

DEEP LEARNING BASED ACCURATE HEPATIC STEATOSIS QUANTIFICATION FOR HISTOLOGICAL ASSESSMENT OF LIVER BIOPSIES

Mousumi Roy¹, Fusheng Wang^{1,2}, Hoang Vo¹, Dejun Teng¹, George Teodoro³,
Alton B. Farris⁴, Eduardo Castillo Leon⁵, Miriam B. Vose^{5,6}, Jun Kong⁷

1 Department of Computer Science, Stony Brook University

2 Department of Biomedical Informatics, Stony Brook University

3 Department of Computer Science, Federal University of Minas Gerais

4 Department of Pathology and Laboratory Medicine, Emory University

5 Division of Gastroenterology, Hepatology, and Nutrition, Department of Pediatrics, Emory
University

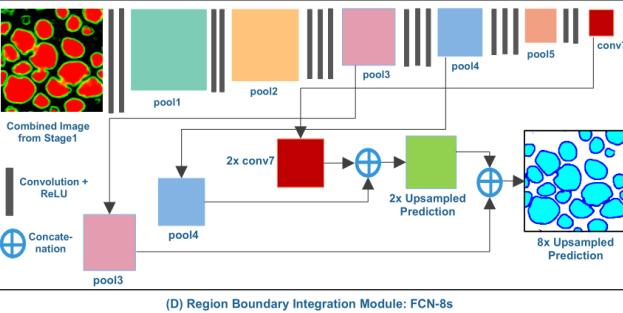
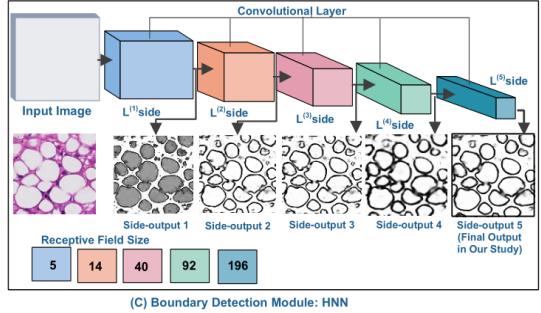
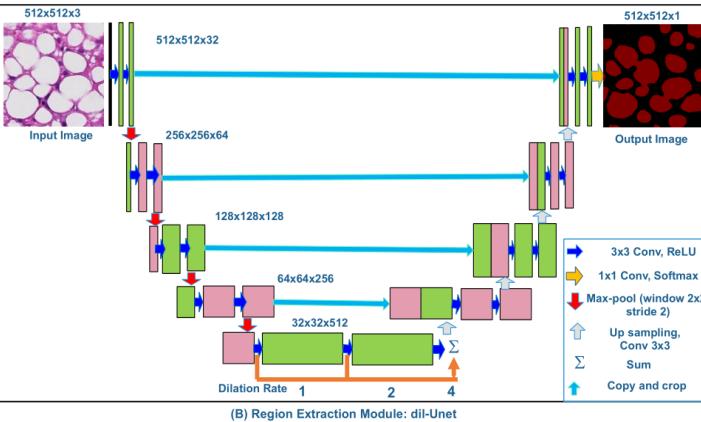
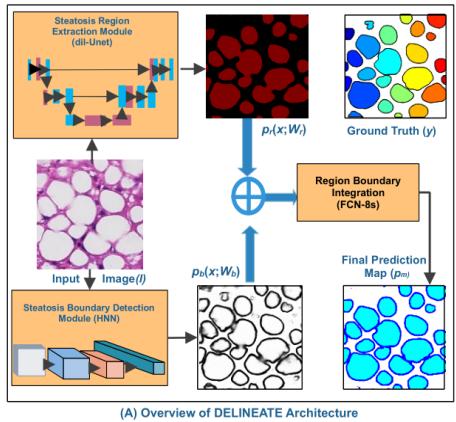
6 Children's Healthcare of Atlanta

7 Department of Mathematics and Statistics, Georgia State University

Introduction

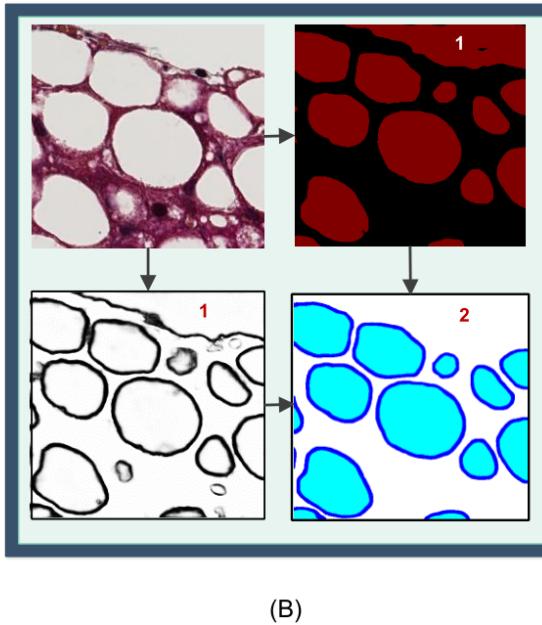
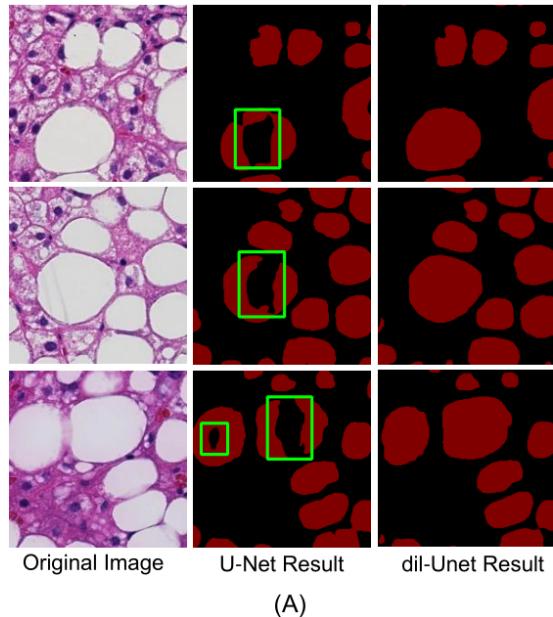
- Hepatic steatosis droplet quantification with histology biopsies has high clinical significance for risk stratification and management of patients with fatty liver diseases
- This process is challenging as there is a large number of overlapped steatosis droplets with either missing or weak boundaries
- We propose a deep learning-based region-boundary integrated network for precise steatosis quantification with whole slide liver histopathology images
- The proposed model consists of two sequential steps: a region extraction and a boundary prediction module for foreground regions and steatosis boundary prediction, followed by an integrated prediction map generation
- Missing steatosis boundaries are recovered from the predicted map and assembled from adjacent image patches to generate results for the whole slide histopathology image
- The resulting steatosis measures both at the pixel level and steatosis object level present strong correlation with pathologist annotations, radiology readouts and clinical data

DELINATE Model



- The DELINATE model first identifies regions and boundaries of steatosis droplets individually (A).
- The resulting two output predictions are combined for generating an integrated prediction map where the clumped steatosis regions are separated.
- The region extraction module detects steatosis regions with a dil-Unet module (B).
- The steatosis boundary detection module is based on a Holistically-Nested Network (HNN) (C).
- By detecting the boundary in an additional module, we can delineate the hidden boundaries of overlapped steatosis regions, and therefore, improve steatosis segmentation accuracy.
- The region-boundary integration network generates the final prediction output from the integrated region and boundary information (D).

Comparison of Segmentation Results



- Comparison of segmentation results between dil-Unet and standard U-Net (A)
- Left: original images; Middle: steatosis segmentation by U-Net; Right: steatosis segmentation by the proposed dil-Unet
- dil-Unet can recover steatosis regions with a substantially improved accuracy
- Comparison of results from the DELINEATE model (B)
- Top-Left: input image; Top-Right: output from the region extraction module; Bottom-Left: output from the boundary detection module; and Bottom-Right: final output of the integration module.
- “1” labels the false positive steatosis region captured by the region prediction module, and “2” labels the corrected steatosis regions by the final integration module

Evaluation of DELINEATE Segmentation Accuracy

- Dataset : whole slide liver histopathology images of 36 children with Nonalcoholic Fatty Liver Disease (NAFLD) collected from Children's Hospital of Atlanta and Emory University
- The segmentation accuracy of DELINEATE model was evaluated by five-fold cross-validation method
- The dataset was randomly partitioned into a training and a testing set by a ratio of 80:20
- Steatosis quantification accuracy was measured both at the object level and the pixel level
- The object-level measures include F_1 score, Precision, Recall, object level Dice index and Hausdorff Distance
- Object level Dice score $D(G, P) = \frac{1}{2} \left[\sum_{i=1}^{N_G} w_i D(G_i, P_i) + \sum_{j=1}^{N_P} \bar{w}_j D(\bar{G}_j, \bar{P}_j) \right]$

Where, G : ground truth set for steatosis instances and

P: set of machine segmented steatosis instances

For the i -th steatosis instance G_i in the ground truth set, we found the maximally overlapped segmented steatosis instance P_i in the same image and computed the Dice index $D(G_i, P_i)$. Similarly, for the j -th segmented steatosis instance \bar{P}_j , we detected the maximally overlapped ground truth steatosis instance \bar{G}_j and computed the Dice index $D(\bar{G}_j, \bar{P}_j)$

w_i is the ratio of the number of pixels in the i -th ground truth instance to the sum of all pixels in all steatosis components in the ground truth image

\bar{w}_i is the ratio of the number of pixels in the i -th instance from deep learning model to the sum of all pixels in all steatosis components in the automated segmented image

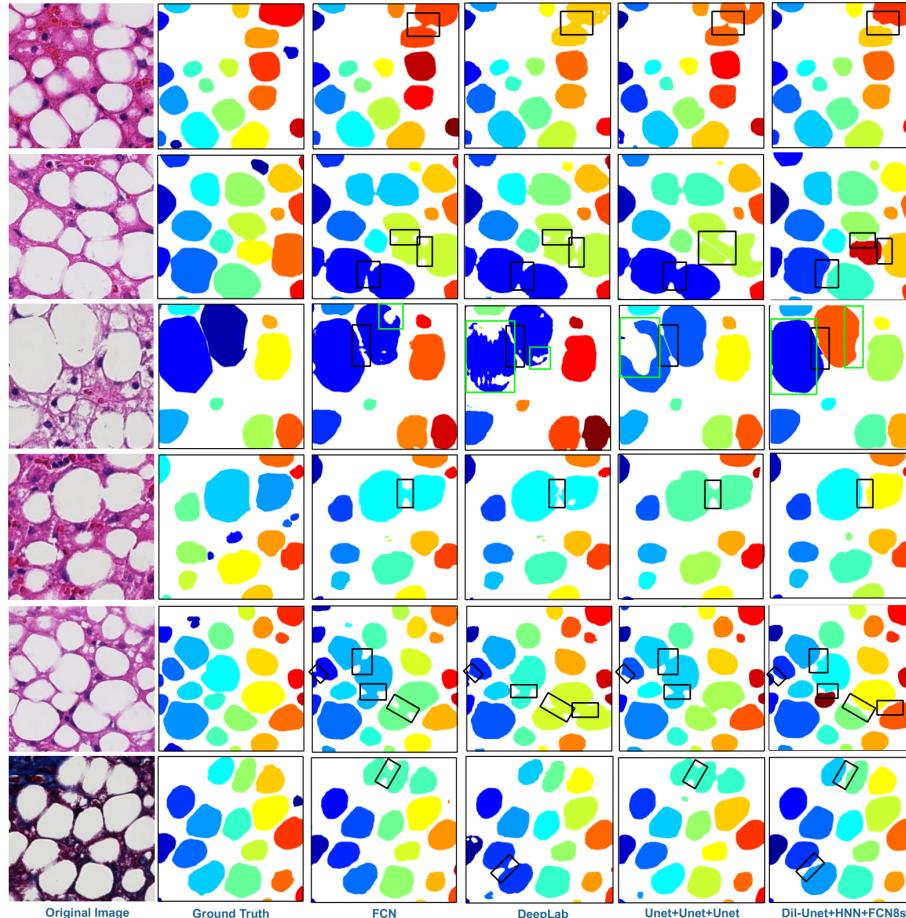
N_G and N_P : numbers of steatosis instances in the ground truth set and the corresponding machine segmented result set

Comprehensive Performance Comparison of Steatosis Segmentation Methods

Models	Approach	Precision	Recall	F1-Score	Object wise Dice Index	Object wise Hausdorff Distance
Standard Models	FCN	0.99 0.01	0.86 0.06	0.92 0.04	0.8338	3.8521
	DeepLab V2	0.99 0.01	0.83 0.08	0.90 0.05	0.9083	5.3179
Variations of Our Models	dil-Unet + HNN + FCN-8s	0.98 0.01	0.91 0.06	0.94 0.03	0.9492	3.4591
	dil-Unet + HNN + FCN-4s	0.97 0.01	0.91 0.06	0.94 0.03	0.9480	3.5753
	Unet + HNN + FCN-4s	0.97 0.01	0.91 0.06	0.94 0.03	0.9489	3.4685
	dil-Unet + HNN + dil-FCN	0.97 0.01	0.91 0.06	0.94 0.03	0.9459	3.6658
	Unet + Unet + Unet	0.97 0.04	0.83 0.07	0.90 0.05	0.9247	5.8289
	Unet + Unet + FCN-8s	0.96 0.03	0.90 0.06	0.93 0.04	0.9458	3.8773

- By contrast to other methods, Delineate dil-Unet+HED+FCN-8s achieves the best overall performance, as indicated by F1-score, Recall, object-wise Dice index and object-wise Hausdorff distance
- Delineate model substantially outperforms state-of-the-art FCN and DeepLab models and achieves better performance on delineating overlapped steatosis droplets

Evaluation of DELINEATE Segmentation Accuracy



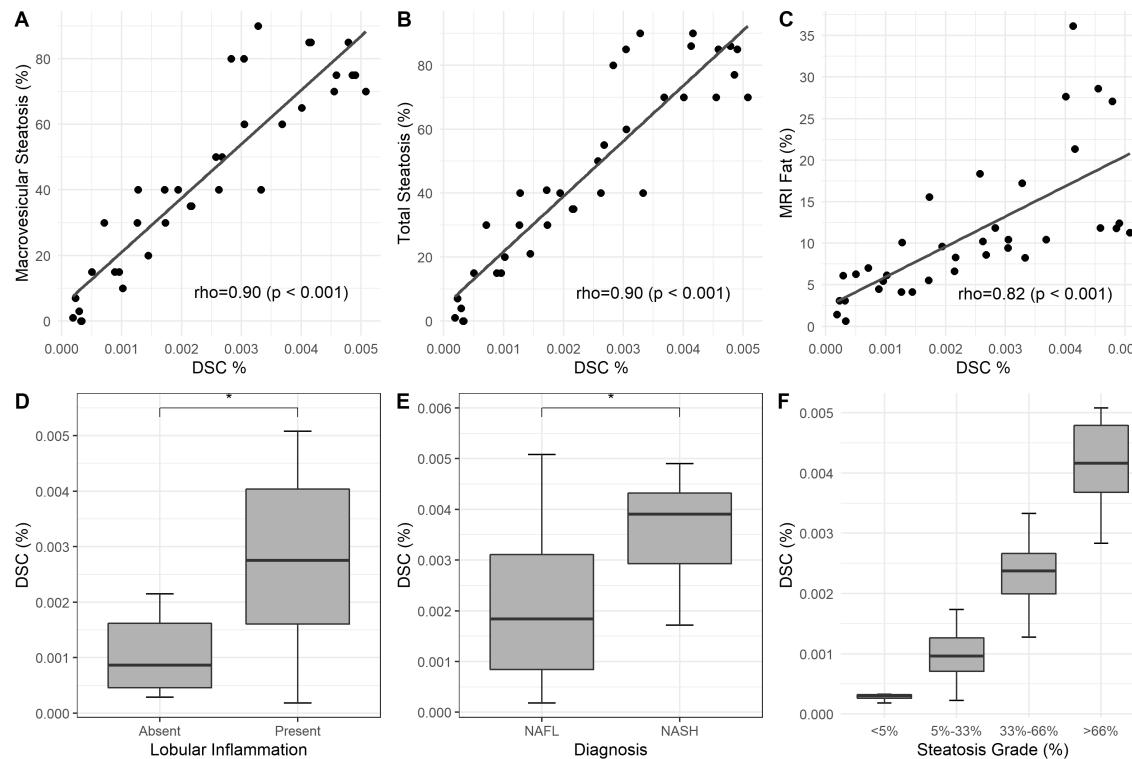
- The clumped steatosis regions indicated by black boxes in all images are well separated by Delineate model but failed by other methods in the comparison study. Problematic regions in green boxes are only fully recovered by Delineate model
- Touching regions highlighted by black boxes are well separated by Delineate model, whereas they are incorrectly segmented by other methods for comparison
- Neither FCN nor DeepLab can process touching steatosis droplets accurately with missing region-boundary integrative information, resulting in lower performance scores

DELINATE Correlation with Pathological Grading, Radiology, and Clinical Data



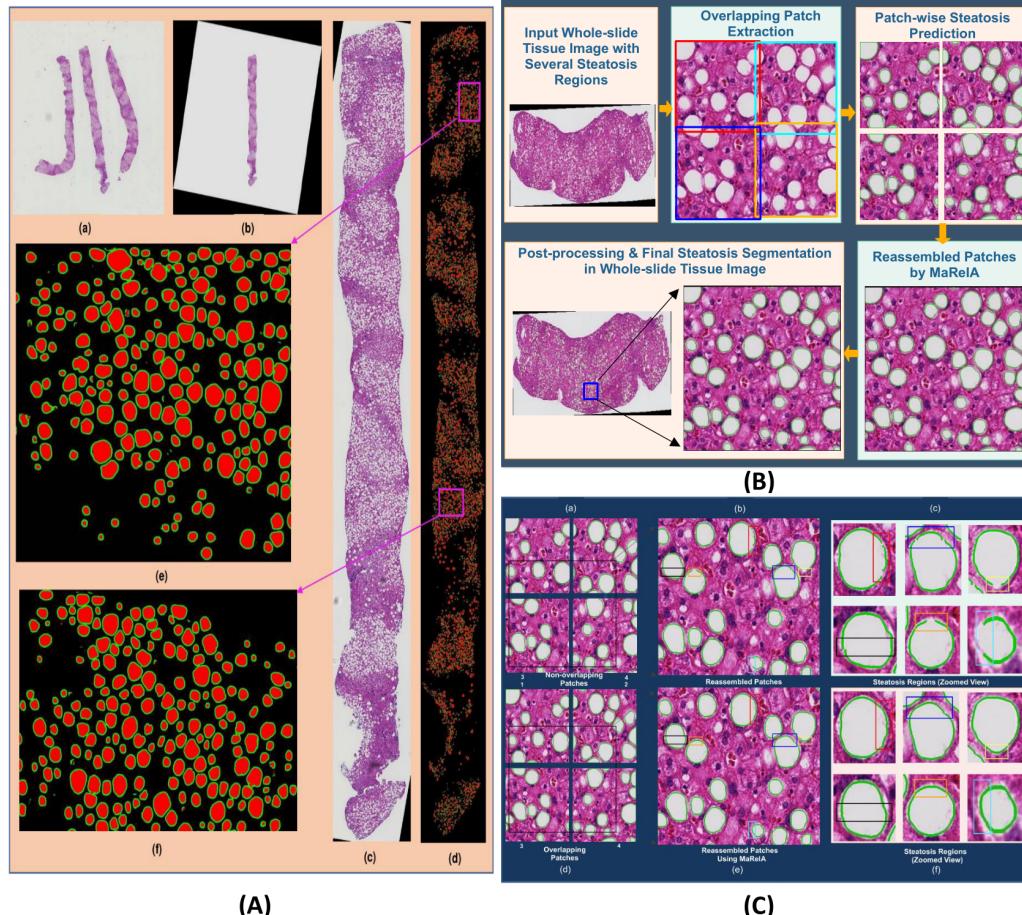
- The results produced by the DELINATE model present strong correlations with liver tissue pathological grading, fat quantity from MRI data, and patient clinical information
- Spearman's correlation was used to analyze the correlation between two variables
- Mann-Whitney test was used to compare the difference between two groups
- For comparisons across diverse steatosis groups, DELINATE was logarithmically transformed before analysis
- Analysis of Variance (ANOVA) was used to study the difference among four histological steatosis grading groups
- This was followed by Tukey's multiple comparison post-test
- We assessed the difference of steatosis measurements across four histological steatosis percentage grades by DSP%, DSC%, and ASP%
- Both DELINATE steatosis measures demonstrate statistically significant difference across grades, with the proposed steatosis count measure DSC% presenting the least p-value

DELINATE Correlation with Pathological Grading, Radiology, and Clinical Data



- Pair-wise scatter plots with correlation coefficients (p-values) are illustrated for DELINATE Steatosis Count% (DSC%) at individual droplet level and manual macrovesicular steatosis measures (A); manual total steatosis measures (B); manual fat readout from MRI images (C).
- We applied Mann-Whitney test to DSC% measures between Lobular Inflammation presence and absence (D); and between NAFL (i.e. Non-NASH) and NASH (E), respectively.
- We applied ANOVA to DSC% measurements of tissue samples among four manually graded histological steatosis percentage groups with p-value less than $5.25e^{-14}$ (F).

Whole Tissue Analysis and Visualization



Analysis & Conclusion

- DELINEATE method enables steatosis pixel to object level quantitation with recovered contours of individual steatosis droplets in clumps
- It allows for individual steatosis size measure and research investigations exploring new steatosis object level morphology features with diagnostic and prognostic value
- DELINEATE can also alleviate inter- and intra-observer variability in steatosis assessment
- Once the DELINEATE model is trained, it can be used as a tool to improve agreement among pathologists and across multiple clinical sites
- It is fully automated. Leveraging parallel computation, DELINEATE model can generate and aggregate steatosis analysis results from whole slide images efficiently