Machine Learning-Assisted Prediction of Hepatic Steatosis using 3D-CNN Auto-segmentation of Contrast-enhanced Portal Venous Phase CT Examinations of the Abdomen

Poster No.:	C-3419
Congress:	ECR 2019
Type:	Scientific Exhibit
Authors:	S. L. Mihalcioiu[1], R. Remtulla[2], O. ciga[2], M. C. H. Mo[3], M. D. A. Attarian[3], P. Savadjiev[3], S. BHATNAGAR[3], C. Reinhold[1], J. J. R. Chong[1]; [1]Montreal, QC/CA, [2]Montreal, Quebec/CA, [3]Montreal/CA
Keywords:	Cirrhosis, Segmentation, Computer Applications-Detection, diagnosis, CT, Liver, Artificial Intelligence, Abdomen
DOI:	10.26044/ecr2019/C-3419

Aims and objectives

Hepatosteatosis refers to the abnormal accumulation of triglycerides in more than 5% of hepatocytes (Lee DH, 2017; Torres, Williams, & Harrison, 2012). Significant etiologies of hepatosteatosis include consumption of alcohol and Non-Alcoholic Fatty Liver Disease (NAFLD) (Miyoshi & Kamada, 2015). NAFLD is a spectrum ranging from benign Isolated Fatty Liver (IFL) to Non-Alcoholic Steatohepatitis (NASH), which involves hepatocellular inflammation and the potential to develop fibrosis, cirrhosis and hepatocellular carcinoma (Torres et al, 2012). NAFLD and NASH have prevalences of 46% and 12.2% respectively among asymptomatic middle aged Americans (Hamer, Aguirre, Casola, Lavine, Woenckhaus, & Sirlin, 2006). NASH is now the third most common indication for liver transplant in the U.S (Torres et al, 2012). One-third to one-half of potential donor livers possess some level of steatosis, which may result in primary graft non-function in liver transplants (Torres et al, 2012).

The clinical importance of early NAFLD diagnosis relates to the increasing prevalence of the condition, the potential reversibility of the disease, the possibility to prevent further hepatic damage and its implications in liver transplantation. Currently a large proportion of hepatic steatosis may be under-reported due to limitations of diagnosing steatosis on single-phase contrast-enhanced abdominal CT scans. Contrast-enhanced portal venous phase CT, the most common CT abdomen exam type, is inaccurate at detecting hepatic steatosis due to the potential for reading errors in attenuation values created by the contrast media inside the liver and spleen as well as by the variability in bolus timing (Lee SS & Park, 2014). Unenhanced CT Abdomens can be accurate at detecting moderate hepatic steatosis but are less frequently used (Lee SS et al, 2014). If a more objective and reliable process could be developed to detect and quantify hepatic steatosis when it appears incidentally on Contrast-enhanced venous phase CT, NAFLD may be diagnosed earlier and more consistently in the general patient population, allowing for earlier patient lifestyle or therapeutic intervention.

Contemporary advancements in machine learning and deep learning brought about a generation of deep convolutional neural networks (DCNN) that are capable of creating algorithmic models, apt to perform many image analysis tasks that approach human-level vision abilities, given sufficient training data and computation. In addition, the machine learning paradigm of utilizing transfer learning, by fine-tuning a pre-trained CNN with generic images, has shown considerable promise in producing neural networks that offer clinical utility, with dataset sizes on the order of thousands of images. These are within the realm of manual label annotation with manual or Natural Language Processing (NLP) supervised learning experiments. Amongst the many neural network variants now publicly available, there are 3D CNN's capable of performing fully automated large solid organ segmentation (e.g. Liver segmentation) as well as synthesizing multiphasic datasets (e.g. Plain, Arterial, Portal Venous, and Delay).

The incidental detection of hepatic steatosis is a novel application of 3D/2D DCNN methods that will permit the simultaneous (1) fully automated segmentation of the liver, and (2) the fully automated classification and grading of steatotic liver disease from contrasted multiphasic CT examinations. The goal of this study is to apply a series of DCNN's to facilitate image recognition of hepatic steatosis on multiphasic CT examinations, on both Plain and Portal Venous phase acquisitions, utilizing both fully automatic liver organ segmentation, and image classification to permit the automatic detection and quantification of hepatic steatosis on all routinely performed examinations.

Methods and materials

Study Design:

Patients were selected and a database was created of unenhanced and Portal Venous CT images through selection of mild, moderate and severe cases of hepatic steatosis as defined by the unenhanced phase appearance. Normal control cases of normal density healthy livers were used for each of the unenhanced CT and portal venous CT groups. The unenhanced CT phase images were used as an internal reference standard for hepatic steatosis. 1.5 cm^2 Regions of Interest (ROI) were taken from hepatic segments V, VI, VII, and VIII from the unenhanced and portal venous phases. To measure splenic attenuation three 1.5 cm^2 ROIs were taken from the superior, middle, and inferior segments of the spleen. Images selected in the Portal Venous and Unenhanced phase were used in the training dataset. A pre-trained transfer-learning 'V-NET' 3D convolutional neural network was utilized to segment the whole liver organ boundary masking the image to just liver parenchyma. This masked 3D volume served as the the training and validation image dataset for training a final 3D CNN classifier to predict for normal/mild/ moderate/severe steatosis. The final algorithm was validated on a reserved set of test examinations, composed of entirely independent patients from the training and validation sets. Standard statistical methods were used to generate the classification accuracy and receiver operating characteristic (ROC) curve of the final trained network with the Unenhanced CT image appearance used as the reference standard. The CNN was evaluated under two use cases unknown ***unenhanced*** and unknown ***portal venous*** CT abdomen and pelvis volumes to determine if the algorithm could successfully segment the liver parenchyma, and distinguish between grades of A-none, B-mild, C-moderate, and D-severe steatosis. The optimal ROC sensitivity and specificity of the CNN algorithm's hepatic steatosis detection were calculated and compared with similar prior studies on human inter-rater variability on hepatic steatosis. A polynomial regression model was fit on the 40 Training cases to predict Average Plain Density from Average PVP Density. The overall training and validation methodology is depicted in Fig. 1 on page 6 and Fig. 2 on page 6.

Inclusion and Exclusion Criteria:

Inclusion: All axial abdominal CT scans of the Portal Venous phase as well as the Unenhanced phase from patients with hepatic steatosis and normal livers as a control dating from 2006-2018.

Exclusion: All examinations with confounding dense artifact, insufficient technical quality or confounding liver/spleen pathology were excluded from the training and validation sets. These comprised hemosiderosis, focal solid liver lesions, liver mets or a prior history of malignancy. Small incidental likely benign lesions less than 2 cm were permitted,

but larger lesions that significantly alter the density of the liver organ as a whole were excluded.

Sample Size:

80 multiphasic single-energy CT examinations were selected, an even balance of mild/moderate/severe steatosis cases, and 20 normal controls. All volumes underwent standardized multi-ROI evaluation to determine an average Plain and PVP liver density, and were split into 40 Training and 40 Test cases.

Images for this section:

FIGURE 1: Training, Calibration, and Evaluation Stages for Segmentation and Average Plain Density Prediction

Overall experimental design composed of 3 stages: (1) 3D VNET training, (2) Segmentation & Regression Calibration, and (3) Final Combined Segmentation/Regression Evaluation. Initial VNET training was pre-trained using contour information from the MICCAI LITS competition with augmentation of steatotic liver contours from the Training set to improve contouring on severely steatotic livers. A multi-stage segmentation and regression process was validated in Stage 2 as well as determination of a polynomial spline regression fit on Training set data. Final evaluation occurred on a held-out Test set testing end-to-end performance of the system.

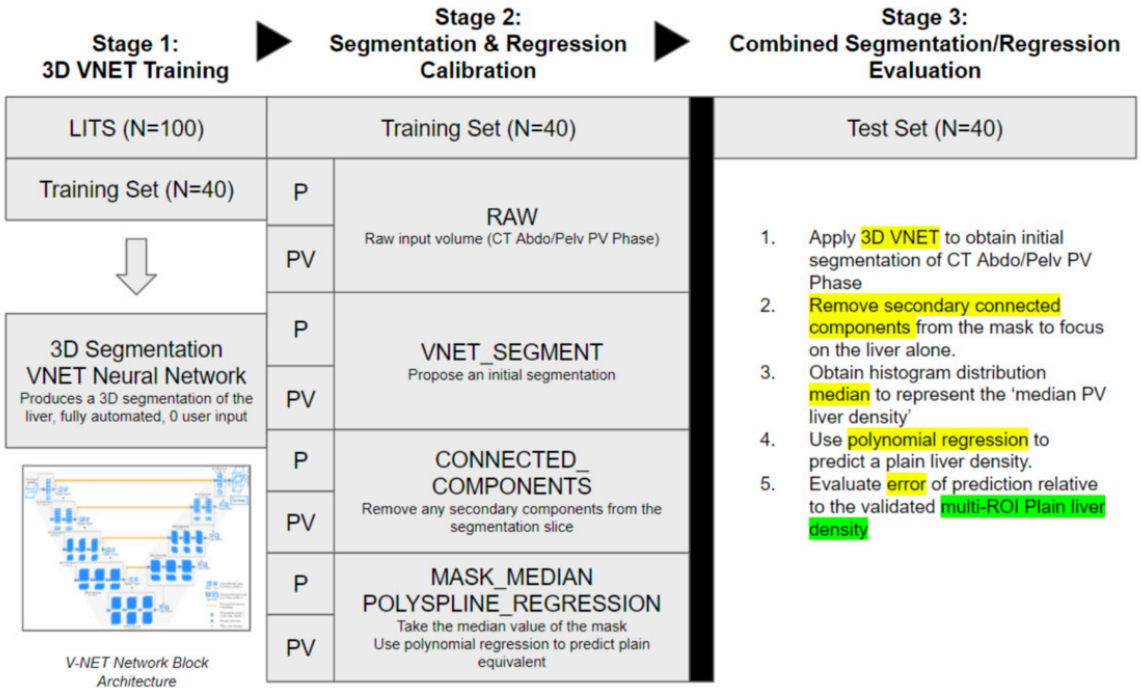

Fig. 1: Training, Calibration, and Evaluation Stages for Segmentation and Average Plain Density Prediction

FIGURE 2: 3D Auto Segmentation, Histogram, & Regression Algorithm Pipeline
The pre-trained 3D CNN ('V-NET') is applied to an unknown CT abdomen volume to obtain a contoured volume of the full liver organ. Voxel histogram analysis is performed to obtain the Portal Venous Phase (PVP) volume density median, thereby limiting the effect of outlier hyperdense (i.e. vascular) or hypodense (i.e. pneumobilia) structures. Using the polynomial regression derived from the Training set, a reverse formula lookup is performed, using the median PVP density as the input variable, and the Simulated Plain density as the output.

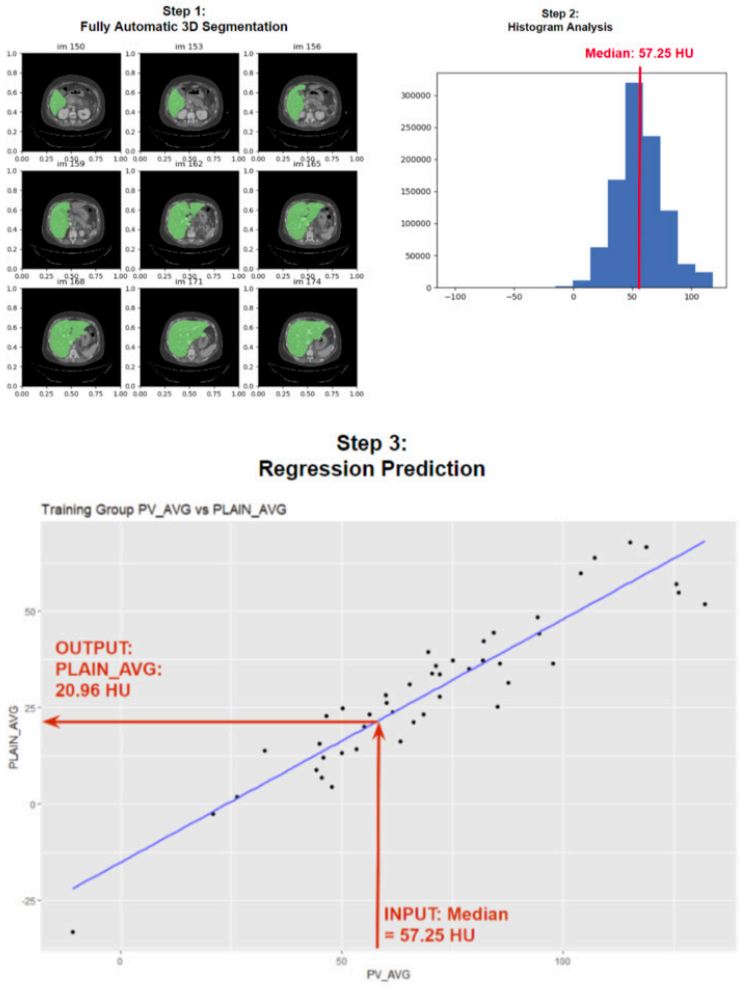

Fig. 2: 3D Auto Segmentation, Histogram, & Regression Algorithm Pipeline

Results

Evaluation of 3D-CNN segmentation on the held-out local Test set yielded a Mean Average Error (MAE) of 2.55 HU for the auto-segmented Median PVP density compared to the manual ROI PVP density (Fig. 3 on page 9). When evaluating the True versus the Simulated Plain Density on the Training dataset, the linear regression model demonstrated a R^2 of 0.873, RMSE of 6.979, and MAE of 5.82HU. The full end-to-end algorithm was evaluated on the 40 Test cases, which yielded a Simulated Average Plain Density MAE of 4.78HU.

Images for this section:

FIGURE 3: Algorithm Performance on Training and Test Sets, Portal Venous and Unenhanced Plain Density Results

Table demonstrating comparable ROI mean density measurements between the Training and Test sets. The polynomial regression model fit was good at 0.8308 for the test set. Simulated Predicted Plain Phase densities to ROI Plain Phase densities demonstrated a Mean Average Error of between 4-5 HU.

	Training Set (N=40)		Test Set (N=40)	
	PVP	**PP**	**PVP**	**PP**
ROI Mean Density	70.73 HU	29.47 HU	71.66 HU	31.13HU
R^2 of Training Set Polynomial Regression to ROI Measurements	0.8733		0.8308	
Mean Average Error (MAE) of Simulated Predicted Plain Phase Density to ROI Plain Phase Density		5.82 HU		4.78 HU

Fig. 3: Algorithm Performance on Training and Test Sets, Portal Venous and Unenhanced Plain Density Results

Conclusion

In our study, a hybrid 3D-CNN/statistical approach is able to predict Average Plain Density of a liver from contrasted examinations at a level comparable with that of human radiologist's assessment to within 4-5HU from standardized measurements of Uncontrasted CT. The use of Plain CT as an internal reference standard provided sufficient data for training both Unenhanced and PV DCNN's to identify hepatic steatosis to the degree possible from Unenhanced CT. Of note, this task is performed in a fully-automated fashion with zero user interaction from the interpreting radiologist requiring neither segmentation or user input.

The successful development of this fully automated, trained algorithm may allow for the quick, objective, and validated confirmation of steatosis on routine CT. Fundamentally, this network could provide value-added information to a reporting radiologist for any contrast CT abdomen case interpreted, who would then be free to agree or disagree with the conclusions drawn from the algorithm using routine clinical radiology analysis. The results provided by the AI would help direct the attention of an interpreting radiologist, particularly in milder or less obvious cases of hepatic steatosis as well as improve the sensitivity and specificity of Portal Venous characterization of steatosis, something that can be conventionally quite difficult or tedious to perform in everyday clinical radiological workflow.

References

Hamer, O. W., Aguirre, D. A., Casola, G., Lavine, J. E., Woenckhaus, M., & Sirlin, C. B. (2006). Fatty liver: imaging patterns and pitfalls. Radiographics, 26(6), 1637-1653.

Lee, D. H. (2017). Imaging evaluation of non-alcoholic fatty liver disease: focused on quantification. Clinical and molecular hepatology, 23(4), 290.

Lee, S. S., & Park, S. H. (2014). Radiologic evaluation of nonalcoholic fatty liver disease. World journal of gastroenterology: WJG, 20(23), 7392.

McFadden S, Roding T, de Vries G, Benwell M, Bijwaard H, Scheurleer J. Digital imaging and radiographic practise in diagnostic radiography: An overview of current knowledge and practice in Europe. Radiography (Lond). 2018. May;24(2):137-141.

Miyoshi, E., & Kamada, Y. (2015). Hepatosteatosis and Primary Hepatoma. In Glycoscience: Biology and Medicine(pp. 1365-1371). Springer, Tokyo.

Torres, D. M., Williams, C. D., & Harrison, S. A. (2012). Features, diagnosis, and treatment of nonalcoholic fatty liver disease. Clinical Gastroenterology and Hepatology, 10(8), 837-858.