

Liver Segmentation

A Weakly End-to-End Supervised Model

<https://doi.org/10.3991/ijoe.v16i09.15159>

Youssef Ouassit ^(✉), Soufiane Ardchir, Reda Moulouki
Faculty of Sciences Ben M'sick, Casablanca, Morocco
ouassit.youssef@gmail.com

Mohammed Yassine El Ghoumari
International School of Marketing and Management
Casablanca, Casablanca, Morocco

Mohamed Azouazi
Faculty of Sciences Ben M'sick Casablanca, Casablanca, Morocco

Abstract—Liver segmentation in CT images has multiple clinical applications and is expanding in scope. Clinicians can employ segmentation for pathological diagnosis of liver disease, surgical planning, visualization and volumetric assessment to select the appropriate treatment. However, segmentation of the liver is still a challenging task due to the low contrast in medical images, tissue similarity with neighbor abdominal organs and high scale and shape variability. Recently, deep learning models are the state of art in many natural images processing tasks such as detection, classification, and segmentation due to the availability of annotated data. In the medical field, labeled data is limited due to privacy, expert need, and a time-consuming labeling process. In this paper, we present an efficient model combining a selective pre-processing, augmentation, post-processing and an improved Seg Caps network. Our proposed model is an end-to-end learning, fully automatic with a good generalization score on such limited amount of training data. The model has been validated on two 3D liver segmentation datasets and have obtained competitive segmentation results.

Keywords—Segmentation, Capsules Network, Deep Learning, Medical images and Liver volumetry.

1 Introduction

Liver segmentation refers to the delimitation of the hepatic zone in the abdomen, it plays an important role in the diagnostic process of the liver. Precise liver segmentation is essential for many clinical applications such as computer-aided diagnosis and computer-assisted surgery. However, manually segmenting the liver is subjective, not very reproducible, takes time and can only be done by experts [1]. Therefore, automated segmentation is desired to achieve precise hepatic segmentation on a large

scale. Despite advances in computer vision, the automatic segmentation of the abdominal organs and especially of the liver remains a difficult problem due to the complex shape, high shape variation between patients, the low boundaries between the adjacent organs, the complexity of the background and the low contrast in this type of images. Recently deep learning has been very successful in several areas more precisely on image processing tasks, this success is due to the high availability of labeled data necessary for these algorithms. In the medical field it is difficult to have a labeled dataset with sufficient amount of data, and this represents a critical issue for applying deep learning models. To meet these challenges, we are introducing an end-to-end model for liver segmentation using the capsules networks. Capsules networks have shown better generalization in the tasks of images classification in a limited amount of annotated data, they have also shown their robustness in the face of view and scale changes[2]. This document is organized as follows. In section 2, we describe the related work of learning the characteristics and frameworks proposed for the segmentation of the hepatic organs. In section 3, we present the proposed model for liver segmentation. In section 4, we report the results of the experience of our approach. Then our conclusion is presented in section 5.

2 Related Works

Many segmentation methods [3] and strategies have been proposed, we classify them in three categories: manual, semi-automatic and full-automatic.

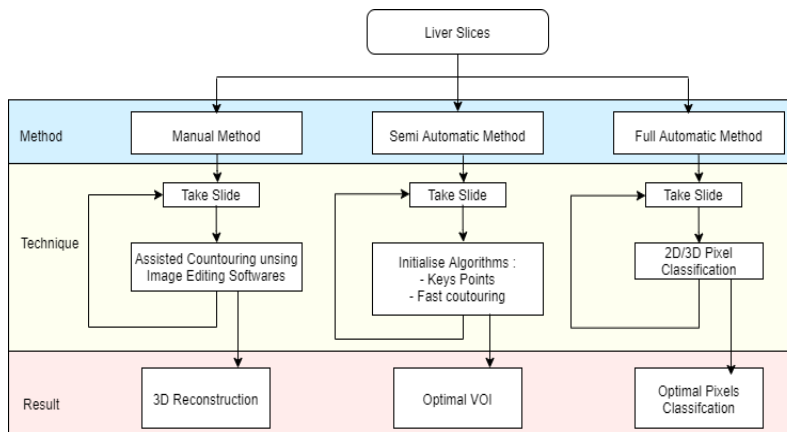


Fig. 1. Segmentation strategies

Manual strategy relies on user interaction via drawing and contouring the liver boundaries for each scan slice, semi-automatic strategy relies on developed hand-craft features extractors and require user initialization of the algorithm and the full-automatic strategy most case are based on artificial intelligence to calculate the likelihood that a given pixel belongs to the object of interest.

Table 1. Strategies strengths and weakness

	Manual	Semi-automatic	Full-automatic
Strengths	High Accuracy, Robustness	Fast Reproducibility	Fast Reproducibility
Weakness	User interaction, Time consuming, Subjectivity	Coarse initialization, Operator dependent	Huge labeled data, Complexity

In Full-automatic strategy Deep learning models have resulted the high accuracy in many application such as classification, recognition and segmentation. The most popular deep learning-based segmentation algorithms are CNN wise pixels classification or auto-encoders architecture. [4] introduced U-Net an alternative CNN-based pixel label prediction and was the first designed and applied in 2015 to process biomedical images. Following this, a lot of works have been done based encoder-decoder structure, with dense connections, skip connections, residual blocks [5], SegNet [6], RefineNet [7], and DeepLab [8]. However, the success of this approaches is in the huge availability of labelled data due the big number of trainable parameter over 130 Million.

3 Capsule Networks Intuition

Works based on the CNN architecture has shown high performance and widely used in computer vision,. However CNN uses max pooling layers to reduce the receptive field, these layers make the network invariant to small translations of objects but they cause the loss of spatial information in the deep layers, while spatial information is important in the task of segmentation. In addition, CNN networks require a huge amount of tagged data for learning, which is not available in the medical field due to privacy and the need for a of experts. Recently Sabour and Hinton [9] introduced the idea of capsule networks that work well on limited labeled data, where a capsule is a set of neurons that learn the pose vector (spatial information, magnitude, prevalence, ...), the length of the vector is the probability of the object's presence. The capsules are equivariant and consist of a neural network which accepts and delivers vectors as opposed to the scalar values of CNNs.

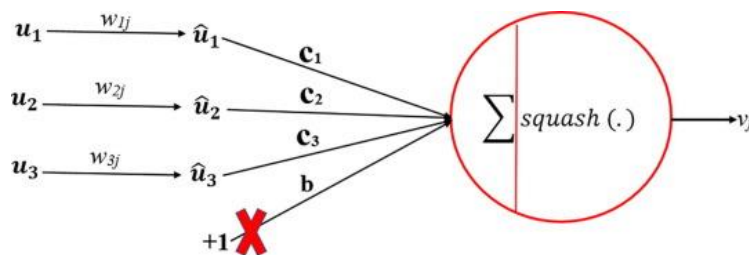


Fig. 2. Representation of a capsule from Sabour paper [9].

Several pros motivate us to use capsules: Viewpoint Invariance, Fewer Parameters, Better generalization to new viewpoints and Defense against white-box adversarial

attacks [2], [10]. Three general methods of capsule implementations exist in literature. These are transforming auto-encoders [11], vector capsules based on dynamic routing [9] and matrix capsules based on expectation-maximization routing [12].

4 Contribution

In our work we proposed a model for ct liver segmentation based on SegCaps [13] network using convolution layers, max-pooling layers, transposed convolution layer, dropout, primary capsule layer, digital capsule layer and fully connected layers layers. To avoid spatial information loss, skip connections are used.

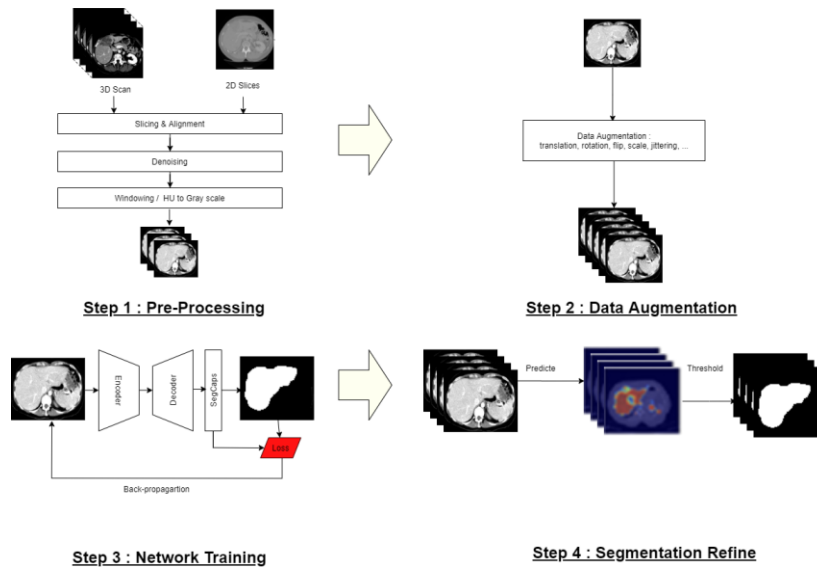


Fig. 3. Model workflow

4.1 Dataset

The datasets contain two collections of CT images. The first database is from CHAO challenge [14] and contains 40 different patients around 90 slices for each patient. Patients in this collection are potential liver donors, who have healthy (no tumors, lesions or any other diseases) liver. The patient orientation and alignment is the same for all slices. Each data set has a resolution of 512x512, x-y spacing between 0.7-0.8 mm and having 3 to 3.2 mm inter-slice distance. The second collection is from SLiver07 challenge contains 20 training 3D CT scans, data is stored in Meta format containing an ASCII readable header and a separate raw image data file, to fit this collection into the network each subject was sliced into 2D images based on axial direction. In order to enhance the generality of model, these two datasets are aligned

and randomly merged to create a new dataset which will be used in the proposed model. Example image data from these three datasets are shown in Fig. 4.

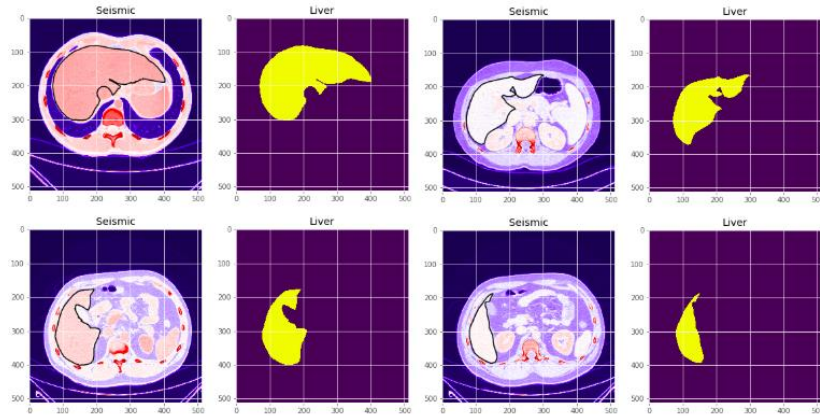


Fig. 4. Dataset images samples.

4.2 Preprocessing

In order to get a high hit rate of liver segmentation in CT images, many preprocessing techniques on the input images has been adopted. First we cleaned the images from noise and artifacts, a median filter is used to remove the outliers and anisotropic diffusion filter is applied to each slice for noise reduction and boundary preservation then we exclude the pixels out of range $[-1000, 2000]$. For normalization and contrast enhancement, CT values are mapped from the soft tissue range to the gray scale $[0, 255]$ and to improve the viewing of images with low contrast stretching is applied to windowed images. The DICOM images come from two different challenges and comprise an heterogeneous sets of Liver CT images that were captured by several technicians with different machines a linearization process is applied to have same orientation.

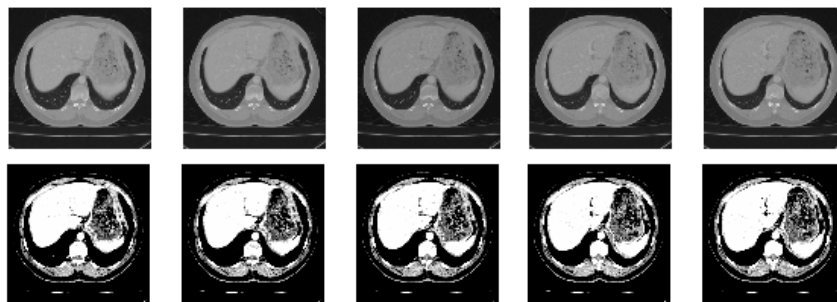


Fig. 5. Images preprocessing results

4.3 Data augmentation

To avoid the overfitting issue and deal with the limited data, one of the most used techniques is the data augmentation. Our model counts over than 1 Million of trainable parameters which need a big dataset for efficient generalization. For this purpose, we suggest to combine two databases CHAOS[14] and Sliver07, to obtain more than 7000 images. However, this number of images is not enough, then we artificially generate new slices. We used classical augmentation techniques on gray-scale images include mostly affine transformations such as crop, addition of noise and elastic deformation (translation, rotation, flip, scale, jittering, sharpening filters, adding mean-images based on clustering techniques), we avoided transformations that cause shape deformation like shearing to preserve the liver characteristics.

To decide on reasonable parameters, such as rotation angle and scaling percentage, images were augmented and visually evaluated. Two conditions were set: all anatomical structures had to be preserved inside the borders of the image and the anatomical structures should not be deformed unrealistically.

The ranges of the parameters are summarized below.

- Random translation in both x-axis and y-axis in the ranges limited by the image borders.
- The rotation angle in the range (-15,15) degrees.
- Scaling from 0.6 to the maximum scaling that could be conducted without breaking the border condition, but never exceeded 1.2.
- The elastic deformation was made randomly to a degree that the anatomical structures of the liver remained realistically intact.

4.4 Model and training

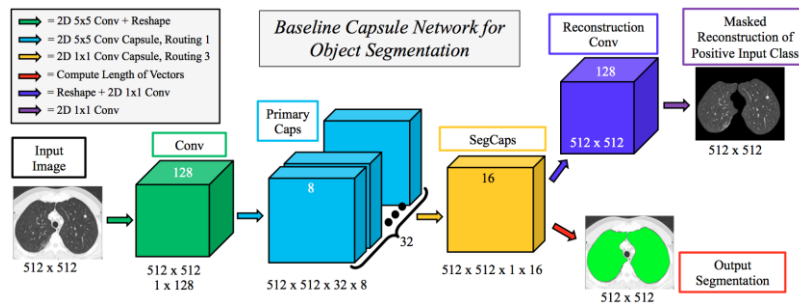


Fig. 6. Seg Caps architecture [13].

The input to the SegCap network is a 512x512 pixel image, in this case, a slice of a CT Scan. This image is passed through a 2D convolutional layer which produces 16 feature maps of the same spatial dimensions. This output forms our first set of capsules, where we have a single capsule type with a grid of 512x512 capsules, each

of which is a 16 dimensional vector. This is then followed by our first convolutional capsule layer. Then this process is generalized to any given layer L in the network.

The training of this network with a 0.1 learning rate and 150 epochs result an overfitting in left learning curve in figure 7. Two modifications has been proposed to increase the network accuracy in such limited data and avoid overfitting:

- Dropout Layers have included after convolutional layers to decrease trainable parameter number.
- Initialize filter by hand crafted filters.

The learning transfer is an efficient way to avoid overfitting and to speed up the training algorithm. It consist of initializing the trainable parameters by the learning result of other models. In our work we proposed to initialize these parameters, in convolutional layers filters, using the most suitable hand-craft features extractors for the liver segmentation task. We essentially used:

- **Laplacian filter:** We proposed to initialize certain filters with the edge detector based on the Laplacian, this operator detects the edges of the image by calculating the expression derived from the second order of the image. The Gaussian Laplacian is very useful because the second derivative is very sensitive to noise and this is useful for filtering image noise [15], [16].

$$\text{LoG}(x,y) = \frac{1}{\pi a^4} \left[1 - \frac{x^2+y^2}{2a^2} \right] e^{-\frac{x^2+y^2}{2a^2}} \quad (3)$$

Gaussian $a = 1.0$ is used.

- **Sobel filter:** The Sobel filter is an operator used for edge detection. It uses convolution matrices. The matrix of size 3×3 undergoes a convolution with the image to calculate approximations of the horizontal and vertical derivatives.

The right learning curve in figure 7, show how the model accuracy has been improved with new filters.

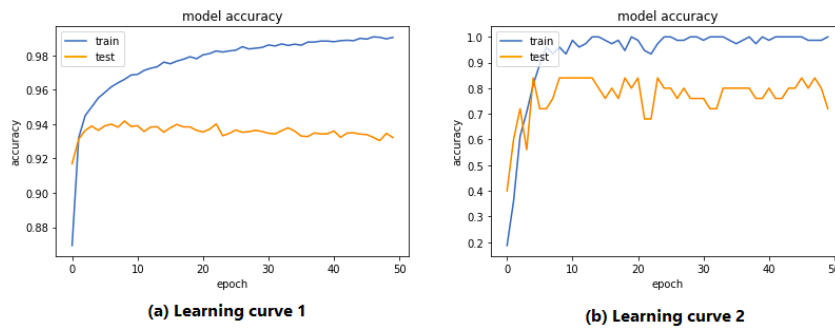


Fig. 7. Learning Curves.

4.5 Post-processing

Based on the result of the network, a frequencies map where each element is the likelihood for pixel to belong the liver, a threshold has been selected to binarize the result of segmentation we checked a threshold between 0.3 and 0.7, the optimal threshold selected was 0.6.

5 Evaluation and Results

For quantitative performance evaluation, several evaluation measures will be used as follows:

The accuracy metric is the ratio of the true positive and true negative pixels over all available pixels :

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

The Dice coefficient (DICE), known also as the overlap index, it's the metric the most used to validate medical volume segmentation [17]:

$$\text{Dice} = \frac{2TP}{2TP+FP+FN} \quad (5)$$

The Jaccard index (JAC) define the intersection between two sets divided by their union:

$$\text{Jac} = \frac{TP}{TP+FP+FN} \quad (6)$$

Table 2. Metrics comparison results

Model	Accuracy	Dice	Jaccard
Proposed Model	0.82	0.66	0.67
UNet	0.99	0.42	0.40
FCN	0.98	0.39	0.35

Next some segmentation results on test images is presented, first column represents the input image, the second column is the truth segmentation, the third column is the predicted segmentation and the last column is the predicted segmentation after post-processing:

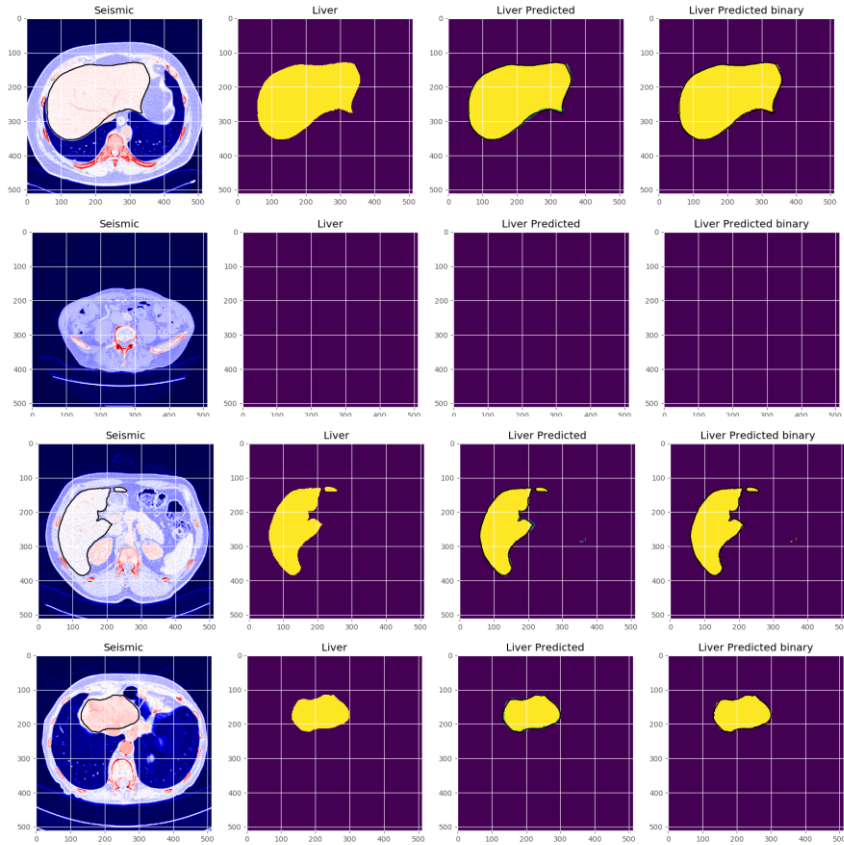


Fig. 8. Predictions results on test images.

6 Conclusion

In this paper, we propose a simple yet effective and efficient capsule neural network-based model to segment the liver in CT images. The network is built up SegCaps network. The dice metric of our method is 0.66, which is a highly promoting result in the limited annotated data. The proposed method also shows significantly good resistance to noise and acceptable inference. As a direction of future work, we envisage the extension of the model by new techniques to reduce free parameters to enhance performance and enable the use of different imaging modalities simultaneously.

7 References

- [1] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, (2019). Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges, *Journal of Digital Imaging*, vol. 32, no. 4, pp. 582–596, <https://doi.org/10.1007/s10278-019-00227-x>
- [2] M. Kwabena Patrick, A. Felix Adekoya, A. Abra Mighty, and B. Y. Edward, (2019). Capsule Networks – A survey, *Journal of King Saud University - Computer and Information Sciences*, <https://doi.org/10.1016/j.jksuci.2019.09.014>
- [3] Y. Ouassit, M. Azzouazi, and M. Y. EL Ghoumari, (2020). Segmentation Features for CT Scans: A Taxonomy, *International Journal of Psychosocial Rehabilitation*, vol. 24, no. 03, pp. 1044–1056, <https://doi.org/10.37200/ijpr/v24i3/pr200856>
- [4] O. Ronneberger, P. Fischer, and T. Brox, (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation, arXiv:1505.04597 [cs].
- [5] F. Yu, V. Koltun, and T. Funkhouser, (2017). Dilated Residual Networks, arXiv:1705.09914 [cs].
- [6] V. Badrinarayanan, A. Kendall, and R. Cipolla, (2015). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, arXiv:1511.00561 [cs].
- [7] G. Lin, A. Milan, C. Shen, and I. Reid, (2016). RefineNet: Multi-Path Refinement Networks for High-Resolution Semantic Segmentation, arXiv:1611.06612 [cs]. <https://doi.org/10.1109/cvpr.2017.549>
- [8] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, <https://doi.org/10.1109/tpami.2017.2699184>.
- [9] S. Sabour, N. Frosst, and G. E. Hinton, (2017). Dynamic Routing Between Capsules, arXiv:1710.09829 [cs].
- [10] R. Mukhometzianov and J. Carrillo, CapsNet comparative performance evaluation for image classification, p. 14.
- [11] G. E. Hinton, A. Krizhevsky, and S. D. Wang, (2011). Transforming Auto-Encoders, in *Artificial Neural Networks and Machine Learning – ICANN 2011*, vol. 6791, T. Honkela, W. Duch, M. Girolami, and S. Kaski, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 44–51. https://doi.org/10.1007/978-3-642-21735-7_6
- [12] G. Hinton, S. Sabour, and N. Frosst, (2018). MATRIX CAPSULES WITH EM ROUTING, p. 15.
- [13] R. LaLonde and U. Bagci, (2018). Capsules for Object Segmentation, arXiv:1804.04241 [cs, stat].
- [14] Ali Emre Kavur, M. Alper Selver, Oğuz Dicle, Mustafa Barış, and N. Sinem Gezer, (2019). CHAOS - Combined (CT-MR) Healthy Abdominal Organ Segmentation Challenge Data. Zenodo, 2019, 10.5281/zenodo.3431873.
- [15] D. Vikram Mutneja, (2015). Methods of Image Edge Detection: A Review, *Journal of Electrical & Electronic Systems*, vol. 04, no. 02, <https://doi.org/10.4172/2332-0796.1000150>.
- [16] G. Kumar and P. K. Bhatia, (2014). A Detailed Review of Feature Extraction in Image Processing Systems, pp. 5–12, 10.1109/ACCT.2014.74.
- [17] A. A. Taha and A. Hanbury, (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool, *BMC Medical Imaging*, vol. 15, no. 1, <https://doi.org/10.1186/s12880-015-0068-x>.

8 Authors

Youssef Ouassit is a PhD student in computer sciences at University Hassan 2, Casablanca, Morocco. He works also as a teacher in Preparatory Classes for High Schools. Email: ouassit.youssef@gmail.com

Soufiane Ardchir is a PhD student in computer sciences at University Hassan 2, Casablanca, Morocco. He works also as a teacher in Preparatory Classes for High Schools.

Reda Moulouki is a PhD student in computer sciences at University Hassan 2, Casablanca, Morocco.

Mohamed Azouazi is a professor of computer engineering at the University Hassan 2, Casablanca, Morocco. His area of expertise is big data and machine learning algorithms. Dr. Azouazi is a contributing author for more than 50 journal articles and conference papers.

Mohammed Yassine El Ghoumari is a professor of computer sciences at the National School of Marketing and Management, University Hassan 2, Casablanca, Morocco. His area of expertise is big data and machine learning algorithms.

Article submitted 2020-04-22. Resubmitted 2020-05-30. Final acceptance 2020-05-30. Final version published as submitted by the authors.