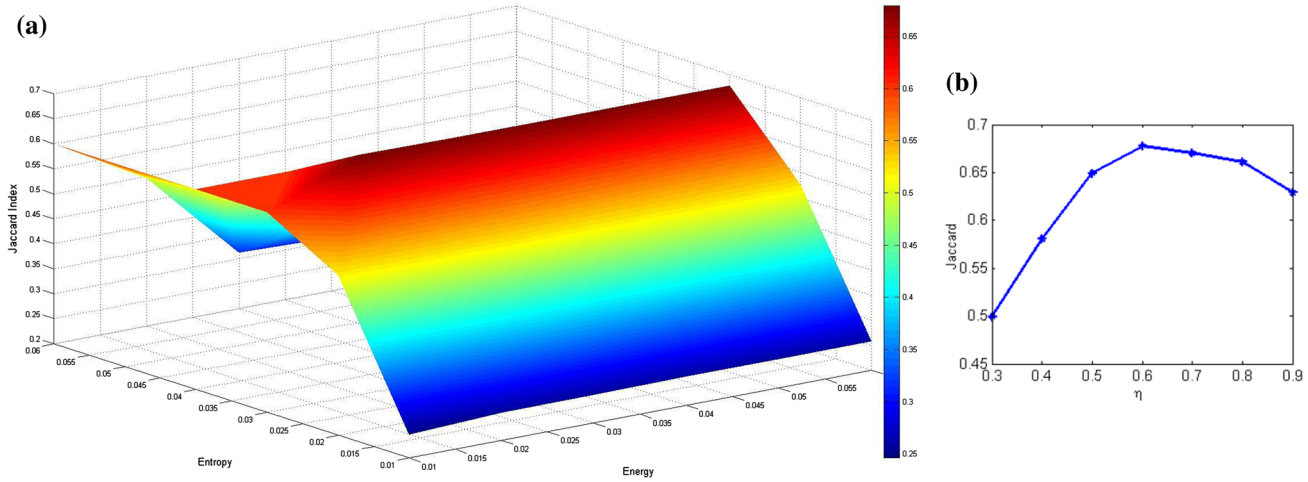

Fig. 3 **a** Plot the effects in segmentation Jaccard Index with respect to both energy and entropy thresholds for knee data set. **b** The effect in Jaccard Index (JI) with respect to the parameter (η)

Table 4 Quantitative measures for all three datasets

	Knee joint dataset			DBT-TU-JU			DMR-IR		
	JI	DS	RVD	JI	DS	RVD	JI	DS	RVD
Prposed method	**0.68**	**0.79**	**−15**	**0.54**	**0.66**	**−2.84**	**0.56**	**0.69**	**−13**
FCM	0.38	0.54	161	0.45	0.58	67	0.24	0.36	293
K-means	0.33	0.47	325	0.36	0.47	100	0.28	0.41	966
RG	0.57	0.71	−63	0.50	0.61	−20	0.08	0.13	−219
Otsu's thresholding	0.35	0.50	196	0.31	0.44	124	0.17	0.28	426
Mean-shift	0.09	0.16	686	0.07	0.14	470	0.03	0.06	−242
FO-DPSO	0.19	0.31	376	0.08	0.15	448	0.08	0.14	945

The bold represent the accuracy of the proposed method

Table 5 Quantitative measures for all three datasets

	Knee joint dataset			DBT-TU-JU			DMR-IR		
	RC	PRC	ACC	RC	PRC	ACC	RC	PRC	ACC
Prposed method	**0.81**	**0.83**	**98.07**	**0.60**	**0.79**	**96.40**	**0.72**	**0.81**	**98.25**
FCM	0.38	1	92.06	0.46	0.92	91.32	1	0.24	89.9
K-means	0.34	0.99	83.22	0.36	0.93	82.70	1	0.08	92.1
RG	0.87	0.65	97.84	0.64	0.68	96.02	0.08	0.83	97.14
Otsu's thresholding	0.35	1	90.77	0.31	0.93	83.82	1	0.17	83.96
Mean-shift	0.09	1	52.55	0.07	0.99	19.73	1	0.03	17.87
FO-DPSO	0.19	1	79.02	0.08	0.97	26.20	1	0.08	64.27

The bold represent the accuracy of the proposed method

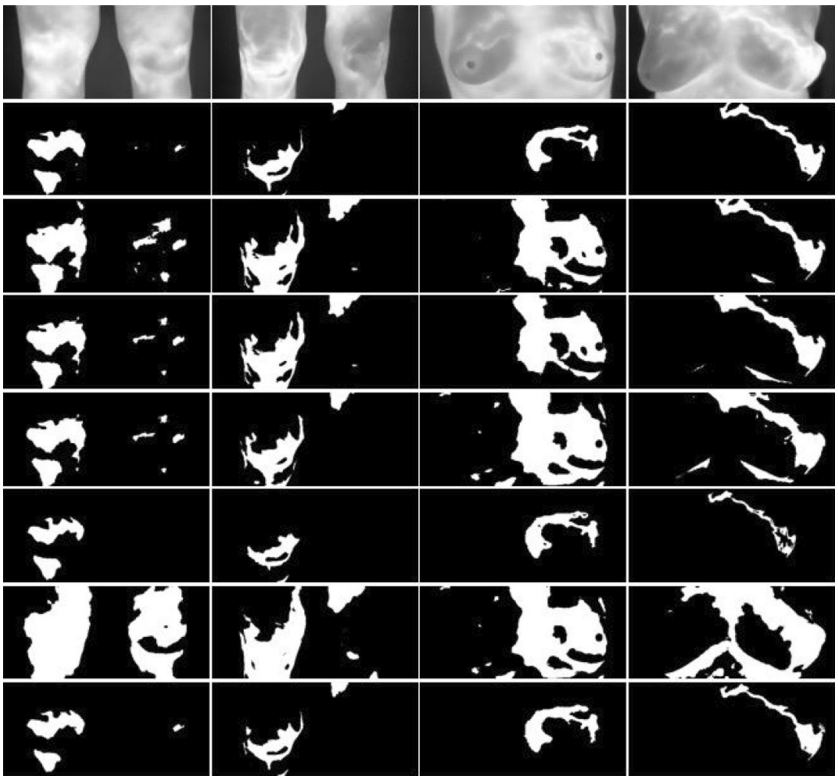

Fig. 4 Comparison of ROI segmentation in knee and breast data set. Row 1: original image, Row 2: ground truths, Row 3: result of K-Means, Row 4: result of FCM, Row 5: Otsu's thresholding, Row 6: region growing (RG), Row 7: FO-DPSO segmentation, Row 8: proposed segmentation method TIDE

including our proposed method perform better on the knee joint dataset.

A statistical significance test was also performed to access the difference of results among these mentioned methods. A Kruskal–Wallis P-test, a nonparametric significance proof for the all the samples of the three datasets was conducted. The analysis is performed considering a 5% significance level over all the performance measures including JI, DS, RC, PRC, RVD,and Accuracy. Table 6 reports the p-values produced by the Kruskal–Wallis test for a pairwise comparison of all the performance measure values of the proposed method with the other existing methods. As a

Table 6 Kruskal Wallie's significant test

Proposed method vs	JI	DS	RC	PRC	RVD	ACC
Knee joint dataset						
FCM	5.5e – 09	5.5e – 09	2.9e – 14	7.2e – 17	3.4e – 14	2.3e – 12
K-means	2.3e – 13	2.3e – 13	5.2e – 16	1.e – 16	1.4e – 11	3.1e – 16
RG	2.1e – 17	1.4e – 17	2.5e – 10	1.2e – 17	2.1e – 10	1.1e – 12
Otsu's thresholding	1.4e – 14	1.4e – 14	1.5e – 16	2.3e – 17	1.2e – 11	9.3e – 17
FO-DPSO	8.2e – 18	8.2e – 18	7.3e – 18	2.0e – 17	4.3e – 15	8.8e – 18
DBT-TU-JU dataset						
FCM	2.6e – 13	2.6e – 13	4.4e – 13	6.7e – 13	3.4e – 04	8.5e – 11
K-means	3.3e – 13	3.3e – 13	7.4e – 13	4.3e – 13	0.0120	4.6e – 10
RG	2.9e – 10	2.9e – 10	4.1e – 07	1.4e – 11	0.8400	0.0028
Otsu's thresholding	4.8e – 13	4.8e – 13	6.9e – 13	2.0e – 11	6.7e – 06	3.2e – 12
FO-DPSO	2.4e – 12	2.4e – 12	6.4e – 13	5.4e – 09	3.9e – 12	4.7e – 14
DMR-IR dataset						
FCM	1.9e – 08	1.9e – 08	4.3e – 13	6.6e – 15	3.2e – 13	1.8e – 12
K-means	5.8e – 07	5.8e – 07	1.9e – 12	1.2e – 14	1.1e – 11	2.6e – 11
RG	3.6e – 17	3.6e – 17	1.5e – 16	4.2e – 17	8.5e – 05	6.9e – 06
Otsu's thresholding	1.0 e – 10	1.0e – 10	3.9e – 14	3.5e – 15	4.9e – 14	3.1e – 14
FO-DPSO	4.3e – 14	4.3e – 14	2.0e – 15	3.5e – 15	2.4e – 15	9.2e – 16

null hypothesis, it is assumed that there are no differences between the values of the two objective functions. The alternative hypothesis considers an existent difference between the values of both approaches. All p-values reported in Table 6 have less than a 0.05 (5% significance level), which is strong evidence against the null hypothesis, indicating that the performance of the proposed method is statistically better and it has not occurred by chance.

Conclusion

There is a lack of algorithms that can automatically segment the hotspot region from the medical thermal images of different diseases. In this paper, we have proposed a method, a new scheme to efficiently segment the hotspot region in the thermal images of abnormal breasts and inflammatory knee joints. The proposed method lends itself to an automatic approach with a manually adjusted threshold value that is simple to implement and quite generalized for other thermal medical image applications. The proposed approach substantially balances the over-, under- and missed-segmentations, thus providing the segmenter a tool to attain acceptable segmentation quality. We have already discussed the fact that the proposed method is an extension of the RG technique, so we expect that it will get better results than the RG method. The proposed method stood up to expectations by obtaining higher accuracy in both the knee and breast datasets. It also performs competitively with the recent state-of-the-art segmentation techniques.

An application of this article could be in the provision of a thermography-guided method to find the source of a respective abnormality using a prior estimation, such as the tumour's position in the breast or the source of knee pain. Another application of this paper could be a dosimetric quantification for patients, suffering from inflammatory knee joint pain related diseases [11]. Depending on the intensity of the knee inflammation, real time monitoring of the dose rate can be performed [12]. Thermography can be a reliable guide for the prognosis of a disease's activity through gradual reduction or an increase in thermal emissions over time with treatment. The proposed scheme may be useful in developing health abnormality detection systems based on thermal images.

Acknowledgements The research work was supported by the Grant No. 5/7/1516/2016-RCH Dated: 20/06/2017 from the Indian Council of Medical Research (ICMR), Government of India.

Compliance with ethical standards

Conflict of interest The authors declare there is no potential conflict of interestwith respect to the authorship and/or publication of this article.

Ethical approval Ref. No.4(6-11)-AGMC/Medical Education/Ethics Com/2018/15136, Dated 31st December, 2018.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Lee J, Lee J, Song S, Lee H, Lee K, Yoon Y (2008) Detection of suspicious pain regions on a digital infrared thermal image usingthe multimodal function optimization. In: 30th annual international conference of the IEEE engineering in medicine and biology society 2008
2. Nola IA, Kolanc D (2015) Thermography in biomedicine. In: 57th international symposium ELMAR (ELMAR)
3. Gogoi UR, Bhowmik MK, Bhattacharjee D, Ghosh AK, Majumdar G (2015) A study and analysis of hybrid intelligent techniques for breast cancer detection using breast thermograms. In: Dutta P, Chakraborty S, Bhattacharyya S (eds) Hybrid soft computing approaches studies in computational intelligence. Springer, New Delhi, pp 329–359
4. Brioschi M, Teixeira ML, Silva MT, Colman FM (2010) Medical thermography textbook: principles and applications. Andreoli, Sao Paulo
5. Qi H, Kuruganti PT, Snyder WE (2012) Detecting breast cancer from thermal infrared images by asymmetry analysis. In: Medicine and Medical Research, vol 38
6. Liu C, Heijden FVD, Klein ME, Baal JGV, Bus SA, Netten JJV (2013) Infrared dermal thermography on diabetic feet soles to predict ulcerations: a case study. In: Mahadevanj-Jansen A, Vo-Dinh T, Grundfest WS (eds) Advanced Biomedical and Clinical Diagnostic Systems. SPIE, Bellingham, p 9
7. Ring F (2010) Thermal imaging today and its relevance to diabetes. J Diabetes Sci Technol 4(4):857–862
8. Ring EFJ (1978) Thermographic terminology. Acta Thermo-Grap 2:1–30
9. Jones B (1998) A reappraisal of the use of infrared thermal image analysis in medicine. IEEE Trans Med Imaging 17:1019–1027
10. González RC, Woods RE (2002) Digital image processing. Prentice-Hall, Upper Saddle River
11. Palmer M (2016) TH-E-209-00: radiation dose monitoring and protocol management. Med Phys 43:3902
12. Takeda T (2016) Treatment strategy of elderly rheumatoid arthritis. Jpn J Clin Immunol 39:497–504
13. Gogoi UR, Majumdar G, Bhowmik MK, Ghosh AK, Bhattacharjee D (2015) Breast abnormality detection through statistical feature analysis using infrared thermograms. In: Proceedings of the international symposium on advanced computing and communication (ISACC) 2015
14. Bardhan S, Bhowmik MK, Nath S, Bhattacharjee D (2015) A review on inflammatory pain detection in human body through infrared image analysis. In: Proccedings of the international symposium on advanced computing and communication (ISACC) 2015
15. Bhowmik MK, Gogoi UR, Majumdar G, Bhattacharjee D, Datta D, Ghosh AK (2017) Designing of ground truth annotated DBT-TU-JU breast thermogram database towards early abnormality prediction. IEEE J Biomed Health Inform 22:1238–1249
16. Silva LF, Saade DCM, Sequeiros GO, Silva AC, Paiva AC, Bravo RS, Conci A (2014) A new database for breast research with infrared image. J Med Imaging Health Inform 4:92–100
17. Bhowmik MK, Bardhan S, Das K, Bhattacharjee D, Nath S (2016) Pain related inflammation analysis using infrared images.

In: Zalameda JN, Bison P (eds) Thermosense: thermal infrared applications XXXVIII. SPIE, Bellingham, p 14
18. Adams Leanne (1994) Seeded RG. IEEE Trans Pattern Anal Mach Intell 16:641–647
19. Bhowmik MK, Bardhan S, Das K, Bhattacharjee D, Nath S (2016) Pain related inflammation analysis using infrared images. In: Zalameda JN, Bison P (eds) Thermosense: thermal infrared applications XXXVIII. SPIE, Bellingham, p 14
20. Paulano F, Jiménez JJ, Pulido R (2014) 3D segmentation and labeling of fractured bone from CT images. Vis Comput 30:939–948
21. Zaproudina N, Varmavuo V, Airaksinen O, Närhi M (2008) Reproducibility of infrared thermography measurements in healthy individuals. Physiol Meas 29:515–524
22. Denoble AE, Hall N, Pieper CF, Kraus VB (2010) Patellar skin surface temperature by thermography reflects knee osteoarthritis severity. Clin Med Insights 45:78–96. https://doi.org/10.4137/CMAMD.S5916
23. Herry CL, Frize M, Goubran RA (2008) Search for abnormal thermal patterns in clinical thermal infrared imaging. In: IEEE international workshop on medical measurements and applications 2008
24. Leroy P (1995) Update in interventional digitalized infra-red thermal imaging in pain management. In: Proceedings of 17th international conference of the engineering in medicine and biology society
25. Ijzerman RG, Serne EH, Weissenbruch MMV, Jongh RTD, Stehouwer CDA (2003) Cigarette smoking is associated with an acute impairment of microvascular function in humans. Clin Sci 104:247
26. Steketee J (1973) Spectral emissivity of skin and pericardium. Phys Med Biol 18:686–694
27. Disasi AF (2003) Early effects of cigarette smoking in hypertensive and normotensive subjects An ambulatory blood pressure and thermographic study. J Min cardioangiol 51:387–393
28. http://www.iact-org.org/professionals/thermog-guidelines.html
29. Herry CL, Frize M, Goubran RA (2008) Search for abnormal thermal patterns in clinical thermal infrared imaging. In: Proceedings of the 2008 IEEE international workshop on medical measurements and applications. https://doi.org/10.1109/memea.2008.4542999
30. Shareef N, Wang D, Yagel R (1999) Segmentation of medical images using LEGION. IEEE Trans Med Imaging 18:74–91
31. Oliver E, Ruiz J, She S, Wang J (2006) The software architecture of the GIMP, 2006
32. Farag A, Lu L, Roth HR, Liu J, Turkbey E, Summers RM (2017) A bottom-up approach for pancreas segmentation using cascaded superpixels and (Deep) image patch labeling. IEEE Trans Image Process 26:386–399
33. Li C, Wang X, Eberl S, Fulham M, Yin Y, Chen J, Feng DD (2013) A likelihood and local constraint level set model for liver tumor segmentation from CT volumes. IEEE Trans Biomed Eng 60:2967–2977
34. Philipsen RHHM, Maduskar P, Hogeweg L, Melendez J, Sanchez CI, Ginneken BV (2015) Localized energy-based normalization of medical images: application to chest radiography. IEEE Trans Med Imaging 34:1965–1975
35. Chiang J, Birla S, Bedoya M, Jones D, Subbiah J, Brace CL (2015) Modeling and validation of microwave ablations with internal vaporization. IEEE Trans Biomed Eng 62:657–663
36. Ghamisi P, Couceiro MS, Martins FML, Benediktsson JA (2014) Multilevel image segmentation based on fractional-order darwinian particle swarm optimization. IEEE Trans Geosci Remote Sens 52:2382–2394
37. Sadri AR, Zekri M, Sadri S, Gheissari N, Mokhtari M, Kolahdouzan F (2013) Segmentation of dermoscopy images using wavelet networks. IEEE Trans Biomed Eng 60:1134–1141
38. Bi L, Kim J, Kumar A, Fulham M, Feng D (2017) Stacked fully convolutional networks with multi-channel learning: application to medical image segmentation. Vis Comput 33:1061–1071
39. Chung F, Schmid J, Magnenat-Thalmann N, Delingette H (2010) Comparison of statistical models performance in case of segmentation using a small amount of training datasets. Vis Comput 27:141–151
40. 6.4 homomorphic filtering using a low pass filter. In: 2.2 Basic Principles of MRI. https://www.cs.sfu.ca/~stella/papers/blairthesis/main/node37.html. Accessed 21 Jun 2018
41. Jadin MS, Taib S (2012) Recent progress in diagnosing the reliability of electrical equipment by using infrared thermography. Infrared Phys Technol 55(4):236–245
42. Etehadtavakol M, Sadri S, Ng EYK (2008) Application of K- and Fuzzy C-means for color segmentation of thermal infrared breast images. J Med Syst 34(1):35–42
43. Snekhalatha U, Anburajan M, Sowmiya V, Venkatraman B, Menaka M (2015) Automated hand thermal image segmentation and feature extraction in the evaluation of rheumatoid arthritis. Proc Inst Mech Eng Part H J Eng Med 229(4):319–331
44. Shahari S, Wakankar A (2015) Color analysis of thermograms for breast cancer detection. International Conference on Industrial Instrumentation and Control (ICIC)
45. Etehadtavakol M, Ng E, Lucas C, Sadri S (2012) Fuzzy C means segmentation and fractal analysis of the benign and malignant breast thermograms. In: Medical infrared imaging, pp 1–20
46. Etehadtavakol M, Chandran V, Ng E, Kafieh R (2013) Breast cancer detection from thermal images using bispectral invariant features. Int J Therm Sci 69:21–36