

SEGMENTATION OF KNEE THERMOGRAMS FOR DETECTING INFLAMMATION

Kakali Das, Mrinal Kanti Bhowmik, Dipti Prasad Mukherjee*

Department of Computer Science & Engineering, Tripura University, India

*Indian Statistical Institute, Kolkata, India

ABSTRACT

Rheumatologists determine treatment plan based on the inflammation of knee joints affected by arthritis. Extraction of the inflamed region or *hotspot* from the knee thermogram is the prerequisite for grading of inflammation and classification of different arthritis. In this paper, we propose an automatic method for extracting the inflamed region from the knee thermograms. We propose an ensemble technique to arrive at a consensus segmentation of the *hotspot* region. We have used variation of information based information theoretic approach to generate consensus segmentation. The fusion of multiple segmentation maps is achieved using local search based greedy iterated conditional modes algorithm to obtain final segmentation result. Experiments show that our proposal scores significantly better in detecting *hotspots* in more than 50 inflammatory knee thermograms.

Index Terms– Thermal Image, Consensus Segmentation, Variation of Information

1. INTRODUCTION

The treatment plan for a patient suffering from arthritis depends on inflammation at the knee [1]. Grade of inflammation indicates the severity of joint damage [2]. Clinicians try to reduce the inflammation to prevent further joint damage [3][4]. Inflammation of joints increase the temperature of the skin surface of the joint. Thermal medical imaging is a relatively inexpensive quick tool to diagnose increased temperature at inflamed region [5]. In a recent work, inflammation was monitored using thermal imaging [6]. Fig. 1 shows typical knee thermograms in pseudocolor (first row) and in gray scale(second row).

The prerequisite for detecting inflamed region at higher temperature or *hotspots* in the knee thermograms is a stable image segmentation technique. This identification of *hotspots* has major clinical importance to predict patient's prognosis. Expert guided manual segmentation of inflamed region is expensive and time consuming and suffers from observer variability. The automatic identification of *hotspots* should act as the major catalyst for wider acceptability of thermograms for arthritis management. To the best of our knowledge, the proposal in this paper is first of its kind suggesting automatic use of thermograms for arthritis management or similar such applications.

The prior works on the segmentation of thermal image includes *k*-means [7], FCM [8], Otsu thresholding [9] or region growing [10]. The performance of these methods is sensitive to prior knowledge of the number of clusters. Region growing requires preset threshold for region merging. In contrast, we propose an energy function based method for extraction of inflamed

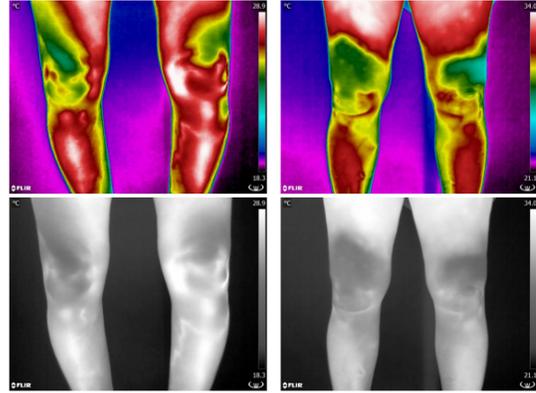


Fig. 1. Knee thermograms: (first row) in pseudocolor (second row) in gray palette.

region from thermal images. The energy function integrates information received from multiple segmentations. This consensus selection of segmentations from multiple sources relies on information theoretic measure of variation of information. Note that our method does not use any threshold or parameter to control segmentation. Local search based greedy optimization similar to iterated conditional modes is used for optimization of the energy function. The process pipeline of our approach is shown in Fig. 2.

The segmentation of *hotspots* has following specific clinical advantage:

1. *Early non-invasive estimation of inflammation (sub-clinical condition).*
2. *Dosimetric quantification:* A measure of the spread of the *hotspot* is linked to severity of inflammation. This can help clinicians to determine dosimetric quantification during follow up.
3. *Observer independence:* Automatic segmentation can reduce the observer variability in the detection of the area of the *hotspot*.
4. *Wider acceptability:* The proposed prototype should act as a major proof of concept for the use of thermal imaging for diagnostic purpose.

Next we discuss the proposed methodology. Section 3 presents results and discussions followed by conclusions.

2. METHODOLOGY

As mentioned in the introduction, the primary goal for the proposed segmentation is to fuse information from multiple segmentations of the input thermogram. The motivation being to

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.

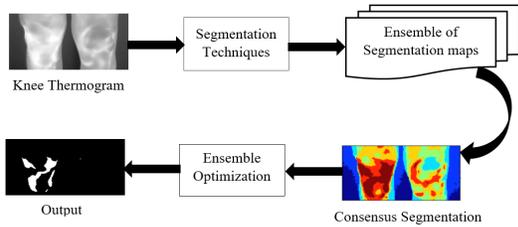


Fig. 2. Block diagram of the proposed approach.

find the core and stable *hotspot* region(s) quickly arriving at a consensus from multiple information sources. We propose to introduce information theoretic measure to achieve at consensus image segmentation. Variation of information (VoI) is the information theoretic measure used in our approach. VoI is used to fuse an ensemble of primary image segmentation techniques.

Why consensus approach? The use of consensus in segmentation is a well-researched topic [11]. It is already well accepted in data clustering community [12]. A reasonable goal for the consensus answer is to seek a clustering that shares the most information within different groupings. Each of the clusterings may suffer from individual initialization issues. Each of the clusterings may employ different optimization approaches ending in solution to the minimization of intra-cluster variance and maximization of inter-cluster distances. A consensus is a solution that is closest to all groupings available.

The solution of consensus segmentation using exhaustive comparison between every pair of segments is a computationally challenging problem. A possible approach could be to declare a base segmentation and compare all other segmentations with respect to the base segmentation using a distance metric [13]. The underlying assumptions being that the distance should be a perfect metric. The variation of information (VoI) that we have used is a perfect metric.

Next we present the generation of segmentation ensemble followed by the description of VoI.

2.1. Generation of Segmentation Ensemble

For ease of understanding, we can assume that the initial segmentation maps are the result of a clustering approach where number of clusters is varied. In this context, it is well known that finding number of clusters for a data set is a difficult problem. And in that sense our VoI based ensemble approach achieves better segmentation compared to any individual segmentation or clustering approach with pre-determined number of clusters. However, note that selection of clustering algorithm or a set of specific segmentation techniques is not a prerequisite to our approach. As the cluster ensemble is populated with the segmentation maps obtained by clustering the image with different number of clusters. Therefore, any clustering method may be used.

Assuming there are Γ_i clusterings, $i = 1, 2, \dots, n$, the thermal image is clustered n times. For ease of understanding, assume that each of n clusterings are obtained by varying the number of clusters from at least two to a certain preset value.

Assume Γ_i clustering produces a segmentation of the thermal image, $S(\Gamma_i) = \{S_1(\Gamma_i), S_2(\Gamma_i), \dots, S_m(\Gamma_i)\}$, generating m numbers of segments in the thermal image. As i varies, differ-

ent number of segments will be produced for the same thermal image. Our objective is to fuse information from these segments using variation of information.

Fig. 3 shows an example of segmentation maps (first three rows) generated using FCM algorithm [14] for different numbers of clusters. Different segments have been represented by different colors. The final result is shown in the last row of the image along with the original thermogram.

For the specific problem at hand, we have two specific advantages. First, to detect *hotspot*, we are looking for the segment having highest intensity values. These regions with highest intensity values represent skin surface area having maximum temperature. Second, the *hotspot* regions have noticeable contrast. We utilize this information in designing the integrated energy function to detect *hotspots*. But before that, we present the information theoretic measure to generate consensus segmentation.

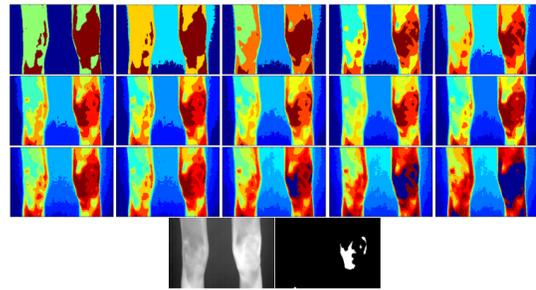


Fig. 3. Example of segmentation ensemble and fusion result. First three rows of the image (15 segmentation maps) show FCM [14] clustering results. First image of the last row is the thermogram and the second image is the result of proposed method.

2.2. Variation of Information

Variation of information is a well known method for comparing lattice of partitions. The lattice of partitions in this problem are $S(\Gamma_i)$, $i = 1, 2, \dots, n$. To compare partitions $S_k(\Gamma_i)$ and $S_l(\Gamma_j)$ where k th segment obtained through clustering Γ_i and l th segment obtained through clustering Γ_j , the information measured from k th and l th segments are combined at the expense of mutual information obtained by any one partition by the other. Entropy is used as a standard measure of information for both k th and l th segments. Therefore, the variation of information is given by,

$$VoI(S_k(\Gamma_i), S_l(\Gamma_j)) = H(S_k(\Gamma_i)) + H(S_l(\Gamma_j)) - 2MI(S_k(\Gamma_i), S_l(\Gamma_j)). \quad (1)$$

where $H(S_k(\Gamma_i))$ and $H(S_l(\Gamma_j))$ represent the general entropy of the two segments $S_k(\Gamma_i)$ and $S_l(\Gamma_j)$ respectively. $MI(S_k(\Gamma_i), S_l(\Gamma_j))$ is the mutual information between segments $S_k(\Gamma_i)$ and $S_l(\Gamma_j)$.

The mutual information between two clusterings is the joint distributions of the random variables associated with respective clusterings. Let α and α' are random variables for $S_k(\Gamma_i)$ and $S_l(\Gamma_j)$ respectively. The joint probability $P(\alpha, \alpha')$ is given by,

$$P(\alpha, \alpha') = \frac{|S_k(\Gamma_i) \cap S_l(\Gamma_j)|}{D}. \quad (2)$$

where $|S_k(\Gamma_i) \cap S_l(\Gamma_j)|$ represents points belonging to segment S_k that is contained in segment S_l and D being the total number of data points. Given this the mutual information between the clusterings $S_k(\Gamma_i)$ and $S_l(\Gamma_j)$ is given by,

$$MI(S_k(\Gamma_i), S_l(\Gamma_j)) = \sum_{\alpha} \sum_{\alpha'} P(\alpha, \alpha') \log \frac{P(\alpha, \alpha')}{P(\alpha)P(\alpha')}. \quad (3)$$

Intuitively, we can think of $MI(S_k(\Gamma_i), S_l(\Gamma_j))$ as follows: We are given a random point in D . The uncertainty about its cluster in $S_k(\Gamma_i)$ is measured by $H(S_k(\Gamma_i))$. Suppose now that we are told which cluster the point belongs to in $S(\Gamma_i)$. How much does this knowledge reduce the uncertainty of the point belonging to $S(\Gamma_j)$? This reduction in uncertainty, averaged over all points, is equal to the mutual information [15].

2.2.1. Is VoI enough?

We have earlier mentioned that *hotspots* have significant image contrast. Therefore, in addition to minimizing mutual variation of information to detect stable core *hotspots* from thermal images, the consensus region should have sharp image contrast. In order to incorporate image contrast feature, we have evaluated the gradient of image.

$$E(S(\Gamma_i)) = \sum_{j=1}^m \Delta(S_j(\Gamma_i)). \quad (4)$$

The gradient of the segment $\Delta(S_j(\Gamma_i))$ is evaluated using 2×2 Robert's cross operator mask [16]. Next we present the algorithm for detecting *hotspots* considering variation of information between segments and their gradient based contrasts.

2.3. Minimization for Consensus Segmentation

Consensus segmentation is obtained minimizing the variation of information between segments. At the same time gradient based contrast should be maximized for the *hotspot* regions. Therefore, the energy function for *hotspot* detection is given by,

$$T = VoI(S(\Gamma_i), S(\Gamma_j)) - \lambda E(S(\Gamma_i)). \quad (5)$$

The idea of consensus segmentation is to generate a base measure of variation of information and then compare all other segmentations with the base measure. The common empirical base measure is the mean of variation of information measure between different segmentations, in this case, $S(\Gamma)$. The parameter λ normalizes contrast value with respect to VoI measure.

Given that there are n different clusterings applied to thermal image, there are n different segmentations. To compare a segmentation $S(\Gamma_i)$ with other segmentations, a total of $(n-1)$ number of VoI measures may be calculated following (1). Therefore, for the i th segmentation $S(\Gamma_i)$, the base measure of VoI is the average of $(n-1)$ VoI measures between $(S(\Gamma_i), S(\Gamma_j))$, where $j = 1, 2, \dots, n, j \neq i$. Given all n segmentations, there are n average VoI measures for each of the $S(\Gamma_i)$. The segmentation representing the minimum of these n average VoI measures, S^* , is the seed segmentation to initiate detection of the *hotspot* region.

In the next phase, each pixel of S^* is checked against its neighbourhood. If the change of segmentation labels of the

neighbors of each pixel of S^* ensures lower T following (5), the change of label is accepted. The algorithmic steps of this proposal is presented next.

1. The iterated greedy local search based minimization of (5) starts from the top-left pixel position of the thermogram.
2. A vector N_8 is populated by the segmentation labels of 8-neighbors of the thermogram pixel. The segmentation labels are extracted from corresponding pixel locations of S^* .

Let y_p^t is the distinct label of a pixel p in the thermogram at t th iteration. The cost for assigning a new label to y_p at iteration t is given by,

$$C(y_p, y_q) = T(1 - \delta(y_p, y_q)), y_q \in N_8 \quad (6)$$

where y_q is the labels of pixels neighbor to p . The Kronecker delta is given by $\delta(a, b) = 1$ if $a = b$ else $\delta(a, b) = 0$ if $a \neq b$.

3. The pixel label is updated when the cost is minimized in a greedy fashion. It is a local search as only the 8-neighbors of a pixel are considered.

$$y_p^{t+1} \leftarrow \arg \min_{y_q^t} C(y_p^t, y_q^t). \quad (7)$$

The iteration is terminated when less than 5% pixels change their labels in two consecutive iterations. Fig. 4 shows the decrease in number of pixels that changes label with the increase in number of iterations. The segment(s) having maximum average intensity in S^* is taken as *hotspots*. The average intensity of the segment in S^* is calculated from the intensity values of corresponding pixels from the original thermogram. We have used clinical validation to determine *hotspot* segmentation as mild, moderate or severe. A related analysis is given in Section 3.1. Next we present experimental results.

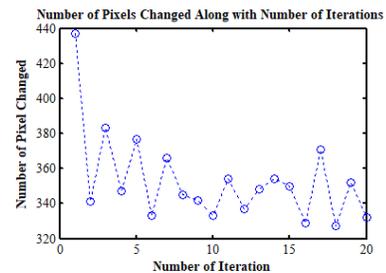


Fig. 4. Plot of the number of pixels changing labels with the increasing iterations.

3. RESULTS AND DISCUSSIONS

The proposed method is evaluated on the thermograms of arthritis patients. The dataset is created in PMR Department of Agartala Government Medical college, Tripura, India. The dataset is captured maintaining the protocols related to thermogram capturing. The FLIR T-650sc camera is used for capturing the thermograms with sensitivity of $0.02^\circ C$ at $30^\circ C$. Fifty thermograms of arthritis patient with resolution of 640×480 pixels have been

used in this experiment. The result is compared with the ground truth to quantify the performance of the method. The ground truths are annotated using the GIMP software by expert rheumatologists.

VoI based Initialization for consensus segmentation:

Fig. 5 shows the efficacy of the VoI based consensus segmentation. Instead of using S^* , the segmentation is initialized with the brightest segment (having maximum temperature) and segment with minimum contrast. These two initializations are taken from the segment ensemble. The results shown in Fig. 5 have established the importance of VoI based selection of initial segmentation compared to adhoc initialization for consensus segmentation.

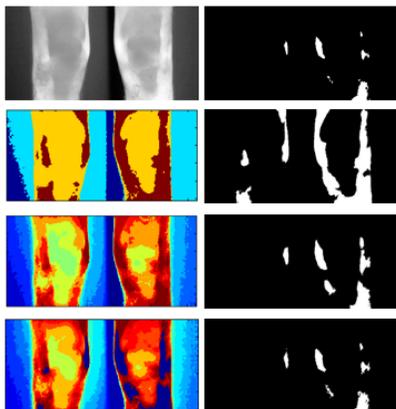


Fig. 5. Example of consensus segmentation. First row: Original image and corresponding ground truth. Second row: Initialization with brightest segment and corresponding segmentation result (30 iteration). Third row: Initialization with segment having minimum contrast and the corresponding segmentation result (25 iterations). Final Row: Initialization with S^* and corresponding segmentation result (15 iterations).

Comparison with existing methods: Proposed method is compared with the baseline methods such as K-means, FCM, Region Growing (RG) and Otsu’s Thresholding (OT). FO-DPSO is a recent algorithm proposed for segmentation of visual images [17]. The proposed approach is also compared with [17]. The proposed method is mentioned as PM in Table 1. Different quantitative measures are employed for the comparison. They are: (1) Jaccard index (JI) [18], (2) Recall (RC) [18], (3) Precision (PRC) [18], (4) Over-segmentation (O_Seg) [19], (5) Under-Segmentation (U_Seg) [19] and (6) Accuracy (ACC). The higher the value of JI, the better is the performance of the algorithm. RC and PRC are considered collectively and for good segmentation both the parameters should be close to 1. Likewise, O_Seg and U_Seg should be close to 1 to represent improved segmentation compared to over and under segmentation. The comparative results are shown in Table 1.

Fig. 6 shows that the proposed method is sensitive to the size of ensemble used. In terms of JI values, the size of ensemble between [10,15] indicates better performance.

3.1. Clinical Validation

In this section we have compared the features of segmented ROI with the clinical grading of inflammation. Clinicians grade the

Table 1. Experimental results and comparisons. For details refer text.

Methods	JI	RC	PRC	ACC	O_Seg	U_Seg
PM	0.98	0.74	0.54	96.06	0.03	0.04
FCM	0.38	0.38	1	92.06	0	0.61
K-means	0.33	0.34	0.99	83.22	0	0.66
RG	0.57	0.87	0.65	97.84	0.39	0.09
OT	0.35	0.35	1	90.77	0	0.10
FO-DPSO	0.19	0.19	1	79.02	0	0.81

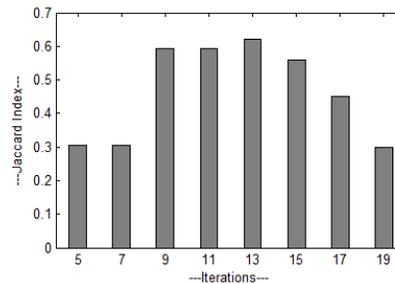


Fig. 6. Evaluation of JI as a function of the size of the ensemble.

inflammation into three categories: mild, moderate, and severe based on the available clinical and pathological investigation. ESR, CRP are the pathological investigations whereas, swelling, tenderness, redness, restriction of movement and warmth of the surface are the clinical investigations. According to the grading of the clinician, the dataset can be divided into two categories, mild and moderate. We calculate the average intensity of the segmented ROI and plot them as shown in Fig. 7. The average intensity values from ROI of moderate category is indicated using red diamonds while the same for mild category is indicated using green rectangles. The plot validates that based on the average intensity of the ROI, detected using our proposed segmentation technique, it is possible to identify the mild from the moderate arthritic category.

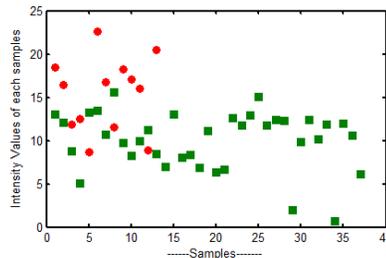


Fig. 7. Classification of mild and moderate knee thermograms from the average intensity of extracted ROI

4. CONCLUSIONS

We propose a quick, reliable and inexpensive tool for detecting the knee surface area affected by arthritis. We believe that uses like this should initiate wider acceptability of thermal imaging in medical diagnosis. The clinical validation of our result is promising and opens up possibility of future use of our result in arthritis grading.

5. REFERENCES

- [1] Iain B. McInnes and Georg Schett, "The pathogenesis of rheumatoid arthritis," *New England Journal of Medicine*, vol. 365(23), 2011.
- [2] J.-F. Chen Y.-C. Chen and T.-T. Cheng, "Sat0092 prevalence of subclinical inflammation in patients with rheumatoid arthritis who have low disease activity: Table 1," *Annals of the Rheumatic Diseases*, vol. 75, 2016.
- [3] H. Daghestani and V. Kraus, "Inflammatory biomarkers in osteoarthritis," *Osteoarthritis and Cartilage*, vol. 23, 2015.
- [4] A. Vargas A. Lopez-Macay M. S. Olmedo A. L. Reyes C. Pineda C. H. C. A. Daz, M. R.-S. Roman and L. V. Rios, "Sat0180 subclinical inflammation in rheumatoid arthritis (ra) in clinical remission, lack of association between cytokines level and ultrasound-defined synovitis:," *Annals of the Rheumatic Diseases*, vol. 73, 2014.
- [5] B. F. Jones, "A reappraisal of the use of infrared thermal image analysis in medicine," *IEEE Transactions on Medical Imaging*, vol. 17, 1998.
- [6] T. Jayakumar B.B. Lahiri, S. Bagavathiappan and John Philip, "Medical applications of infrared thermography: A review," *Infrared Physics Technology*, vol. 55, 2012.
- [7] V Sowmiya B Venkatraman U Snekhalatha, M Anburajan and M Menaka, "Automated hand thermal image segmentation and feature extraction in the evaluation of rheumatoid arthritis," *Inst Mech Eng H: J Eng Med*, vol. 229, 2015.
- [8] Therace Teena B. Venkatraman M. Menaka U. Snekhalatha, M. Anburajan and Baldev Raj, "Thermal image analysis and segmentation of hand in evaluation of rheumatoid arthritis," *2012 International Conference on Computer Communication and Informatics*, vol. 1, 2012.
- [9] M. S. Jadin and S. Taib, "Infrared image enhancement and segmentation for extracting the thermal anomalies in electrical equipment," *Electron Electr Eng*, vol. 120, 2012.
- [10] Sangeetha Vivek N.Selvarasu and N.M.Nandhitha, "Performance evaluation of image processing algorithms for automatic detection and quantification of abnormality in medical thermograms," *International Conference on Computational Intelligence and Multimedia Applications*, 2012.
- [11] GOLDBERGER J. ALUSH A., "Ensemble segmentation using efficient integer linear programming," *IEEE Transactions on PAMI*, vol. 34, 2012.
- [12] GHOSH J. STREHL A., "Cluster ensembles - a knowledge reuse framework for combining multiple partitions," *J. of Machine Learning Research*, vol. 3, 2002.
- [13] D. Cremers E. Rodol, S. Rota Bul, "Robust region detection via consensus segmentation of deformable shapes," *Eurographics Symposium on Geometry Processing*, 2014.
- [14] William Full James C.Bezdek, Robert Ehrlich, "Fcm: The fuzzy c-means clustering algorithm," *Computers Geosciences*, vol. 10, 1984.
- [15] Marina Meila, "Comparing clusterings by the variation of information," *Learning Theory and Kernel Machines*, vol. 2777, 2003.
- [16] E. Davies, "Machine vision: Theory, algorithms and practicalities," *Academic Press*, 1990.
- [17] G. Pajares D. Zaldivar D. Oliva, E. Cuevas and V.O.-Enciso, "A multilevel thresholding algorithm using electromagnetism optimization," *Neurocomputing*, vol. 139, 2014.
- [18] H. R. Roth J. Liu E. Turkbey A. Farag, L. Lu and R. M. Summers, "A bottom-up approach for pancreas segmentation using cascaded superpixels and (deep) image patch labeling," *IEEE Transactions on Image Processing*, vol. 26, 2017.
- [19] S. Sadri N. Gheissari M. Mokhtari A. R. Sadri, M. Zekri and F. Kolahdouzan, "Segmentation of dermoscopy images using wavelet networks," *IEEE Transactions on Biomedical Engineering*, vol. 60, 2013.