I can-cer vive : a comparative study of machine learning algorithms for the segmentation of medical images in order to perform breast cancer screening

ÉCOLE NATIONALE DES SCIENCES DE L'INFORMATIQUE

Ichrak HAMDI Ines BOUHELAL
Supervised by
DR.MOHAMED AMINE MEZGHICH
University of Manouba
National School of Computer Sciences

جامعة منوبة
Université de la Manouba

Abstract

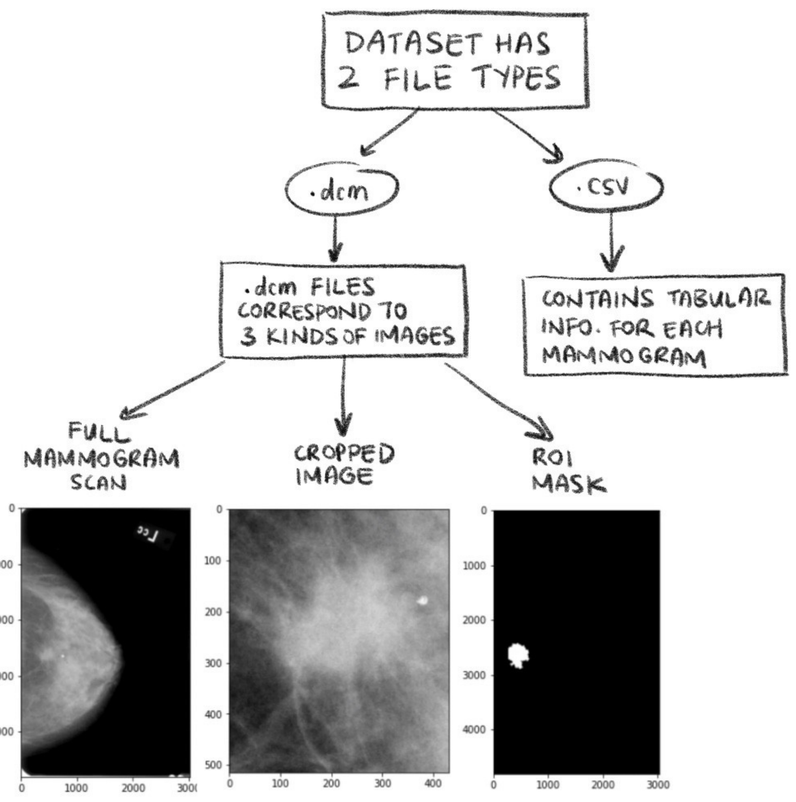
In 2020, there were 2.3 million women diagnosed with breast cancer and 685 000 deaths globally. As of the end of 2020, there were 7.8 million women alive who were diagnosed with breast cancer in the past 5 years, making it the world's most prevalent cancer.
Early breast cancer detection is one of the most important issues that need to be addressed worldwide as it can help increase the survival rate of patients. Mammograms have been used to detect breast cancer in the early stages; if detected in the early stages, it can drastically reduce treatment costs. The detection of tumours in the breast depends on segmentation techniques. Segmentation plays a significant role in image analysis and includes detection, feature extraction, classification, and treatment. Segmentation helps physicians quantify the volume of tissue in the breast for treatment planning.

Objective

Our objective is to do a comparative study of machine learning algorithms for the segmentation of medical images in order to perform breast cancer screening so the doctors can detect the tumor faster with some algorithms to know which one is more accurate.

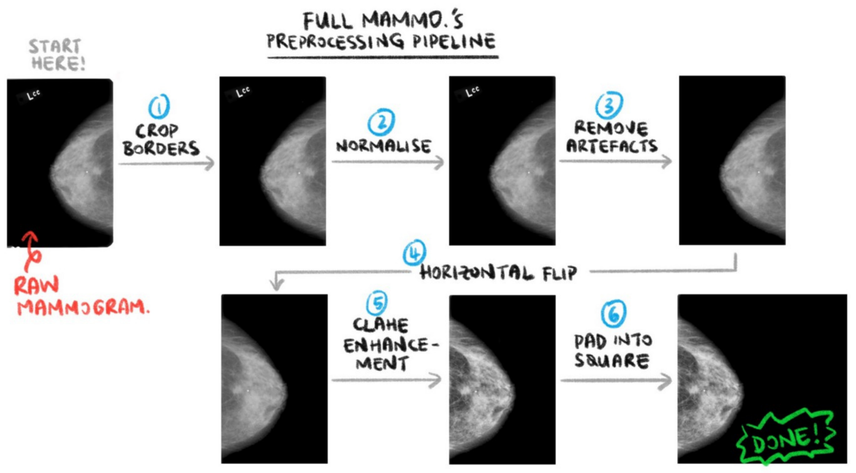
Data Collection

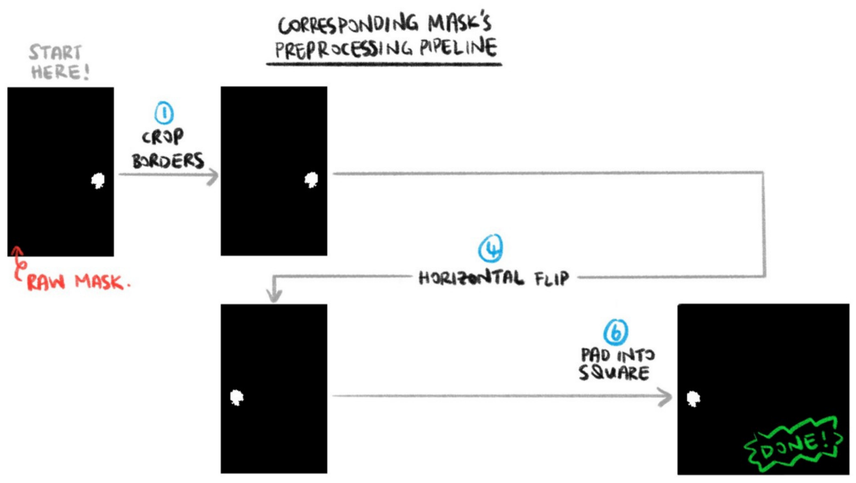
We chose to use the CBIS-DDSM dataset because it is suitable for computer vision projects of intermediate complexities. With 2,620 scanned film mammography images, it is large enough to conduct decent model training. Furthermore, since the CBIS-DDSM dataset contains real world mammogram scans, they are "messy" enough such that robust and intentional image preprocessing needs to be done in order to achieve decent results at the task at hand.



Methodology

1- Image Preprocessing Pipeline



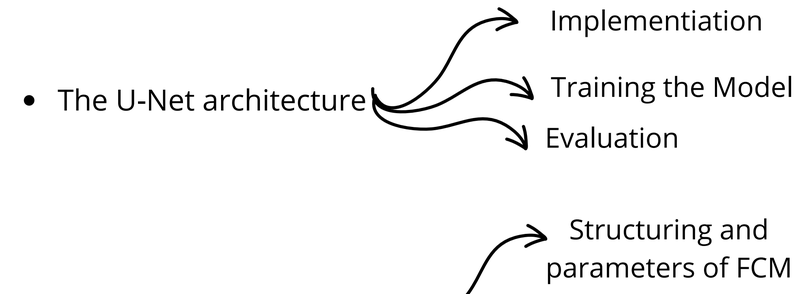


2- Applying 3 machine learning algorithms to visualizase the the differences between them

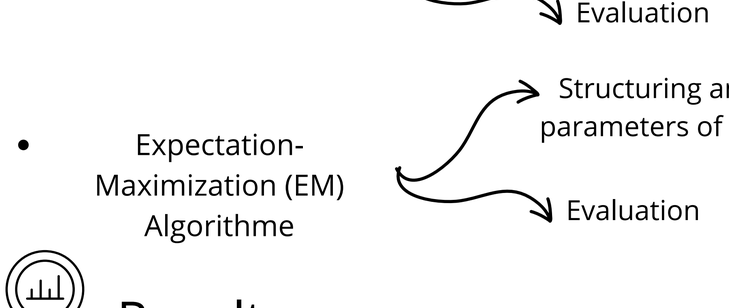
- The U-Net architecture → Implementiation; Training the Model; Evaluation
- Fuzzy C-Means Algorithme → Structuring and parameters of FCM; Evaluation
- Expectation-Maximization (EM) Algorithme → Structuring and parameters of EM; Evaluation

Results

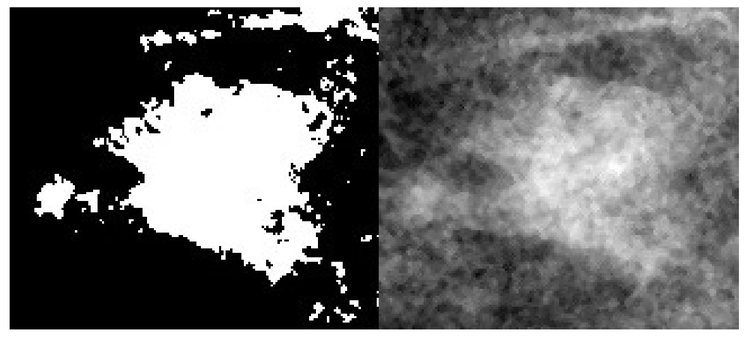
1- The U-Net architecture



2- Fuzzy C-Means Algorithme



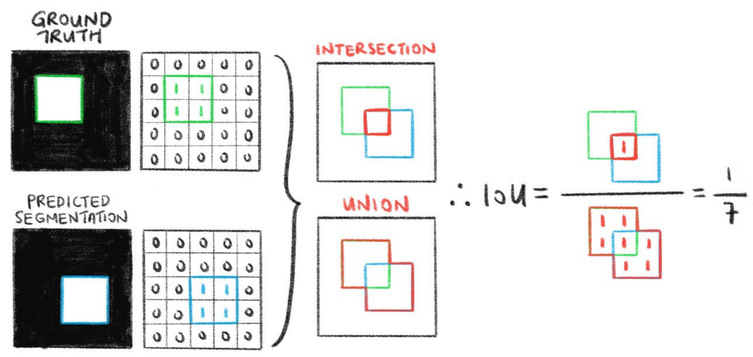
3- Expectation-Maximization (EM) Algorithme



4- Comparison

Metrics	U-Net	FCM	EM
IOU	0,6789	0,5795	0,5445
Dice	0,5321	0,4810	0,7748
Speed	Training for 25h executed in some secondes	0:0:0:88761	0:13:28:180522

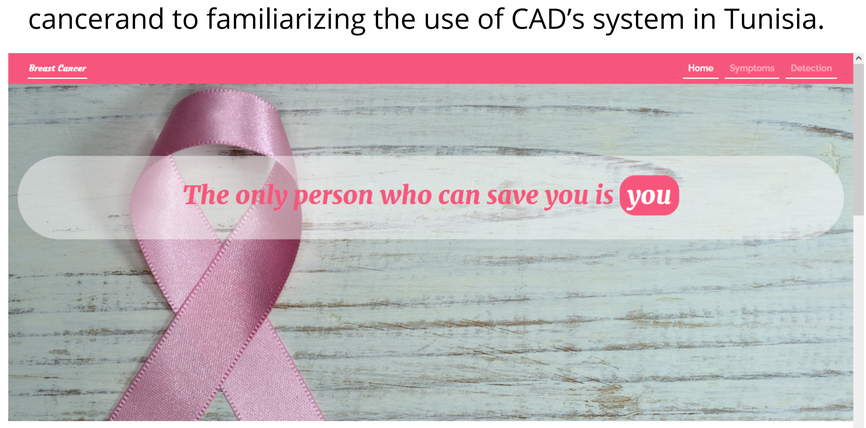
The custom IOU metric: is the percentage overlap between the predicted segmentation and the ground truth, divided by the area of union between the predicted segmentation and the ground truth.



Dice : the similarity of two samples

Conclusion

A free Web application that allows to make a comparative study of machine learning algorithms for the segmentation of medical images in order to learning more about breast cancerand to familiarizing the use of CAD's system in Tunisia.



Let's Talk Statistics

Prospectives

- App version
- Classification of breast cancer types. Integration of a part allowing the user to communicate with a doctor thus making him aware of the dangers of breast cance.
- Trying new algorithmes.

References

[Nayak et al., 2015 Nayak,, Naik, B., and Behera, H. (2015). Fuzzy c-means (fcm) clustering algorithm : a decade review from 2000 to 2014. Computational intelligence in data mining-volume 2, pages 133–149
Sharma, N., & Aggarwal, L. M. (2010). Automated medical image segmentation techniques. Journal of medical physics/Association of Medical Physicists of India, 35(1), 3.
Gefen, S., Tretiak, O. J., Piccoli, C. W., Donohue, K. D., Petropulu, A. P., Shankar, P. M., ... & Brookeman, J. R. Image Segmentation A Shape-Based Approach to the Segmentation of Medical Imagery Using Level Sets.A. Tsai, A. Yezzi, Jr., W. Wells, C. Tempany, D. Tucker, A. Fan, WE Grimson, and A. Willsky 137 Ultrasonography Combining Low-, High-Level and Empirical Domain Knowledge for Automated Segmentation of Ultrasonic Breast.

Contact

https://github.com/ichrakhamdi
ichrak.hamdi@ensi-uma.tn
https://www.linkedin.com/in/hamdi-ichrak-740bb81b9/
+21629144540