# Semantic Segmentation of Mice Wounds


## ABSTRACT

Semantic segmentation has been successfully explored in biological studies to handle various applications, such as identifying wounds. This study explores two image segmentation approaches to identify mice wounds, specifically the U-Net and Random Forest algorithms. The latter was combined with features extracted from the first two layers of VGG16, which was used as a feature extractor. Experiments were performed with a real dataset developed by the Pain, Neuropathy, and Inflammation Laboratory at the State University of Londrina with the approval of the University Ethics Committee on Animal Research and Welfare. The experimental results were promising, showing that both alternatives can provide accurate predictions for most images regarding FSCore and IoU evaluation measures. Statistical tests were also applied, showing that U-Net obtained statistically better results with an average FScore of 0.72 and IoU of 0.58.



## KEYWORDS

Semantic Segmentation, wound identification, U-Net, Random Forest.


## 1 INTRODUCTION

Image Segmentation is the process of dividing an image into regions [8], often the first step in the image analysis process. It highlights the portion of the image with relevant information, called the Region of Interest (ROI). However, different portions of images can reference the same object (parts), or the same objects can appear in different regions (plurality), thus having different ROI with the same aspect of meaning. For these portions to be understood as the same object or parts of the same region, the Semantic Segmentation adds meaning to the pixel [16], resulting in a pixel-wise classification problem. Badrinarayanan et al. [3] brings up a series of recent studies exploring pixel labeling classification.

This pixel-level precision enables machines to discern objects, their boundaries, and their relationships within the scene. The applications of Semantic Segmentation are wide-ranging, from autonomous vehicles navigating complex environments to medical image analysis for disease detection and treatment planning, as well as in augmented reality, where it enriches user experiences by accurately identifying and interacting with objects in the real world. Moreover, it facilitates the development of advanced Artificial Intelligence (AI) systems for tasks like object recognition, scene understanding, and even human-computer interaction.

In studies in the biological area, semantic segmentation has been used in skin wounds and is an important research area for evaluating the healing process in animals and humans [25]. The use of Machine Learning (ML) algorithms has been shown to be effective in identifying and delimiting wound areas, allowing more accurate and objective analysis of the healing process. Strategies have been proposed using traditional ML and Deep Learning (DL) algorithms [1].

Although DL models have achieved remarkable results across various domains, it is crucial to acknowledge that they can be computationally intensive and time-consuming, mainly when working with large datasets and complex architectures. This computational burden can hinder rapid prototyping and model iteration. Hence, exploring alternatives with traditional ML algorithms remains valuable. Traditional methods, such as Decision Trees (DTs), Support Vector Machines (SVMs), and Random Forests (RF), often offer faster training times, better interpretability, and require fewer computational resources. Moreover, they can be highly effective for tasks with limited data or when the interpretability of results is a priority. In many practical scenarios, a judicious blend of DL and traditional ML algorithms can yield efficient and effective solutions, balancing performance and computational demands.

This study explores Semantic Segmentation to solve a mice wound pixel classification problem. Experiments were carried out with two different approaches: i) a DL architecture - U-Net, and ii) a traditional ML algorithm - Random Forest (RF) using latent representations of a pre-trained Convolutional Neural Network (Visual Geometry Group - `VGG` [24]). The image dataset used in the experiments is composed of wounds in mice, generated by the Laboratory of Research in Pain, Neuropathy and Inflammation of the State University of Londrina (UEL).

This article is organized as follows: Section 2 presents the theoretical concepts necessary for understanding the work and related works found in the literature; Section 3 describes the proposed methodology for comparing and evaluating the segmentation algorithms; Section 4 presents the experimental results; and finally, Section 5 presents conclusions and suggestions for future work.

## 2 BACKGROUND

Semantic segmentation is a vital aspect of computer vision, involving identifying and classifying individual pixels within an image. Its applications span diverse fields like autonomous driving and medical image analysis. The conventional methods for this task demand substantial amounts of meticulously annotated data, a time-consuming and expensive process. Most semantic segmentation networks commonly utilize cross-entropy as their loss function and assess network performance using the intersection-over-union (IoU) metric [9].

### 2.1 U-Net

The U-Net algorithm is a Convolutional Neural Network (CNN) architecture developed for image segmentation, which has been shown to be efficient in several applications in the medical context [2, 21]. The network design consists of a symmetric network with descending *pooling* layers to encode low-level information into a high-dimensional representation, ascending convolution (or transposed convolution), and up-sampling layers to reconstruct the segmented image.

The technique of concatenating information from the decoding layer with the corresponding encoding layer allows the combination of low and high-level information, which provides a more

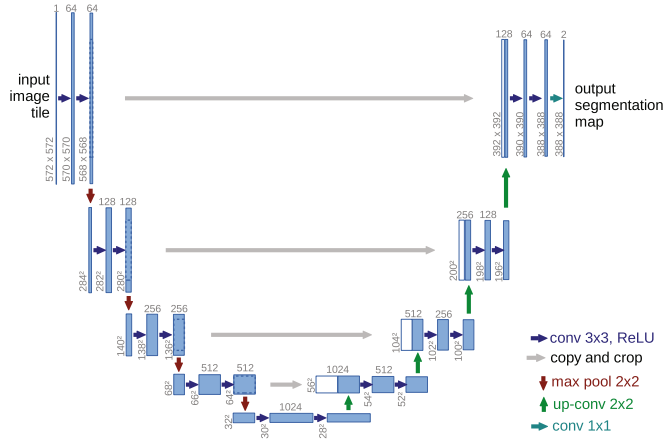



**Figure 1: U-Net architecture. The half of the figure (*encoder*) reduces the original images to some latent representation. The second half (*decoder*) reconstructs images using latent features to classify pixel labels. Figure from Ronneberger et al. [23].**

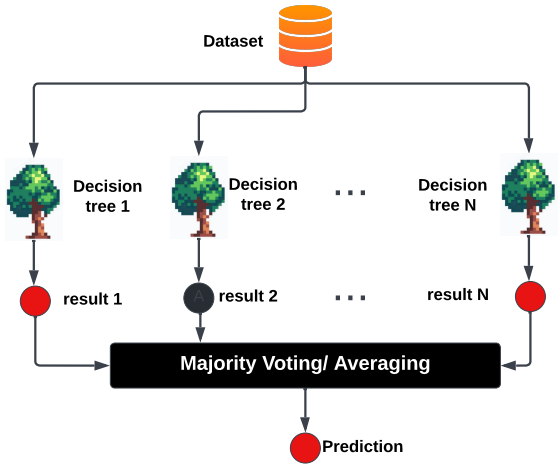



**Figure 2: Random Forest ensemble. The *forest* comprises several random decision trees. Each tree is slightly different, handling a subset of features and examples. The prediction returned by the ensemble is the agreement of the committee.**

accurate representation of the original image. Figure 1 shows the U-Net architecture, composed by a *encoder* and *a decoder*. The *encoder* is usually a pre-trained classification network like VGG/ResNet models applying convolution blocks followed by a maxpool down-sampling to encode the input image into feature representation at multiple levels.

The decoder is the second part of the architecture. The goal is to semantically project the discriminative (latent) features with lower resolution learned by the encoder onto the pixel space to get a dense classification. The decoder consists of upsampling and concatenation operations followed by regular convolutions. U-Net also uses ReLU-like activation functions for all convolutional layers and a sigmoid activation layer at the end to produce the binary output of the segmentation.

## 2.2 Random Forest

Random Forest (RF) [5] is a traditional ML algorithm that uses multiple Decision Trees (DTs) [6, 22] to generate predictions. Figure 2 presents a schematic idea of the algorithm. Each tree is trained on a random sample of the training data and a random subset of the attributes, creating different data views. The final prediction is made by polling the predictions of all the trees, making the model more robust and less susceptible to overfitting. RF is widely used in classification and regression applications such as image recognition, sentiment analysis, and fraud detection [4].

Semantic image segmentation involves identifying and labeling each pixel in an image based on its semantic content, such as identifying each object or region present in the image. This is important in many areas, such as computer vision, robotics, and medical image processing. RF can be trained to classify each image pixel into different classes, such as object, background, or edge, using texture, color, and shape information. The result is precise image segmentation that can be used in different applications.

## 2.3 Related works

A recent overview of the field of Semantic Segmentation is provided by Liu et al. [15], emphasizing its significance in image processing and computer vision. In the study, the authors highlight the impact of Deep Neural Networks (DNNs) on the progress of segmentation methods and categorize methods into traditional ML and recent DL-based approaches.

Most of the studies in the literature have been using DL algorithms to solve semantic segmentation of wounds [14, 17]. In Wang et al. [26], the authors performed semantic segmentation of ulcers' images. Experiments were carried out with state-of-the-art models that allow pixelwise classification, such as U-Net, FCN-VGG16, Mask-RCNN, and MobileNetV2. Results reported that U-Net obtained an accurate FScore value (0.915) and the highest recall among all models (0.912).

Fully Connected Networks (FCN) were also studied in Kaymak et al. [12] for segmentation of skin lesions, an experimental study was carried out for segmentation for cancer detection that was compared with other networks such as U-Net and SegNet, same study object as in Mishra and Daescu [18], which, in addition to CNNs, tests the performance of the segmentation using Otsu's threshold, being a much simpler technique than neural networks.

In another study [19], the authors performed the task of segmenting diabetic foot ulcers (DFU) in small datasets. They evaluated different DL models, such as the U-Net, V-Net, and SegNet. They were evaluated by accuracy, IoU, and FScore metrics. The results were positive for U-Net, considering that the model outperformed all the other models on the three metrics (0.949, 0.948, and 0.972, respectively).

Also, aiming to segment DFU images, Wang et al. [27] compared the DL ConvNet model with the classic ML algorithm Support Vector Machine (SVM), using RGB values as features, evaluating

pixel-wise accuracy and IoU. Results were positive for ConvNet in both metrics. The model obtained a value of 0.95 and 0.473 for pixel accuracy and IoU, respectively, against 0.776 and 0.264 for SVM.

Innani et al. [10] substitute U-Net encoding blocks with other networks (ResNet, InceptionResNetV2 and EfficientNets) in an encoder-decoder strategy for wound segmentation also for cancer. However, as the objective of the work is to perform a multiclass classification task, there is a second step after the segmentation. The proposed approaches are evaluated quantitatively in terms of the accuracy, mean IoU, and Dice Similarity Coefficient measures. The proposed cascaded end-to-end DL-based approach was the best overall, with the classification accuracy of the lesion significantly improved because of prior segmentation.

Kang and Nguyen [11] developed a hybrid framework combining RF and DL that learns flexible filters using an iterative optimization algorithm, and segments input images using the learned representations. Experiments were performed in a hand segmentation dataset for hand-object interaction and using two other semantic segmentation datasets. The results show that the proposed method achieves real-time semantic segmentation using limited computational and memory resources.

## 3 EXPERIMENTAL METHODOLOGY

In this section, we detail the experimental methodology adopted in this article. An overview of the flow of experiments, including sub-steps, is shown in Figure 3. The following sub-sections give additional details regarding them: the image dataset, data preparation, classification algorithms used, and models' training and evaluations.

### 3.1 Image Dataset

The dataset for this study consists of 71 images of mice with wounds on their backs. This is a real dataset developed by the Pain, Neuropathy, and Inflammation Laboratory at the State University of Londrina (UEL) with the approval of the UEL Ethics Committee on Animal Research and Welfare (process number 15654.2019.33). Data acquisition process had a protocol, but different cameras obtained images. Thus, 36 images have a resolution of $1024 \times 768$ pixels, while the remaining 35 images have $4032 \times 3024$ pixels.

It is important to mention that assessing the wound progression is outside the scope of this work, as this step requires more information than just detecting the wound region. Hence, this study focuses on pixelwise classification (semantic segmentation) between wounds and non-wounds. Figure 4 shows an example of an image and its corresponding labels and predictions for an animal. The classes (pixel labels) were defined manually using the `Labelme` tool[1].

### 3.2 Data Preprocessing and Augmentation

Since images may have different resolutions and neural network input layers require a specific image size, all the images were rescaled to a resolution of $256 \times 256$ pixels. The rescale operation was applied to original and mask (label) images. A normalization step followed it, modifying pixel labels from $[0, 255]$ to interval $[0, 1]$.

[1] https://github.com/wkentaro/labelme

Due to the low number of images in the dataset (71), a Data Augmentation (DA) process was also required. DA is a technique widely used in Computer Vision to increase the amount and variety of data [28]. The technique involves applying transformations to the original data and generating new instances. The main advantages of using DA are:

- (i) models' precision improvement since it generates more significant variability in the data;
- (ii) cost reduction in data collecting and labeling, given that collecting medical images is challenging and complex work, as well as labeling them, this task is reduced as transformations are applied to the original images and masks simultaneously; and
- (iii) overfitting prevention, since the model tends to learn specific patterns present in the available training data when there is low variability in the data, losing its generalization power to new data.

A total of five transformations were considered in DA: (1) horizontal flip, (2) vertical flip, (3) rotation with a maximum angle of 35 degrees, (4) salt and pepper noise, and (5) image translation. With the original images, the dataset with DA has 426 images available for model induction. It is important to highlight that augmented images are used only in the training fold to induce the predictive models. Only the original images in the testing fold were assessed when evaluating these models.

### 3.3 Algorithms

We explored two different strategies for wound recognition:

i) U-Net: a pure DL architecture widely adopted for Semantic Segmentation problems; and
ii) RF + VGG16: a hybrid and cheap alternative that extracts VGG16 latent features to train a Random Forest model.

Initial experiments were done using different Fully Convolutional Network (FCN) [16] architectures, but none of them recognized wound regions in the testing images. We only obtained good results when testing U-Net as the DL pure model. In fact, U-Net is one of the most widely used architectures for Semantic Segmentation of medical images [21].

U-Net was trained with Adam optimizer [13] and a learning rate of $\alpha = 0.0001$. The loss function optimized was the binary cross entropy since the network was used to perform binary classification. The U-Net architecture uses ReLu activation functions between convolutional layers and the sigmoid activation function as a classifier in its last layer. U-Net was trained for 100 epochs using batch size 2. These values were defined empirically.

In the RF algorithm, the default value for the number of trees is between $[100, 500]$ depending on the coding language and correspondent libraries/packages. However, all of them follow the same idea: the higher the number of trees, the better the generalization power of the induced model. In our experiments, the RF was implemented with the scikit-learn Python library, which uses $t = 100$ trees in the ensemble and '*Gini*' index as the attribute evaluation criterion as default values.

Some initial experiments were performed using RBG pixel values for the RF algorithm, but the predictions did not identify any pixel as a wound. Thus, an alternative explored is using 64 feature maps

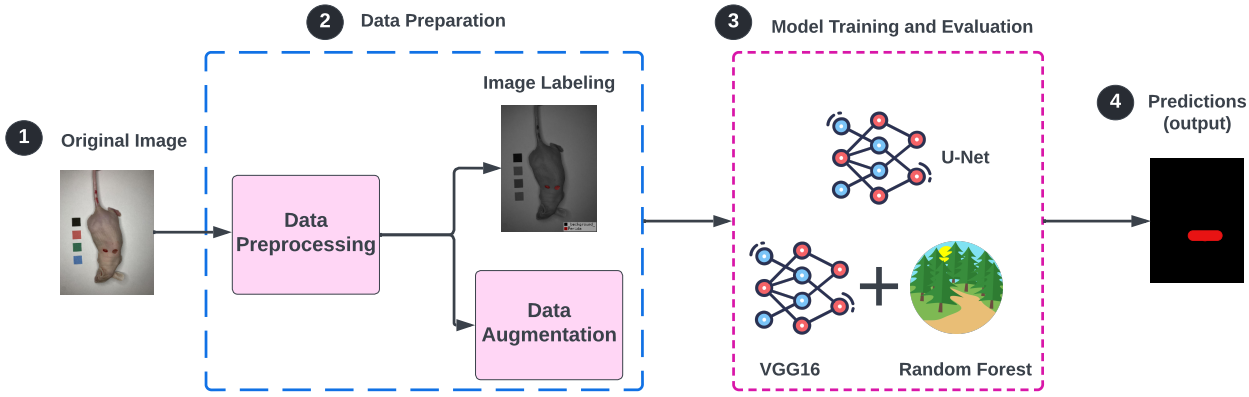
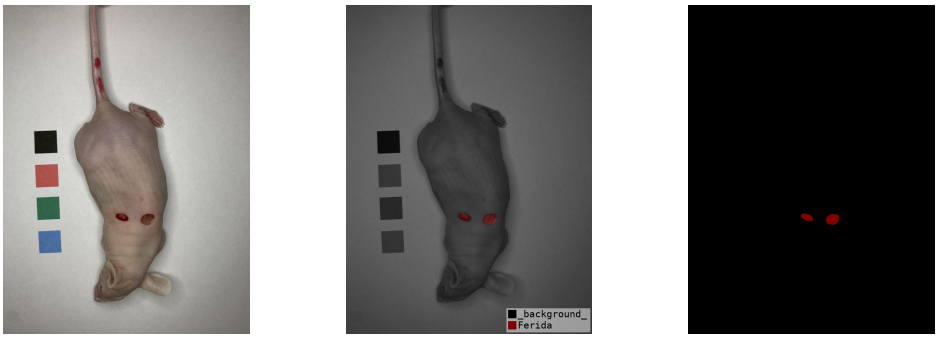



**Figure 3: Experimental methodology for semantic segmentation of mice wounds.**




**(a) Original image.** **(b) Classes generated with `Labelme`.** **(c) Pixel mask.**

**Figure 4: Example of a dataset image instance and its correspondent labels.**

extracted from the first two convolutional layers of the VGG16[2] neural network [24] to feed the algorithm. This DL model is pre-trained in the `ImageNet` dataset. VGG16 is well known for accurately performing computer vision tasks, including image recognition [20]. By extracting these feature maps, it is possible to obtain general information about the images that can be used to train machine learning models, such as the RF, with a lower computational cost.

### 3.4 Experimental Setup

The resampling strategy adopted in experiments was a simple hold-out with five repetitions using different seeds. The dataset was split into training, testing, and validation sets, with proportions of 60%, 30%, and 10%, respectively. It is worth mentioning that the validation data was extracted as a part of the training set and not used in the testing set.

The results obtained were evaluated for precision and efficiency in the segmentation of wound images by F-Score, an evaluation metric that combines precision and recall to provide an overall measure of the model's accuracy, and IoU [7]. All the code was written in Python, with U-Net coded using `PyTorch`[3]. The U-net implementation used no pre-trained weights. The RF+VGG16 strategy used `scikit-learn`[4] and `Keras`[5] libraries. Experiments were executed on a desktop with a Ryzen 5 5600g processor, using a normal CPU environment, with 16GB of RAM, and motherboard b550m. All the algorithms used their default hyperparameter values, and no hyperparameter tuning process was conducted. The code repository of this study is publicly available[6].

## 4 RESULTS

Figure 5 illustrates each algorithm's FScore and Intersection over Union (IoU) values and suggests that both algorithms present similar result distributions. The U-Net median values for FScore (0.75) and IoU (0.60) were slightly higher than RF ones (0.70 and 0.53). The figure also shows that most images have FScore values above 0.6. The overall results indicate that both algorithms could identify wounds regions and estimate the shape and wound filling even with a small dataset.

[2]Visual Geometry Group, 16 layers.
[3]https://segmentation-modelspytorch.readthedocs.io/en/latest/docs/api.html#unet
[4]https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html
[5]https://keras.io/api/applications/vgg/
[6]https://github.com/BrunoMarcato/MiceWoundSegmentation

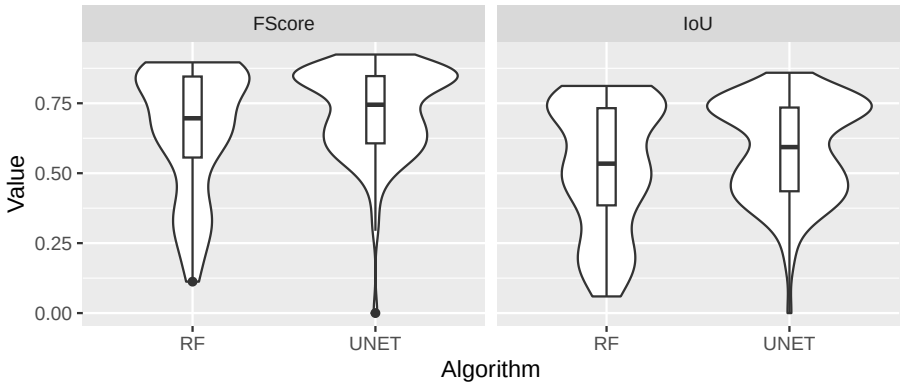



**Figure 5: Overall results with FScore values obtained on test images.**

However, due to the nature of the distributions represented by the violins, some images presented difficulty for the algorithms. U-Net could not perform generalization for some images with tiny wounds. These images have a highly unbalanced pixel distribution, as few pixels compose wound regions. It may affect the FScore values but not the IoU. IoU only measures hits and misses on the object but not on the background, which is why it is usually used as an image segmentation evaluation metric. RF, on the other hand, has more extended distributions for both metrics. Violins indicate a considerable number of images with lower values of IoU and FScore, more than U-Net. A superficial analysis could identify that this occurred in some images when it fails to fill in the wound region or even incorrectly classifies other parts of the animal as a wound, such as regions of the paws or tail.

All the algorithms were executed in the same training and testing partitions and repeated with different seeds (varying from 1 to 5). Thus, the non-parametric Wilcoxon test with $\alpha = 0.05$ (95% of significance) was applied to assess the statistical significance of these results. The null hypothesis states that there is no difference in performance between the algorithms. The test was applied for both metrics (FScore and IoU) and obtained a p-value = 0.00101, which suggests a difference in terms of performance in favor of the U-Net.

Although U-Net is statistically better than RF regarding predictions, the computational cost for training both algorithms differs. RF took 20 minutes in total to execute the five repetitions, while U-Net needed 9 hours. That is, the 64 feature maps extracted from the VGG16 and fed into the RF can be a viable alternative to speed the process of wound recognition rather than using a deep architecture, such as U-Net.

## 4.1 Prediction analysis

Figure 6 presents a heatmap with the obtained predictions of each algorithm in each image. The X-axis lists all the images, while the y-axis lists the evaluated algorithms (U-Net and RF). Heatmaps were generated for both performance measures (FScore and IoU). The higher the values, the more blue the cell is (and better). The lower the values, the more red the cell is (and worse). It is desired heatmaps with as much blue as possible.

Both measures show the same behavior: FScore values are higher in the same images where the correspondent IoU values are high. Good predictive values were obtained with images with ids = {1, 2, 3, 5, 6, 7, 66, 68, 69, 70, 71}. All of them are images with bigger wounds. An example can be seen in Figures 7a and 7b. Those images illustrated the U-Net predictions on the image with id = 5. Obtaining predictions are close to the desired area identified in the class mask (Figure 7b).

On the other hand, predictions were not accurate for images with ids = {28, 29, 38, 50, 52}, to name a few. Wounds in these images are tiny, with fewer pixels, and thus more challenging to identify. An example is depicted in Figures 7c and 7d, with the RF predictions on the image with id = 50 and its desired mask. Here it is possible to see that besides being unable to recognize the small regions of wounds, the RF algorithm wrongly identifies other parts of the animal as a wound.

The fact that RF uses initial layers of VGG ends up using high-level features such as color, thus giving importance to red regions, which is expected for wounds. However, this color is also present in other regions, as shown in the image (7c) with prediction generated by RF, which can be one of the causes of the loss of performance concerning U-Net.

# 5 CONCLUSIONS

This study investigated U-Net and RF with VGG latent features to solve a semantic segmentation problem of mice wounds. Experiments were carried out with an image dataset generated by the Pain, Neuropathy, and Inflammation Laboratory at the State University of Londrina, composed of 71 images showing wounds in mice. U-Net and RF+VGG were executed five times with different seeds splitting data into training and testing folds. Predictions were evaluated by F-Score and IoU metrics.

Results were promising: U-Net obtained F-Score and IoU average values of 0.75 and 0.60, respectively. Most images generated accurate predictions of the wound regions, with some exceptions in images with tiny wounds. The RF+VGG alternative was faster

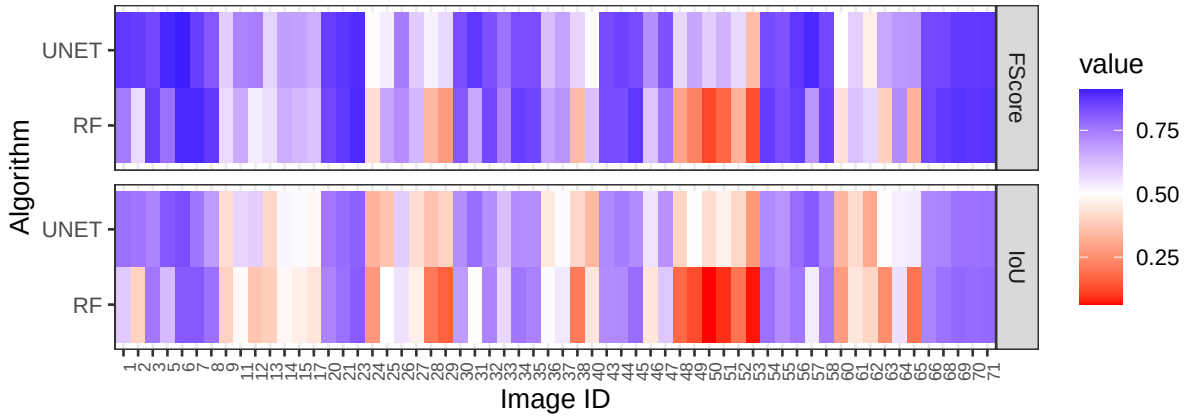
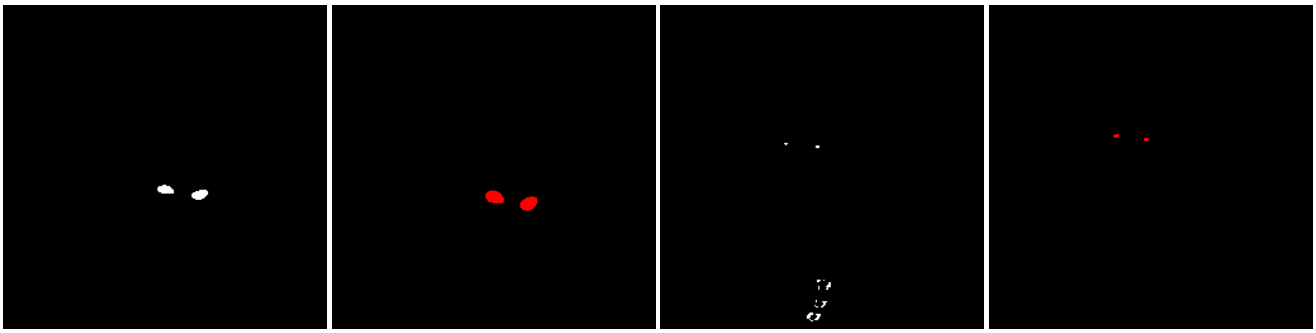



**Figure 6: Predictions obtained by each algorithm in each image.**



**(a) U-Net - Image 5** **(b) Mask - Image 5** **(c) RF - Image 50** **(d) Mask - Image 50**

**Figure 7: Examples of good and bad segmentation wound areas obtained in experiments.**

regarding the computational cost but statistically lower in performance, with lower average F-Score and IoU (0.70 and 0.53). However, it can be improved with some post-processing techniques to fill regions not wholly identified.

An alternative might be replacing VGG16 with a more robust DL architecture, such as ResNet, or setting up the U-Net and RF with DL architectures with pre-trained weights by Transfer Learning. Other alternatives for future works include evaluating more and different DL architectures, ML algorithms and automating the entire pipeline.

## REFERENCES


[1] Charu C. Aggarwal. 2018. *Neural Networks and Deep Learning: A Textbook.* Springer International Publishing, Cham. 105–167 pages. https://doi.org/10.1007/978-3-319-94463-0

[2] Reza Azad, Ehsan Khodapanah Aghdam, Amelie Rauland, Yiwei Jia, Atlas Haddadi Avval, Afshin Bozorgpour, Sanaz Karimijafarbigloo, Joseph Paul Cohen, Ehsan Adeli, and Dorit Merhof. 2022. Medical Image Segmentation Review: The success of U-Net. arXiv:2211.14830 [eess.IV]

[3] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. 2017. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39, 12 (2017), 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615

[4] Gérard Biau and Erwan Scornet. 2016. A random forest guided tour. *TEST* 25 (2016), 197–227. https://doi.org/10.1007/s11749-016-0481-7

[5] Leo Breiman. 2001. Random forests. *Machine learning* 45, 1 (2001), 5–32.

[6] L. Breiman, J.H. Friedman, R.A. Olshen, and C.J. Stone. 1984. *Classification and Regression Trees.* Chapman & Hall (Wadsworth, Inc.).

[7] Francisco J García-López, P B García-Allende, Olga M Conde, and José M López-Higuera. 2019. Comparison of evaluation metrics for hyperspectral image segmentation. *Sensors* 19, 7 (2019), 1501.

[8] R.C. Gonzalez and R.E. Woods. 2010. *Processamento Digital De Imagens.* Pearson, São Paulo.

[9] Yifeng Huang, Zhirong Tang, Dan Chen, Kaixiong Su, and Chengbin Chen. 2020. Batching Soft IoU for Training Semantic Segmentation Networks. *IEEE Signal Processing Letters* 27 (2020), 66–70. https://doi.org/10.1109/LSP.2019.2956367

[10] Shubham Innani, Prasad Dutande, Bhakti Baheti, Ujjwal Baid, and Sanjay Talbar. 2023. Deep Learning Based Novel Cascaded Approach for Skin Lesion Analysis. In *Computer Vision and Image Processing*, Deep Gupta, Kishor Bhurchandi, Subrahmanyam Murala, Balasubramanian Raman, and Sanjeev Kumar (Eds.). Springer Nature Switzerland, Cham, 615–626.

[11] Byeongkeun Kang and Truong Q. Nguyen. 2019. Random Forest with Learned Representations for Semantic Segmentation. *IEEE Transactions on Image Processing* 28, 7 (2019).

[12] Ruya Kaymak, Cagri Kaymak, and Aysegul Ucar. 2020. Skin lesion segmentation using fully convolutional networks: A comparative experimental study. *Expert Systems with Applications* 161 (2020), 113742. https://doi.org/10.1016/j.eswa.2020.113742

[13] Diederik P. Kingma and Jimmy Ba. 2017. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs.LG]

[14] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. 2017. A survey on deep learning in medical image analysis. *Medical Image Analysis* 42 (2017), 60–88. https://doi.org/10.1016/j.media.2017.07.005

[15] Xiaolong Liu, Zhidong Deng, and Yuhan Yang. 2018. Recent progress in semantic image segmentation. *Artificial Intelligence Review* (27 jun 2018). https://doi.org/

10.1007/s10462-018-9641-3

[16] Jonathan Long, Evan Shelhamer, and Trevor Darrell. 2015. Fully convolutional networks for semantic segmentation. In *Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Boston, MA, USA, 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965

[17] Zahra Mirikharaji, Kumar Abhishek, Alceu Bissoto, Catarina Barata, Sandra Avila, Eduardo Valle, M. Emre Celebi, and Ghassan Hamarneh. 2023. A survey on deep learning for skin lesion segmentation. *Medical Image Analysis* 88 (2023), 102863. https://doi.org/10.1016/j.media.2023.102863

[18] Rashika Mishra and Ovidiu Daescu. 2017. Deep learning for skin lesion segmentation. In *International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 1189–1194. https://doi.org/10.1109/BIBM.2017.8217826

[19] Rania Niri, Douzi Hassan, Lucas Yves, and Sylvie Treuillet. 2020. Semantic Segmentation of Diabetic Foot Ulcer Images: Dealing with Small Dataset in DL Approaches. https://doi.org/10.1007/978-3-030-51935-3_17

[20] Alashiri Olaitan, Adeyinka Adewale, Sanjay Misra, Akshat Agrawal, Ravin Ahuja, and Jonathan Oluranti. 2022. Face Recognition Using VGG16 CNN Architecture for Enhanced Security Surveillance—A Survey. In *Futuristic Trends in Networks and Computing Technologies*, Pradeep Kumar Singh, Sławomir T. Wierzchoń, Jitender Kumar Chhabra, and Sudeep Tanwar (Eds.). Springer Nature Singapore, Singapore, 1111–1125.

[21] Narinder Singh Punn and Sonali Agarwal. 2022. Modality specific U-Net variants for biomedical image segmentation: a survey. *Artificial Intelligence Review* 55 (2022), 5845–5889. Issue 7. https://doi.org/10.1007/s10462-022-10152-1

[22] J. Ross Quinlan. 1993. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[23] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer.

[24] Karen Simonyan and Andrew Zisserman. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *International Conference on Learning Representations*. CoRR, San Diego, CA, USA.

[25] Eli Stevens, Luca Antiga, and Thomas Viehmann. 2020. *Deep Learning with PyTorch*. Manning Publications Co, Shelter Island, NY.

[26] Chuanbo Wang, DM Anisuzzaman, Victor Williamson, Mrinal Kanti Dhar, Behrouz Rostami, Jeffrey Niezgoda, Sandeep Gopalakrishnan, and Zeyun Yu. 2020. Fully Automatic Wound Segmentation with Deep Convolutional Neural Networks. arXiv:2010.05855 [eess.IV]

[27] Changhan Wang, Xinchen Yan, Max Smith, Kanika Kochhar, Marcie Rubin, Stephen M. Warren, James Wrobel, and Honglak Lee. 2015. A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks. In *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, Milan, Italy, 2415–2418. https://doi.org/10.1109/EMBC.2015.7318881

[28] Suorong Yang, Weikang Xiao, Mengcheng Zhang, Suhan Guo, Jian Zhao, and Furao Shen. 2022. Image Data Augmentation for Deep Learning: A Survey. arXiv:2204.08610 [cs.CV]