



Automated NASH CRN Scoring Pipeline Through Lesion Segmentation for Alcohol-Associated Liver Disease

Iljung Kim¹, Sangwoong Yoon², Yung-Kyun Noh^{1,2}, Chady Meroueh³, Joseph C. Ahn⁴, Resham Ramkissoon⁴, Peyman Nejat⁵, Jason D. Hipp³, Hamid R. Tizhoosh⁵, Vijay H. Shah⁴

¹Department of Computer Science, Hanyang University, Seoul, Republic of Korea,
²School of Computational Sciences, Korea Institute for Advanced Study, Seoul, Republic of Korea
³Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA
⁴Department of Gastroenterology, Mayo Clinic, Rochester, MN, USA
⁵Department of Artificial Intelligence and Informatics, Mayo Clinic, Rochester, MN, USA

iljung0810@hanyang.ac.kr swyoon@kias.re.kr nohyung@hanyang.ac.kr Meroueh.Chady@mayo.edu
Ahn.Joseph@mayo.edu Ramkissoon.Resham@mayo.edu Nejat.Peyman@mayo.edu Hipp.Jason@mayo.edu
Tizhoosh.Hamid@mayo.edu Shah.Vijay@mayo.edu



Introduction

The prevalence of alcohol-associated liver disease is increasing, and diagnosis usually relies on invasive and non-invasive methods, including liver biopsies.

The quantification of disease activity of ALD relies on scores derived from the NAFLD activity score NAS based on NASH-CRN scoring system.

There are significant inconsistencies in SLB scoring among human experts, which are considered a major issue.

Most machine learning (ML) NAS scores relies on NAFLD cases and are not sufficiently reproducible in ASH.

Objective

1. Develop an automated pipeline using artificial intelligence that provides a consistent NAS score for patient with ALD.

Method

1. Obtained liver biopsy 132 ASH slides.
2. Slides were digitalized with partial annotations.
3. Cropped 1024x1024 patches from region of interesting.

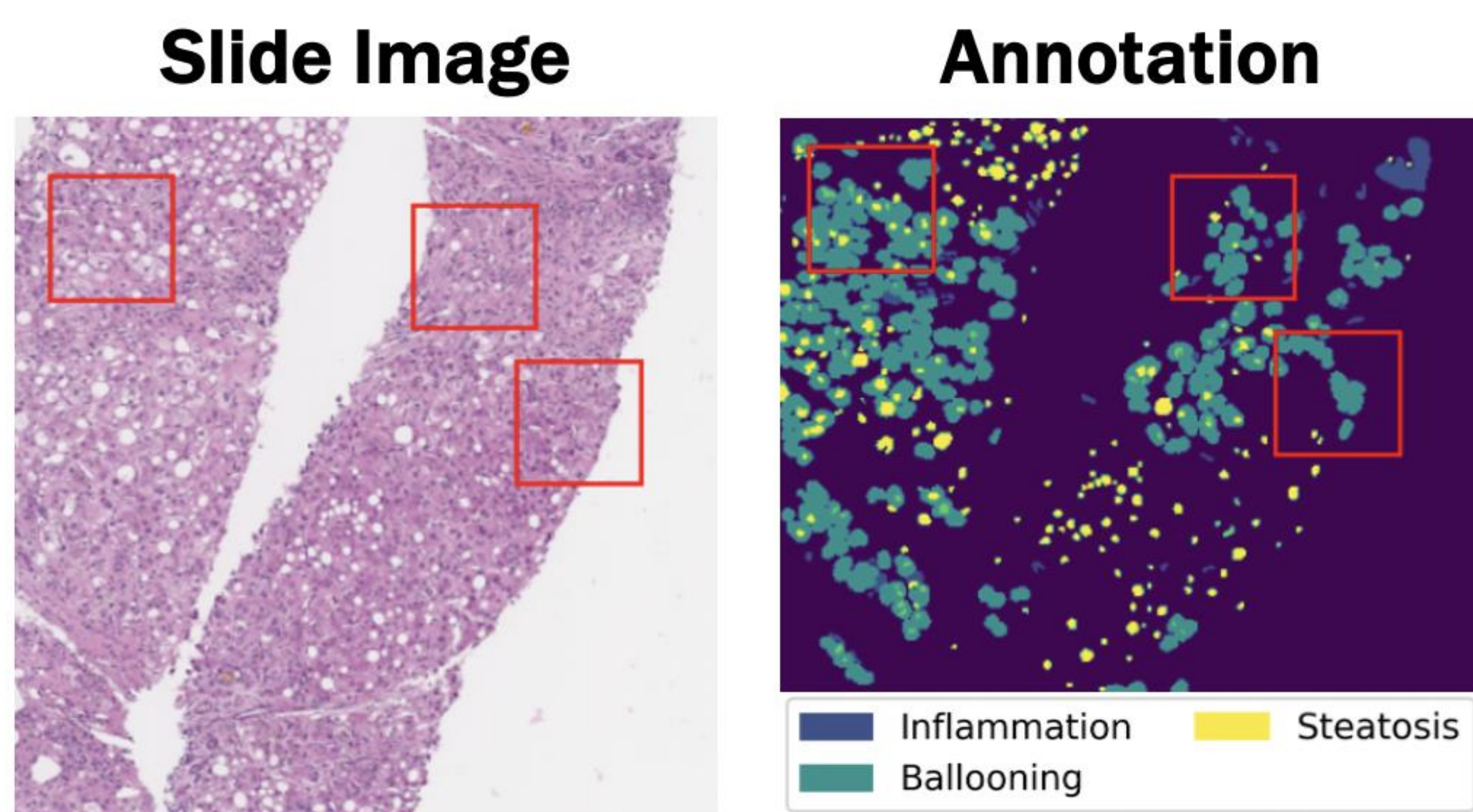


Figure 1. Example of patch sampling.

4. Train fully convolutional network (FCN) model with patches using focal loss.

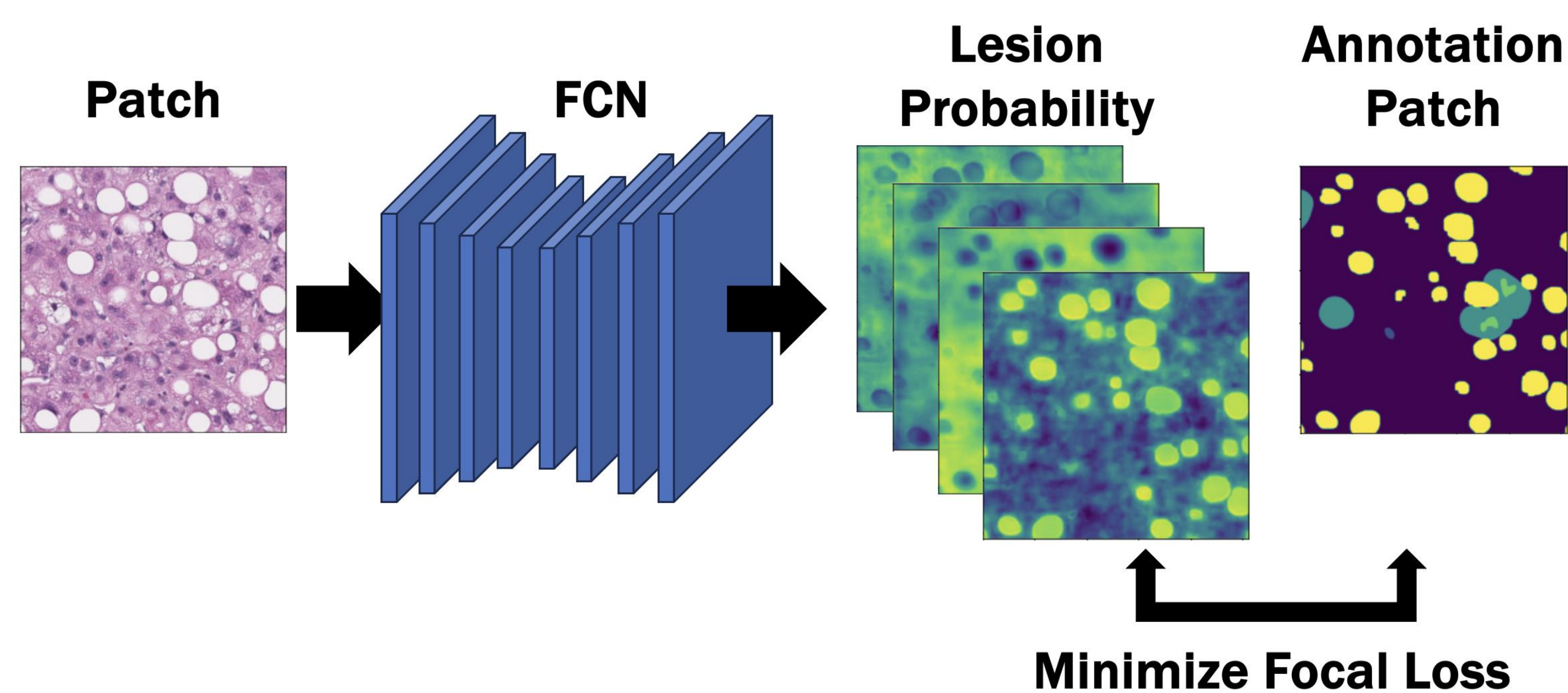


Figure 2. Processing of patch to probability prediction.

$$\text{Cross Entropy} = -\log(P_t)$$

$$\text{Focal Loss} = -(1 - P_t)^\gamma \log(P_t)$$

Using focal loss resolves class imbalance problem.

5. Count contour or area of lesions for scoring.

Results

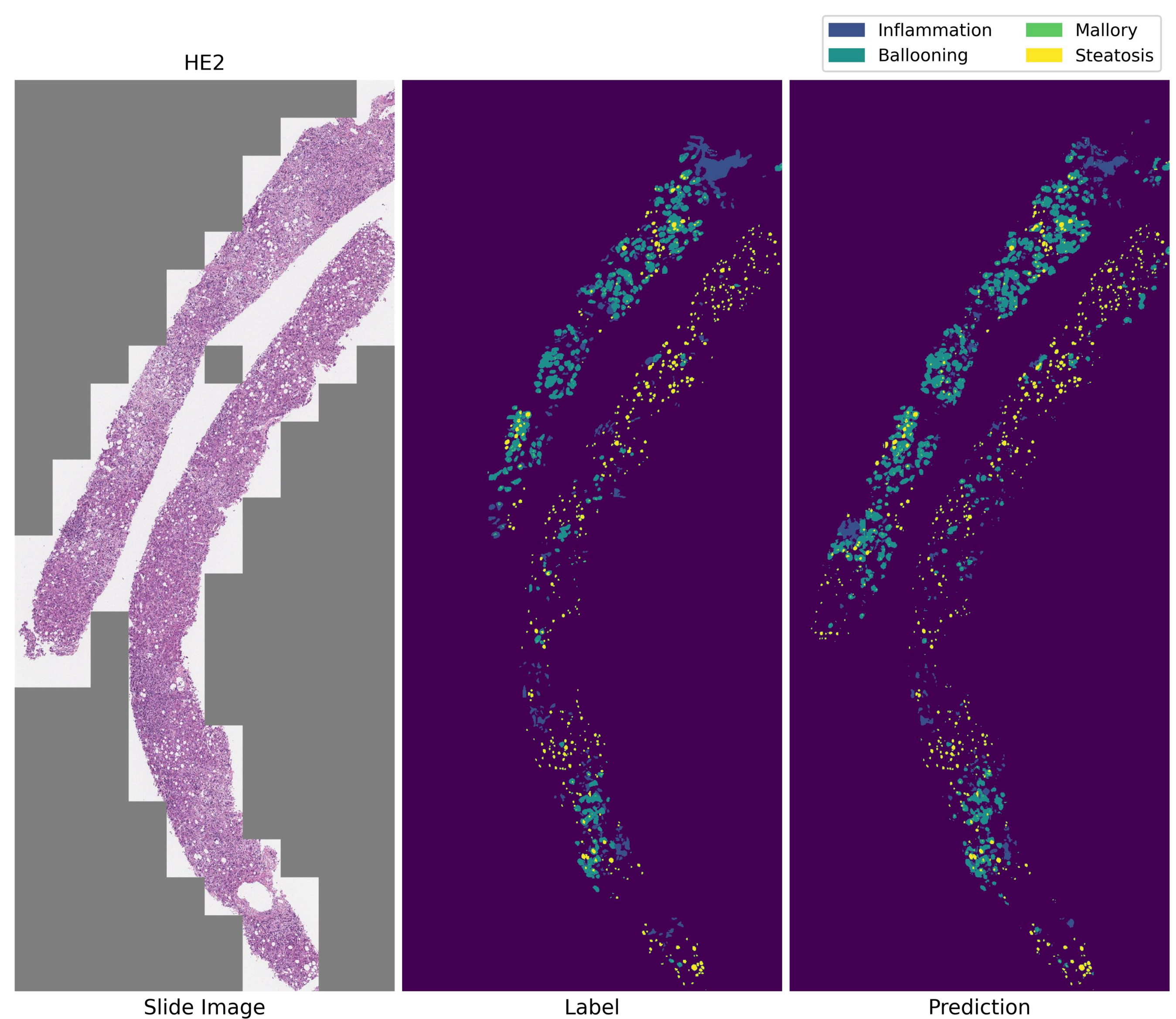


Figure 3. The FCN model's prediction and Label.

		Predicted Steatosis Score					
		0	1	1.5	2	2.5	3
Given Steatosis Score	0	5	3	-	-	-	-
	1	8	26	-	5	-	1
	1.5	-	3	-	4	-	-
	2	-	15	-	20	-	1
	2.5	-	1	-	7	-	1
	3	-	2	-	16	-	14

Accuracy = 65/132 = 49.2%
Correlation = 0.715

		Predicted Lobular Inflammation Score				Predicted Ballooning Score				
		0	1	2	3					
Given Lobular Inflammation Score	0	-	6	-	-	Given Ballooning Score	0	6	4	2
	1	4	62	3	5		1	10	37	17
	2	-	19	5	5		2	2	9	45
	3	1	6	3	13					

Accuracy = 80/132 = 60.6%
Correlation = 0.491

Accuracy = 88/132 = 66.7%
Correlation = 0.533

Table 1 –3. Performance of the model's prediction.

Conclusion

The demonstrated fully automated scoring system consistently provided accurate **automated NAS scoring pipeline** using artificial intelligence.

Using partially annotated lesions, FCN successfully segmented most of the lesion regions, yielding robust and trustworthy results NAS scoring.

The model can even identify slides that are obviously mislabeled by human experts.