Software Impacts 11 (2022) 100228

Contents lists available at ScienceDirect

Software Impacts

journal homepage: www.journals.elsevier.com/software-impacts

Original software publication

CGFFCM: A color image segmentation method based on cluster-weight and feature-weight learning

Amin Golzari Oskouei, Mahdi Hashemzadeh *

Faculty of Information Technology and Computer Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran
Artificial Intelligence and Machine Learning Research Laboratory, Azarbaijan Shahid Madani University, Tabriz, Iran

ARTICLE INFO

Keywords:
Clustering
Color image segmentation
Cluster weighting
Feature weighting

ABSTRACT

CGFFCM (Cluster-weight and Group-local Feature-weight learning in Fuzzy C-Means) is a clustering-based color image segmentation approach. It applies an automatic cluster weighting strategy to mitigate the initialization sensitivity and a group-local feature weighting technique to improve the clustering accuracy. In addition, it exploits an efficient combination of image features, consisting of eight features from three different groups (i.e., local homogeneity, CIELAB color space, and texture), to increase the image segmentation quality. CGFFCM also utilizes the imperialist competitive algorithm to optimize its feature weighting process. An open-source Matlab implementation of CGFFCM is available.

Code metadata

Current code version	*V1.0*
Permanent link to code/repository used for this code version	https://github.com/SoftwareImpacts/SIMPAC-2021-197
Permanent link to reproducible capsule	https://codeocean.com/capsule/8809672/tree/v1
Legal Code License	MIT license
Code versioning system used	Created with Matlab R2016b
Software code languages, tools, and services used	Matlab
Compilation requirements, operating environments & dependencies	
If available Link to developer documentation/manual	https://github.com/Amin-Golzari-Oskouei/CGFFCM
Support email for questions	a.golzari@tabrizu.ac.ir, hashemzadeh@azaruniv.ac.ir

1. Introduction

Image segmentation is the process of segmenting a digital image into several parts. The aim is to simplify or change the representation of an image into something more meaningful and easier to analyze. In computer vision, the image segmentation process is used in various applications such as object detection [1], people counting [2], fire detection [3], medical image analysis [4], and other industrial and monitoring applications [5,6].

A large number of different techniques appear in the literature for image segmentation which can be classified into three main categories: (1) threshold-based methods [7,8], (2) region-based methods [9], and (3) clustering-based methods [9–15]. Among these approaches, the clustering-based methods, especially k-partitional algorithms such as k-means and fuzzy c-means (FCM), are trendy and useful because of preserving more image information, easy and fast implementation [9], and the good results they provide [13,16].

In this work, we provide an open-source Matlab implementation of the CGFFCM (Cluster-weight and Group-local Feature-weight learning in Fuzzy C-Means) method for the color image segmentation task, which has been recently published in [16].

2. CGFFCM

The basis of CGFFCM [16] is inspired by our clustering algorithm presented in [17]. In CGFFCM, an automatic cluster weighting scheme is performed to reduce the sensitivity to the clustering initialization, and a group-local feature weighting strategy is applied to better image segmentation. In addition, the clustering process is combined with the Imperialist Competitive Algorithm (ICA) [18] to optimize the

DOI of original article: https://doi.org/10.1016/j.asoc.2021.108005.

The code (and data) in this article has been certified as Reproducible by Code Ocean: (https://codeocean.com/). More information on the Reproducibility Badge Initiative is available at https://www.elsevier.com/physical-sciences-and-engineering/computer-science/journals.

* Correspondence to: Azarbaijan Shahid Madani University, Tabriz-Azarshahr Road, Tabriz, Post Code: 5375171379, Iran.
E-mail addresses: a.golzari@tabrizu.ac.ir (A. Golzari Oskouei), hashemzadeh@azaruniv.ac.ir (M. Hashemzadeh).

https://doi.org/10.1016/j.simpa.2022.100228
Received 23 December 2021; Received in revised form 2 January 2022; Accepted 9 January 2022

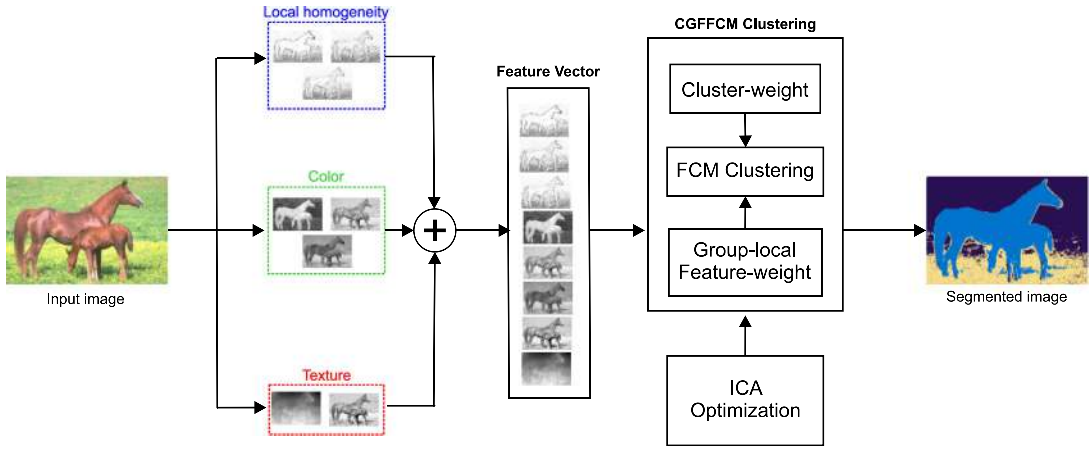

Fig. 1. The structure of CGFFCM.

feature weighting process. Moreover, an efficient combination of image features, consisting of eight features from three groups (i.e., local homogeneity, CIELAB color space, and texture), are used to have better segmentation results. Fig. 1 illustrates the structure of CGFFCM. As illustrated in this figure, the process begins by extracting the features from the input image. Then, the extracted features are entered into the CGFFCM clustering algorithm. The ICA method is also utilized by the CGFFCM to optimize its feature weighting process. The output is the segmented image.

In CGFFCM, the basic idea is that the importance of the extracted features is different for each image. A group of features may be important in general, but the importance of a sub-feature in a group may be less or more than other sub-features. The goal of the CGFFCM is to improve image segmentation quality by emphasizing important groups of features and sub-features, simultaneously. To this end, a group-local feature weighting technique is utilized during the clustering process. In this technique, unlike other feature weighting methods such as the local weighting techniques (feature has different weights in different clusters) [17,19] or global weighting ones(feature has same weights in all clusters) [17,19], the groups and the sub-features in each group get different weights depending on their importance degree in the clusters. Thus, weighting is done more effectively and efficiently, yielding better segmentation results.

In addition, to reduce the sensitivity to initial centers, a cluster weighting technique is used. The weight of the clusters is extracted dynamically while taking the importance of the features and sub-features into account. More details regarding the feature weighting and cluster weighting mechanisms are available in [16,17].

3. How to use CGFFCM

The implementation of CGFFCM has been done in Matlab. The modules and submodules that existed in the implemented program are shown in Fig. 2. As shown in this figure, it includes the main function called *Demo* and three modules (i.e., the *FeatureExtractor*, *CGFFCM*, and *Evaluator*), each of which also has some submodules.

The main modules of the program are described below:

- **Demo:** this module contains a main script in which other modules are called. By running the demo, the input image (or images) is loaded first. Then, two general parameters, i.e., the *maximum number of iterations* and the *number of algorithm restarts*, are entered by the user.
- **FeatureExtractor:** this module is responsible for extracting the features from the input image. The output of this module is an $N \times 8$ matrix, where N indicates the number of image pixels and 8 is the number of extracted features. The first three features are for local homogeneity, the following three are for color space, and the last two are for texture features.

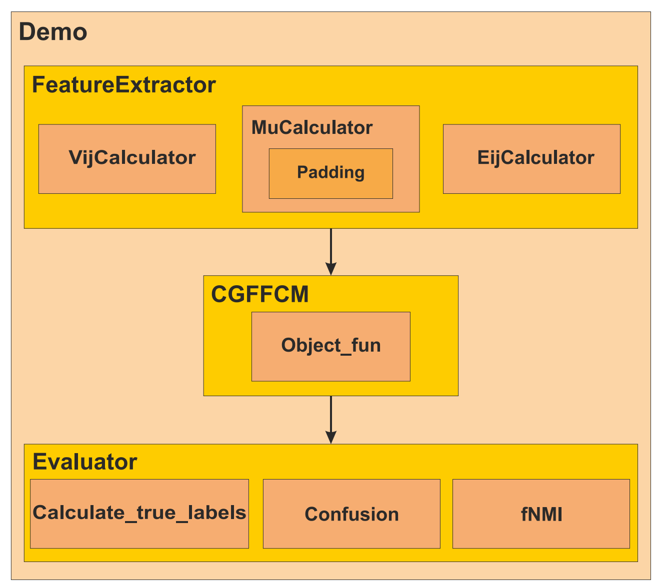

Fig. 2. The modules and submodules that existed in the implemented program.

- **CGFFCM:** this function implements the CGFFCM method. Its input comes from the *FeatureExtractor* module, and its outputs are as follows:
  - **Cluster_elem:** is an $N \times K$ matrix that represents the membership of nth pixel to the kth region in the input image;
  - **M:** is a matrix of cluster centers ($K \times D$), where K and D are the number of clusters and the number of features, respectively;
  - **EW_history:** is the value of the objective function stored during the algorithm iterations;
  - **Z:** is a $K \times D$ matrix that represents the weight of the dth feature in the kth cluster;
  - **W:** is a vector of length K that represents the weights of clusters.
- **Evaluator:** this function is also called by the main function *Demo*. The output of this function is the performance metrics calculated on the result segmented image.

4. Impact

CGFFCM emerged from our last research work in [16,17] and has been proving to be a robust tool for different image segmentation

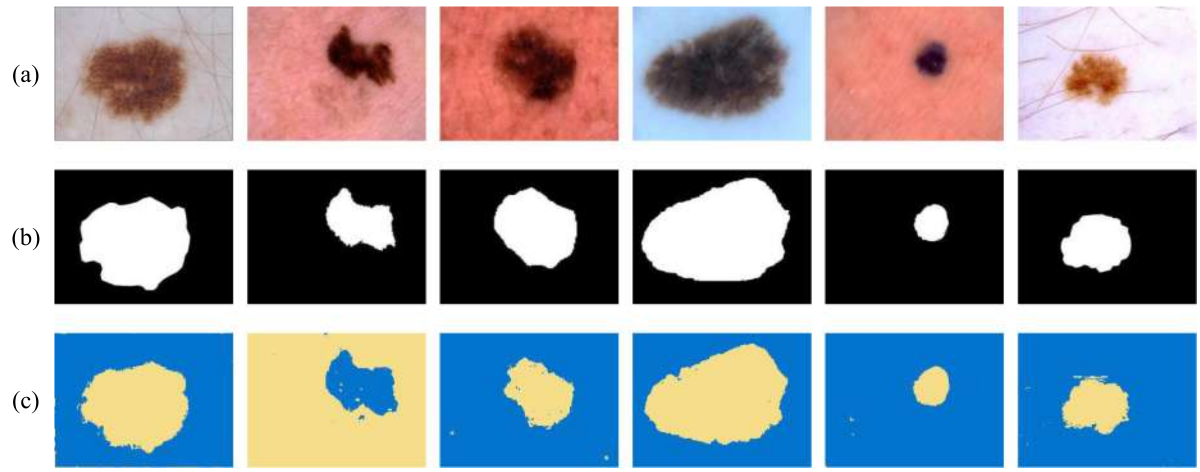

Fig. 3. Example outputs of CGFFCM on lesion segmentation application. (a) Input image, (b) ground truth, and (c) segmented image by CGFFCM.

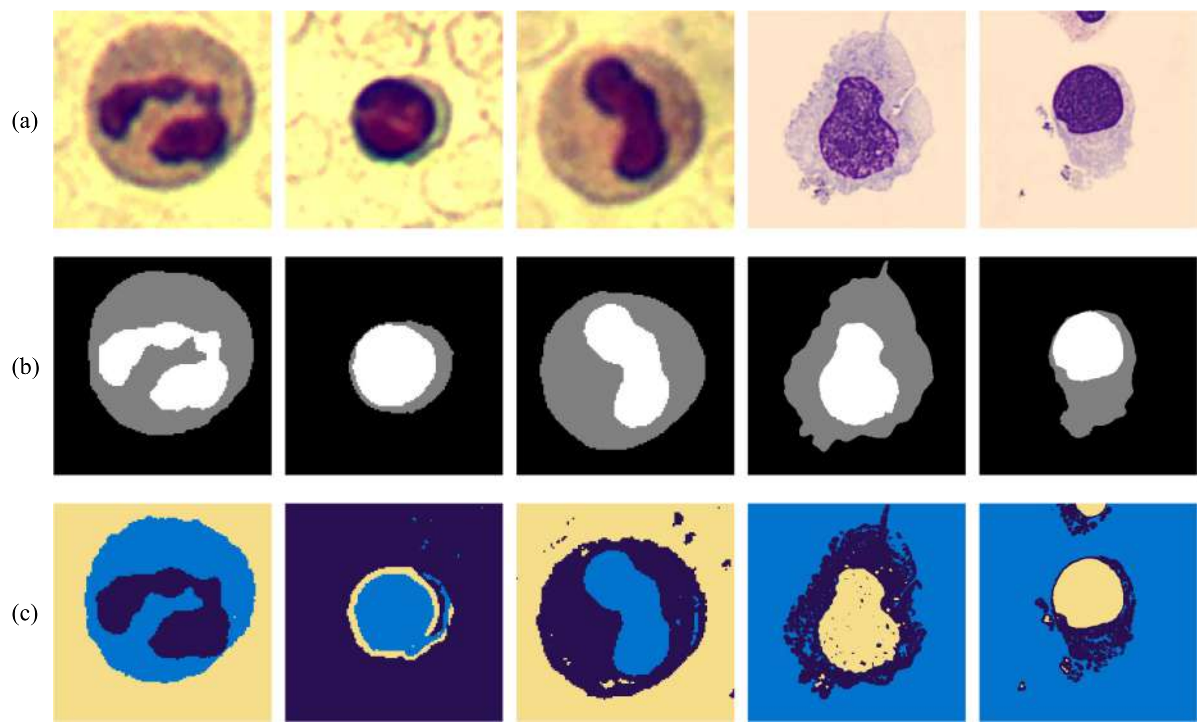

Fig. 4. Example outputs of CGFFCM on white blood cells segmentation application. (a) Input image, (b) ground truth, and (c) segmented image by CGFFCM.

tasks. In this section, some applications of implemented software are introduced.

Skin lesion segmentation: Skin lesion segmentation, which is one of the medical image segmentation tasks, is essential for detecting melanoma. Melanoma, the most life-threatening skin cancer, can suddenly occur on normal skin without warning and develop on a pre-existing lesion. Therefore, lesions must be carefully monitored. The automatic segmentation of skin lesions in the dermoscopic images accelerates melanoma diagnosis [20]. We run the CGFFCM software on the ISIC (International Skin Imaging Collaboration) 2017 dataset [21]. Some of the images in this dataset, their grand-truth, and the obtained results are shown in Fig. 3. As shown in this figure, the skin lesions are detected well. The detection accuracy obtained in this experiment and the output images show that the implemented software is a suitable tool for this application.

White blood cells segmentation: Determination of white blood cells is vital for diagnosing many diseases. The number of white blood cells and morphological changes provide valuable information for the positive results of the diseases such as Acute Lymphocytic Leukemia (ALL) [22]. We run the CGFFCM software on the WBC (White Blood Cell) dataset [23]. Some of the images in this dataset, their grand-truth, and the obtained results are shown in Fig. 4. As shown in this figure, the implemented program is also helpful for detecting white blood cells. The results obtained in this experiment and the outputs indicate that CGFFCM software is a suitable tool for this application.

Aerial image segmentation: Today, the automatic detection of objects in satellite imagery is of interest to machine vision researchers. Automatic detection of buildings, natural areas, roads, etc, from aerial images, has been an essential task for many applications such as mapping, GIS (Geographical Information Systems) database generation, land use analysis, change detection, and urban monitoring [24]. We run the CGFFCM software on some sample aerial images, some of which are shown in Fig. 5. As shown in this figure, from these complex images, well-segmented images are obtained. The results obtained in

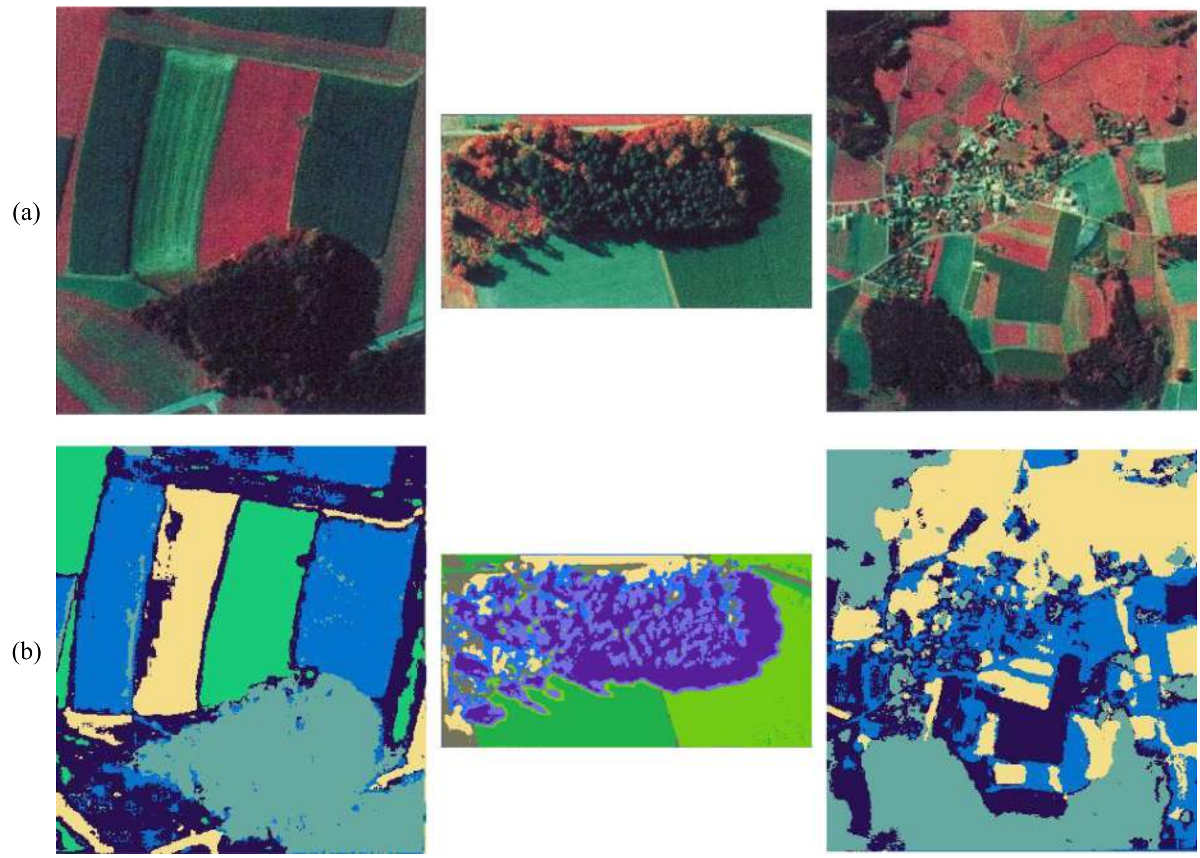

Fig. 5. Example outputs of CGFFCM on aerial image segmentation application. (a) Input image, (b) segmented image by CGFFCM.

this experiment show that CGFFCM software is also a suitable tool for this application.

The other applications of CGFFCM and the experimental results reported in [16] confirm that the CGFFCM achieves higher performance than the state-of-the-art methods for the color image segmentation tasks. The provided source code is highly efficient, extensible, easily understandable, modular, and specially designed to deal with the high-dimension challenge in clustering tasks. It is helpful for beginners as well as advanced researchers to explore new ideas and for machine learning professionals. To adapt the code to deal with other cluster-based tasks, only minor changes to the source code are required (such as modification of the distance function). In the future, the authors plan to further improve CGFFCM by focusing on implementing some modules, such as a function for defining various distance metrics.

CRediT authorship contribution statement

Amin Golzari Oskouei: Conceptualization, Implementation, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Mahdi Hashemzadeh:** Conceptualization, Implementation, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] F. Yang, H. Fan, P. Chu, E. Blasch, H. Ling, Clustered object detection in aerial images, in: Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 8311–8320.

[2] M. Hashemzadeh, N. Farajzadeh, Combining keypoint-based and segment-based features for counting people in crowded scenes, Inform. Sci. 345 (2016) 199–216, http://dx.doi.org/10.1016/j.ins.2016.01.060, 2016/06/01/ 2016.

[3] M. Hashemzadeh, A. Zademehdi, Fire detection for video surveillance applications using ICA K-medoids-based color model and efficient spatio-temporal visual features, Expert Syst. Appl. 130 (2019) 60–78, http://dx.doi.org/10.1016/j.eswa.2019.04.019, 09/15/ 2019.

[4] M. Hashemzadeh, B. Adlpour Azar, Retinal blood vessel extraction employing effective image features and combination of supervised and unsupervised machine learning methods, Artif. Intell. Med. 95 (2019) 1–15, http://dx.doi.org/10.1016/j.artmed.2019.03.001, /04/01/ 2019.

[5] M. Hashemzadeh, N. Farajzadeh, A machine vision system for detecting fertile eggs in the incubation industry, Int. J. Comput. Intell. Syst. 9 (5) (2016) 850–862.

[6] S. Wazarkar, B.N. Keshavamurthy, A survey on image data analysis through clustering techniques for real world applications, J. Vis. Commun. Image Represent. 55 (2018) 596–626, http://dx.doi.org/10.1016/j.jvcir.2018.07.009, /08/01/ 2018.

[7] L. He, S. Huang, Modified firefly algorithm based multilevel thresholding for color image segmentation, Neurocomputing 240 (2017) 152–174, http://dx.doi.org/10.1016/j.neucom.2017.02.040, /05/31/ 2017.

[8] S. Pare, A. Kumar, V. Bajaj, G.K. Singh, A multilevel color image segmentation technique based on cuckoo search algorithm and energy curve, Appl. Soft Comput. 47 (2016) 76–102, http://dx.doi.org/10.1016/j.asoc.2016.05.040, /10/01/ 2016.

[9] L. Feng, H. Li, Y. Gao, Y. Zhang, A color image segmentation method based on region salient color and fuzzy C-means algorithm, Circuits Systems Signal Process. 39 (2) (2020) 586–610, http://dx.doi.org/10.1007/s00034-019-01126-w, 02/01 2020.

[10] K.S. Tan, N.A. Mat Isa, W.H. Lim, Color image segmentation using adaptive unsupervised clustering approach, Appl. Soft Comput. 13 (4) (2013) 2017–2036, http://dx.doi.org/10.1016/j.asoc.2012.11.038, /04/01/ 2013.

[11] H.-J. Xing, M.-H. Ha, Further improvements in feature-weighted fuzzy C-means, Inform. Sci. 267 (2014) 1–15, http://dx.doi.org/10.1016/j.ins.2014.01.033, /05/20/ 2014.

[12] Z. Zhou, X. Zhao, S. Zhu, K-harmonic means clustering algorithm using feature weighting for color image segmentation, Multimedia Tools Appl. 77 (12) (2018) 15139–15160, http://dx.doi.org/10.1007/s11042-017-5096-9.

[13] T.R. Farshi, J.H. Drake, E. Özcan, A multimodal particle swarm optimization-based approach for image segmentation, Expert Syst. Appl. 149 (2020) 113233, http://dx.doi.org/10.1016/j.eswa.2020.113233, 2020/07/01.

[14] F. Zhao, Z. Zeng, H. Liu, R. Lan, J. Fan, Semisupervised approach to surrogate-assisted multiobjective kernel intuitionistic fuzzy clustering algorithm for color image segmentation, IEEE Trans. Fuzzy Syst. 28 (6) (2020) 1023–1034, http://dx.doi.org/10.1109/TFUZZ.2020.2973121.

[15] L.H. Son, T.M. Tuan, Dental segmentation from X-ray images using semi-supervised fuzzy clustering with spatial constraints, Eng. Appl. Artif. Intell. 59 (2017) 186–195, http://dx.doi.org/10.1016/j.engappai.2017.01.003, /03/01/2017.

[16] A.G. Oskouei, M. Hashemzadeh, B. Asheghi, M.-A. Balafar, CGFFCM: CLuster-weight and group-local feature-weight learning in fuzzy C-means clustering algorithm for color image segmentation, Appl. Soft Comput. (2021) 108005, http://dx.doi.org/10.1016/j.asoc.2021.108005, 2021/10/26.

[17] M. Hashemzadeh, A. Golzari Oskouei, N. Farajzadeh, New fuzzy C-means clustering method based on feature-weight and cluster-weight learning, Appl. Soft Comput. 78 (2019) 324–345, http://dx.doi.org/10.1016/j.asoc.2019.02.038, 05/01/2019.

[18] E. Atashpaz-Gargari, C. Lucas, Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition, in: 2007 IEEE Congress on Evolutionary Computation, Ieee, 2007, pp. 4661–4667.

[19] A. Golzari Oskouei, M.A. Balafar, C. Motamed, FKMAWCW: CAtegorical fuzzy k-modes clustering with automated attribute-weight and cluster-weight learning, Chaos Solitons Fractals 153 (2021) 111494, http://dx.doi.org/10.1016/j.chaos.2021.111494, 2021/12/01.

[20] R. Kaymak, C. Kaymak, A. Ucar, Skin lesion segmentation using fully convolutional networks: A comparative experimental study, Expert Syst. Appl. 161 (2020) 113742, http://dx.doi.org/10.1016/j.eswa.2020.113742, 2020/12/15.

[21] N.C.F. Codella others, Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC), in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), 4-7 2018, 2018, pp. 168–172, http://dx.doi.org/10.1109/ISBI.2018.8363547.

[22] M. Habibzadeh, M. Jannesari, Z. Rezaei, H. Baharvand, M. Totonchi, Automatic White Blood Cell Classification using Pre-Trained Deep Learning Models: ResNet and Inception (Tenth International Conference on Machine Vision), SPIE, 2018.

[23] X. Zheng, Y. Wang, G. Wang, J. Liu, Fast and robust segmentation of white blood cell images by self-supervised learning, Micron 107 (2018) 55–71, http://dx.doi.org/10.1016/j.micron.2018.01.010, /04/01/2018.

[24] F. Dornaika, A. Moujahid, Y. El Merabet, Y. Ruichek, Chapter 32 - a comparative study of image segmentation algorithms and descriptors for building detection, in: P. Samui, S. Sekhar, V.E. Balas (Eds.), Handbook of Neural Computation, Academic Press, 2017, pp. 591–606.