Tuberculosis detection using optical flow and the activity description vector

Fernando Llopis, Andrés Fuster-Guilló, Juan Ramón Rico-Juan, Jorge Azorín-López, and Irene Llopis

University of Alicante. Carretera San Vicente del Raspeig s/n 03690 San Vicente del Raspeig - Alicante, Spain informacio@ua.es
http://www.ua.es {fernando.llopis, jazorin, juanramonrico, fuster}@ua.es, ilq2@alu.ua.es

Abstract. Early detection of tuberculosis can save many lives as it remains one of the leading causes of death, half a century after its discovery. The analysis of chest CT scanned images can be a quick and economic mechanism for detecting not only the type of tuberculosis, but also differentiating whether or not the disease is multi-drug resistant. These are two of the objectives of the ImageClef Tuberculosis task of 2018, and are the ones studied by the group of the University of Alicante in this edition. We have carried out two work approaches, one based exclusively on the use of Deep Learning techniques on a sequence of 2D images extracted from a 3D tomography and on a second approach using Optical Flow to convert the 3D tomography into a motion representation in order to calculate the ADV (a previous descriptor provided by the group). This descriptor is able to synthesize the information of a sequence into one image. This article presents the experiments carried out and the results obtained within the task.

Keywords: Tuberculosis · Optical Flow · Activity Description · Deep Learning.

1 Introduction

ImageClefTuberculosis is one of the tasks of ImageClef 2018 [11]. The ImageCLEFtuberculosis task 2018 [5] includes three independent subtasks.

1. Subtask 1: MDR detection The goal of this subtask is to assess the probability of a TB patient having resistant form of tuberculosis based on the analysis of chest CT scan.
2. Subtask 2: TBT classification The goal of this subtask is to automatically categorize each TB case into one of the following five types: (1) Infiltrative, (2) Focal, (3) Tuberculoma, (4) Miliary, (5) Fibro-cavernous.
3. Subtask 3: Severity scoring This subtask is aimed at assessing TB severity score based on chest CT image. The Severity score is a cumulative score of severity of TB case assigned by a medical doctor.

In this first participation our initial objective was to compare two models, Deep learning and Optical Flow to check their results in task 1. Finally we made a delivery about task 2 using the second model that had given us better results in the experimentation of the task 1.

This document is structured as follows: in sections 2 we present the architectures of the models used: Deep Learning and Optical Flow. In section 3 we show the experimetntation done with both models. Section 4 presents the official results of the experiments and Section 5 summarizes the document and offers a series of proposals for future work.

2 Our approaches to the solution

2.1 Deep Learning

Deep neural networks have managed to solve problems or increase efficiency in problems related to image processing [7]. On the one hand, the convolutional layers manage to extract discriminative characteristics from the images so that they can be evaluated by subsequent layers [12]. On the other hand, recurrent neural networks have also evolved in their approach and are mainly used in sequence analysis [15].

To address the issue of the first task (resistant form of tuberculosis), 3D chest images of CT scan are used. In a first stage, these 3D images are transformed into a sequence of 2D images each one that represent the entry of the neural networks.

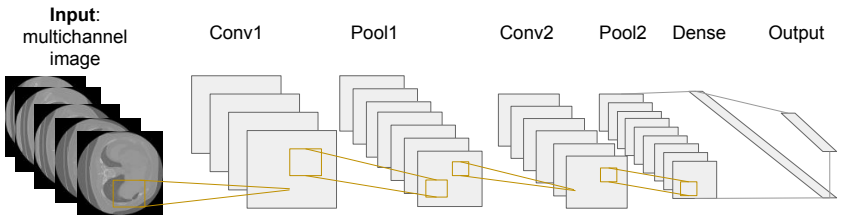

Fig. 1. Basic scheme of a deep convolutional neural network for classification task.

Different approaches will be proposed to address the problem of tuberculosis detection:

1. Convolutional neural network (CNN) with data augmentation: The main idea is to use the advantages of convolutional layers for a single multi-channel image. In this case, each channel would be a 2D gray image (see Figure 1).
2. Convolutional layers combined with a recurrent neural network: The natural way to combine the advantages of convolutional layers and sequential

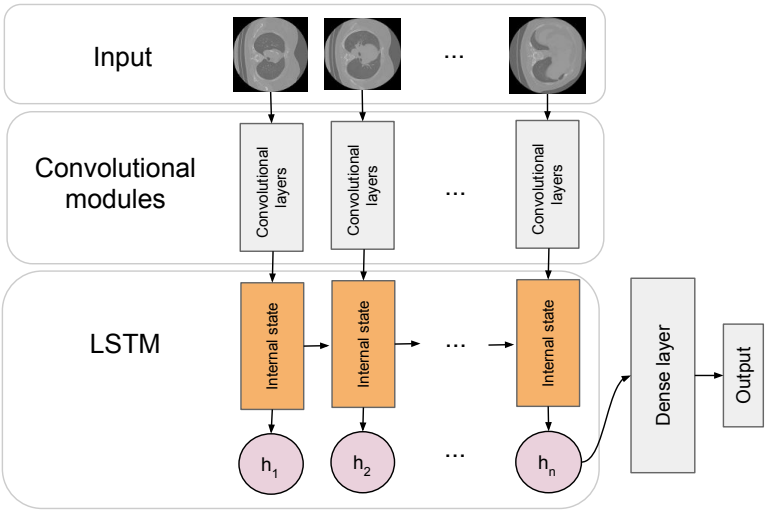

Fig. 2. Basic scheme of a combination of a convolutional layer with a Long-Short Term Memory network for classification task.

treatment is to combine it in networks with multiple inputs per tomography. Figure 2 shows a basic scheme about this approach.

3. Pretrained network and classification: As first approximation to extract features from an image VGG16 deep convolutional neural network [16] is used with the ImageNet [12] weights learned (4096 features per image). The main idea is concatenate the features of each input image belonging to the same tomography to get the final features vector to classify in a classical way.
4. Pretrained network and classification as a sequence using recurrent neural network. In this case, the extraction of features is similar to the one described in the previous paragraph and each feature vector would be considered as a component of a sequence to be treated by a well knowns recurrent neural network called Long-Shot Term Memory (LSTM) [10].

2.2 Optical Flow plus ADV

In this sections we propose a combined method based on optical flow and a characterization method called ADV, to deal with the classification of chest CT scan images affected by different types of tuberculosis. The key point of this method is the interpretation of the set of cross-sectional chest images provided by CT scan, not as a volume but as a sequence of video images. We can extract movement descriptors capable of classifying tuberculosis affections by analyzing deformations or movements produced in these video sequences.

The concept of optical flow refers to the estimation of displacements of intensity patterns. This concept has been extensively used in computer vision in

different application domains: robot or vehicle navigation, car driving, video surveillance or facial expression [6]. In biomedical context optical flow has been used to analyze organ deformations [9,17]. We can find different methods in the literature to obtain the optical flow [3]. One of the most used method to estimate motion at each pixel is Lucas Kanade [13]. In this work we will use Lucas Kanade method to extract optical flow comparing sequences of consecutive images. Nevertheless, we need not only to estimate motion but describe this motion.

In order to describe motion there are several methods used in different computer vision context like human behavior recognition [8]. A successful method to describe human behavior based on trajectory analysis is presented in [1]. The paper proposes a description vector called (ADV Activity Description Vector) tested in several contexts [2]. In summary, the ADV vector describes the activity in image sequence by counting for each region of the image the movements produced in four directions of the 2D space. A detailed description of the method can be found in [1]. In this paper we propose the use of ADV to describe motion in the optical flow obtained from sequences of cross-sectional chest images provided by CT scan.

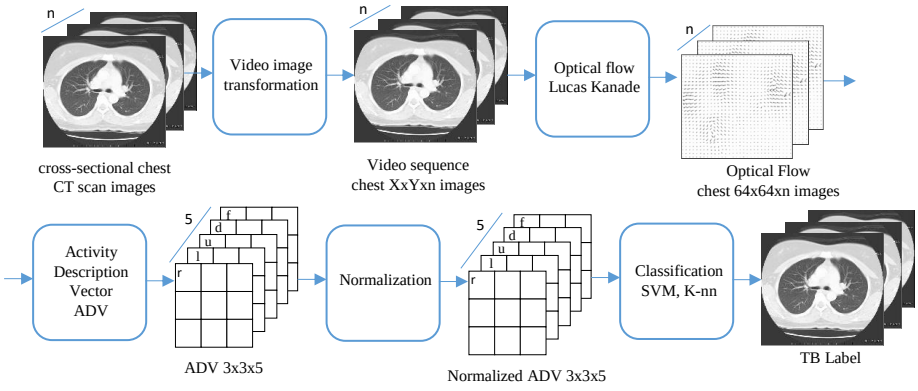

Fig. 3. Optical flow plus ADV process stages

The figure summarizes the successive stages of the process for extracting the activity descriptors (optical flow+ADV) that will be the input of a classifier. In the first stage a transformation over the cross-sectional chest images provided by the CT scan is performed in order to transform image formats into video sequences adapted to calculate optical flow. The second stage implements the Lucas Kanade method to obtain optical Flow. The third stage calculates the activity description vector ADV (3x3x5) accumulating within each 3x3 region of the image, the displacements of the optical flow in four directions of a 2D space (right, left, up, down). The fifth component of the ADV calculates the frequencies in direction changes. In the fourth stage a normalization of the ADV

vector in performed. Finally, the last stage uses the ADV vector normalized as the input for a generic classifier in order to evaluate the results.

3 Experimentation

3.1 Preliminary experiments using Deep Learning

In order to validate the results the wide 10-fold cross validation (10-CV) technique are used and 7 images of 2D are extracted from the original 3D tomography. For the experiments *Keras* v2.1.6 [4] and *scikit-learn* v0.19.1 [14] Python software are used, in order to build deep neural networks and apply classifiers, respectively.

Table 1 shows the first approach using CNN. The results are close to 50% which means that the network has not learned the difference between the two classes.

The second approach consists of combination of CNN with RNN. In this case, 2 layers are used and the filters are: 32 (3x3), 64 (3x3). The accuracy is 0.50 and individual proofs are [0.54 0.58 0.42 0.50 0.48 0.52 0.52 0.44 0.48 0.52]. The results are also unsatisfactory and we will try a new approach.

Conv. layers	**Detail filters x kernel**	**Accuracy mean**	**10-CV results**
4	64x7, 64x3, 64x3, 64x3	0.52	[0.46 0.54 0.54 0.54 0.52 0.52 0.52 0.52 0.52 0.52]
2	64x7, 64x3	0.51	[0.46 0.54 0.54 0.50 0.52 0.52 0.52 0.52 0.48 0.52]

Table 1. Classification results using CNN.

The third try using a pretrained network (VGG16) with ImageNet weights configuration. In this case, VGG16 is used to extract the weights of the penultimate layer as descriptors of image. These features extracted from the latest layers of the neural network are called neural codes. The number of final characteristics is 28672 corresponding to 7 images per times 4096 neural codes per image. Table 2 summarize the experiments using classifiers belonging different families of algorithms attending to neural codes directly or normalizing with L2 function.

Last approach using deep learning architectures consists of get neural codes as previous try and classify the sequence of 7 images with a recurrent neural network (LSTM). Again, the accuracy is 0.49 and detailed fold results are [0.62 0.54 0.42 0.46 0.28 0.48 0.48 0.48 0.56 0.56].

In general, the results per folder (10-CV) are very different probably due to the nature of neuronal networks with random initialization of neurons, the optimizers that have to adjust thousands of parameters that finally find local minimums and also due to the small amount of images available to train a neuronal network where small differences between the training and test sets

Neural Codes	Classifier algorithms	Accuracy mean	10-CV results
	Nearest Neighbors	0.48	[0.54 0.46 0.65 0.35 0.36 0.6 0.32 0.56 0.44 0.52]
	Linear SVM	0.48	[0.46 0.46 0.69 0.54 0.4 0.6 0.24 0.52 0.44 0.48]
	RBF SVM	0.53	[0.54 0.54 0.54 0.54 0.52 0.52 0.52 0.52 0.52 0.52]
	Decision Tree	0.54	[0.62 0.54 0.58 0.5 0.44 0.44 0.48 0.72 0.56 0.52]
original	Random Forest	0.41	[0.27 0.54 0.5 0.38 0.4 0.56 0.28 0.4 0.4 0.4]
	AdaBoost	0.49	[0.46 0.65 0.58 0.42 0.28 0.52 0.52 0.64 0.36 0.48]
	Naive Bayes	0.47	[0.54 0.58 0.5 0.27 0.48 0.52 0.36 0.48 0.48 0.44]
	Logistic Regression	0.48	[0.5 0.46 0.69 0.54 0.44 0.52 0.28 0.48 0.4 0.44]
	XGBoost	0.52	[0.46 0.65 0.65 0.46 0.32 0.6 0.4 0.6 0.6 0.44]
	Nearest Neighbors	0.50	[0.62 0.5 0.69 0.42 0.4 0.48 0.32 0.64 0.4 0.48]
	Linear SVM	0.53	[0.54 0.54 0.54 0.54 0.52 0.52 0.52 0.52 0.52 0.52]
	RBF SVM	0.45	[0.42 0.54 0.54 0.46 0.24 0.44 0.28 0.6 0.44 0.56]
	Decision Tree	0.52	[0.58 0.5 0.5 0.58 0.56 0.4 0.52 0.6 0.36 0.56]
L2	Random Forest	0.52	[0.42 0.73 0.42 0.42 0.6 0.48 0.48 0.52 0.52 0.6]
	AdaBoost	0.47	[0.54 0.58 0.54 0.31 0.4 0.56 0.24 0.52 0.44 0.56]
	Naive Bayes	0.47	[0.54 0.58 0.5 0.27 0.48 0.52 0.36 0.48 0.48 0.44]
	Logistic Regression	0.47	[0.42 0.62 0.58 0.46 0.2 0.48 0.28 0.6 0.56 0.52]
	XGBoost	0.50	[0.65 0.54 0.62 0.42 0.32 0.56 0.4 0.6 0.44 0.48]

Table 2. Classification results applied to features extraction with VGG16 pretrained network.

allow generating sets easier to classify in some cases than in others. On the other hand, no preprocessing has been applied to 2D images which could also influence the high variations in results.

3.2 Preliminary experiments using Optical Flow plus ADV

For this experiments, the wide 10-fold cross validation (10-CV) technique have been used again. All images of the original 3D tomography are used to calculate the optical flow for each patient. For the experiments *Matlab* R2013b has been used to calculate the optical flow, the ADV and the classifiers.

Table 3.2 shows the performance results of the proposed method.

Classifier	OF size	ADV	Accuracy MDR	Accuracy DS	Accuracy
SVM	64x64	3x3	0,5097	0,312	0,6567
3-knn	64x64	3x3	0,5135	0,52	0,4627

Table 3. Classification results using Optical flow plus ADV

3.3 Frequency Matrix with Deep Learning

A modification of the Optical Flow experiment was to use the frequency matrices generated as input to a neural network.

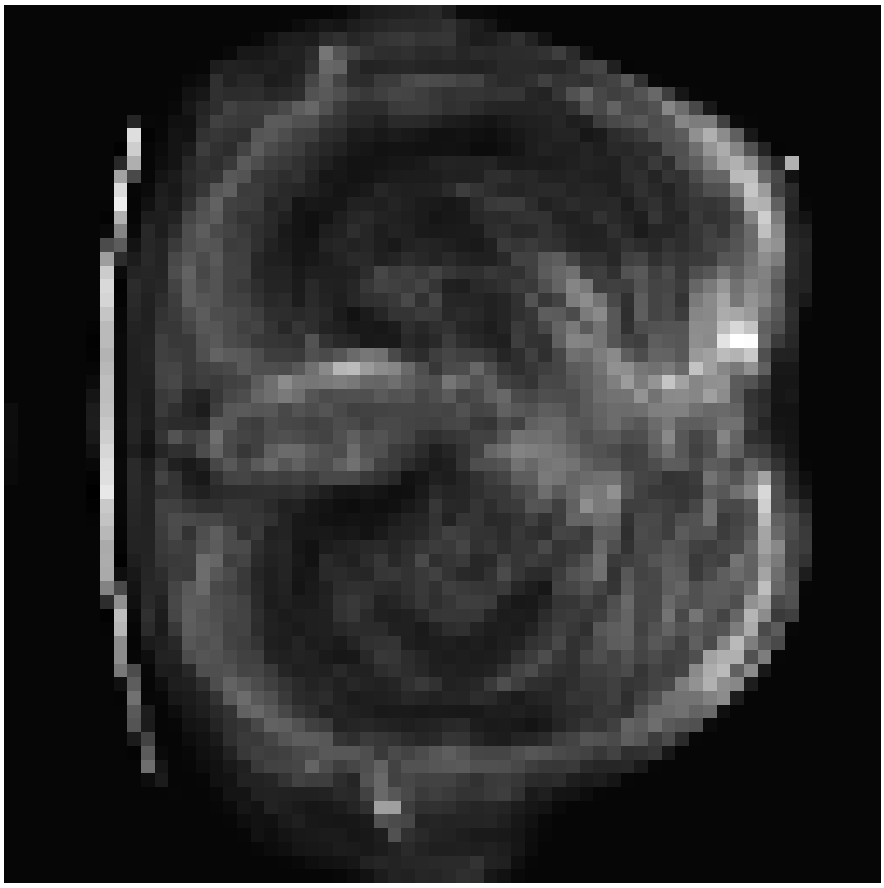

Fig. 4. Frequency Matrix.

In figure 1 you can see an example of Frequency Matrix..

4 Results

1. Run 1: MDR Baseline. The Baseline is a probabilistic model in which the image was not analyzed and only the data of sex and age have been taken into account.
2. Run 2: ADV 3x3, SVM, 1000 SMOTE upsampling

3. Run 3 Frequency Normalized. In this model we apply Deep Learning techniques on the normalized frequency matrix obtained through the Optical Flow.
4. Run 4 In this model we apply Deep Learning techniques (Decision Tree) on a subset of 2D images of the tomography.
5. Run 5 In this model we apply Deep Learning techniques (Decision Tree) on a subset of 2D images of the tomography.

As can be see in the table 4 the model of Optical Flow SVM obtains the best results, for the sake of using only selected images.

Table 4. Results of University of Alicante vs better results at SubTask 1

Run	AUC	Rank AUC	ACC	Rank ACC
VISTA@UEvora	0.6178	1	0.5593	8
San Diego VA HCS/UCSD	0.6114	2	0.6144	1
MDRBaseline0	0.5669	10	0.4873	32
testSVMSMOTE	0.5509	15	0.5339	20
testOpticalFlowwFrequencyNormalized	0.5473	16	0.5127	24
DecisionTree25v2	0.5049	26	0.5000	29
testOFFullVersion2	0.4971	29	0.4958	31
testOpticalFlowFull	0.4845	32	0.5169	23
testFrequency	0.4781	34	0.4788	34
testflowI	0.4740	35	0.4492	39

Due to we had little time available for second task, we only present the two models of Optical Flow, SVM and 3nn.

Run 1: ADV 3x3, SVM, 1000 SMOTE upsampling Run 2: ADV 3x3, 3-nn, 1000 SMOTE upsampling

The results were significantly better using the 3-nn but very far from the rest of the participants 5.

Table 5. Results of University of Alicante vs better results at SubTask 2

Run	AUC	Rank AUC	ACC	Rank ACC
UIIP BioMed	0.2312	1	0.4227	1
T23nnFinal	0.0204	32	0.2587	31
T2SVMFinal	-0.0920	38	0.1167	38

5 Conclusions and future work

Early detection of tuberculosis is a major social challenge, given the devastating effects of the disease. On the other hand, it represents a scientific challenge of

the highest level. As the organizers claim, "you have to work to get methods that allow a correct detection of the disease that kills thousands and thousands of people". In this paper we have proposed two different approaches to face the problem. The first one is based on the use of Deep Learning techniques on a sequence of 2D images extracted from a 3D tomography. The second approach uses Optical Flow to convert the 3D tomography into a motion representation in order to calculate the ADV (a previous descriptor provided by the group). This descriptor is able to synthesize the information of a sequence into one image. The experiments carried out in these two approaches allow us to confirm the interest of these lines of research and encourage us to seek improvements in the proposed methodologies.

References

1. Azorin-Lopez, J., Saval-Calvo, M., Fuster-Guillo, A., Garcia-Rodriguez, J.: Human behaviour recognition based on trajectory analysis using neural networks. In: Proceedings of the International Joint Conference on Neural Networks (2013). https://doi.org/10.1109/IJCNN.2013.6706724
2. Azorin-Lopez, J., Saval-Calvo, M., Fuster-Guillo, A., Garcia-Rodriguez, J., Cazorla, M., Signes-Pont, M.T.: Group activity description and recognition based on trajectory analysis and neural networks. In: 2016 International Joint Conference on Neural Networks (IJCNN). vol. 2016-Octob, pp. 1585–1592 (2016). https://doi.org/10.1109/IJCNN.2016.7727387
3. Chao, H., Gu, Y., Napolitano, M.: A survey of optical flow techniques for robotics navigation applications. Journal of Intelligent and Robotic Systems: Theory and Applications **73**(1-4), 361–372 (2014). https://doi.org/10.1007/s10846-013-9923-6
4. Chollet, F., et al.: Keras. `https://keras.io` (2015)
5. Dicente Cid, Y., Liauchuk, V., Kovalev, V., , Müller, H.: Overview of ImageCLEFtuberculosis 2018 - detecting multi-drug resistance, classifying tuberculosis type, and assessing severity score. In: CLEF2018 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Avignon, France (September 10-14 2018)
6. Fortun, D., Bouthemy, P., Kervrann, C.: Optical flow modeling and computation: A survey. Computer Vision and Image Understanding **134**, 1–21 (2015). https://doi.org/10.1016/j.cviu.2015.02.008
7. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press (2016), `http://www.deeplearningbook.org`
8. Gowsikhaa, D., Abirami, S., Baskaran, R.: Automated human behavior analysis from surveillance videos: a survey. Artificial Intelligence Review **42**(4), 747–765 (2014). https://doi.org/10.1007/s10462-012-9341-3, `https://doi.org/10.1007/s10462-012-9341-3`
9. Hata, N., Nabavi, A., Wells, W.M., Warfield, S.K., Kikinis, R., Black, P.M.L., Jolesz, F.A.: Three-dimensional optical flow method for measurement of volumetric brain deformation from intraoperative MR images. Journal of Computer Assisted Tomography **24**(4), 531–538 (2000). https://doi.org/10.1097/00004728-200007000-00004
10. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural computation **9**(8), 1735–1780 (1997)

11. Ionescu, B., Muller, H., Villegas, M., de Herrera, A.G.S., Eickhoff, C., Andrearczyk, V., Cid, Y.D., Liauchuk, V., Kovalev, V., Hasan, S.A., Ling, Y., Farri, O., Liu, J., Lungren, M., Dang-Nguyen, D.T., Piras, L., Riegler, M., Zhou, L., Lux, M., Gurrin, C.: Overview of ImageCLEF 2018: Challenges, datasets and evaluation. In: Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the Ninth International Conference of the CLEF Association (CLEF 2018), vol. 11018. LNCS Lecture Notes in Computer Science, Springer, Avignon, France (September 10-14 2018)
12. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. pp. 1097–1105 (2012)
13. Patel, D., Saurahb, U.: Optical flow measurement using Lucas Kanade method. Int J Comput Appl **61**(10), 6–10 (2013)
14. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research **12**, 2825–2830 (2011)
15. Rumelhart, D., Hinton, G., Williams, R.: Learning sequential structure in simple recurrent networks. Parallel distributed processing: Experiments in the microstructure of cognition **1** (1986)
16. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
17. Xavier, M., Lalande, A., Walker, P.M., Brunotte, F., Legrand, L.: An adapted optical flow algorithm for robust quantification of cardiac wall motion from standard cine-MR examinations. IEEE Transactions on Information Technology in Biomedicine **16**(5), 859–868 (2012). https://doi.org/10.1109/TITB.2012.2204893