Human Pathology (2014) **45**, 785–792

ELSEVIER

Human PATHOLOGY

www.elsevier.com/locate/humpath

Original contribution

Automatic classification of white regions in liver biopsies by supervised machine learning☆,☆☆

Scott Vanderbeck BS, MS[a], Joseph Bockhorst PhD[a], Richard Komorowski MD[b], David E. Kleiner MD, PhD[c], Samer Gawrieh MD[d,*]

[a]*Department of Electrical Engineering and Computer Science, University of Wisconsin, Milwaukee, WI 53211, USA*
[b]*Department of Pathology, Medical College of Wisconsin, Milwaukee, WI 53226, USA*
[c]*Laboratory of Pathology, National Cancer Institute, Bethesda, MD 20892, USA*
[d]*Division of Gastroenterology and Hepatology, Medical College of Wisconsin, Milwaukee, WI 53226, USA*

Received 29 July 2013; revised 12 November 2013; accepted 13 November 2013

Keywords:
NAFLD;
Steatosis;
Variability;
Sensitivity and specificity;
Digital image analysis

Summary Automated assessment of histological features of non-alcoholic fatty liver disease (NAFLD) may reduce human variability and provide continuous rather than semiquantitative measurement of these features. As part of a larger effort, we perform automatic classification of steatosis, the cardinal feature of NAFLD, and other regions that manifest as white in images of hematoxylin and eosin–stained liver biopsy sections. These regions include macrosteatosis, central veins, portal veins, portal arteries, sinusoids and bile ducts. Digital images of hematoxylin and eosin–stained slides of 47 liver biopsies from patients with normal liver histology (n = 20) and NAFLD (n = 27) were obtained at 20× magnification. The images were analyzed using supervised machine learning classifiers created from annotations provided by two expert pathologists. The classification algorithm performs with 89% overall accuracy. It identified macrosteatosis, bile ducts, portal veins and sinusoids with high precision and recall (≥82%). Identification of central veins and portal arteries was less robust but still good. The accuracy of the classifier in identifying macrosteatosis is the best reported. The accurate automated identification of macrosteatosis achieved with this algorithm has useful clinical and research-related applications. The accurate detection of liver microscopic anatomical landmarks may facilitate important subsequent tasks, such as localization of other histological lesions according to liver microscopic anatomy.

1. Introduction

Non-alcoholic fatty liver disease (NAFLD) is the most common liver disease in children and adults in the United States [1,2]. NAFLD has a spectrum that starts with a mild phenotype simple steatosis where only steatosis is present, and ends with a severe phenotype non-alcoholic steatohepatitis where steatosis is present with hepatic necro-inflammation and fibrosis [3]. Accurate distinction of mild from severe phenotypes of the disease is essential because simple steatosis rarely progresses, whereas non-alcoholic steatohepatitis can

☆ Funding statement: This research was supported by the University of Wisconsin-Milwaukee Research Foundation (J.B. and S.G.) grant number PRJ-39IB and in part by the Intramural Research Program of the NIH, National Cancer Institute (D.E.K.).

* Corresponding author. Division of Gastroenterology and Hepatology, Indiana University School of Medicine, Indianapolis, IN 46202, USA.
E-mail address: sgawrieh@iu.edu (S. Gawrieh).

0046-8177/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.humpath.2013.11.011

progress to cirrhosis, liver failure, and hepatocellular carcinoma [4-7]. Phenotyping of NAFLD is currently based on a pathologist's evaluation of the "gold standard" diagnostic test, liver biopsy [3].

The state-of-the-art scoring system of liver biopsies performed for NAFLD is based on manual pathologist assessment and semi-quantification of four key histological features: steatosis grade, lobular inflammation, fibrosis stage and degree of hepatocyte ballooning [8]. It is therefore critical that these features of injury are reliably and reproducibly recognized.

Our group and others [8-15] have demonstrated that there is intra- and inter-observer variability in pathologists' assessment of NAFLD histological features and assignment of a final disease diagnosis / phenotype. Automation of this assessment may reduce variability, increase accuracy, and offer a continuous rather than semi-quantitative grading of these histological features of NAFLD.

The overall aim of our ongoing project is to test the hypothesis that decision support systems for pathologists, which include computational methods for quantification of the key histological features of NAFLD, will lead to continuous, more accurate and less variable scores and ultimately better phenotyping and improved patient outcome. Key liver biopsy features—including macrosteatosis, central veins (CV), portal veins (PV), portal arteries (PA), sinusoids (SN) and bile ducts (BD)—oftentimes contain a white region, or manifest entirely as white regions, in liver biopsies when stained with a hematoxylin-eosin (H&E) stain as shown in Fig. 1. In this paper, we present an important subtask of our project—the accurate categorization of the white regions in liver biopsy images.

2. Materials and methods

This research protocol was reviewed and approved by the Internal Review Board of the Medical College of Wisconsin. A dataset consisting of 59 unique liver biopsy scans is included in our dataset and represents patients with a full range of different diagnosis and steatosis grades. Of our 59 patients, pathologist steatosis grading was available from two different pathologists for 47 patients (20 with normal liver histology and 27 with NAFLD of varying severity). Using a custom-built, Web-based Java Applet, our study pathologists manually annotated biopsy images and categorized 1969 different white regions.

Supplementary Table 1 provides counts for the total number of annotations for each feature type in our dataset, and Supplementary Fig. 1 shows the count of patients in our study by steatosis grade. There is overall good agreement between pathologists on the grade for each patient with D.E.K. having a slightly lower threshold for calling small degrees of steatosis as grade 1 than R.K. (Supplementary Fig. 1).

We carried out the following steps to classify white regions (Fig. 2):

1. Obtain Image: liver biopsy samples are obtained and stained with a H&E stain. They are then scanned at 20× magnification using a NanoZoomer scanner (Hamamatsu) and stored as RGB images with 8 bits per color channel.
2. Convert RGB image to gray scale: 2 separate 8-bit gray scale images are created. The first is used to distinguish between tissue sample and background. The second is created with higher contrast and used for identifying white regions within the tissue sample.
3. Conversion to black and white image: Otsu's method [16] is used to determine a global threshold value from our first gray scale image. This threshold is then applied to convert the image to black and white.
4. Identify biopsy (image foreground versus background): a region growing algorithm is used to separate image foreground from image background. Any small artifacts outside of the primary biopsy tissue region are removed from the image foreground based on their small size.
5. Image Adjustment: the gray scale image created from the green channel is smoothed using a 3×3 average

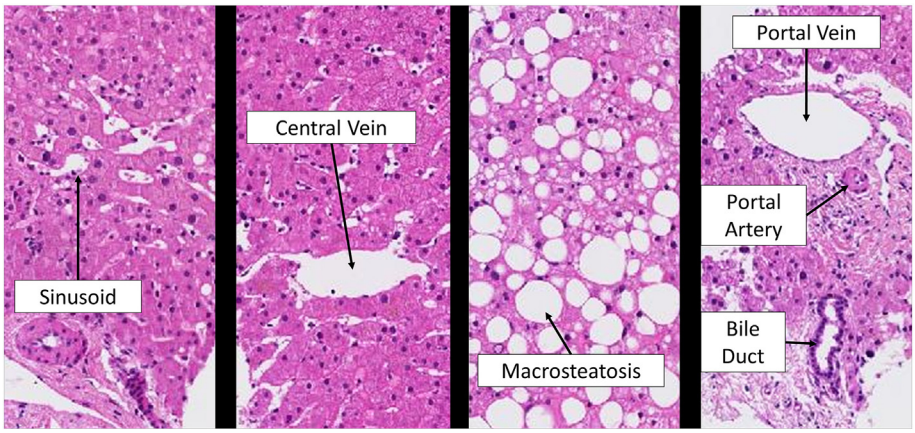

Fig. 1 White regions in liver biopsies.