

Wound image segmentation using clustering based algorithms

Ihor Farmaha

Lviv Polytechnic National University, **Ukraine**

Marian Banaś

AGH University of Science and Technology, **Poland**

Vasyl Savchyn, Bohdan Lukashchuk

Lviv Polytechnic National University, **Ukraine**

Taras Farmaha

Danylo Halytsky Lviv National Medical University, **Ukraine**

Date of submission to the Editor: 05/2019

Date of acceptance by the Editor: 08/2019

INTRODUCTION

One of the branches of medicine where it is advantageous to use automated solutions is an analysis of wounds. The precise assessment and measurement of wounds is an important aspect in determining the effectiveness of treatment (Jansen, S., Rachel, K. 2016). Without techniques, firstly to assess the size and extent of a wound and then to measure its progress, it is difficult to formulate an effective wound management strategy.

In clinical practice, one of the most frequently used non-automated methods of measurement is determining length and width using a disposable ruler (WoundEducators.com, 2012). The disadvantages of this technique are obvious: the shape of the wound is greatly simplified to obtain results. Measurement of surface area using this technique, although simple and affordable, is inaccurate for measuring wounds with complex form, large or hollow (Flanagan M., 2003). There are also improved methods with a slightly more complex formula for calculating the area of wounds with an elliptical form (Jansen, S., Rachel, K. 2016).

Importance of this topic is mentioned and proved in papers "Computer software for vulnerometry" (Lukavetskyi, O., et al., 2017) and "Implementation features of wounds visual comparison subsystem" (Jaworski, N. et al., 2018) where, topic from medical point of view and computer program for calculation of wounds area are described. This program adds user an ability to encircle wound on image, using computer mouse, after that, wound area is calculated automatically.

Automation of wound analysis on the human body surface using algorithms of machine learning will not only save time and other resources of doctors, patients and hospitals, but will also be a significant step towards the development of telemedicine.

Main goal of this paper is to describe an image segmentation method, based on clustering algorithms, adopted to be used on bio-medical images with wounds, that can improve and automate wound area selection. Other objective is to achieve improvement of segmentation results within the process of data (images with wounds) accumulation.

Wound image segmentation belongs to the field of image analysis. Many studies are being carried out in this field. Two major problems can be highlighted:

- image classification
- image segmentation

Paper “Fine-grained wound tissue analysis using deep neural network” (Nejati, H., et al., 2018) describes development of the algorithm for 7 types of chronic wounds recognition. It is based on deep neural networks (DNN). Two main steps are:

- dimensionality decrease using DNN with AlexNet architecture
- images classification using support vector machines (SVM) method

Dataset LSVRC-2010ImageNet (Nejati, H., et al., 2018) is used for model training.

This method belong to an unsupervised learning techniques, thus annotated data is needed (it is needed to place labels on wound areas (on of 7 classes, according to tissues types)).

One more research “Medical Image Classification Based on Deep Features Extracted by Deep Model and Statistic Feature Fusion with Multilayer Perceptron” (Lai, Z. and Deng, H., 2018) is also focused on image classification and proposes their own method that is based on existing ones but is aimed to eliminate their disadvantages.

Main steps include:

- dimensionality decrease and feature extraction using autoencoder neural network architecture
- extraction of image features using traditional techniques (based on knowledge about medical images feature)
- combination of features, extracted on two previous steps, for classification, based on neural network

Datasets HIS2828 and ISIC2017 are used (Lai, Z. and Deng, H., 2018)]

Paper “Wound Image Analysis Classifier for Efficient Tracking of Wound Healing Status” (Kumar, K., Reddy, B., 2014) also describes method for solving classification problem, but is focused on prediction of wound healing process.

All researches, described above are focused on the classification and widely use neural networks, which are great for this kind of problem. But there are some disadvantages, that were slightly mentioned:

- neural networks need large amount of training data, this data needs to comply to many rules, to be used for training (thus it is common to use filters, etc.). In bio-medical field there are a lot of legal regulations, because of sensitive information and it is not easy to gather images with wounds.

- classification also needs this data to be annotated (labeled), which increase complexity in many times, because medical images can be annotated only by doctors, for sufficient precision of the results.

Segmentation, on the other hand is form of unsupervised learning, where minimum of human interaction is needed and algorithm is training `on its own` or just with small help.

Approach for solving of the segmentation task is described in paper “Active Contour Based Segmentation Techniques for Medical Image Analysis” (Hemalatha, R., Thamizhvani, T. and Dhivya, A., 2018). Active contours method needs preliminary human interaction to encircle area, after that it is reduced to create more precise segment of area inside. It is widely used in medical sphere (cardio images, brain MRI, etc.). Advantage of this method is that it can create much smoother contours than other methods. Also, the initial contour can be set programmatically.

From materials above can be assumed that usage of neural networks provides results improvement on data accumulation, typical deep neural networks architectures can't be used, because they solve classification problem. It is needed to use clustering neural network model, with preliminary preparations, because of use of bio-medical images.

MATERIALS AND METHODS

To achieve improvement of the segmentation results during data (images with wounds) accumulation, algorithm (Deep Embedded Clustering – DEC), described in “Unsupervised Deep Embedding for Clustering Analysis” (Xie, J., Girshick, R. and Farhadi, A., 2016) was chosen. It contains preliminary step of model training, where model refers to a neural network model. After that, this trained model is used in further real time clustering.

Algorithm consist of two main steps:

- image feature extraction
- extracted features clustering

Each of this steps uses neural network models.

This algorithm is used to cluster input image from some dataset (MNIST for example, was used in paper) into one of few clusters (10, if it is MNIST dataset). It does not exactly cover needs of wound image segmentation, because the main purpose is to divide image into segments (clusters), but not to assign whole image into some of the predefined clusters. But this approach was adopted to solve the current task.

If images are of one type or have some similarities, every pixel of every image can be assigned to some cluster. As a result of this pixels clustering – set of clusters, each filled with pixels from every image in the dataset, will be created. This seems as a decision, but one pixel can not contain any semantic information, consequently clustering will be inefficient.

To solve the problem it was decided to perform preliminary segmentation of the original image, using traditional machine learning algorithms.

Taken from directory in file system, image is segmented into number of segments, which, after processing are saved to another directory, to be used in model training.

Two main criterias of segmentation algorithm, that were taken into account:

- speed performance, applicable for segmentation of images with high resolution (3104×1746 px)
- is able to create large amount of segments, with rectangle-like shape (it is obvious that segments of only rectangular shape will not cover our needs, because edges of the wounds can be hardly covered with rectangles, but if form of segments differs strongly from segment to segment and is of some complex form, sizes unification step, that is needed for further feature extraction will significantly corrupt segment).

The most suitable algorithm was chosen between algorithms that need preliminary human interaction, and those that do not need it. Algorithms were tested on notebook Asus VivoBook X510U with CPU Intel i7 – 8550U/BQA, 16 GB DDR4 RAM, using programming language Python and programming library scikit-image.

Tested algorithms:

- Felzenszwalb
- Quickshift
- SLIC (number of cluster centers is needed to be set)
- Watershed (number of markers is needed to be set)
- For Watershed and SLIC segmentation it was decided to define number of markers/cluster centers, using formula 1:
- $n = \max(\text{width}, \text{height})$ (1)
- where width refers to the image width resolution in pixels, and height – to image height resolution in pixels.

Based on the Table 1, from Results section, it was decided to select SLIC algorithm, because it meets requirements – computational time does not increase (or even decrease in some cases) for images with large resolution and number of created segments is enough for algorithm needs. Also it's based on the idea to create clusters-super pixels. Super pixels are groups of pixels that are connected with some semantic meaning, and it can be helpful for further clustering.

After image is segmented, its segments are to be unified in size. Because segments are of rectangular shape, area of each segment is calculated and after that average value is taken using formula 2 (Fig. 1):

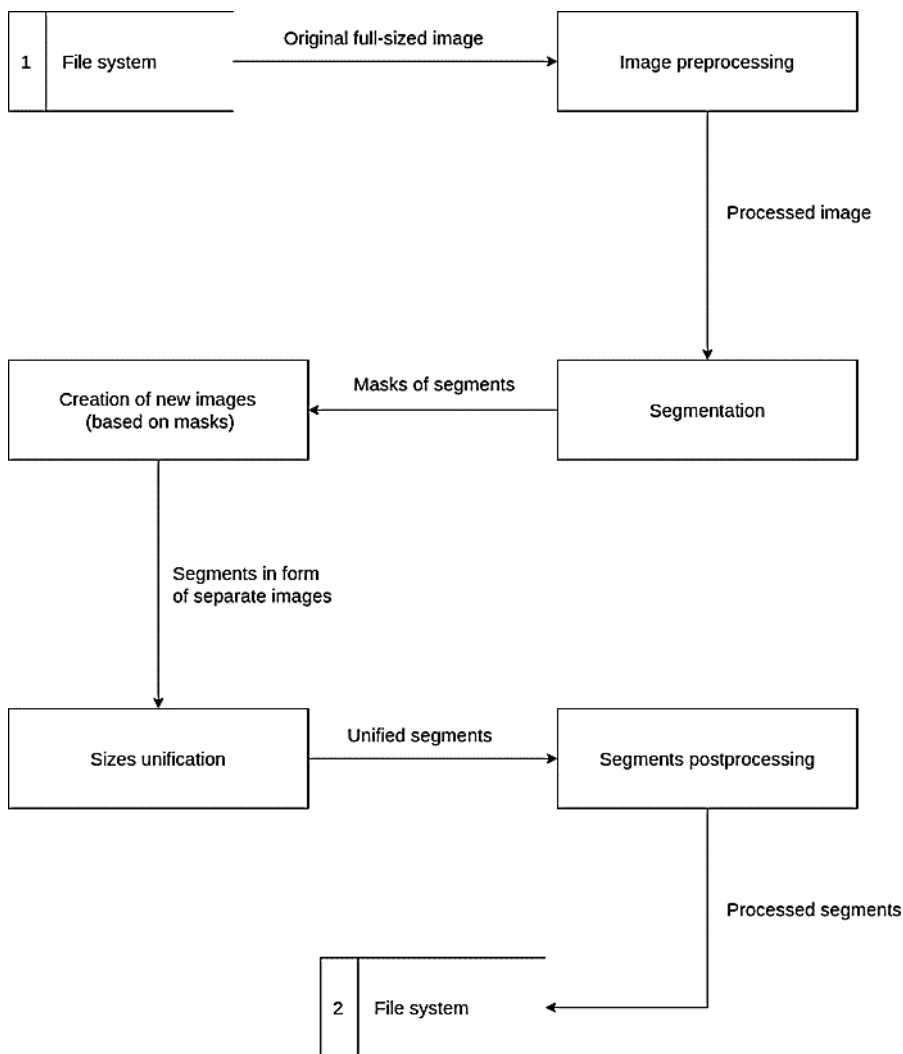


Fig. 1 DFD of preliminary image segmentation process

$$avg = \sum_0^{n-1} \frac{S_i}{n} \tag{1}$$

where :

n – number of segments

S_i – area of i segment,

After that segments are resized to square shape and dimensions are calculated according to avg.

The next step – clustering of received segments, between some (6) numbers of clusters, that represent different states of skin on and near wound and background.

To introduce improvements of the results on data accumulation, autoencoder model is used to extract features from segments.

Output of encoder part represents image features. These features are enough to re-create original image with sufficient accuracy and are used further to perform clustering. As mentioned above, segments dimensions are decreased using autoencoder and result of middle layer is passed to clustering layer.

On input autoencoder received image-vector of pixels of size 4356 (average area for segments from images with resolution 3104 × 1746 px). Output of encoder decreased vector to 90.

This vector of 90 features was clustered, using K-means method and clustering into 6 clusters, and cluster centroids were used as input for clustering layer, that performed clustering using DEC method – form of unsupervised learning, kind of `self-training`. Model training is made by refining the clusters, by learning from their high confidence assignments.

RESULTS

Table 1 show results of testing of different segmentation algorithms.

Table 1 Testing of segmentation algorithms

Algorithm	Created segments	Execution time	Predefined markers/ cluster centers
SLIC	224	6.87	250
SLIC	934	7.85	1000
SLIC	3557	7.15	3800
SLIC	2863	7.12	3104
Watershed	252	13.22	250
Watershed	3108	16.14	3104
Felzenswalb	3715	25.98	-
Quickshift	18212	67.54	-

Figure 2 show image with wound, segmented, using SLIC method, with highlighted segments border (left image) and clustered (right image).

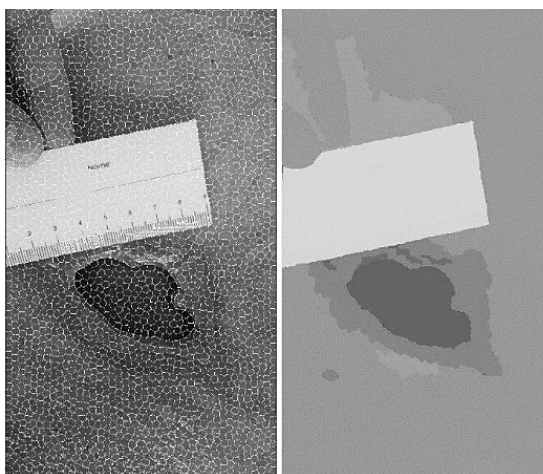


Fig. 2 Segmented (left) and clustered (right) image with wounds

Results are achieved after training of model on segments, created from 150 images.

Segments on the left image cover different areas with good precision. Right image has been marked with different shades of grey clusters. There are some

collisions with created clusters. Figure 3 shows segmentation of the same image, using only SLIC method.



Fig. 3 Segmentation using SLIC method

DISCUSSION

Based on the results it can be stated that neural networks can be used while solving image segmentation problem. Usage of neural networks or DNN, as a partial case, provides improvement of the results during data accumulation. That distinguishes described approach from traditional ones, where intermediate results can't be saved in some sort of trained model. But, unlike classification case, there are not so many DNN architectures, that can be used in unsupervised learning. Autoencoder, as one of the simplest, plays an important role in solving clustering problem and is present as an important part of many methods. DEC – is not an exception, it also adds clustering layer that is trained in `self-training` way. This can cause errors (for example if some mismatch occurs at the early beginning of prediction, it will lead to further errors), but reduce human efforts – there is no need for creation of labeled datasets. There are also some other approaches that can be tested, despite of the `self-training` neural network layer. Adaptation step is mandatory, clustering of super pixels, but not just pixels leads to a reduction of a required size of training dataset.

CONCLUSION

In the conclusion I would like to highlight some drawbacks, which were not mentioned in the Discussion section of the described method, and possible steps for improvement:

- high dependence on training data. The heterogeneity of training images affects the allocation of features and segmentation. For better results, more homogeneous images are needed,

- fixed size and shape segments are needed for autoencoder. Usage of images with lower or higher resolutions than those that were used while model training may lead to errors, because segments may be corrupted during resizing,
- low speed performance.

Further steps for method improvement:

- usage of GPU-parallelized computations to increase speed performance,
- data accumulation not only with new images, but also using modification of existing ones (applying some transformations on existing images, etc.),
- introduction of human interaction into method (for example usage of active contours method).

Also some other approaches of clustering can be investigated, to be used after feature extraction step.

Method can be integrated into applications for doctors, this can lead to a speed up of data accumulation and facilitate doctors work.

ACKNOWLEDGEMENTS

This research has been carried out within the framework of the project NAWA No PPI/APM/2018/1/00049/U/001, which is financed by the Polish National Agency for Academic Exchange.

REFERENCES

- Jansen, S., Rachel, K. (2016). The Evolving Field of Wound Measurement Techniques: A Literature Review. [online] Available at: <https://www.woundsresearch.com/article/evolving-field-wound-measurement-techniques-literature-review> [Accessed 31 June 2019].
- WoundEducators.com (2012). Wound measurement techniques. [online] Available at: <https://woundeducators.com/wound-measurement-techniques/> [Accessed 31 June 2019].
- Flanagan M. (2003) Improving accuracy of wound measurement in clinical practice. *Ostomy Wound Manage.*, pp. 28-40.
- Lukavetskyi, O., Stoianovskyi, I. and Farmaha, T. (2017). Computer program for vulnerometry (in Ukrainian), Kharkiv Surgical School: Medical Scientific and Practical Journal, pp. 145-147.
- Jaworski, N., Farmaha, I., Farmaha, T., Savchyn, V. and Marikutsa, U. (2018). Implementation features of wounds visual comparison subsystem. XIVth International Conference perspective technologies and methods in mems design., pp. 114-117.
- Nejati, H., Ghazijahani, H. and Abdollahzadeh, M. (2018). Fine-grained wound tissue analysis using deep neural network [online] Available at: <https://arxiv.org/pdf/1802.10426.pdf> [Accessed 31 June 2019].
- Lai, Z. and Deng, H. (2018) Medical Image Classification Based on Deep Features Extracted by Deep Model and Statistic Feature Fusion with Multilayer Perceptron. [online] Computational Intelligence and Neuroscience Available at: <https://www.hindawi.com/journals/cin/2018/2061516/> [Accessed 31 June 2019].
- Kumar, K., Reddy, B. (2014) Wound image analysis classifier for efficient tracking of wound healing status. [online] An International Journal. Available at: <http://aircconline.com/sipij/V5N2/5214sipij02.pdf> [Accessed 31 June 2019].
- Hemalatha, R., Thamizhvani, T. and Dhivya, A. (2018). Active Contour Based Segmentation Techniques for Medical Image Analysis. [online] Available at: <https://www.intechopen.com/books/medical-and-biological-image->

analysis/active-contour-based-segmentation-techniques-for-medical-image-analysis [Accessed 31 June 2019].

Xie, J., Girshick, R. and Farhadi, A. (2016). Unsupervised Deep Embedding for Clustering Analysis. [online] Available at: <https://arxiv.org/pdf/1511.06335.pdf> [Accessed 31 June 2019].

Abstract.

Classic methods of measurement and analysis of the wounds on the images are very time consuming and inaccurate. Automation of this process will improve measurement accuracy and speed up the process. Research is aimed to create an algorithm based on machine learning for automated segmentation based on clustering algorithms Methods. Algorithms used: SLIC (Simple Linear Iterative Clustering), Deep Embedded Clustering (that is based on artificial neural networks and k-means). Because of insufficient amount of labeled data, classification with artificial neural networks can't reach good results. Clustering, on the other hand is an unsupervised learning technique and doesn't need human interaction. Combination of traditional clustering methods for image segmentation with artificial neural networks leads to combination of advantages of both of them. Preliminary step to adapt Deep Embedded Clustering to work with bio-medical images is introduced and is based on SLIC algorithm for image segmentation. Segmentation with this method, after model training, leads to better results than with traditional SLIC.

Keywords: clustering, segmentation, machine learning, neural networks, wounds