Deep Learning to Assess Hepatic Steatosis using Ultrasound B-mode Images

Authors:

Pedro Vianna[1,2], MSc (pedro.vianna@umontreal.ca)

Shanel Gauthier[3], MSc

Emmanuel Montagnon[2,4], PhD

An Tang[2,4,5], MD, MSc

Guy Cloutier[1,2,5,6], PhD

Affiliations:

[1] Laboratory of Biorheology and Medical Ultrasonics (LBUM), CRCHUM.

[2] Centre de Recherche du Centre hospitalier de l'Université de Montréal (CRCHUM).

[3] MILA – Québec Artificial Intelligence Institute

[4] Laboratoire Clinique du Traitement de L'Image (LCTI), CRCHUM

[5] Department of Radiology, Radio-Oncology and Nuclear Medicine, Université de Montréal.

[6] Institute of Biomedical Engineering, Université de Montréal.

February 3rd, 2022

Virtual Seminar 2022

Département de radiologie, radio-oncologie
et médecine nucléaire
Faculté de médecine
Université de Montréal

Disclosures

Funding for this project was supported by:

- Programme IVADO Recherche Fondamentale, 2020-2022, Ultrasound classification of chronic liver disease with deep learning, A. Tang, I. Rish, G. Wolf, G. Cloutier, S. Kadoury, M. Chassé, B. Nguyen.

- Bourse de Mérite – Intelligence artificielle de la Faculté de médecine, 2021-2022, P. Vianna, G. Cloutier.

- Dr. Cloutier’s research fund, Laboratoire de biorhéologie et d’ultrasonographie médicale du CRCHUM.

1. Introduction

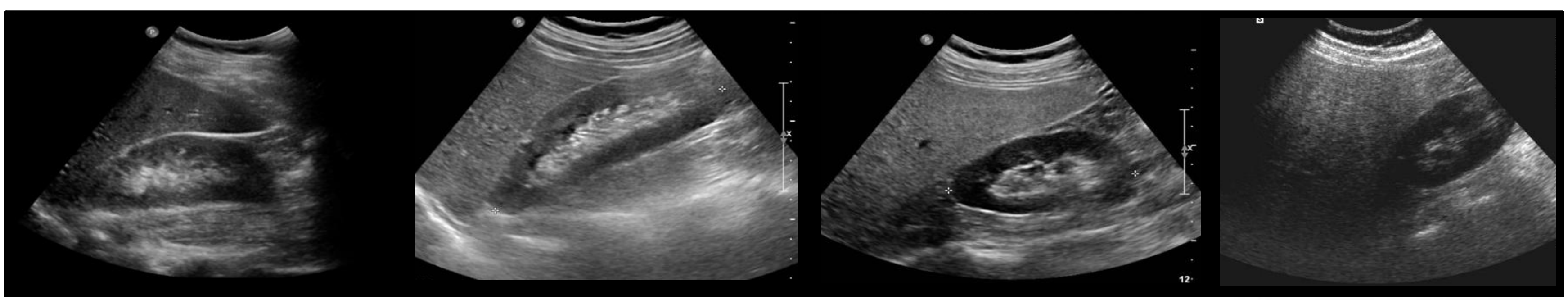
1.1 Hepatic Steatosis

- Excessive and pathologic intracellular accumulation of fat in the liver, which can lead to nonalcoholic fatty liver disease (NAFLD).

- Liver biopsy is the reference standard for the diagnosis and grading, however non-invasive alternatives are desirable. Ultrasound is cost-effective, portable, and available at point of care.

- We propose the use of deep learning techniques to assess the severity of hepatic steatosis from ultrasound images.

Normal Mild Moderate Severe



Département de radiologie, radio-oncologie
et médecine nucléaire
Faculté de médecine
Université de Montréal

1. Introduction

1.2 Deep Learning

- Unlike traditional machine learning techniques, deep learning methods can learn the main features presented in an image by itself, disregarding handmade features extractors designed by a specialist in the field.

- Convolutional neural networks (CNN) to perform the steatosis grading could be introduced in a computer assisted diagnosis system. Determining which CNN accomplishes better results for the task in ultrasound images should help the evolution of this field of study.

2. Methods

2.1 Dataset

- The dataset consists of 2621 images from 102 patients, obtained during routine procedures at CHUM using Philips iU22 Ultrasound System.

- The steatosis grade gold standards were acquired from biopsy reports. 29% of the images (18 patients) are associated with severe steatosis, 19% (16 patients) are from moderate cases, 26% are from (23) patients with mild grading and, finally, 26% are from (45) patients with absence of steatosis. For this work, we tried to detect the presence of steatosis, i.e. classify between absence of steatosis or presence of steatosis (mild, moderate and severe).

- We generated a five-fold cross-validation split to evaluate the performance of different deep learning architectures on 85% of the dataset, with the remaining 15% being used as a held-out test set. Training, validation and testing sets contain different patients.

Département de radiologie, radio-oncologie
et médecine nucléaire
Faculté de médecine
Université de Montréal

2. Methods

2.2 Deep Learning models

- In order to compare the performance of two “off-the-shelf” convolutional neural network (CNN) architectures on this task, we used our images to train the VGG16 and the ResNet50. Both networks are well-established within the artificial intelligence field, being functional on different tasks, including medical applications[1].

- Transfer Learning was also implemented, to investigate a possible improvement[2]. Models trained on natural images (from ImageNet) and fine-tuned on our dataset are also evaluated on the present work.

[1]George C Kagadis et al 2020 Phys. Med. Biol. 65 215027
[2]Tajbakhsh N et al 2016 IEEE Trans Med Imaging. 35(5):1299-1312.

Département de radiologie, radio-oncologie
et médecine nucléaire
Faculté de médecine
Université de Montréal

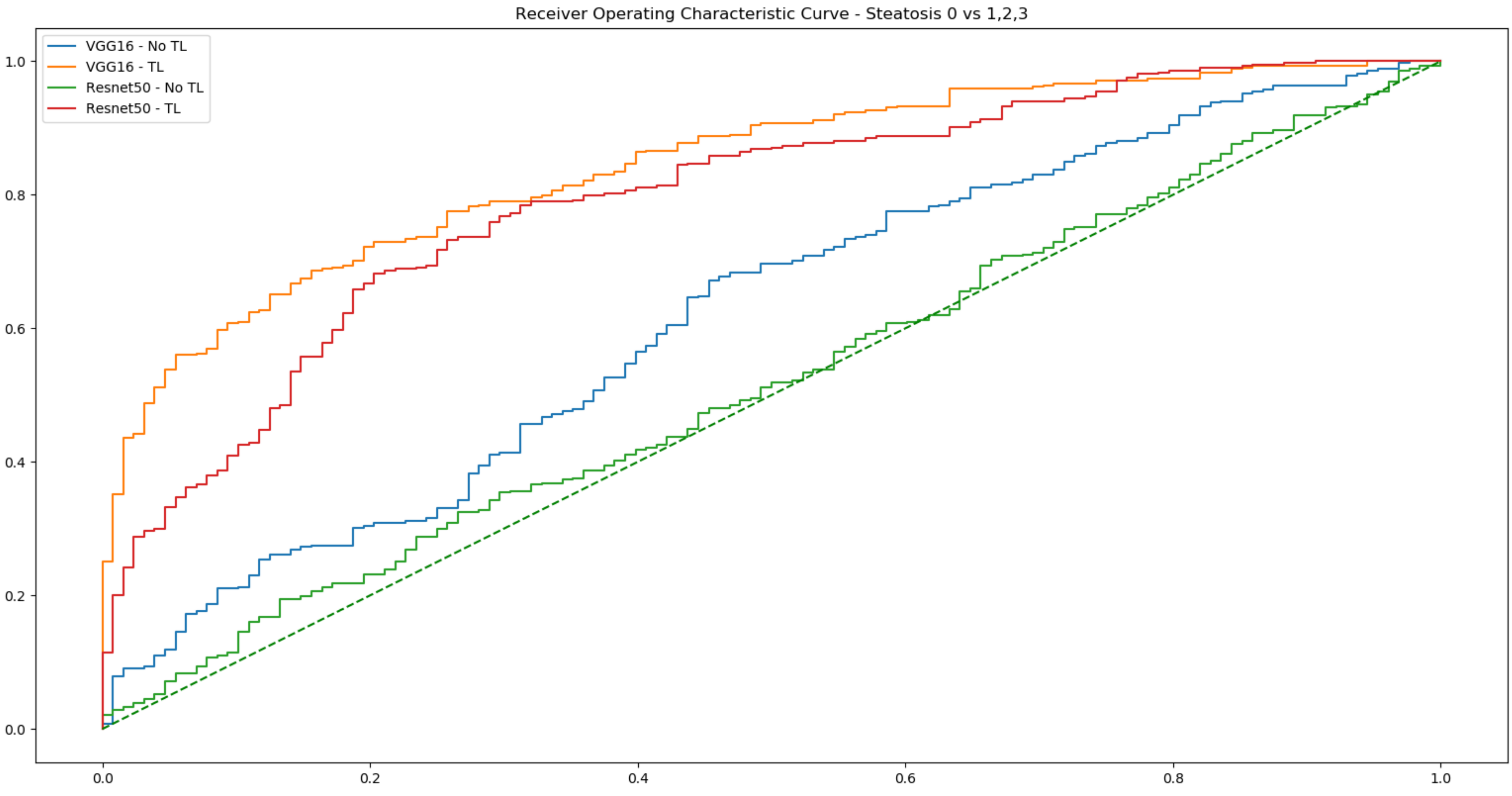
3. Results

Model	AUROC Curve	Sensitivity	Specificity	Accuracy	F1-Score	PPV	NPV
VGG16	0.618	0.847	0.281	0.714	0.819	0.794	0.360
VGG16 – Transfer Learning	**0.843**	**0.959**	**0.359**	**0.819**	**0.890**	**0.830**	**0.730**
ResNet50	0.520	0.876	0.148	0.705	0.820	0.771	0.268
ResNet50 – Transfer Learning	0.793	0.923	0.328	0.784	0.867	0.818	0.568

AUROC: Area Under Receiver Operator Characteristic
PPV: Positive Predictive Value
NPV: Negative Predictive Value

Département de radiologie, radio-oncologie
et médecine nucléaire
Faculté de médecine
Université de Montréal

3. Results



Receiver Operating Characteristic Curve - Steatosis 0 vs 1,2,3
VGG16 - No TL
VGG16 - TL
Resnet50 - No TL
Resnet50 - TL
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0

4. Discussion

- VGG16 yielded a better performance than ResNet50 for this task. It is worth noticing that both networks were trained using their original configurations, therefore a thorough hyperparameter tuning can alter this comparison.

- Using Transfer Learning greatly improved the metrics for both networks, and significantly decreased the time-to-train.

- Both networks demonstrated to be capable of achieving relevant results, compared with other works[1,2] in the field. It is difficult to directly compare the results with those of previous papers due to differences in the datasets, enrolment criteria, and reference standards.

- The lower specificity and NPV reveal that the models are predicting a smaller number of true negatives. This may be explained by the imbalance between the two classes.

[1]Michael Byra et al 2018 Int J Comput Assist Radiol Surg; 13, 1895–1903.
[2]Michal Byra et al 2022 J. Ultrasound Med. 41,175-184

Département de radiologie, radio-oncologie
et médecine nucléaire
Faculté de médecine
Université de Montréal

5. Conclusion

- Ultrasound can be used as a non-invasive modality in order to assess hepatic steatosis.
- Convolutional Neural Networks are useful for fatty liver prediction on ultrasound images, especially when using Transfer Learning.
- This could be a step forward to reduce the number of liver biopsies in near future.

- For future work, it is worth investigating different dichotomizations of steatosis grades (i.e. detect moderate-severe cases, and detect severe cases), as well as a four-way classification.
- While the present work focused on using "off-the-shelf" models, an optimization of the hyperparameters for the networks needs to be addressed, as well as the performances of different CNNs.