Australasian Physical & Engineering Sciences in Medicine (2019) 42:259–277
https://doi.org/10.1007/s13246-019-00726-9

TECHNICAL PAPER

2-Stage classification of knee joint thermograms for rheumatoid arthritis prediction in subclinical inflammation

Shawli Bardhan[1] · **Mrinal Kanti Bhowmik**[1]

Received: 20 January 2018 / Accepted: 17 January 2019 / Published online: 31 January 2019

Abstract

Presence of inflammation in knee joint is the early indication of arthritis. In this paper, we performed the inflamed region segmentation from knee joint thermograms for structural feature extraction based knee abnormality prediction. Existing four popular segmentation techniques are investigated, namely K-means, Fuzzy C-means, Otsu, Single seeded region growing. We proposed modified multi-seeded region growing method that generates 98.6% accurate segmentation rate compared to ground truth of inflammation. Based on the spread of the inflammation oriented structural feature analysis, in the first stage of classification we classified arthritis affected knee joint thermograms, and all other types of thermograms (non-arthritis) with 91% accuracy. Among different types of arthritis, the most damaging type that causes disability of joints in long run is known as rheumatoid arthritis (RA). Early diagnosis of RA in subclinical stage enormously helps clinicians to decrease the disease affect. In second stage of classification, we integrated the RA and non-RA categorization by extracting texture, shape and frequency level features. Experiment shows that the combination of all features decreases the accurate detection rate of RA classification. To increase the classification rate, we incorporated the accuracy based feature selection procedure. The RA classification rate obtained with accuracy based feature selection is 73% whereas existing support vector machine-recursive feature elimination (SVM-RFE) and RELIEF methods provide 67% and 71% correct classification rate respectively. The area under the curve (AUC) of accuracy based feature selection, SVM-RFE, and RELIEF for RA classification are 0.72, 0.65 and 0.67, respectively and it indicates better classification outcome of the accuracy based feature selection method.

Keywords Arthritis · RA · Segmentation · Feature extraction · Feature selection · Classification

Introduction

Rheumatoid arthritis (RA) is an autoimmune chronic disease that causes inflammation in affected joints, generates pain and deformity which leads to long-term disability. Worldwide RA occurrence rate is in between 0.5 and 1% and in India, the rate is 0.9% [1–3]. The analysis shows that mortality rate due to RA is 22% of all deaths under arthritis and other rheumatic conditions (AORC) [4]. Thermal image based inflammation analysis of arthritis may help in diagnosis of RA in subclinical stage. Depending on the fact, this study addresses the problem of predicting RA affected knee joints with the help of thermal image captured from arthritis affected patients, normal subjects and pain related subjects with unknown reason.

Existing studies shows that globally 100 different kinds of arthritis are present among which RA is the commonly occurring inflammatory arthritis which causes damage to body joints and internal organs [5]. The other commonly occurring types are osteoarthritis (OA), reactive arthritis (ReA), polyarthritis (PA) etc. RA may affect several joints of the body like knee, ankle, wrist, finger joint, elbow, and hip. The most commonly affected large joint by RA is knee [6]. Diagnosis of RA depends on the subjective evaluation of tenderness, swelling, restriction of movement and pathological tests. In case of subclinical condition, factors associated with subjective evaluation are mild in nature which results imprecise diagnosis of arthritis. The diagnosis of arthritis depending on structural image

✉ Mrinal Kanti Bhowmik
mrinalkantibhowmik@tripurauniv.in

Shawli Bardhan
shawli.cse@gmail.com

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

of joints are popularly performed using X-ray plate, histological image, and Magnetic Resonance Image. Observer dependent subjective evaluation based diagnosis of RA by medical expert provides higher sensitivity (above 0.9) compare to X-ray image and histological image oriented accurate RA diagnosis (0.55 and 0.38, respectively) [7]. X-ray and MRI generate a structural diagnosis of arthritis through structural imaging that is able to detect an abnormality in the acute stage of the disease. Clinical investigation of RA highly depends on erythrocyte sedimentation rate (ESR). Higher the value of ESR indicates the presence of inflammation in the body which is related to RA. The normal range of ESR is 0–15 mm/h for male and 0–22 mm/h for female [8]. In some cases, the ESR range for RA and for other kind of arthritis may become similar as shown in Fig. 1. The Fig. 1 also shows that ESR rate of some RA patients is under normal range. In such situation, clinical test based diagnosis of RA is impalpable for medical experts. Thermal imaging may serve as image based non-invasive potential inflammation detecting tool in subclinical stage of RA diagnosis. Inflammation oriented diagnosis of RA using thermogram may also help medical experts in the treatment of disease depending on the type of arthritis.

Investigation of previous works shows that, Snekhalatha et al. [9] analyzed the hand thermograms of the rheumatoid patients and normal subjects depending on temperature oriented first order statistical features and gray level co-occurrence matrix features. In their study, features are extracted from the segmented region generated through the application of K-means clustering on hand thermograms. Borojević et al. [10] also used first order statistical features for temperature based analysis of RA and OA. The temperature oriented study of thermograms by Frize et al. [11] reports that thermal imaging can detect the presence of RA

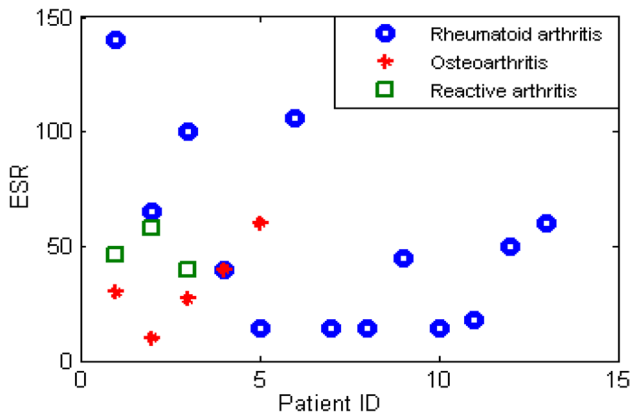

Fig. 1 ESR rate for type based arthritis patients

in the human body. The study also shows that compared to other joints, the metacarpal phalangeal joint of the middle and index finger, and the knee joint best indicate the presence RA in human body. In the area of feature-oriented analysis of thermogram, Acharya et al. [12] performed textural feature-based classification of normal and malignant breast thermogram from corresponding co-occurrence and run-length matrix of thermal breast image. Classification of normal and abnormal breast thermogram was executed by Francis et al. [13] by extracting the statistical and texture features from the curvelet domain of thermogram. Milosevic et al. [14] also detected pattern based abnormality through the extracted gray level co-occurrence Matrices from breast thermogram. Classification of breast thermogram in between three groups, i.e., breast with benign tumor, breast with malignant tumor, and healthy breast has been carried out by EtehadTavakol et al. [15] using the extracted bi-spectral invariant features from breast thermogram intensity distribution.

The thermogram analysis based review in medical domain shows a vast application of thermal imaging in breast thermogram classification. The review also shows that image oriented thermogram based analysis of RA is present in past research, but contains lack of automated classification of RA. In this paper, we focused on the area of automated classification of thermogram for detection of RA. The knee joint is selected as the large joint for classification as the knee is most commonly affected joint by arthritis and clearly indicates the presence of RA [6, 16, 17]. Focusing on RA diagnosis from thermogram, contributions of our analysis includes,

(a) segmentation of inflamed region of interest from knee thermograms through hybrid technology,
(b) structural feature based classification of arthritis and other types (normal knee or knee pain with unknown reason) of knee thermograms,
(c) classification of RA affected thermograms among all types of arthritis affected knee thermograms by selection of relevant features.

Methodological description

The classification of RA from all types of knee thermogram is performed using two-stage classifier as shown in Fig. 2. The knee thermograms are arthritis affected or not is determined in the 1st stage of classification using structural features as described in "1st stage classification: arthritis affected knee vs. other types of knee (non-arthritis)". Arthritis affected thermograms are RA or non-RA will be determined in the second stage of classification (described

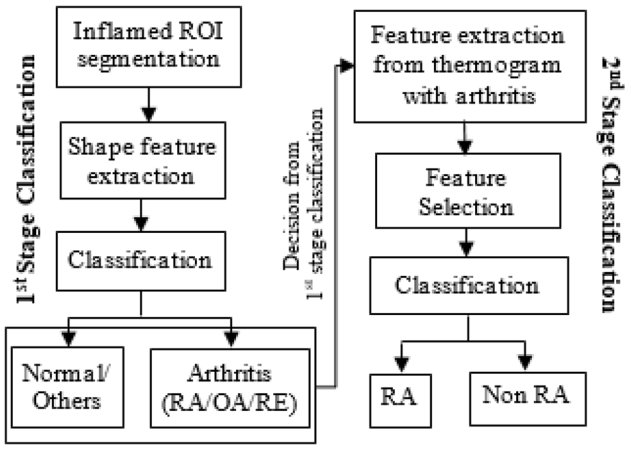

Fig. 2 Overall flow-diagram of RA classification using knee thermogram

in "2nd stage classification: RA vs. non RA"). Pre-processing of all the thermogram are performed to increase the accuracy of classification as detailed in "Thermogram pre-processing".

1st stage classification: arthritis affected knee vs. other types of knee (non-arthritis)

Inflammation in knee region indicates increase of temperature due to abnormality. In case of arthritis, the body defense system turns against the healthy tissues by considering those as foreign substances which causes inflammation in joints [18]. Arthritis in knee region results tenderness, swelling, redness and synovitis [19]. In subclinical stage, those symptoms may not be clearly detected by medical experts where thermography helps in detection of subclinical inflammation. Inflammation in knee region may also occurs due to increase in body temperature, external pressure in the knee area or excessive work and may cause pain in the knee region. Therefore, the 1st stage of classification aims to classify in between the thermograms of arthritis affected knees and other knees (knee with pain due to unknown reason/healthy knee) by analyzing the shape of inflammation spreading. The classification is followed by inflamed region segmentation and shape feature extraction from the segmented regions.

Inflamed region of interest (ROI) segmentation

Segmentation of inflamed region from knee thermogram extracts the suspicious regions that may contain temperature more than the average temperature of the knee. The suspicious regions in thermogram are brighter than surrounding pixels with fuzzy boundaries and low contrast. Therefore, segmentation of inflamed regions with maximum rate of true positive (TP) and true negative (TN) is a complex task to perform. In past research, Fuzzy C-means (FCM) [9, 20, 21], K-means [21–23], Otsu thresholding (Otsu) [24–26] and single seeded region growing (RG) [27, 28] methods are popularly used for segmenting inflamed region. Performance of all the methods mainly depends on user input parameters like cluster number, cluster center, initial seed point, threshold level. The existing region growing method for thermogram inflamed region extraction is mainly single seeded and threshold oriented where maximum intensity pixel is selected as the seed point. But, thermogram may contain multiple disconnected inflamed region of interests. Therefore, in our analysis, we use Modified Multi-seeded region growing method. The Algorithm 1 and 1.1 represents the procedure of the modified multi-seeded region growing used for thermogram segmentation in this study. The steps are elaborately detailed below.

Algorithm input and variable initialization The output of the pre-processing step (described in "Thermogram pre-processing") in gray scale format is used as the input of the Algorithm and represented by variable I of size $x \times y$. Here, x is the total numbers of rows and y is the total number of columns of the input image. The threshold value T is fixed to 0.01 for growing the region i.e., for considering similar types of pixels compare to seed pixel under inflamed region and *COND* is initialized as 0. The detailed of fixing T as 0.01 is discussed in "T and ϕ value selection for modified multi-seeded region growing". The *COND* is the variable used as iteration stopping criteria. A matrix B' of size $x \times y$ is initialized with zero value for storing the binary mask of region growing outcome.

ALGORITHM 1 : Modified Multi-Seeded Region Growing

Input: $I_{x\times y}$

Variable Initialize: $T \leftarrow 0.01$; $COND \leftarrow 0$; $B^{'}_{x\times y} \leftarrow 0$;

Step 1: $I^{'} = I - 1; P = I^{'}/255; R = P; \phi = 238/255$;

Step 2: $M, u, v \leftarrow$ *Maximum intensity value of* P *and its position from* Algorithm 1.1

Step 3: *if* $M < \phi$ *then* $COND \leftarrow 1$

Step 4: *while* $COND \neq 1$

Step 4.1: *for* $i = -1, \cdots, 1$

for $j = -1, \cdots, 1$

if $|P(u+i, v+j) - \mathrm{M}| < T$ then $P(u+i, v+j) \leftarrow 1$

Step 4.2: $B = Binary\ mask\ generation\ from\ P\ using\ equation\ 1$

Step 4.3: $B^{'} = B^{'} \cup B$

Step 4.4: $k = Transpose\ the\ binary\ mask\ using\ equation\ 2$

Step 4.5: $P = R \otimes k$

Step 4.6 $M, u, v \leftarrow$ *Maximum intensity value of* P *and its position using* Algorithm 1.1

Step 4.7: *if* $M < \phi$ *then* $COND \leftarrow 1$

else $COND \leftarrow 0$

Step 5: $F = I \otimes B^{'}$

Output: $B^{'}, F$

ALGORITHM 1.1: Maximum intensity value and its position extraction.

Input: $P_{x\times y}$; $x =$ total number of rows in I; $y =$ total number of columns in I;

Initialize: $M \leftarrow 0; u \leftarrow 0$; $v \leftarrow 0$;

Step 1: *for* $i = 1, \cdots, x$

Step 1.1: *for* $j = 1, \cdots, y$

Step 1.1.1: *if* $M < P(i, j)$ *then* $M \leftarrow P(i, j)$

Step 2: $u \leftarrow i$; $v \leftarrow j$;

Output: M, u, v ;

Step 1 The input image I is stored in I' by decrementing all the pixel intensities with value one. Deduction of pixel value helps in correct seed point selection in further steps. The I' is normalized by dividing with 255 (highest gray level pixel intensity) and stored in P and R for computational simplicity. The variable ϕ is fixed by taking the mean of histogram peaks with highest gray level of all the thermograms and used for conditional checking for iteration progress. The detailed of fixing ϕ is discussed in "T and ϕ value selection for modified multi-seeded region growing". Depending on ϕ, the *COND* value is determined for condition checking.

Step 2, 3, 4 The selection of seed point is performed by extracting the highest pixel intensity and its position in the input image using Algorithm 1.1. The variable M stores the maximum intensity value within the image and u, v stores the corresponding row and column position of M. If the value of M is less than ϕ, then *COND* will set to 1 and Algorithm processing will terminate. In first iteration $COND = 1$ indicates that, the themogram does not contain any inflamed region of interest. If step 3 does not satisfies, then *COND* will remain with value 0 and iteration process will progress. The process of region growing is iterative in manner and step 4 mainly controls the progress of iteration.

Step 4.1 In this step, the region growing will take place considering the selected seed (M) in previous step. Each eight neighborhood pixel of the seed pixel is compared with the seed value depending on absolute difference. The neighborhood pixel (s) whose values are very near to the seed value (allowing a minimum threshold, T) are considered under inflamed region of interest and also set with value 1.

Step 4.2,4.3 To represent the inflamed region of interest, a binary mask is created depending on the output of step 4.1 and using Eq. 1. The binary mask based outcome of nth iteration is merged with the binary mask of $(n-1)$th iteration in step 4.3 for extraction of optimal inflamed region of interest:

$$B\,(m,n) = \begin{cases} 1 \;\; if \;\; P(m,n) = 1 \\ 0 \;\; if \;\; P(m,n) \neq 1 \end{cases}, \tag{1}$$

here, (m, n) indicates the each pixel position in P.

Step 4.4,4.5 The binary mask based outcome of step 4.3 is transposed using Eq. 2. This step mainly applied to avoid the pixels which are already considered under inflamed region of interest in the next seed selection process of Algorithm 1.1. The step 4.5 performs pixel wise multiplication of the step 4.4 outcome and R. The outcome of step 4.5 considers the pixels with value 1 of B' as 0. Through this, in step 4.5, the 0 intensity pixels of k generates corresponding 0 values in P and not considered as seed in Algorithm 1.1:

$$k(m,n) = \begin{cases} 1 \;\; if \;\; B'(m,n) \neq 1 \\ 0 \;\; if \;\; B'(m,n) = 1 \end{cases}, \tag{2}$$

B' is the outcome of step 4.3. Here, (m, n) indicates the each pixel position in B'.

Step 4.6, 4.7 The step 2 is repeated here for selection of seed in each iteration using Algorithm 1.1. In step 4.7, the COND is fixed to 0 or 1 depending on the condition $M < \phi$. The value of COND is used for termination of iteration in step 4.

Step 5 In this step, the final binary mask (B') represented the inflamed ROI is pixel wise multiplied with the original input image (I). The output of the multiplication (F) provides original intensity of the inflamed region.

Algorithm output The binary mask B' and the inflamed region F is the final outcome of the modified multi-seeded region growing algorithm. The mask mainly represents the area of inflammation with value 1 and used for shape feature extraction of inflamed region. The F consists of inflamed region pixel intensity values and used for statistical analysis in classification step.

Shape feature extraction

In case of arthritis affected knee, maximum inflammation present in patella region and its surrounded synovial membrane in connected cluster manner. Therefore, after inflamed ROI segmentation, a set of shape features are extracted from the ROI which contains information regarding the area and boundary of inflammation spreading [29]. The arthritis affected knee contains more inflammation compared to the normal knee or pain containing knee due to general causes [30]. So, in case of knee arthritis, the spread of inflammation is clearly visible with a large area of spreading, whereas inflammation spreading to other knees are blurry and not in patella region as shown in Fig. 3b, d. Thus we use seven shape features to classify the knee thermograms in between arthritis affected group and normal/others group as given in Table 1.

Arthritis and non-arthritis classification

The 1st stage classification is performed to assign the thermogram in a certain class in between arthritis affected knee or other knee (normal or knee pain for unknown reason). The classification is executed by a classifier that takes shape features of inflamed ROI, given in Table 1 as input. Both the 1st and 2nd stage of classification used in this study was performed by using the support vector machine (SVM) classifier. The SVM classifier is the well-known classifier for medical image classification [12, 29, 38]. We use SVM classification with linear kernel in this study. The linear kernel is the simplest one compare to others for SVM classifier and requires minimum training and testing time, and

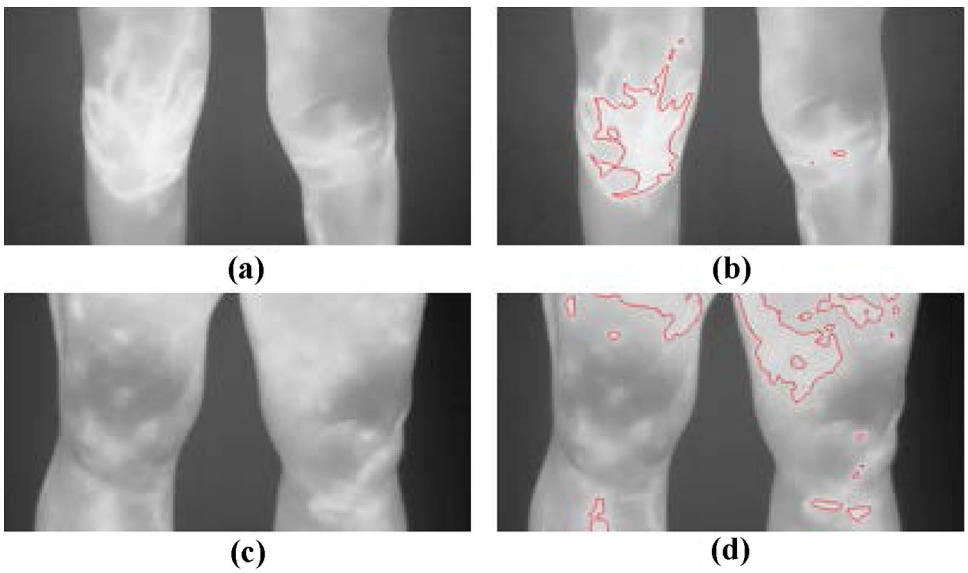

Fig. 3 Examples of inflamed region oriented ground truth of knee thermograms; **a** knee thermogram of arthritis patient; **b** ground truth of inflamed region of arthritis related thermogram (**a**) marked by red boundary; **c** knee thermogram related to knee pain; **d** ground truth of inflamed region of knee pain related thermogram (**c**) marled by red boundary

Table 1 Shape features extracted from ROI

Feature	Description
Area	Total number of pixels inside the extracted ROI
Euler number	Relationship in between the number of contiguous parts and number of holes present in an inflamed ROI
Perimeter	Total number of points present in the in the boundary of all the disjoint inflamed region
Convexity	Average of the ratio in between smallest convex polygon of each connected inflamed ROI and contour of that inflamed ROI
Eccentricity	Summation of the ratio in between the major axis and minor axis of each connected inflamed region
Orientation	Summation of the angle between the major axis and horizontal axis for each connected inflamed region
Solidity	Solidity defines whether the inflamed region is convex or concave by dividing the inflamed area with convex hull. Solidity of Ideal convex shaped inflamed region is 1. Summation of solidity of each connected inflamed region in a thermogram represents the overall solidity

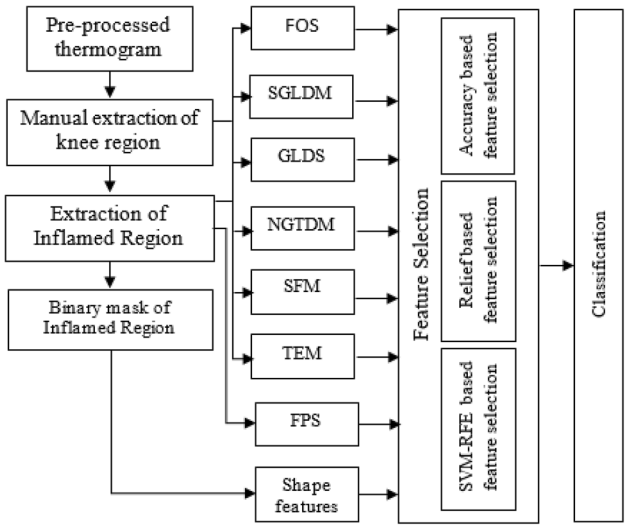

Fig. 4 Flow chart of 2nd stage classification (RA classification) using *FOS* first order statistical features, *SGLDM* special gray level dependence matrix feature, *GLDS* gray level difference statistics feature, *NGTDM* neighbourhood gray tone difference matrix feature, *SFM* statistical feature matrix, *TEM* texture energy measure, *FPS* fourier power spectrum features, and shape features

still generates good accuracy of classification compare to other kernels. The accuracy of classification is calculated by taking the mean of threefold cross validation accuracy through SVM classifier. In the next stage of classification, the rheumatoid arthritis (RA) affected knee thermograms are detected from the group of arthritis affected knees as described in the following section.

2nd stage classification: RA vs. non RA

In our analysis, classification in between RA and other types of arthritis using knee thermograms is performed through the analysis of extracted texture, frequency, and shape level features of inflamed regions and texture features of holistic knee regions. The above mentioned features extracted from arthritis affected knee thermograms are used as the input of 2nd stage classifier for detection of RA affected knee. The Fig. 4 shows the flowchart of 2nd stage classification for RA detection and detailed in next subsections.

Feature extraction

After preprocessing and inflamed region of interest segmentation from arthritis affected knee thermograms, three types of features are extracted, namely, texture feature, frequency level feature and shape feature from inflamed ROI and holistic knee thermograms. The preprocessing step of thermogram is detailed in "Thermogram pre-processing" and procedure for ROI extraction is already detailed in "Inflamed region of interest (ROI) segmentation".

Texture feature extraction Analysis of texture feature on arthritis-related knee thermogram and inflamed ROI can be used to identify RA. Figure 5 shows the histogram of knee thermogram and ROI for the preliminary assumption of a textural difference between RA and other arthritis. Visually histograms of RA and other kind of arthritis consists of differences related to intensity range as shown in Fig. 5b, f. The maximum intensity of RA thermogram reaches near to highest gray level pixel intensity as shown in Fig. 5b compare to other types of arthritis (shown in Fig. 5f). Histogram of the intensity distribution of inflamed ROI extracted from RA and other arthritis is also analyzed and shown in Fig. 5d, h, respectively. Visually the histogram of ROI has a minimum difference in between RA and other arthritis. As histogram wise thermogram of RA and other arthritis contains visual differences, therefore textural features are extracted from thermograms and ROI for RA classification. In our analysis, 55 texture features are extracted from arthritis-related thermograms and their inflamed ROIs by grouping them into 6 categories as detailed in Table 2.

Frequency level feature extraction Coarse and fine texture analysis of inflamed ROI for RA classification is performed based on Fourier power spectrum analysis of the inflamed regions. In case of coarse texture, the extracted radial sum

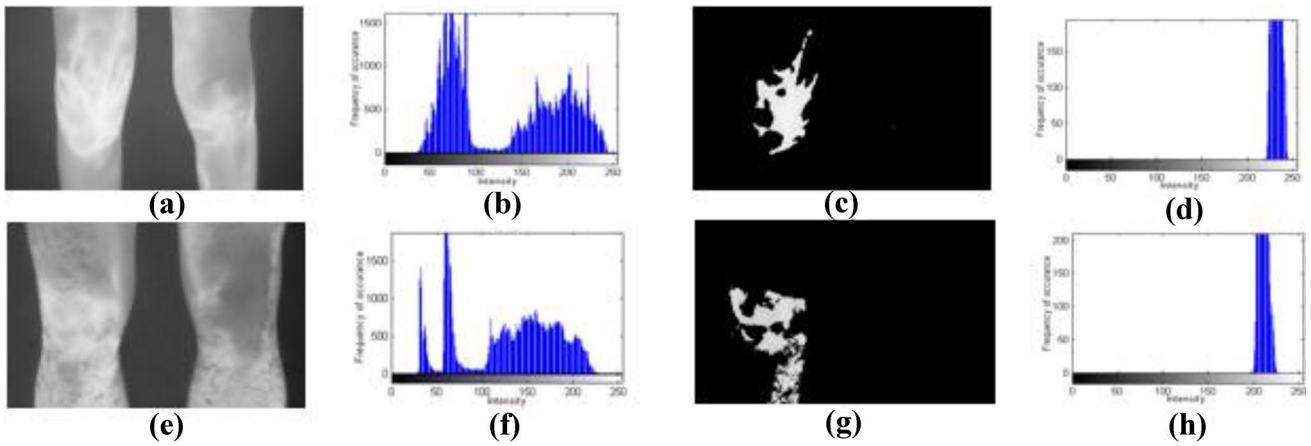

Fig. 5 **a** RA affected knee; **b** histogram of RA affected knee (**a**); **c** inflamed ROI extracted from RA affected knee; **d** histogram of ROI shown in **c; e** arthritis affected knee (not RA); **f** histogram of **e; g** inflamed ROI extracted from **e; h** histogram of **g**

Table 2 Features extracted for 2nd stage classification (RA classification)

Features ID	Features group	Abbreviation	No. of features	Features name	Applied on
Texture level features					
F1–F9	First order statistical features [31]	FOS	9	Mean, variance, median, mode, skewness, kurtosis, energy, entropy, standard deviation	Knee thermogram, inflamed ROI
F10–F35	Spatial gray level dependence matrices [32]	SGLDM[a]	26	Mean and difference based extracted features energy, contrast, correlation, variance, homogeneity, sum average, sun of variance, sum entropy, entropy, difference variance, difference entropy, information measures of correlation (from x axis and y axis individually)	Knee thermogram, inflamed ROI
F36–F40	Gray level difference statistics [33]	GLDS	5	Homogeneity, contrast, energy, entropy, mean	Knee thermogram, inflamed ROI
F41–F45	Neighborhood gray-tone difference matrix [34]	NGTDM	5	Coarseness, contrast, busyness, complexity, texture strength	Knee thermogram, inflamed ROI
F46–F49	Statistical feature matrix [35]	SFM	4	Coarseness, contrast, periodicity, regularity/roughness	Knee thermogram, inflamed ROI
F50–F55	Texture energy measure [36, 37]	TEM	6	Texture energy using LL, EE, SS, LE,ES, LS kernel (5 × 5)	Knee thermogram, inflamed ROI
Frequency level features					
F56–F57	Fourier power spectrum [37]	FPS	2	Radial sum, angular sum	Inflamed ROI
ROI shape features					
F58–F64	Shape features [29]	SF	7	Area, Euler number, perimeter, convexity, eccentricity, orientation, solidity	Binary mask of inflamed ROI

[a]In SGLDM, all the mentioned 13 features are calculated in 0°, 45°, 90° and 135° angle. In the first mean group, mean of each feature in all four angle are calculated. So, the mean group of SGLDM generates 13 features. In the difference group, difference between maximum and minimum value of a feature in its 4 angles are calculated. So the difference group of SGLDM also generates 13 features. The SGLDM contains total 26 features by combining the mean and difference group

from Fourier transformed ROI is higher compared to fine texture. Similarly, the angular sum will be higher for the ROI texture containing many lines or edges in a particular direction.

Shape feature extraction from inflamed ROI Extraction of shape features is performed on the binary mask of inflamed ROIs. These features extraction are implemented by following the shape feature extraction procedure as detailed in 1st stage classification given in "Shape feature extraction". As generated inflammation in case of all knee arthritis is similar in nature and also present in the patellar region, therefore shape feature may not be able to classify RA from a group of arthritis affected knee thermograms.

Feature selection for 2nd stage classification (RA classification)

In the 2nd stage of classification, extraction of features has been performed from knee thermograms and inflamed ROIs for classification of RA.Total 119 features are extracted from each knee thermogram, where 55 features are extracted from holistic knee thermogram, 57 features are extracted from inflamed regions of knee thermogram and seven features are extracted from the binary mask of inflamed regions. Among those extracted features of size D, some features are positively correlated with each other. The optimal subset of features, d (d < D) may increase the accuracy of classification [40]. In our analysis, we use the support vector machines-recursive feature elimination (SVM-RFE) [38], and RELIEF [39] methods for selection of important features as input of classifier. The SVM-RFE is a rank based iterative method of feature reduction wherein each iteration the rank of the feature is determined depending on the effectiveness of that feature for classification of data. The selected features using SVM-RFE is fed as the input of SVM classifier for detection of RA. In RELIEF, the weight of each feature is iteratively determined according to the ability to discriminate in between groups depending on neighbor pattern and distance margins. But the choice of rank threshold and weight threshold for selection of relevant features is the main limitation of SVM-RFE and RELIEF respectively. Therefore, we performed group wise higher accuracy based feature selection for proper detection of RA in our analysis. The significance of features selection and procedure of accuracy based feature selection is detailed in "Feature selection for RA classification".

RA vs. non-RA classification

Classification of RA from a group of arthritis affected knee thermograms is performed after extracting features and selecting optimal features from those. The SVM with linear kernel classifier is used for RA classification. In the second stage of classification also, the accuracy of the classifier is determined by taking the average of threefold classification rate.

Thermogram acquisition, dataset description, data pre-processing and ground truth generation

Classification of RA has been implemented on our arthritis affected captured knee thermogram dataset. The acquisition of the knee data was performed at Physical Medicine and Rehabilitation (PMR) Department, Agartala Government Medical College (AGMC) of Govind Ballav Pant (GBP) Hospital, Agartala, Tripura (W), by following the standardized method of thermogram acquisition under the supervision of a medical expert. All the captured thermograms consist of unwanted information like company logo, temperature scale, background region, temperature flow related information in whole leg etc. That information increase the processing time of analysis and may increase the error rate of classification. Therefore, all the thermograms are gone through the pre-processing before segmentation and feature extraction. Also in this study, for feature-based classification of RA, the inflamed region is extracted from knee thermograms. Accurate extraction of the inflamed region increases the accuracy of the inflammatory area oriented shape based feature effect. Validation of the accurate inflamed region segmentation has been performed depending on the ground truth data related to the inflamed area of each thermogram prepared by medical experts. In the rest of the section we detailed the dataset along with acquisition details, thermogram preprocessing as per requirement and procedure of ground truth generation.

Thermogram acquisition and dataset description

Standardized protocols for acquisition of knee thermogram mainly depends on environmental condition, camera specification, camera positioning and patient preparation [3]. The most influencing factors related to environmental conditions are room temperature, room lightning condition and humidity. Depending on the review of medical thermogram acquisition [43], the room temperature is fixed to 22 °C and the room has been kept dark during acquisition to avoid unnecessary artifacts from light source. The humidity is not considered during the acquisition as the variation is low due to constant air conditioning effect of the room. The FLIR T-650sc thermogram system is used for capturing the data with 640 × 480 resolution and 0.02 °C sensitivity

at 30 °C. The distance between the thermal camera and knee region has been kept within 2-m distance for anterior view of knee. During patient preparation, 15 min rest in controlled environment has been followed for each subject which stabilizes of temperature distribution in human body. During the rest and acquisition time, no clothing, lotion, contact of other body parts or material are allowed in the knee region. Standardized methods of thermogram acquisition decreases poor precision of measurements during feature extraction.

Following the standard protocols of thermogram acquisition, the created dataset contains total 110 knee thermograms where 60 thermograms are from arthritis group and rest are from healthy or knee pain affected group. The age of most of the subjects from both the groups are within the range of 40–65 years. Among those 60, 20 thermograms belong to rheumatoid arthritis (RA), 20 belong to osteoarthritis (OA), and rests are of reactive arthritis (ReA) group. The knee thermograms used for analysis are stored following the gray palette form of color representation as shown in Fig. 6a. In this palette, the relation between temperature and pixel intensity is proportional to each other. The 24-bit gray palette consists of the red (R), green (G) and blue (B) channels. The values of R, G and B channels for any pixel is equal to each other and this characteristic is responsible for visually gray representation of thermogram with 24-bit. Transformation of a thermogram from 24-bit to 8-bit gray scale is performed by taking the R or G or B channel values for every pixel. This transformation simplifies the analysis without losing the properties of original thermogram as the 8-bit representation is identical with any of the R, G, B channels of the image. In our study, we take the R channel pixel intensity value of each pixel.

Thermogram pre-processing

In our captured dataset, the temperature range of thermograms varies from each other depending on the body core temperature of patient and disease severity. The maximum temperature in a thermogram is represented by higher gray level intensity and minimum is by lower. So, temperature representation of each gray shade pixel intensity varies for each of the thermograms. For normalization of the temperature range of all the thermograms and their corresponding grayscale pixel intensity, the range of each thermogram is fixed into 20 (minimum temperature) to 40 (maximum temperature) degree centigrade. After normalization, manual cropping of knee region from whole leg thermogram is performed to decrease the complexity of unwanted region analysis. Figure 6b, c shows an example of the stepwise output of thermogram preprocessing.

Ground truth generation

The region of interest based classification of knee thermogram depending on the structure of inflammation

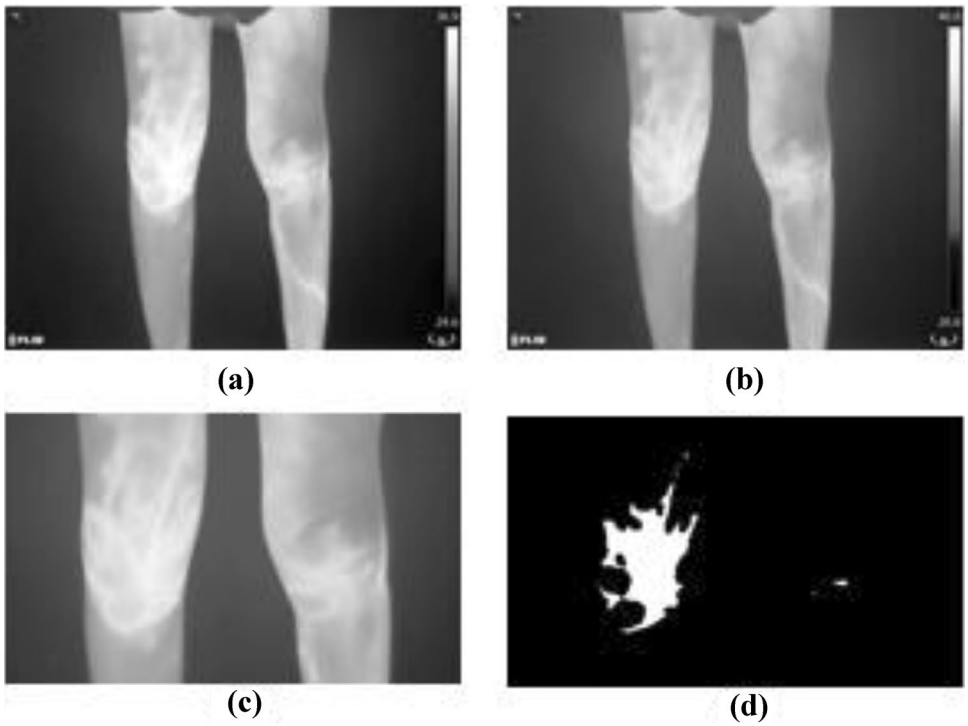

Fig. 6 **a** Original thermogram; **b** temperature adjusted thermogram, **c** manually cropped knee thermogram; **d** ground truth representing inflamed ROI

spreading is highly influenced by the accurate extraction of the inflamed areas from the pre-processing output. The accuracy of the ROI segmentation is determined by the ground truth data related to inflammation spreading generated by medical expert. For our dataset, ground truth generation has been performed in two steps. In the first step, five technicians individually marked the inflamed region of each knee thermogram. The final ground truth is generated from the five sets created by technicians using voting policy for each pixel. In the voting policy, one pixel is marked as part of inflamed ROI, if at least three technicians consider that pixel under inflamed ROI. In the second step, the output of voting policy based final ground truth is further evaluated by medical expert. If the medical expert rejects the output as appropriate inflamed region than marking process by technicians will be performed again under expert's supervision for generation of accurate ground truth through voting policy. Figure 6d shows the binary mask of the corresponding final ground truth region of Fig. 6c.

Experimental results

In our analysis, classification of RA from knee thermograms mainly depends on segmentation of inflamed region, feature extraction and feature selection. This section details the outcome of each factors depending on accuracy and includes comparative study.

Performance analysis of ROI segmentation

Accurate segmentation of inflamed region helps in shape analysis of inflammation spreading. In our analysis, first phase of thermogram classification in between arthritis affected knee and other knee is performed depending on the shape of inflammation spreading. The modified multi-seeded region growing is applied in our captured knee thermogram dataset for extracting inflamed ROI as detailed in "Inflamed region of interest (ROI) segmentation". Popularly used existing methods of inflamed region extraction from thermograms are K-means clustering (K-means) [19–21], Fuzzy c means clustering (FCM) [9, 18, 19], Otsu threshold based clustering (Otsu) [24–26], and single seeded region growing [27, 28]. For comparative study, selection of the number of clusters of all the existing clustering methods for maximum accuracy are chosen through trial and error method. The optimal accuracy is achieved with cluster number 7 for K-means clustering, and 10 for Fuzzy c means clustering with fuzzy weighting exponent value 1.5. In case of Otsu threshold based clustering, the optimal segmentation accuracy is found with threshold level 9 and the threshold value for single seeded region growing is fixed to 0.01 for accurate segmentation of inflamed ROI. In our modified multi-seeded region growing method also we use 0.01 as the threshold value. Except region growing, the cluster representing the maximum intensity value of the input image is considered as the final inflamed region oriented output for comparative study. The Fig. 7 shows an example of the segmentation outcome of existing and proposed methods along with ground truth data in binary format. Table 3 represents the intensity matching oriented quantitative analysis base comparative study of segmentation outcome of the proposed and existing methods. The intensity matching based analysis is performed depending on true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values generated through comparing the ground truth (GT) and segmentation outcome (SO). TP is the total number of pixels that were considered both by GT and SO under inflamed region and represented with pixel value 1. On the other hand, TN indicates the pixels which were mutually considered as non-inflamed region in both the GT and SO, and represented with pixel value 0. Therefore, $TP = GT \cap SO$ (considering pixel value 1) and $TN = \overline{(GT \cup SO)}$ [42]. FP stands for those pixels which are under non-inflamed region in GT (with intensity value 0), but considered as inflamed region in SO (with intensity value 1). FN is the reverse condition of FP, where the inflamed region in the GT are considered as non- inflamed region in the SO. The FP and FN can be represented as, $FP = SO \cap \overline{GT}$ and $FN = GT \cap \overline{SO}$ [42]. Depending on total number of TP, TN, FP and FN, five measuring factors are calculated for comparing the area of GT and SO. The factors are sensitivity (SENS), specificity (SPEC), accuracy (ACC), similarity (SIM) and border error (BE). Except border error, higher the value of all the factors indicates better segmentation. The Table 3 detailed the formulations [42] and measuring factor based segmentation outcomes in our knee thermogram dataset. Observation of Fig. 7 and Table 3 indicates that the modified multi-seeded region growing method provides better segmentation outcome compared to the existing method of inflamed region segmentation from thermograms.

Experimental result on classification

In our analysis, the first stage of the classification categorizes arthritis affected knee and other knee thermograms using SVM classifier. Shape feature based first stage classification achieves 91% correct detection rate with sensitivity 0.93 and specificity 0.9. For RA classification, the extracted features from arthritis affected knees are fed to SVM classifier. Inflamed ROI of arthritis affected thermogram are used for texture, frequency and shape level feature extraction. Also, texture features are extracted from holistic knee

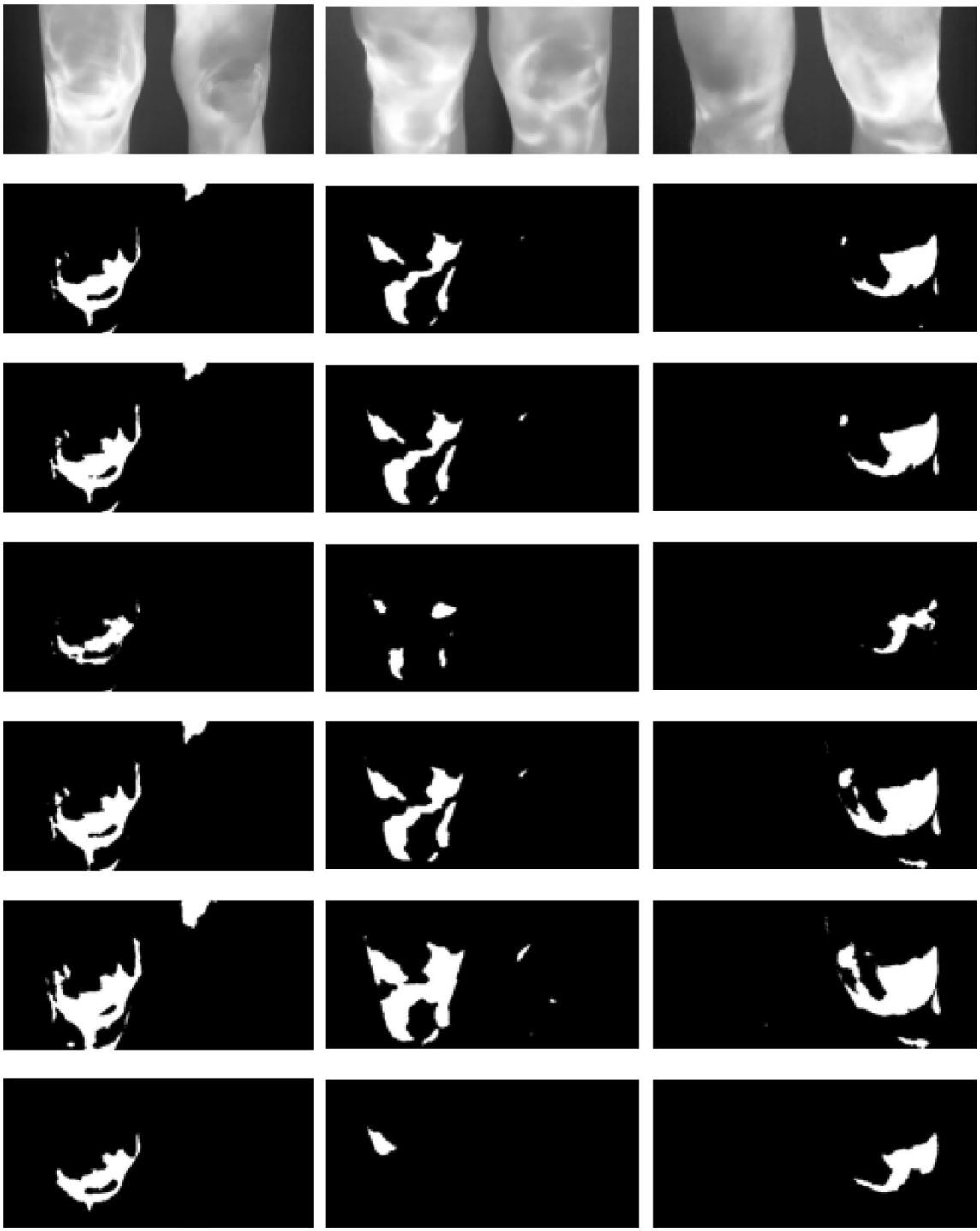

Fig. 7 Thermogram segmentation outputs: Row 1: input thermogram; Row 2: ground truth; Row 3: modified multi-seeded region growing; Row 4: K-means; Row 5: FCM; Row 6: Otsu thresholding ; Row 7: single seeded region growing

thermogram of arthritis affected patients. For effectiveness of individual feature group, we first analyzed the group-wise accuracy of RA classification. The texture feature group contains six intra-groups as given in Table 2. The Fig. 8 shows the accuracy based outcome of the analysis. Observation on Fig. 8 are as follows:

Table 3 Performance measure of segmentation techniques

Measuring factors	Formula	Used methods for thermogram inflamed region segmentation				
		K-means	FCM	Otsu	Single seeded region growing	Modified multi seeded region growing
Sensitivity [42]	$\frac{TP}{TP+FN}$	0.352	0.994	0.997	0.395	0.978
Specificity [42]	$\frac{TN}{TN+FP}$	0.999	0.973	0.949	0.998	0.986
Accuracy (%)	$\frac{TP+TN}{TP+TN+FP+FN} \times 100$	97	97.4	95	97.3	98.6
Similarity [42]	$\frac{2TP}{2TP+FN+FP}$	0.467	0.731	0.607	0.494	0.83
Border error [42]	$\frac{FP+FN}{TP+FN}$	0.753	1.756	2.701	0.77	0.508

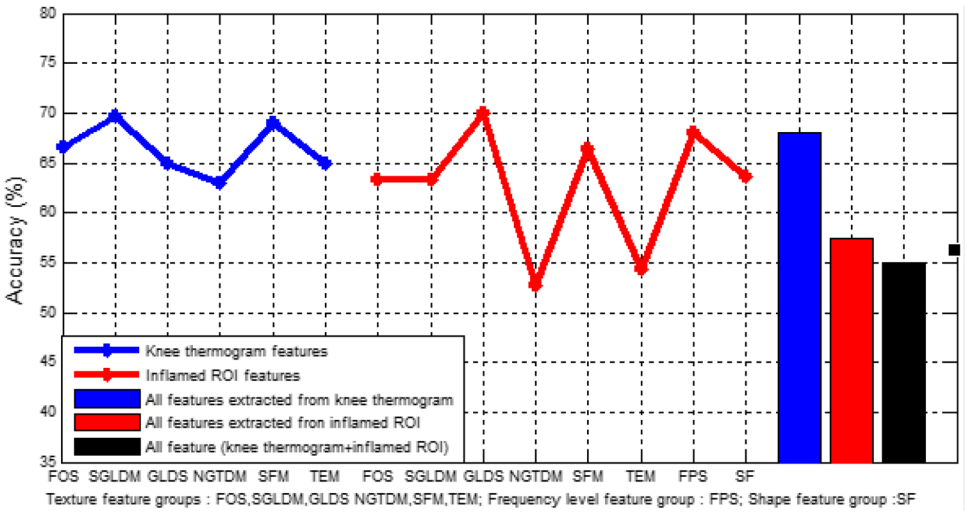

Fig. 8 Feature group base accuracy of RA classification using SVM classifier

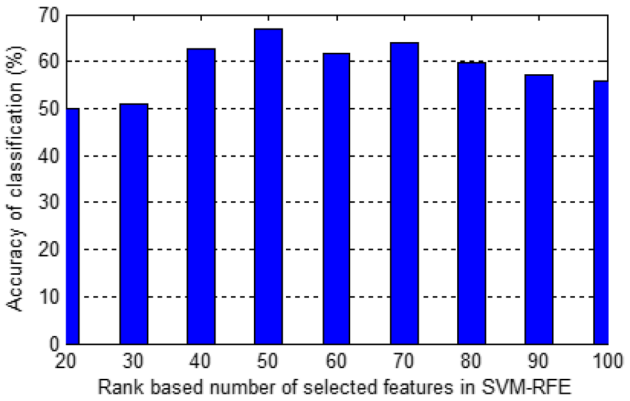

Fig. 9 Rank based accuracy of classification using SVM-RFE

(a) depending on accuracy, extracted knee thermogram features are more important compared to inflamed ROI features for RA classification,
(b) combination of all extracted features significantly decreases the rate of classification. So feature selection method is essential for discriminating feature extraction from all the extracted features,
(c) individually, extracted gray level difference statistics (GLDS) features from inflamed ROI generates the best accuracy (70%) of RA classification.

Feature selection for RA classification

The accuracy of classification depends on features used as input in the classifier, and selection of appropriate features is a complex task to perform. In fact, accuracy of classification for the same dataset may differ depending on classifier and features. As mentioned in "RA vs. non-RA classification", SVM is considered as the classification method for RA diagnosis. Observations of the previous analysis (given in Fig. 8 and "Experimental result on classification") shows that, use of all extracted features in the classifier decreases the rate of accurate RA detection. Therefore, we perform the selection of features using SVM-RFE and RELIEF as mentioned in "Feature selection for 2nd stage classification (RA classification)". In case of SVM-RFE based feature selection, the linear kernel is used. The weight of each feature in SVM-RFE is negatively correlated with rank value and rank of each feature for our analysis is shown in Fig. 10a. Therefore, feature with minimum rank contains maximum weight value for discriminating the groups. In trial and error

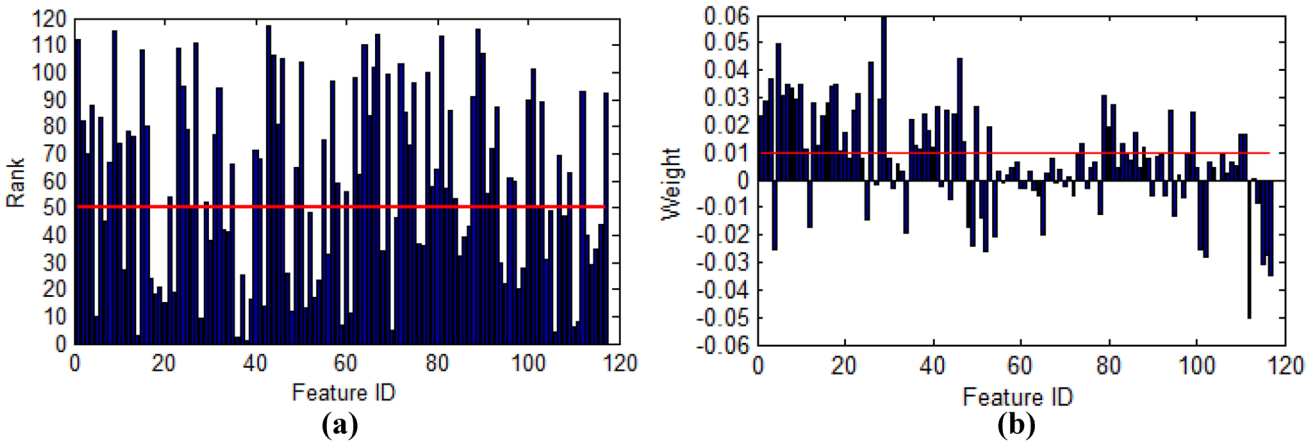

Fig. 10 **a** Rank of features using SVM-RFE; **b** weight of features using RELIEF

for RA classification using SVM-RFE based feature selection, we varied the upper limit of rank threshold from 20 to 100 with an interval of 10 starting from 1. Analysis shows that, rank threshold 50 generates intra highest accuracy of 67% using SVM-RFE by selecting 50 different features as shown in Fig. 9 (Rank limit 50 marked by the horizontal line in Fig. 10a). Depending on the number of features selected using SVM-RFE, the weight threshold of RELIEF method has been fixed to 0.01. 47 features are selected with weight threshold greater than 0.01 whereas use of weight threshold greater than 0.02, 0.015 and 0.005 selects 30, 37 and 64 features respectively. As the weight threshold 0.01 for RELIEF selects maximum nearest numbers of features i.e. 47 compare to SVM-RFE, so 0.01 is considered as the final weight for RELIEF based feature selection as marked in Fig. 10b with horizontal line. Final selected features (47 features with weight 0.01) using RELIEF generates 71% correct classification of RA. The Table 4 mentioned the selected features using SVM-RFE and RELIEF. The SVM-RFE and RELIEF contain the limitation of parameter selection like selection of rank threshold and weight respectively.

Therefore in our experiment, we performed accuracy based selection of feature groups in descending order until maximum nearest number of features are selected compare to SVM-RFE (50 features). The Fig. 8 shows the accuracy of each group of feature. Following the Fig. 8, the GLDS feature group extracted from inflamed ROI is considered which contains five number of features as it gives maximum group based accuracy. To select similar number of features compare to SVM-RFE, the second highest accuracy generated SGLDM (holistic knee thermogram based features) feature group has been selected. Combination of these two groups contains total 31 number of features. In this way, SFM (holistic knee thermogram based features), FPS (inflamed ROI based features), FOS (holistic knee thermogram based features), and SFM (inflamed ROI based features) groups are also selected. Collection of these six groups contains 50 number of features which is exactly equal to SVM-RFE. The accuracy based selected feature set generates 73% correct classification of RA. Observation from Fig. 8 shows that the each of the selected six groups individually generates above 65% and maximum 70% accuracy of classification. Our accuracy based selection of feature group increases the rate of RA classification i.e. 73% compared to existing feature selection methods as shown in Table 4.

Discussion

The merits of thermogram oriented joint inflammation detection in diagnosis of arthritis widely depends on accurate segmentation of inflamed regions, features extraction along with selection and classification as mentioned in the previous sections. Depending on the review works, the primary objective of the study is to classify RA thermograms considering the factors like segmentation, feature selections and classification as discussed below.

Inflamed region segmentation

The segmentation of inflamed region is an important step in inflammation analysis and feature extraction. Although segmentation of inflamed regions from thermograms present in previous research, but contains some short comings. In past, authors performed Otsu threshold based clustering [22–24], fuzzy c means clustering [9, 20, 21], and K-means clustering [21–23] for arthritis related thermogram segmentation. The problem associated with clustering method for segmentation are cluster number selection, cluster center selection and selection of the cluster which represents the inflamed region of interest. The region growing [25, 26] is also applied on joint thermograms for segmenting inflamed

Table 4 Selection of features for 2nd stage classification (RA classification)

Feature selection methods	Applied on	Feature group	Original number of feature	Number of selected features[a]	Selected features	Accuracy
SVM-RFE (number of total selected features: 50)	Knee thermogram	FOS	9	2	Skewness, energy	67%
		SGLDM	26	11	Mean based: contrast, homogenity, sum of entropy, entropy, difference variance, difference entropy, information measures of correlation in y axis. difference based: sum of average, sum of entropy, difference entropy, and information measure of correlation in x axis	
		GLDS	5	4	Homogenity, contrast, energy, entropy	
		NGTDM	5	2	Contrast, busyness	
		SFM	4	2	Contrast, periodicity	
		TEM	6	4	Texture energy using EE, SS, LE, ES kernel (5×5)	
	Inflamed ROI	FOS	9	2	Variance, kurtosis	
		SGLDM	26	8	Mean based: variance, sum of average, sum of variance, information measures of correlation in x and y axis. Difference based: sum of entropy, entropy, difference variance	
		GLDS	5	2	entropy, mean	
		NGTDM	5	2	Busyness, complexity	
		SFM	4	2	Contrast, roughness	
		TEM	6	4	Texture energy using LL, EE, LE, LS kernel (5×5)	
		FPS	2	1	Radial sum	
	ROI mask	SF	7	4	Area, Euler no., perimeter, convexity	
RELIEF (number of total selected features: 47)	Knee thermogram	FOS	9	6	Mean, variance, skewness, energy, entropy, standard deviation	71%
		SGLDM	26	17	Mean based: energy, contrast, correlation, variance, sum of average, sum of variance, sum of entropy, entropy, difference variance, difference entropy, information measures of correlation in x and y axis. Difference based: contrast, correlation, sum of average, sum of entropy, entropy	
		GLDS	5	4	Contrast, energy, entropy, mean	
		NGTDM	5	4	Coarseness, contrast, busyness, texture strength	
		SFM	4	3	Contrast, periodicity, regularity/roughness	
		TEM	6	2	Texture energy using SS, LS kernel (5×5)	
	Inflamed ROI	FOS	9	0	Nil	
		SGLDM	26	7	Mean based: information measures of correlation in x axis Difference based: variance, homogenity Sum of average, sum of entropy, difference of entropy, information measures of correlation in y axis based	

Table 4 (continued)

Feature selection methods	Applied on	Feature group	Original number of feature	Number of selected features[a]	Selected features	Accuracy
		GLDS	5	0	Nil	
		NGTDM	5	1	Coarseness	
		SFM	4	1	Coarseness	
		TEM	6	0	Nil	
		FPS	2	1	Angular sum	
	ROI Mask	SF	7	1	Area	
Accuracy based feature group selection (number of selected features: 50)	Knee thermogram	FOS	9	9	Mean, variance, median, mode, skewness, kurtosis, energy, entropy, standard deviation	73%
		SGLDM	26	26	Mean and difference based: energy, contrast, correlation, variance, homogenity, sum of average, sun of variance, sum entropy, entropy, difference variance, difference entropy, information measures of correlation (x axis, y axis based)	
		GLDS	5	0	Nil	
		NGTDM	5	0	Nil	
		SFM	4	4	Coarseness, contrast, periodicity, regularity/roughness	
		TEM	6	0	Nil	
	Inflamed ROI	FOS	9	0	Nil	
		SGLDM	26	0	Nil	
		GLDS	5	5	Homogenity, contrast, energy, entropy, mean	
		NGTDM	5	0	Nil	
		SFM	4	4	Coarseness, contrast, periodicity, regularity/roughness	
		TEM	6	0	Nil	
		FPS	2	2	Radial sum, angular sum	
	ROI Mask	SF	7	0	Nil	

[a]Feature selection methods are applied on total extracted 119 features from knee thermograms and ROIs as mentioned in Table 2

region of interest in past. The drawback of region growing associated with selection of single seed and threshold value depending on the highest intensity pixel for extracting similar pixels considering minimum threshold difference. As our thermograms, contains several disjoint inflamed region of interest, so multiple seed selection is required. In our analysis, we incorporate threshold based multiple seed selection with existing region growing to increase the accuracy of segmentation. Comparative study of segmentation as given in Table 3 and Fig. 7 shows that, modified multi-seeded region growing provides higher accuracy of segmentation compare to K-means, FCM, Otsu and existing region growing methods. Analysis of Table 3 also shows that, the modified multi-seeded region growing generates maximum amount of sensitivity, specificity and similarity compare to existing segmentation techniques with minimum amount of border error.

The accurate segmentation of inflamed ROI using modified multi-seeded region growing is based on the selection of threshold value and seed point in each iteration. The threshold value and seed selection depends on T and ϕ respectively as mentioned in "Inflamed region of interest (ROI) segmentation" and in Algorithm 1. The following subsections mentioned the procedure of T and ϕ value selection followed by computational complexity based analysis of modified multi-seeded region growing method.

T and ϕ value selection for modified multi-seeded region growing

In modified multi-seeded region growing based inflamed region segmentation, the threshold value (T) is responsible for considering pixels under inflamed ROI comparing with seed point. In each iteration, the eight neighborhood pixels of a seed point are compared with seed pixel and considered under inflamed ROI. As the inflamed region consists of variation in pixel intensity, so minor difference in between seed point and its neighborhood pixels may exists. The neighborhood of the seed pixel is considered under inflamed region of interest if the absolute difference between them is less than T. In this study, the input image is normalized and ranged from 0 to 1. The value of T need to greater than 0, because the absolute difference in between seed pixel and neighborhood is always greater than or equal to 0. With $T = 0$, the condition in step 4.1 of Algorithm 1 will never satisfied. So, T is fixed with a minimum value i.e., 0.01 for satisfying the condition in step 4.1 and to consider the neighborhood pixels of seed pixel with minimum variance as region of interest.

The proposed modified multi-seeded region growing (detailed in Algorithm 1) is an iterative method for extraction of inflamed region of interest from thermogram. The method mainly comprises of two steps: seed selection and seed pixel based comparison of neighborhood pixels of seed to consider those under inflamed region. In each iteration both the steps are performed for optimal inflamed region extraction. With the increase of iteration, all the pixels will be considered under inflamed region. So, the iteration stopping criteria is introduced depending on ϕ. The variableϕ decides whether the maximum intensity pixel in each iteration can be considered as seed pixel or not. In step 3 of Algorithm 1, if the value of maximum intensity pixel (M) is less then ϕ, M will not considered as seed pixel and the variable COND will set to 1 and also progress of iteration will terminate. Selection of the value of ϕ is performed here depending on the histogram peaks of the input thermograms. The inflamed region in the input thermogram is our region of interest for segmentation and it is represented with higher gray level pixel intensity in the image. Each input image consists of multiple peaks related to intensity distribution in the histogram. Therefore, from the histogram of input thermogram, we find out eight optimal peaks as shown in Fig. 11a [41]. As we are aim to segment the inflamed region, so the peak with highest gray level (h) of each thermogram is considered for ϕ value generation. The Fig. 11b plots the h value of all the input thermograms and mean of all h is considered for generation of ϕ here. The mean of h for our dataset is 238 and in step 1 of Algorithm 1, the value is normalized by dividing with 255. The final ϕ for our dataset is, $\phi = mean(h)/255$, i.e., $\phi = 238/255$ as given in step 1 of Algorithm 1.

Computational complexity based analysis of modified multi seeded region growing

The computational complexity of the Modified Multi seeded Region Growing depends on total number of iterations required to reach terminating condition and size of the image. Computation time for searching of the pixel with maximum intensity using Algorithm 1.1 depends on size of the image. Considering the size of the image as N and required number of iteration as T,the computational complexity modified multi seeded region growing method is $O(NT)$. The computational complexity of existing single seeded region growing is $O(N)$. Modified multi seeded region growing provides more accuracy of segmentation compared to existing single seeded region growing (shown in Table 3) at the cost of computational time. The computational complexity of existing clustering techniques, i.e., K-means, Fuzzy C-means, and Otsu method are $O(NndT)$, $O(Nn^2dT)$, and $O(L^n)$ respectively [44, 45]. Here, n stands for cluster number, d indicates number of dimension and L contains the highest intensity level. The K-means, Fuzzy C-means and Otsu method generates clustering outcome. Therefore, in the post processing step, the cluster representing the inflamed region need to be extracted depending on input image highest intensity. The corresponding cluster representing input

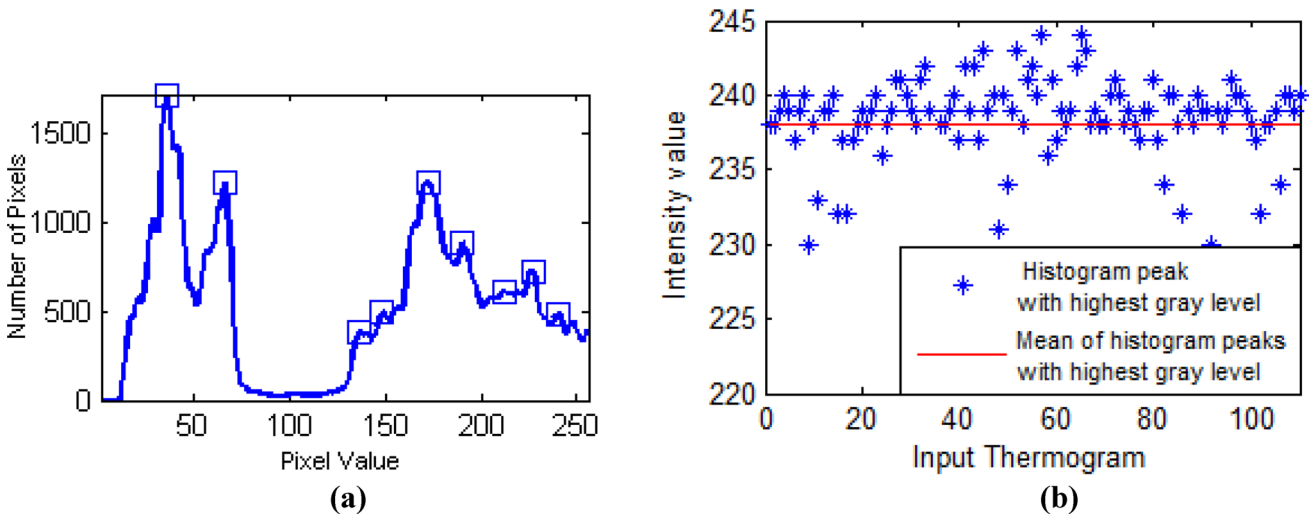

Fig. 11 **a** Histogram peaks of a thermogram marked with rectangular box; **b** histogram peak with highest gray level intensity value of each thermogram

image highest intensity signifies the inflamed region of interest. The computational complexity of final cluster selection in post processing step is $O(N)$ for K-means, Fuzzy C-means and Otsu method and it combines with the method's own computational complexity. Therefore, computational complexity based analysis shows that, modified multi seeded region growing requires less time compare to existing clustering methods of thermogram segmentation.

Feature extraction and selection for classification

In our analysis, classification of RA has been performed depending on extracted features from thermograms. In the first phase of classification, arthritis and non-arthritis thermograms are classified using shape features of inflamed regions and achieved 91% accurate detection rate as mentioned in result section. In the second stage, RA has been classified from arthritis thermograms by extracting features from inflamed regions and holistic knee. Combination of all features decreases the rate of RA classification as shown in Fig. 8. Therefore, feature selection has been performed from the extracted features to increase the classification rate. SVM-RFE and RELIEF methods are used for selection of effective features. The feature selection through SVM-RFE and RELIEF is highly influenced by the selection of input parameter like rank and weight respectively. In case of SVM-RFE with rank up to 50 (staring from 1), number of selected features is 50 and generates 67% accurate detection rate. Analysis shown in "Feature selection for RA classification" implies that decrease of the rank value also decreases the number of features and accuracy for SVM-RFE (shown in Fig. 9). Also with the increase in rank value, more features are selected. But higher the

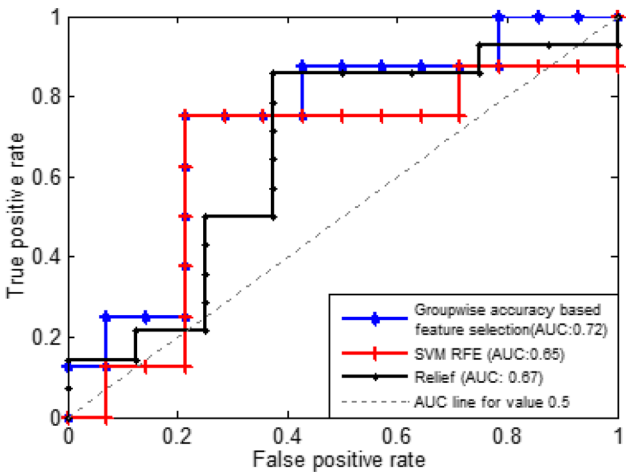

Fig. 12 Receiver operating characteristic (ROC) curve for comparing feature selection methods

number of features decreases the accuracy of classification. Therefore rank value 1–50 based selected features for SVM-RFE generates intra highest accuracy of RA classification. Depending on the number of features selected through SVM-RFE, the weight parameter of RELIEF method is fixed to 0.01 as described in "Feature selection for RA classification" and generates 71% correct detection rate of RA. In our analysis, we increases the correct classification rate of RA by performing accuracy based selection of feature group. From Table 4, it can be observe that thought the number of selected features for SVM-RFE and our accuracy based feature selection methods are similar (50 features), but the correct detection rate differs. With our feature selection method, the accuracy increased by 6% compared to

SVM-RFE because of the selected features using both the methods are not identical. In case of accuracy based feature selection, features from six different groups are selected whereas, for SVM-RFE, features are selected from 14 different groups which result the dissimilar accuracy of classification. Figure 12 shows the receiver operating characteristic (ROC) curve for the feature selection methods. From the correct detection rate of RA given in Table 4 and the ROC curve, we can see that accuracy and area under the curve (AUC) are higher for our accuracy based feature selection method compare to other two existing methods.

Conclusion

Depending on the spread of inflammation analysis from thermogram, prediction of abnormality related to RA in knee has been performed in this study. Segmentation of inflamed ROI is carried out on the thermograms for shape, texture and frequency level feature extraction from ROI regions. For the segmentation of inflamed region, we have integrated modified multi seeded region growing method. The suggested hybrid method of segmentation increases the accuracy compared to state of the art techniques. After segmentation, each ROI is represented by seven structural features for 1st stage of classification input. The classification differentiates in between arthritis and non-arthritis group of thermogram with 91% of correct detection rate in SVM classifier. In the 2nd stage of RA classification, 119 features are extracted from inflamed ROIs and holistic knee therograms. As the combination of all features decreases the accuracy, so feature selection methods are incorporated in RA classification. The RELIEF, SVM-RFE and our accuracy based feature selection provides 71%, 67% and 73% correct detection rate of RA respectively using knee thermograms. Here, our accuracy based feature selection shows better performance (73% accuracy) than the other two feature selection methods. In our future study, we will capture and collect the thermograms of RA affected other joints along with knee and investigate the classification rate by combining different arthritis affected joints.

Acknowledgements This work was supported by the Indian Council of Medical Research (ICMR), Government of India, under Grant Number. 5/7/1516/2016-RCH, Dated: 20/06/2017. The second author would like to thank Dr. S. B. Nath, Assistant Professor, Agartala Government Medical College (AGMC), Agartala for his kind support during the collection and inflammation oriented ground truth generation of knee thermogram dataset. Also authors would like to thank Prof. Dipti Prasad Mukherjee, India Statistical Institute, Kolkata for his valuable suggestions regarding feature selection during the classification stage of this work and Prof. Debotosh Bhattacharjee, Jadavpur University, Kolkata for his guidance during the revision stage of the manuscript.

Compliance with ethical standards

Conflict of interest The author declares no potential conflict of interest with respect to the authorship and/or publication of this article.

Ethical approval The study was approved by the Research cell and Ethics committee of Agartala Government Medical College, Tripura, India with 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. Silman AJ, Hochberg MC (2001) Epidemiology of the rheumatic diseases. 2nd edn. Oxford University Press, New York
2. Milind P, Sushila K (2012) How to live with rheumatoid arthritis? Int Res J Pharm 3(3):115–121
3. Bhowmik MK et al (2016) Pain related inflammation analysis using infrared images. SPIE Commercial + Scientific Sensing and Imaging. International Society for Optics and Photonics
4. Sacks JJ, Helmick CG, Langmaid G (2004) Deaths from arthritis and other rheumatic conditions, United States, 1979–1998. J Rheumatol 31:1823–1828
5. (2018) Arthritis by the numbers/book of trusted facts and figures (vol 2, 4100.17.10445). https://www.arthritis.org/Documents/Sections/About-Arthritis/arthritis-facts-stats-figures.pdf. Accessed 29 Oct 2018
6. Schäfer VS et al (2016) Arthritis of the knee joint in rheumatoid arthritis—evaluation of treatment response by ultrasound in daily clinical practice. Open Rheumatol J 10:81–87 **(PMC Web)**
7. Ropes MW (1959) Diagnostic criteria for rheumatoid arthritis. Anti Rheum Dis 18:49–53
8. Neelima AM (2012) Textbook of oral and maxillofacial surgery, Chapter-2 (art of diagnosis), 3rd edn. Jaypee Brothers Medical Publisher Ltd., Daryaganj, p 15
9. Snekhalatha U, Anburajan M, Teena T, Venkatraman B, Menaka M, Raj B (2012) Thermal image analysis and segmentation of hand in evaluation of rheumatoid arthritis. In: Computer communication and informatics (ICCCI), 2012 international conference, IEEE, pp 1–6
10. Borojević N, Kolarić D, Grazio S, Grubišić F, Antonini S, Nola IA, Herceg Ž (2011) Thermography of rheumatoid arthritis and osteoarthritis. In: ELMAR, 2011 proceedings, pp 293–295, IEEE
11. Frize M, Adéa C, Payeur P, Di Primio G, Karsh J, Ogungbemile A (2011) Detection of rheumatoid arthritis using infrared imaging. In: Medical imaging 2011: image processing, vol 7962. International Society for Optics and Photonics, p 79620M
12. Acharya UR, Ng EYK, Tan JH, Sree SV (2012) Thermography based breast cancer detection using texture features and support vector machine. J Med Syst 36(3):1503–1510
13. Francis SV, Sasikala M, Saranya S (2014) Detection of breast abnormality from thermograms using curvelet transform based feature extraction. J Med Syst 38(4):23
14. Milosevic M, Jankovic D, Peulic A (2014) Thermography based breast cancer detection using texture features and minimum variance quantization. EXCLI J 13:1204
15. EtehadTavakol M, Chandran V, Ng EYK, Kafieh R (2013) Breast cancer detection from thermal images using bispectral invariant features. Int J Thermal Sci 69:21–36

16. Wood AM, Brock TM, Heil K, Holmes R, Weusten A (2013) A review on the management of hip and knee osteoarthritis. Int J Chronic Dis. https://doi.org/10.1155/2013/845015
17. Informed Health Online [Internet] (2006) Rheumatoid arthritis: Overview. 2013 Oct 23 [Updated 2016 Aug 11]. Institute for Quality and Efficiency in Health Care (IQWiG), Cologne, Germany. https://www.ncbi.nlm.nih.gov/books/NBK384455/. Accessed 18 Jan 2018
18. "What is AiArthritis?" (n.d.) Retrieved 18 January 2018, from-https://www.aiarthritis.org/whatisaiarthritis
19. (2018) "What Is Arthritis?" http://www.arthritis.org. http://www.arthritis.org/about-arthritis/understanding-arthritis/what-is-arthritis.php. Accessed 18 Jan 2018
20. Araki S, Nomura H, Wakami N (1993) Segmentation of thermal images using the fuzzy c-means algorithm. In: Fuzzy systems, second IEEE international conference on, IEEE, pp 719–724
21. Golestani N, Tavakol ME, Ng EYK (2014) Level set method for segmentation of infrared breast thermograms. EXCLI J 13:241
22. 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
23. Shahari S, Wakankar A (2015) Color analysis of thermograms for breast cancer detection. In: Industrial instrumentation and control (ICIC), 2015 international conference on, IEEE, pp 1577–1581
24. Jadin MS, Taib S (2012) Infrared image enhancement and segmentation for extracting the thermal anomalies in electrical equipment. Electron Electr Eng 120(4):107–112
25. Font-Aragonés X, Faúndez-Zanuy M, Mekyska J (2013) Thermal hand image segmentation for biometric recognition. IEEE Aerosp Electron Syst Mag 28(6):4–14
26. Dutta T, Sil J, Chottopadhyay P (2016) Condition monitoring of electrical equipment using thermal image processing. In: 2016 IEEE first international conference on control, measurement and instrumentation (CMI), IEEE, pp 311–315
27. Selvarasu N, Vivek S, Nandhitha NM (2007) Performance evaluation of image processing algorithms for automatic detection and quantification of abnormality in medical thermograms. In: Conference on computational intelligence and multimedia applications, 2007. international conference on, IEEE, vol 3, pp 388–393
28. Nandhitha NM, Sheela Rani B, Kalyana Sundaram P, Venkataraman B, Raj B (2007) Performance evaluation of hot spot extraction and quantification algorithms for online weld monitoring from weld thermographs. 24th Int Symp Autom Robot Construct 3:461–466 **(ISBN: 0-7695-3050-8, 13 Dec 2007)**
29. George YM et al (2014) Remote computer-aided breast cancer detection and diagnosis system based on cytological images. IEEE Syst J 8(3):949–964. https://doi.org/10.1109/jsyst.2013.2279415
30. Frize M et al (2009) Preliminary results of severity of illness measures of rheumatoid arthritis using infrared imaging. In: 2009 IEEE international workshop on medical measurements and applications. https://doi.org/10.1109/memea.2009.5167981
31. Bhowmik MK et al (2017) Designing of ground truth annotated DBT-TU-JU breast thermogram database towards early abnormality prediction. IEEE J Biomed Health Inf. https://doi.org/10.1109/jbhi.2017.2740500
32. Haralick RM et al (1973) Textural features for image classification. IEEE Trans Syst Man Cybern SMC-3 6:610–621. https://doi.org/10.1109/tsmc.1973.4309314
33. Weszka JS et al (1976) A comparative study of texture measures for terrain classification. IEEE Trans Syst Man Cybern SMC-6 4:269–285. https://doi.org/10.1109/tsmc.1976.5408777
34. Amadasun M, King R (1989) Textural features corresponding to textural properties. IEEE Trans Syst Man Cybern 19(5):1264–1274. https://doi.org/10.1109/21.44046
35. Wu C-M, Chen Y-C (1992) Statistical feature matrix for texture analysis. CVGIP Graph Models Image Process 54(5):407–419. https://doi.org/10.1016/1049-9652(92)90025-s
36. Laws KI (1979) Texture energy measures. In: DARPA image understanding workshop. DARPA, Los Altos, CA, pp 47–51
37. Wu C-M et al (1992) Texture features for classification of ultrasonic liver images. IEEE Trans Med Imaging 11(2):141–152. https://doi.org/10.1109/42.141636
38. Guyon I, Weston J, Barnhill S, Bapnik V (2002) Gene selection for cancer classification using support vector machines. Mach Learn 46(1–3):389–422
39. Sun Y (2007) Iterative RELIEF for feature weighting: algorithms, theories, and applications. IEEE Trans Pattern Anal Mach Intell 29(6):1035–1051. https://doi.org/10.1109/tpami.2007.1093
40. Shafreen Banu K, Hari Ganesh S (2015) A study of feature selection approaches for classification. In: 2015 international conference on innovations in information, embedded and communication systems (ICIIECS), Coimbatore, pp 1–4
41. De Silva DVSX et al (2010) Adaptive sharpening of depth maps for 3D-TV. Electron Lett 46(23):1546–1548
42. Sadri A, Reza et al (2013) Segmentation of dermoscopy images using wavelet networks. IEEE Trans Biomed Eng 60(4):1134–1141
43. Bardhan S, Bhowmik MK, Nath S, Bhattacharjee D (2015) A review on inflammatory pain detection in human body through infrared image analysis. In: 2015 international symposium on advanced computing and communication (ISACC) 2015. https://doi.org/10.1109/isacc.2015.7377350
44. Ghosh S, Dubey SK (2013) Comparative analysis of K-means and fuzzy c-means algorithms. Int J Adv Comput Sci Appl 4:4
45. Sthitpattanapongsa P, Srinark T (2011) An equivalent 3D Otsu's thresholding method. Pacific-Rim symposium on image and video technology. Springer, Berlin