




Review

# Quantification of Liver Fibrosis—A Comparative Study

Alexandros Arjmand<sup>1,\*</sup>, Markos G. Tsipouras<sup>2,\*</sup>, Alexandros T. Tzallas<sup>1</sup> , Roberta Forlano<sup>3</sup> ,  
Pinelopi Manousou<sup>3</sup> and Nikolaos Giannakeas<sup>1,\*</sup> 

<sup>1</sup> Department of Informatics and Telecommunications, University of Ioannina, GR47100 Arta, Greece; tzallas@uoi.gr

<sup>2</sup> Department of Electrical and Computer Engineering, University of Western Macedonia, GR50100 Kozani, Greece

<sup>3</sup> Liver Unit/Division of Integrative Systems Medicine and Digestive Disease, Department of Surgery and 12 Cancer, Imperial College, London SW7 2AZ, UK; r.forlano@imperial.ac.uk (R.F.); pinelopi.manousou@nhs.net (P.M.)

\* Correspondence: alexarjmd@gmail.com (A.A.); mtsipouras@uowm.gr (M.G.T.); giannakeas@uoi.gr (N.G.)

Received: 10 November 2019; Accepted: 1 January 2020; Published: 8 January 2020



**Abstract:** Liver disease has been targeted as the fifth most common cause of death worldwide and tends to steadily rise. In the last three decades, several publications focused on the quantification of liver fibrosis by means of the estimation of the collagen proportional area (CPA) in liver biopsies obtained from digital image analysis (DIA). In this paper, early and recent studies on this topic have been reviewed according to these research aims: the datasets used for the analysis, the employed image processing techniques, the obtained results, and the derived conclusions. The purpose is to identify the major strengths and “gray-areas” in the landscape of this topic.

**Keywords:** digital image processing; medical image analysis; liver biopsy images; histological staining; collagen proportional area; liver fibrosis

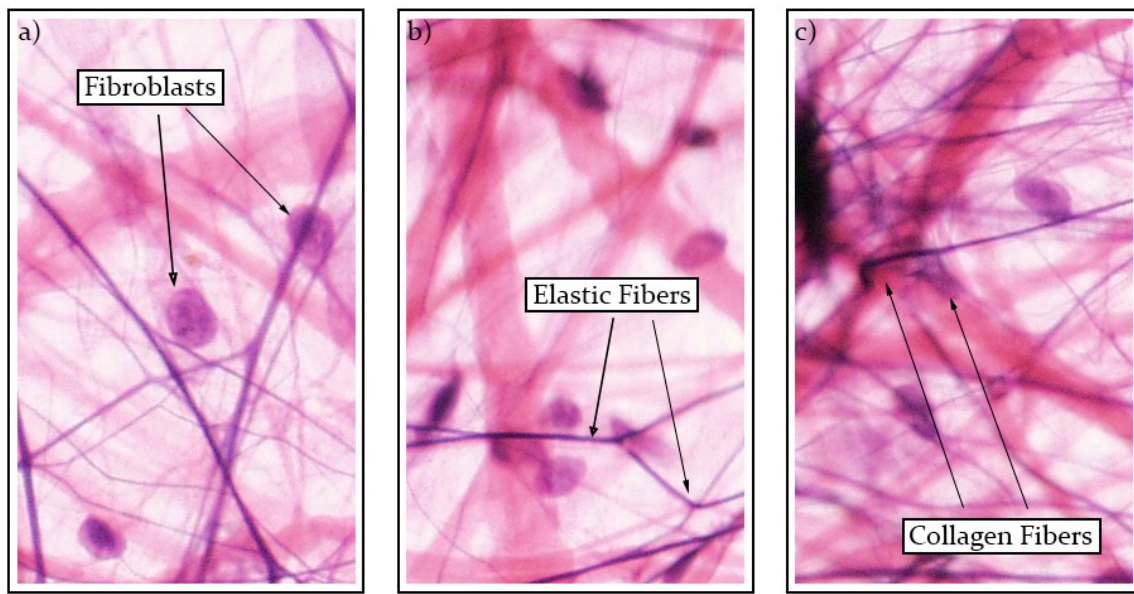
## 1. Introduction

Early diagnosis and accurate staging have always been considered essential for the determination of a treatment strategy. Recently, nearly 35 million people globally have died due to hepatic diseases [1]. During the last few decades, the clinical management of chronic liver disease (CLD) has been increasingly focused on the prevention of the development or progression of fibrosis. Liver fibrosis develops as a result of chronic damage, which eventually leads to cirrhosis, marking the final stage of liver disease. Technically, thick collagen fibers have been targeted as the main source of fibrosis (Figure 1). During chronic hepatic injury, stellate hepatocytes are activated and transformed into a myofibroblast-like phenotype forming an extracellular matrix (ECM). The main causes of stimulating these cells include tissue inflammation, cytokine production from injured parenchymal cells, and disruption of the extracellular matrix [2].

Fibrosis estimation is a standard procedure, where liver biopsies are formalin-fixed and paraffin-embedded. Thereafter, the specimens are commonly cut to a thickness of 4  $\mu\text{m}$  and stained using several dyes. The most common dyes used for fibrosis estimation are the Masson’s trichrome (MT) and Sirius red (SR). The above dyes bind selectively to liver collagen, allowing for fibrosis quantitation. Liver biopsy has become the gold standard in most clinical diagnostic efforts [3]. The needle biopsy specimens are examined from pathologists in order to determine the stage of the liver disease. The specimens are staged and graded according to a scoring system, with respect to the zonation as well as the number and the width of septa. The score represents categories of increasing severity based on a combination of the fibrosis amount and fibrous forms such as septa, bridges and

nodules [4]. There are four scoring systems—histology activity indexes (HAIs) for the staging of liver diseases:

- Knodell HAI [5];
- Scheuer HAI [6];
- METAVIR [7];
- Ishak HAI [8].










**Figure 1.** Visualization of individual findings found in fibrotic liver tissue samples. (a) Fibroblasts as principal cells accompanying almost all connective tissues. (b) Thin elastic fibers produced by the elastin protein. (c) Advanced accumulation of broad collagen fibers, which can lead to the end-stage of hepatic cirrhosis.

The indexes are presented in the following Table 1.

**Table 1.** Histology activity indexes (HAIs) as scoring systems for fibrosis and cirrhosis grading.

Knodell [5]		Scheuer [6]	
Grade	Description	Grade	Description
0	No fibrosis	0	None
1	Fibrous portal expansion	1	Enlarged, fibrotic portal tracts
2	Bridging fibrosis (portal-portal or portal-central linkage)	2	Periportal or portal-portal septa but intact architecture
3	Cirrhosis	3	Periportal or portal-portal septa but intact architecture
		4	Probable or definite cirrhosis
METAVIR [7]		Ishak [8]	
Grade	Description	Grade	Description
0	No fibrosis	0	No fibrosis
1	Stellate enlargement of portal tract, but without septa formation	1	Fibrous expansion of some portal areas, with or without short fibrous septa
2	Enlargement of portal tract with rare septa formation	2	Fibrous expansion of most portal areas, with or without short fibrous septa
3	Numerous septa without cirrhosis	3	Fibrous expansion of most portal areas with an occasional portal to portal (P-P) bridging
4	Cirrhosis	4	Fibrous expansion of portal areas with marked bridging (P-P) as well as portal-central (P-C)
		5	Marked bridging (P-P and/or P-C) with occasional nodules (incomplete cirrhosis)
		6	Cirrhosis, probable or definite

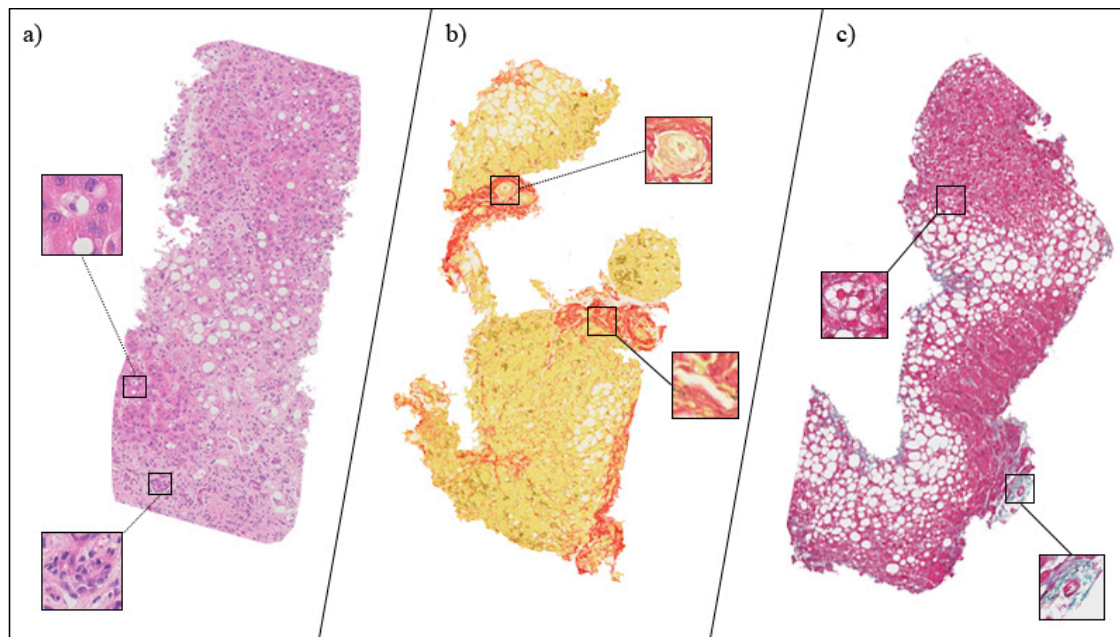
The Knodell HAI has four stages (0 to 3), the Scheuer and METAVIR use five stages (0 to 4), and the Ishak HAI (Figure 2) refers to seven stages (0 to 6). These scoring systems have been as well introduced as standards for all semi-quantitative methods in the classification of the chronic active hepatitis (CAH) [9]. The last stage of the Knodell, Scheuer, and METAVIR scoring systems describes cirrhosis, while in the Ishak HAI, stage 5 indicates incomplete cirrhosis with occasionally detected nodules and stage 6 definite cirrhosis, respectively. The main disadvantage of the diagnosis based on semi-quantitative systems is the subjectivity of the observer, commonly referred to as “intra-observer” variability.

Iskak grade	Categorical description	Histological visualization
0	No Fibrosis	
1	Fibrous expansion of some portal areas, ± short fibrous septa	
2	Fibrous expansion of most portal areas, ± short fibrous septa	
3	Fibrous expansion of most portal areas with an occasional portal to portal (P-P) bridging	
4	Fibrous expansion of portal areas with marked bridging (P-P) as well as portal-central (P-C)	
5	Marked bridging (P-P and/or P-C) with occasional nodules (incomplete cirrhosis)	
6	Cirrhosis, probable or definite	

**Figure 2.** Illustration of the Ishak system stages (0–6) on histologically stained Sirius red (SR) tissue sections.

A quantitative approach for fibrosis scoring is the measurement of the collagen proportional area (CPA), which is the ratio of the collagen area to the hepatic tissue area in microscopy liver biopsy images using digital image analysis (DIA) techniques. Several studies have been presented in the literature in the past three and a half decades, with the majority focusing on the clinical validation of CPA by investigating the correlation of semi-quantitative scoring systems (SSS) with the CPA obtained from DIA.

This study aims to identify the strengths and limitations of DIA techniques and to point out “gray-areas” as diagnostic obstacles in the major findings classification landscape, which tend to make unreliable the interpretation of medical results. The rest of the paper is structured as follows: Section 2 focuses on the search steps followed to collect all reviewed research materials, with image analysis methods aiming at identifying fibrotic areas of collagen accumulation in liver biopsy specimens (Figure 3). Subsequently, Section 3 describes the used methodological approach (divided into subsections). The results produced from each of the studies presented in the literature are discussed in Section 4, with the concluding remarks listed in Section 5.



**Figure 3.** Tissue sections colored with three of the most widely used histological stains: (a) Hematoxylin and Eosin (H&E), (b) Sirius red (SR), and (c) Masson's trichrome (MT) stains. In the H&E sample, Hematoxylin stains the cell nuclei in blue, while Eosin stains the extracellular matrix and the cytoplasm in pink. In SR, most tissue objects are stained in yellow, which helps to emphasize the red collagen fibers that may surround healthy structures, such as the hepatic veins or sinusoids. MT stains most of the structures in dark red/purple (i.e., ballooned hepatocytes), which aims to further highlight the blue areas of collagen accumulation.

## 2. Materials and Methods

### 2.1. Data Sources

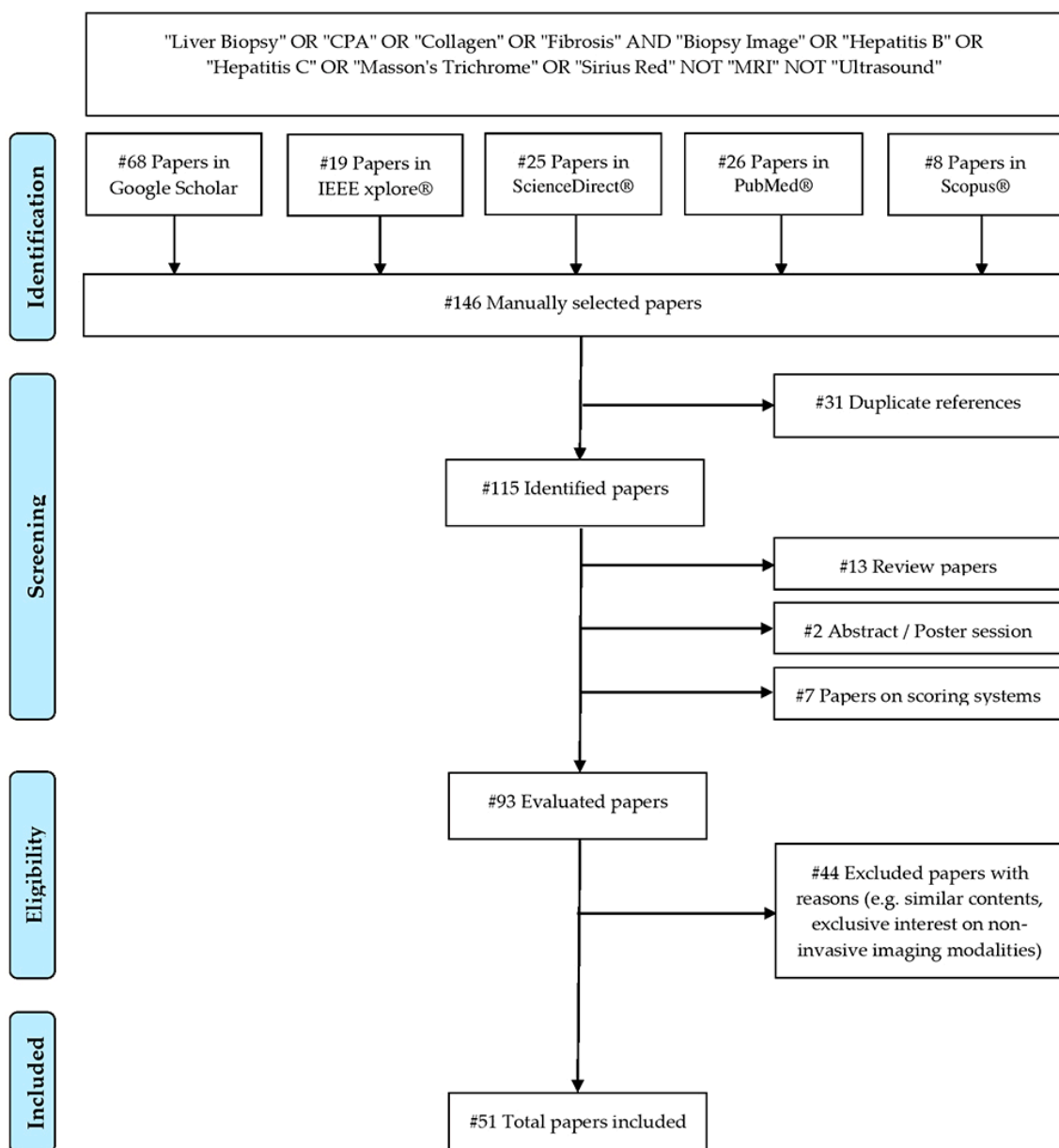
In this context, we provide an analysis of the main publications from the field of histological image analysis, focusing on datasets and key findings to be identified. All the image processing techniques that have been employed from various studies are described as concisely and comprehensively as possible. All publications have been acquired from an efficient number of digital library sources, including

- Google Scholar, a web engine ideal for searching free academic publications along with conference papers;
- ScienceDirect<sup>®</sup>, characterized as the leading web source in the access to a large database, emphatically composed of medical research works;
- PubMed<sup>®</sup>, a free search engine of the National Center for Biotechnology Information (NCBI) which is part of the National Library of Medicine (NLM) in the United States;
- Scopus<sup>®</sup>, the official database of Elsevier<sup>®</sup> and ideal for navigation between many citations, according to their research field;
- IEEE Xplore<sup>®</sup>, a digital library that provides research materials, especially related to computer science and engineering.

### 2.2. Search Terms

Throughout the search for all the works reviewed in this paper, a rigorous process was followed to collect a sufficient number of classic and new methodologies that reflect as much as possible the initial concept of CPA as a novel idea for the fibrosis quantification in liver biopsy specimens. A specific chronological order was also used to select the results of the research.

- In the Google Scholar engine, the keywords “CPA, Fibrosis, Liver, Image” were applied, where the process was completed on the 11th tab as the results included more and more posts not related to the subject of interest.
- In ScienceDirect® and the “Keywords” field, the input of “Liver Biopsy, CPA, Collagen, Fibrosis and Biopsy Image” was included.
- All of the above steps were also followed in the PubMed® database to verify the publications obtained with the research. In Scopus® and the “Author” field, the name of five researchers who have been distinguished in this field of research was declared.
- In IEEE Xplore®, the research process with all the previous search terms plus the addition of other keywords, such as “Hepatitis B, Hepatitis C, Sirius red, Masson’s trichrome.” The full search command is displayed in Figure 4. At the end of the search, a large number of results appeared was limited to less than 120 papers by applying system-proposed checkbox filters, including “conference” and “journal” papers, “liver,” and “medical image processing.”



**Figure 4.** Paper selection flowchart according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.



### 2.3. Study Selection Process

Original full-text articles published in English (except the article of “Nakayashi et al.”) were included in this review, focusing on the use of liver biopsy specimens for the separation of collagen fibers and quantification of liver fibrosis. According to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [10] (Figure 4), duplicate references that were manually identified were excluded from the entire set of collected documents. Afterward, during the screening process, additional items were excluded if they were review articles, abstracts, or even conference poster presentations.

Extra articles were examined and excluded by each author independently based on the following key reasons: (1) they did not report performance evaluation metrics or statistics, (2) they contained similar methodological steps to previously published works, (3) they were exclusively based on different than histology (microscopy) imaging modalities (magnetic resonance, ultrasound, shear wave elastography, etc.), and (4) they did not appear appropriate for inclusion in this review after the reading of title and abstract. In addition, (5) if multiple articles written by the same author had similar content, articles published in journals were included instead of articles presented at conferences. Moreover, (6) if multiple conference papers with similar content were written by the same author, the most recent one was selected. In general, decisions on the inclusion, exclusion, and classification of articles were resolved through meetings and discussions among the authors.

### 2.4. Analysis Method

The articles selected for this review were divided into categories and then grouped into four summary tables based on their methodological approach. For instance, the first category includes studies that used semi-quantitative determinations of the fibrosis and cirrhosis stage with the four previously referred or more HAIs. In the same category, innovative works that revolutionized the field of digital histopathology were added. The performance of these implemented methods had to be compared with those coming from expert histopathologists, who used the four scoring systems (HAIs) to obtain approval for their application in clinical settings.

The second category of articles encompasses works comparing digital microscopy with non-invasive medical imaging methods. However, their efficacy was compared to that of the liver biopsy gold standard. The third category features research works in which second-harmonic generation/two-photon excitation fluorescence (SHG/TPEF) microscopy was employed to display collagen areas in non-stained histological samples. Even here, the performance of this relatively new technology was compared to the traditional and more established light microscopy.

For the fourth and last category, a review on intelligent diagnostic systems is presented, in which they apply a more independent approach to the liver fibrosis quantification problem. As expected at this point, recent methodologies combining image processing techniques with machine learning algorithms, for the automated segmentation and classification of chronic disease findings, are analyzed. Before closing this group, state-of-the-art methods based on deep neural networks are presented. These approaches provide a fully automated analysis of liver specimens without the inclusion of physician-labeled histological data.

## 3. Literature Review

Starting in Section 3.1, the aim of each study presented in the literature is described in detail. Section 3.2 refers to liver biopsies available to research teams and the characteristics of donor disease. The histological stains proposed by specialists are presented in Sections 3.3 and 3.4 the fibrosis grading systems to assess the integrity of diagnostic methods. The digitization equipment, as well as the image analysis steps, are described in Section 3.5. Section 4 focuses solely on the author’s diagnostic results and other observations.

### 3.1. Aim of Study

#### 3.1.1. Early Studies Using Semi-Quantitative CPA Evaluations

Early studies [11–14] used microscopy image analysis as a more sensitive method to assess collagen quantitation in liver biopsies, recognizing the lack of other quantitative indexes and scoring techniques. Yet, these early studies noticed the disadvantages of image analysis as a quantitation technique, since this was time-consuming and needed sophisticated software and hardware equipment so that all small changes in fibrosis could be detected. Part of early studies [14–17] reported a high correlation of the CPA with most of the staging and scoring systems, thus confirming the validity of the image analysis-based CPA assessment as a quantitative scoring system. Having that as a major result of several studies in the 90 s, researchers targeted the image analysis process, trying to develop and validate image analysis tools specifically designed for the CPA assessment.

In this context, Manabe et al. [12] calculated the area of fibrosis through image analysis, which was used to assess the fibrogenic reduction of chronic non-A and non-B hepatitis activity after interferon (IFN)- $\alpha$  treatment. Nakayashi et al. [13] studied the quantitative determination of liver collagen and compared different staining dye combinations. Chevallier et al. [14] relied on the correlation of the collagen density quantification results with those produced by the widespread Knodell scoring system. Nakabayashi et al. [15] demonstrated the relationship between liver collagen content and histological grading of fibrosis based on a modified Knodell's score. Kage et al. [16] studied the correlation between the Scheuer staging system and the area of fibrosis (analogous to the collagen index—CI). Pilette et al. [17] studied the correlation of the serum concentration of several potential markers of hepatic fibrosis, with two different scoring systems, whereas the area of fibrosis was calculated through DIA.

Masseroli et al. [18] and Caballero et al. [19] introduced the “FibroQuant,” an image analysis application for automated quantitation of CPA in histologic images. Specifically, in Masseroli et al., FibroQuant was validated for the automatic segmentation of areas of fibrosis against images manually annotated from expert histopathologists. Caballero et al. aimed at the comparison of semi-quantitative methods with a DIA system for the evaluation of liver fibrosis in biopsies from patients with chronic hepatitis and different responses to IFN treatment.

Colloredo et al. [20] evaluated the impact of the size of biopsy needle samples on the grading and staging of fibrosis. According to them, since liver biopsy size can strongly affect the grading and staging of chronic viral hepatitis, only samples of 2 cm long (1.4 mm) were considered ideal to ensure the minimum number (4–6) of the complete portal tracts for reliable fibrosis quantification. In Tanano et al. [21], the  $\times 40$  magnification proved to be the ideal size for calculating the fibrotic areas. This magnification size was also selected by Wright et al. [22]. They reported that the lower and higher than  $\times 40$  magnifications led to an inferior correlation between the DIA and semi-quantitative Ishak scoring system.

O'Brien et al. [23] aimed for assessing the validity of the calculated fibrosis ratio, computed by a DIA technique, as a measure of fibrosis stage in liver biopsies of patients with chronic hepatitis C (CHC). Ryder et al. [24] aimed to identify the factors leading to histological progression of fibrosis, in a cohort of patients with hepatitis C virus (HCV) without the intervention of therapy. Lazzarini et al. [25] claimed that the O'Brien et al. study [23] could not reliably differentiate between mild fibrosis (Ishak 1–3) and advanced fibrosis or cirrhosis (Ishak 3–6). To address this, the goal was to correlate all Ishak stages with fibrosis rates (%) coming from DIA techniques, to improve the discrimination between varying levels of liver fibrosis.

According to Arima et al. [26], although end-stage liver cirrhosis reflects the histopathological distortion of the lobular architecture and formation of porto-portal bridging, the density quantification of the fibrous septa is usually neglected. With this in mind, the research team focused on the assessment of fibrosis after hepatic cirrhosis eradication, by using a DIA system for the fibrosis extent determination

after interferon treatment. Xie et al. [27] relied on a computer-aided image analysis system so that the hepatic tissue CPA in chronic hepatitis B (CHB)-related decompensated cirrhosis could be evaluated.

Moving a step forward, Manousou et al. [28,29] suggested that CPA could be a better method than the Ishak staging system or hepatic venous pressure gradient (HVPG) to assess fibrosis progression. Calvaruso et al. [30] evaluated the relationship between CPA using DIA and the Ishak grading system in patients with HCV who had undergone liver transplantation (LT). The correlation between the CPA and Ishak staging system was also evaluated in Calvaruso et al. [31] for determining the disease stage and grade of necroinflammatory activity (NIA). In the meantime, Tsochatzis et al. [32] attempted to compare the performance between histological semi-quantitative and quantitative methods specifically developed for sub-classifying cirrhosis with CPA.

Huang et al. [33] compared different parameters for CPA calculation, such as staining methods, biopsy sizes, and magnifications, to determine the optimum method of CPA measurement. In a later study [34], they studied the ability of CPA measurement and METAVIR fibrosis stage to predict clinical outcomes. In this case, the endpoints of the work included several different conditions, such as liver-related death, liver decompensation, or development of hepatocellular carcinoma (HCC).

### 3.1.2. Comparison of Biopsy Imaging with Non-Invasive Imaging Modalities

In Naveau et al. [35], the FibroTest biomarker was validated in alcoholic liver disease (ALD) patients who also had chronic viral hepatitis. They aimed for comparing the diagnostic and prognostic values of FibroTest against other well-known biomarkers, including FibrometerA and Hepascore. Raftopoulos et al. [36] compared the performance of several liver fibrosis biomarkers, developed for CHC (Hepascore, FibroTest, APRI) with the METAVIR fibrosis staging system and the image analysis CPA calculation.

Although the new quantitative DIA technology was successfully applied to the assessment of liver fibrosis in chronic viral hepatitis patients, its clinical use was usually limited to specialized centers. As a solution, Campos et al. [37] proposed the use of well-known and easily accessible hardware and software tools to perform CPA calculations based on image analysis techniques. Zhou et al. [38] considered also the validation of an inexpensive DIA technique for fibrosis quantification in patients with CHB.

The following studies took into account an extra medical imaging modality for the diagnosis of hepatic diseases, namely the liver stiffness measurement (LSM), which can be integrated in an ultrasound or magnetic resonance (MR) machine. Initially, Banerjee et al. [39] aimed at assessing the diagnostic accuracy of a multiparametric-MR method for liver tissue characterization, along with fibrosis, steatosis and iron quantification. In a later stage, they conducted a comparison of the proposed MR technique against transient elastography (TE) and the gold standard of liver biopsy. Marino et al. [40], chose the CPA to detect the expansion of fibrosis at the very early stages after liver transplantation. Based on this preference, their work aimed to evaluate whether sinusoidal fibrosis (SF) in early liver biopsies represents an early and accurate marker for identifying patients with severe HCV recurrence after LT. For Chen et al. [41], their goal was to compare the diagnostic performances of CPA and LSM in chronic hepatitis C biopsy specimens. This comparison would determine the liver fibrosis quantification by using the CPA and would in parallel focus on the necroinflammatory effects of liver stiffness (LS).

In Ding et al. [42], the main purpose of the study was to investigate the relationship between liver stiffness measured with the point shear wave elastography (PSWE) technology and the amount of collagen in CHB biopsy specimens based on digital computer image analysis. The objective of Thiele et al. [43] was to determine the diagnostic accuracy of transient elastography (TE) and real-time two-dimensional shear wave elastography (2D-SWE) techniques in ALD patients, based on liver biopsy, Ishak staging, and CPA determinations. Rastellini et al. [44] recognized the contribution of recent advances in non-invasive fibrosis detection strategies, including serum markers, TE, and other elasticity-based radiological techniques. However, they chose the gold standard of liver biopsy to



investigate the correlations between fibrosis density assessed by CPA and HVPG, in conjunction with the development of portal hypertension in chronic advanced ALD (cALD) patients.

Ali et al. [45] demonstrated the augmented mesenchymal stem cells' (MSCs) ability to repair fibrotic livers in sodium nitroprusside (SNP) pretreated mice, as a consequence of nitric oxide (NO)-induced hepatic stellate cells (HSCs) apoptosis. However, a necessary step was the quantification of fibrosis area, bilirubin, and alkaline phosphatase (ALP) before the onset of the proposed treatment method.

### 3.1.3. Methodologies Based on SHG/TPEF Imaging Microscopy

An important step towards microscopy imaging evolution was achieved beginning with Sun et al. [46], where a second-harmonic generation (SHG) microscopy has been able to measure highly ordered structures, such as type I collagen. With this in mind, fully automated quantification algorithms were developed to evaluate the progression of liver fibrosis and cell necrosis. An expansion of the previous method was attempted by Tai et al. [47] to identify collagen progression patterns and verify the feasibility of the Fibro-C-Index in human tissue diagnosis.

In Guilbert et al. [48], the group proposed a methodology for measuring SHG signals derived from isotype I and III fibrillar collagens, which are the major components involved in liver fibrosis. Sun et al. [49] proposed a new "Beijing" classification approach in both pre-treatment and post-treatment of hepatitis B and in post-treatment of hepatitis C. In this way, the team evaluated the dynamic changes in the quality of fibrosis in three classes—(a) predominantly progressive, (b) predominately regressive, and (c) indeterminate.

### 3.1.4. Intelligent Diagnostic Systems for Automated CPA Detection

We are moving into a new era, where the application of machine learning algorithms is spreading for solving complex problems. Starting with Matalka et al. [50], they presented a novel and intelligent automated quantification system (AQS) using a multilayer feed-forward backpropagation artificial neural network (MLFFBP-ANN). Stanciu et al. [51] relied on a combined approach based on second-harmonic generation/two-photon excitation fluorescence (SHG/TPEF) imaging for detecting changes in collagen deposition.

On the other hand, Xu et al. [52] presented a fully-automated assessment method called "qFibrosis," which was based on extensive image processing and quantification of portal, septal and fibrillar collagen deposition. In Sun et al. [53], the research team employed the fully automated qFibrosis system to distinguish the changes among the portal, septal and fibrillar collagen features.

As stated in Wang et al. [54], small changes in liver fibrosis may be overlooked in clinical trials, leading to invalid diagnostic outcomes. For this reason, they developed a fully automated method for providing objective hepatitis B-related fibrosis quantitative measurements. Collagen was detected using SHG microscopy in unstained liver biopsies, while hepatocyte morphology was recorded using TPEF microscopy.

A method based on DIA was designed by Meejaroen et al. [55] based on image segmentation and Bayesian classification techniques to determine the differences between fibrosis regions and healthy tissue. Thong-on and Watchareeruetai [56] focused on implementing a fibrosis detection system based on an automated feature extractor using linear genetic programming (LGP).

In Giannakeas et al. [57], an automated diagnostic method based on clustering was proposed for the extraction of the image of the tissue and the detection of fibrosis in CHC biopsy specimens. Clustering refers to a more sophisticated process of image segmentation, as opposed to thresholding techniques that were used for separating the background from the tissue pixels, as well as for the detection of fibrosis pixels in liver tissue. Tsipouras et al. [58] introduced a fully automated methodology for the extraction of CPA and the classification of different tissue regions. In Tsouros et al. [59], a modified clustering algorithm was proposed for the classification of CPA through segmentation.

The research community is currently experiencing a revolution in the field of machine learning with the advent of deep neural networks (DNNs). Deep learning, in general, is based on the idea

that algorithms can be trained without the extensive need for the user's interference. This approach has offered the medical field more automated solutions to the identification of complex structures that characterize severe pathological conditions. Many researchers have been inspired by the above, including Vicas et al. [60] for fully automating the fibrosis detection procedure with classical computer vision techniques (image processing, conventional machine learning) and convolutional neural networks (CNNs). A CNN model was also employed in Yu et al. [61] for the identification of fibrotic areas in liver biopsy specimens. At a later stage, the team made a comparison of the deep model's performance with other conventional classification algorithms.

### 3.2. Histological Samples

In this subsection, the methods and materials from the first attempts of liver fibrosis quantification are presented. In most of the works published before 2000, the number of patients included was limited and equal to  $87.9 \pm 77.9$  (mean  $\pm$  std), while the number of overall patients (viral and alcohol-related) was  $96.3 \pm 87.5$  (mean  $\pm$  std). In addition, only one study (Nakabayashi et al. [15]) included normal control subjects ( $n = 5$ ). In later studies [18–61],  $121.8 \pm 102.5$  (mean  $\pm$  std) patients with viral hepatitis were recruited, with the total number of patients being  $122.1 \pm 100.2$  (mean  $\pm$  std). Moreover, seven studies depended on patients with liver transplantation [21,27–31,40], while in more recent studies [45–47,51,52,61], non-human tissues were included.

#### 3.2.1. Number of Histological Samples in Early Methodologies

Jimenez et al. [11] were one of the first to take into account the collagen content in 38 needle liver biopsies including: (a) eight steatosis without fibrosis, (b) eight with chronic hepatitis, (c) seven fibrotic, and (d) 15 alcoholic cirrhotic. Manabe et al. [12] examined 59 needle-biopsy sections with the hepatitis C virus antibody, taken before and after IFN treatment for histological evaluation and collagen quantitation purposes. Nakayashi et al. [13] included 24 samples, some of which were recovered from autopsies and the remainder from various hepatic transfusions.

An exception is attributed to Chevallier et al. [14], in which a larger number of patients (200) were recruited. Exceptions between the early years include also the work of Pilette et al. [17], in which they relied on 243 patients presenting (a) alcoholic liver disease ( $n = 160$ ), (b) chronic hepatitis B or C infection ( $n = 83$ ), and (c) cirrhosis ( $n = 116$ ).

Nakabayashi et al. [15] recruited a total of 65 patients with histologically proven liver disease. Of the 65 patients in total, 11 were already diagnosed with chronic persistent hepatitis (CPH), 19 with chronic aggressive hepatitis 2a (CAH2a), 26 with chronic aggressive hepatitis 2b (CAH2b), and nine of the overall with hepatic cirrhosis occurrence. Kage et al. [16] selected 25 samples from patients suffering from chronic hepatitis C and 20 from patients with CHB (45 in total). Each biopsy was collected within six months of the onset of each disease.

Masseroli et al. [18] decided to implement an automated method for the detection of different samples, recovered from 59 individuals—(a) 11 healthy and (b) 48 diagnosed with different grades of chronic viral hepatitis C activity. In Caballero et al. [19], 49 patients were overseen with chronic elevation of transaminases, including (a) alanine aminotransferase (ALT) in combination with aspartate aminotransferase (AST) and (b) anti-HCV antibodies and HCV-ribonucleic acid (RNA) in serum.

#### 3.2.2. Number of Histological Samples in Later Methodologies

In Colloredo et al. [20], 161 biopsy samples were obtained from patients with chronic viral hepatitis (types B, C or D). Only biopsies of length  $\geq 3$  cm were examined in this study, as they concerned: (a) 130 cases of CHC, (b) 29 cases of CHB (1 with hepatitis D co-infection), and (c) two cases of concomitant chronic hepatitis B and C infection. In Tanano et al. [21], liver biopsies were performed in 46 patients with biliary atresia (BA), some of whom later underwent either hepatic portoenterostomy (HPE), stoma closure, or LT. Wright et al. [22] randomly selected 30 samples from a collection of 300 patients who have been chronically infected with the HCV.

O'Brien et al. [23] included a total of 230 CHC biopsy samples. Ryder et al. [24] included 214 HCV infected patients. In Lazzarini et al. [25], fibrosis rates were calculated in 164 biopsies from untreated patients with CHC, representing all Ishak stages. Arima et al. [26] included 25 patients with hepatitis C infection and a part of the population which presented a complete response to IFN therapy. An exception regarding the number of collected biopsy samples is found in the Xie et al. [27], where they took into consideration a small number of 53 samples from liver transplant patients with CHB-induced decompensated cirrhosis.

Manousou [28] evaluated a total of 135 patients with advanced liver disease, with an overall median follow-up of 76 months. Of the total number of patients, a second separate analysis was performed in 89 patients to predict clinical liver decompensation due to the HVPG measurement. Later, Manousou et al. [29] studied a group of 155 patients with recurrent HCV hepatitis after liver transplantation. In Calvaruso et al. [30], 115 samples with HCV prevalence were evaluated, and all patients underwent transplantation for cirrhosis. Calvaruso et al. [31] evaluated a series of 121 patients with recurrent hepatitis C following LT. All patients had previously undergone LT due to HCV cirrhosis and were evaluated for at least six months after surgery. Tsochatzis et al. [32] included 69 patients diagnosed with cirrhosis after liver biopsy. Alcohol was the major surgical biopsy etiology (38%), followed by hepatitis C (27.5%).

Huang et al. [33] took into account 201 patients with chronic hepatitis C, who underwent liver biopsy and serum fibrosis markers. The same team in a later study [34] selected 533 patients for the final analysis, with a median follow-up time of 10.5 years. Naveau et al. [35] included a total of 218 patients with heavy alcohol consumption, in which liver biopsy and FibroTest were included for fibrosis staging. Raftopoulos et al. [36] conducted fibrosis assessments on 179 CHB patients. In Campos et al. [37], liver biopsies were obtained from 282 patients with CHC infection. Zhou et al. [38] included a total of 142 CHB patients who underwent surgical biopsy.

In Banerjee et al. [39], 79 patients were able to participate in the magnetic resonance analysis. Seven healthy volunteers with a body mass index (BMI) < 25 kg/m<sup>2</sup> were subjected to a check-up test. In Marino et al. [40], a total of 101 HCV patients with early biopsy (< 6 months) and HVPG measurements one year after LT were included. Chen et al. [41] included a total of 137 patients diagnosed with CHC, who participated in a cohort for the analysis of antiviral treatment responses.

In Ding et al. [42], 78 patients with surgically removed hepatic neoplasms underwent an ultrasound scan and elastography, as well as hematological, biochemical, and virological examinations within 3 days before surgery. Thiele et al. [43] conducted their study based on 199 with a significant risk of chronic ALD, where histology was performed for clinical indication. The study population in Rastellini et al. [44] consisted of a total of 70 participants, nine of whom were healthy and 61 suffered from cALD, divided respectively into 20 patients with long-term abstinence from alcohol (> 6 months) and 41 with regular consumption.

The Sun et al. [46] and Tai et al. [47] methodologies were tested in both animal (Wistar rats) and human biopsy samples. In Guilbert et al. [48], only 12 large biopsy samples were collected, due to the lack of biological material with a thickness greater than 50 µm. The patients included in this study presented chronic liver disease (CLD) with various degrees of severity, mainly due to excessive alcohol consumption and hepatitis B or C viruses. Sun et al. [49] applied the Beijing classification to evaluate 71 paired biopsies with chronic hepatitis B, before and after Entecavir-based therapy. All 71 patients were primarily diagnosed with significant fibrosis or cirrhosis (Ishak > 3).

A set of 260 samples (50 normal and 210 fibrotic) was used by Matalka et al. [50] to build an automated system that would determine whether a sample has notable fibrosis content or not. 40 male Wistar rats were used in Stanciu et al. [51]. Thirty-five rats were treated with thioacetamide (TAA) over a period of 14 weeks to develop liver fibrosis.

Xu et al. [52] included 25 non-human tissues (rats with TAA-induced liver fibrosis) and 162 human CHB biopsy specimens. Sun et al. [53] included a total of 162 enrolled CHB patients with paired liver biopsies pre and post-antiviral (Entecavir) treatment. In Wang et al. [54], 175 total biopsy samples from

patients with chronic HBV-infection were taken into consideration. To clarify the ability of SHG/TPEF microscopy imaging in fibrosis, the scanned biopsy images were divided into training ( $n = 105$ ) and testing ( $n = 70$ ) cohorts to develop and evaluate the fibrosis quantification system.

In recent years, the employment of machine learning techniques has flourished as a necessary computer vision tool, but the published methodologies [55–59] were evaluated using a relatively limited dataset (8–79 biopsy specimens). This is due to the fact that the semi-quantitative assessment and the annotation of various histological structures is a time-consuming but also a necessary process for the employment of each method. Notably, Meejaroen et al. [55] conducted an experiment based on a dataset consisting of a total of 34 images. A subset of 10 images was used to train a Bayesian classifier, whilst 24 additional images were used to test the accuracy of the proposed method. Thong-on and Watchareeruetai's [56] biopsy image analysis prioritized the segmentation of fibrous tissue to determine the severity of liver damage. To achieve this goal, all experiments have been conducted with a dataset including 10 training and 22 test images.

On the contrary, CNN deep architectures require a large amount of data samples compared to conventional machine learning algorithms. For this particular reason, about 10 images per patient (107 samples) were exported in Vicas et al. [60]. In Yu et al. [61], the methodology's discrimination capability was tested in 25 Wistar rat biopsy samples, which were randomly distributed in two groups—(a) 21 subjects with TAA treatment and (b) a group of four as controls. In each case, a total of 100 slides coming from 5  $\mu\text{m}$  needle liver samples were analyzed.

### 3.3. Tissue Staining

#### 3.3.1. Histological Stains in Early Year Experiments

The techniques used for staining varied greatly among researchers before 2004, including (a) Sirius red (SR), (b) Fast green (FCF), (c) Gordon-Sweet's silver impregnation, (d) picro-Sirius red, and (e) Azan. Particularly in Jimenez et al. [11], SR and Fast green were defined as the ideal histological dyes, as these were able to bind to collagen and non-collagenous proteins, respectively. In Manabe et al. [12], thicker sections of 4  $\mu\text{m}$  were stained with Hematoxylin and Eosin (H&E) as well as with Masson's trichrome (MT) for standard histological evaluation, while for the fibrosis visualization Gordon-Sweet's silver impregnation and SR were among the favorite dyes. Nakayashi et al. [13] included a double stain of SR and Fast green in 24 samples, while others included Azan or Van Gieson stains.

Subsequently, in Chevallier et al. [14], thicker than usual 3  $\mu\text{m}$  sections were stained in a 0.1% picro-Sirius red solution. Nakabayashi et al. [15] decided to rely only on liver sections colored with SR stain, to produce a more accurate determination of different collagen content types. All reported histological samples in Kage et al. [16], were stained with H&E and Azan for the morphometric fibrosis evaluation. Pilette et al. [17] stained 3- $\mu\text{m}$ -thick sections with Hematoxylin-Eosin-Saffran, Masson's trichrome, and 0.1% picro-Sirius red solutions.

In Masseroli et al. [18] study, some histological samples were sectioned with the gold standard of 4  $\mu\text{m}$  thickness. Afterwards, they were stained with H&E, reticulin, Gomori's trichrome, Perl's, and Periodic Acid Schiff (PAS)-diastase dyes. For the remaining samples, coloring with only a 1% picro-Sirius red stain was applied. Caballero et al. [19] stained the liver biopsy specimens with H&E, reticulin, Gomori's trichrome, Perl's and PAS-diastase dyes. In Colloredo et al. [20], 5- $\mu\text{m}$ -thick sections were cut from each paraffin block and colored with H&E, Periodic Acid Schiff (PAS), along with Masson's trichrome and Perl's stains.

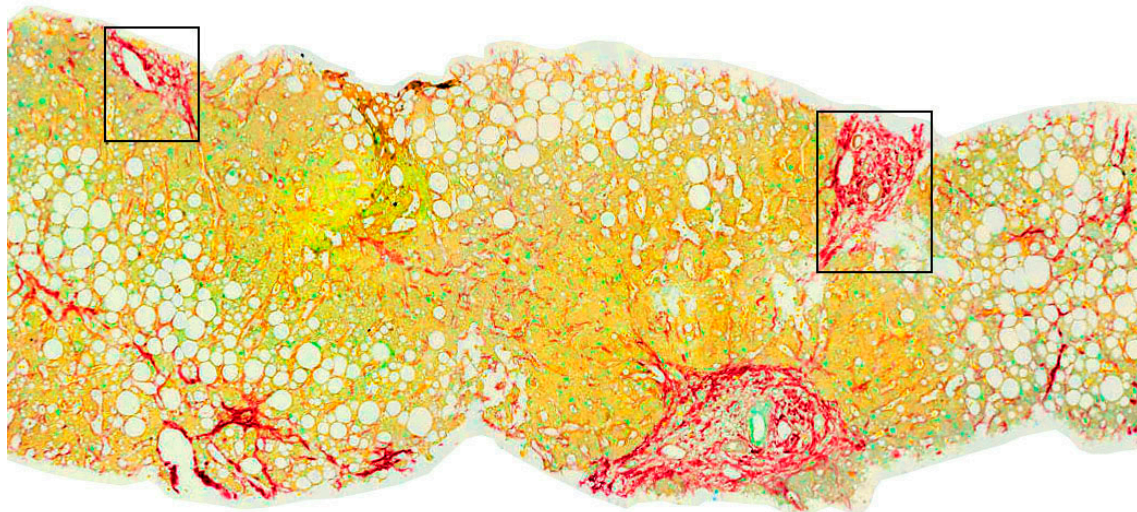
Unlike other studies, Tanano et al. [21] after fixating the samples, 2- $\mu\text{m}$ -thick sections were stained with Azan Mallory's. In Wright et al. [22], a random number of 30 biopsy sections were selected and stained with picro-Sirius red dye. The various experimental options were continued with O'Brien et al. [23], where sections of 5  $\mu\text{m}$  and H&E and Mallory trichrome stains were reviewed. Ryder et al. [24] followed in measuring the progression of liver fibrosis on Perl's stained tissue sections. In Lazzarini et al. [25], fibrosis was assessed using H&E and trichrome-stained specimens.



Arima et al. [26] prepared 2–4  $\mu\text{m}$ -thick sections for H&E and Azan-Mallory staining and silver impregnation. In Xie et al. [27], tissue samples with a thickness of 3  $\mu\text{m}$  were colored with H&E or Masson's trichrome stains.

### 3.3.2. Histological Stains Used in Recent Years

Over the last 15 years, the staining used to detect CPA has increasingly been based on picro-Sirius red (Figure 5) and MT. Manousou et al. [28,29] developed a DIA method based on picro-Sirius red staining for the quantification and determination of CPA. Calvaruso et al. [30,31] emphasized the difference between SR and picro-Sirius red histological stains. Sirius red may remain a coloring technique ideal for morphological fibrosis measurement, but the research team was more focused on picro-Sirius red staining for the tissue collagen regions to be evaluated by DIA. Liver biopsy specimens in Tsochatzis et al. [32] were stained with H&E, Periodic acid-Schiff stain with diastase digestion (DPAS), Orcein, Victoria Blue, and Perl's dyes. In other tissue sections, picro-Sirius red staining was preferred for collagen quantification and determination of CPA by DIA.



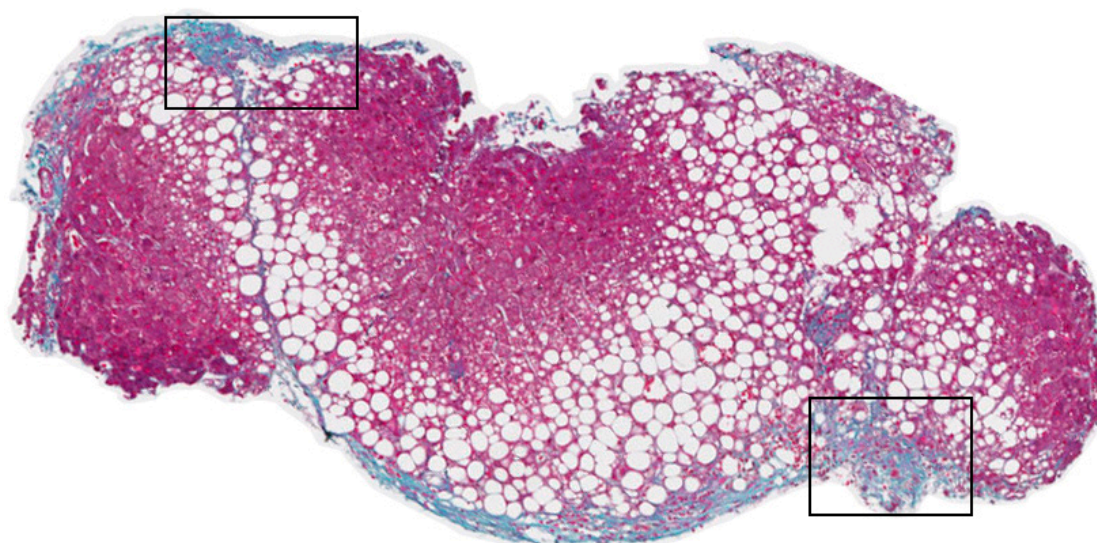
**Figure 5.** Paraffin-embedded sample and stained with Sirius red dye. The bounding boxes indicate areas of high collagen accumulation, where parenchymal nodules are surrounded by fibrous tissue, meeting the diagnostic criteria of cirrhosis.

Even though Sirius red is known to be one of the best techniques for the quantitative determination of collagen in liver fibrosis, more and more studies rely on Masson's trichrome (Figure 6), as it is currently regarded as the gold standard for biopsy tissue coloring. Published research works [34,36,37,55] have conducted experiments on MT colored samples because this staining method is available in most pathology laboratories and an optimal contrast can be obtained between the blue stained fibrous tissue and the red parenchyma.

Huang et al. [33] decided to turn to tissue coloring with both Sirius red and Masson's trichrome. The final results were in the favor of SR stain. The same applies to Zhou et al. [38], where both SR and MT stains were used in 16G biopsy needle samples. On the contrary, Naveau et al. [35] preferred to rely on more than one histological coloring option, as biopsy specimens were stained with Hematoxylin-Eosin-Saffran, MT, and picro-Sirius red dyes.

For Marino et al. [40], two-micron sections were stained with H&E and Masson's trichrome for evaluation by an expert pathologist, while samples for the assessment of sinusoidal fibrosis (SF) and collagen types I and III were stained in SR. Chen et al. [41] applied the MT, H&E, and reticulin stains in specimens for interpretation by an experienced pathologist blinded to the LSM results. On the other hand, selected liver tissue sections were stained using picro-Sirius red for measuring the CPA.





**Figure 6.** Masson's trichrome stained biopsy specimen, carrying a magnification factor of  $\times 40$  and showing thin and broad (bounding boxes) fibrous septa in end-stage cirrhosis.

Thiele et al. [43] chose to stain sections of  $3\ \mu\text{m}$  thickness with SR coloring for the detection of collagen fibers. The same applies to Rastellini et al. [44], with picro-Sirius red finally being the ideal stain in tissue samples recovered from cirrhotic subjects to determine the density of fibrosis using CPA. In Ali et al. [45],  $5\text{-}\mu\text{m}$ -thick liver samples from mice biopsies were cut from different liver lobes and stained only with SR.

In Sun et al. [46], SHG/TPEF microscopy was compared either to SR or MT coloring, whereas Tai et al. [47] were only interested in the comparison with MT staining. In Guilbert et al. [48], the final sections received an SR stain for a METAVIR fibrosis assessment. As for Sun et al. [49], after SGH/TPEF analysis of the  $5\ \mu\text{m}$  samples, these were stained with H&E, reticulin and MT for evaluation by an experienced hepatologist blinded in each patient treatment assignment.

Matalka et al. [50] used digitized images coming from  $5\text{-}\mu\text{m}$  specimens and colored with Van Geison stain. After performing TPEF imaging in Stanciu et al. [51],  $5\text{-}\mu\text{m}$ -thick specimens were sectioned from the hepatic lobe and stained with MT for fibrosis scoring by an experienced pathologist.

In Xu et al. [52], biopsy samples  $> 15\text{-mm}$  and  $5\text{-}\mu\text{m}$ -thick were fixed in formalin and embedded in paraffin for SHG-imaging. They were subsequently stained with MT for histological evaluation by an expert hepatologist. In Sun et al. [53],  $5\text{-}\mu\text{m}$ -thick histological sections were stained with H&E, MT, and Reticulin. All biopsy samples were evaluated independently by pathologists, as they were blinded to treatment assignment, biopsy sequence and other clinical details. Wang et al. [54] gave a detailed analysis of the MT staining process after the automated SHG/TPEF microscopic analysis.

Giannakeas et al. [57] and Tsipouras et al. [58] insisted on the selection of the picro-Sirius red dye for CPA extraction. Yu et al. [61], in addition to the use of SHG microscopy to non-stained TAA-induced fibrotic images, followed the direction of picro-Sirius red liver samples analysis by an expert pathologist.

### 3.4. Fibrosis Staging

Staging of liver fibrosis and thus the following-up of a chronic disease is utilized according to semi-quantitative scoring systems, known as histology activity indexes (HAIs). The current review is focused on four rating systems—(a) Knodell, (b) Scheuer, (c) METAVIR, and (d) Ishak, as described in Table 1. The earliest methods employed the Knodell, the Scheuer, and the METAVIR scoring systems or modifications of these.

#### 3.4.1. Semi-Quantitative Evaluations with the Knodell and Scheuer HAIs

Starting with Manabe et al. [12], all values produced were compared with the results obtained by the morphometric analysis of liver collagen and the Knodell scoring histological index. The same applies to Chevallier et al. [14], who were claiming that the Knodell HAI is reliable for estimating the extent of fibrosis and assessing the activity outcome of viral chronic hepatitis. Nakabayashi et al. [15] investigated the correlation between the percentage of fibrosis and a modified Knodell scoring system. In Kage et al. [16], the degree of fibrosis in each biopsy specimen was evaluated independently by three hepatologists, based on the Scheuer scoring system.

The main objective of Pilette et al. [17] was to determine the correlation between semi-quantitative estimates of the total fibrous area with an image detection method. The semi-quantitative evaluation of fibrosis was assessed by the Knodell and the METAVIR scoring systems. This work was then followed by Masseroli et al. [18], who analyzed the correlation between quantitative image analysis and semi-quantitative fibrosis estimates derived from both Knodell and Scheuer HAIs. In Caballero et al. [19], the grading of necroinflammatory activity in each biopsy specimen was assessed with a modified Knodell HAI (HAI-K) and a modified Scheuer HAI (HAI-S).

The histological activity (grade) and the amount of fibrosis (stage) were evaluated in the [20,22] studies, by employing the Ishak scoring system. In Colloredo et al. [20], the Ishak score was used to evaluate the fibrosis measurements in biopsy specimens with a reduced length of 1 cm and a width of 1 mm. In O'Brien et al. [23], two pathologists evaluated the extent of fibrosis for each biopsy specimen using the Ishak scoring system. Ryder et al. [24] evaluated the fibrosis and necroinflammatory changes according to the Knodell and Ishak staging systems. Lazzarini et al. [25] compared the diagnostic capability of their DIA technique with the Ishak scoring system.

#### 3.4.2. Newer Methodologies Relying on the Ishak HAI

The Ishak HAI began to gain ground concerning its choice in the following studies. Manousou et al. [28,29] aimed at estimating the quantitative measurement of CPA and finally comparing it with the Ishak stage. The goal of Calvaruso et al. [30,31] was to evaluate the relationship between the amount of hepatic collagen measured by the DIA, the Ishak grade of necroinflammatory activity and the HVPG.

In Huang et al. [33,34], the method's reliability in CPA measurement was examined based on its correlation with the METAVIR stage classification system. Moving to the next stage, a correlation between CPA models and serum fibrosis markers was also evaluated. The correlation between CPA and each METAVIR stage was determined by Spearman coefficients, whereas the correlation between CPA and serum fibrosis markers by Pearson coefficients, respectively.

#### 3.4.3. Recent Research Works Employing the METAVIR HAI

The METAVIR HAI has been extensively used in recent works, including Naveau et al. [35], where a scoring system influenced by the METAVIR scoring system was employed. In Raftopoulos et al. [36], all biopsies were interpreted by expert histopathologists who used the METAVIR scoring system. More specifically, a modification was applied in which the presence of METAVIR stage F2, F3, or F4 was named "significant fibrosis," while the term "advanced fibrosis" was maintained for the METAVIR F3 or F4 stage.

In Campos et al. [37], the research team decided to compare the quantitative assessment of hepatic fibrosis coming from the newly developed DIA, with semi-quantitative results derived both from the Ishak and METAVIR scores. In Zhou et al. [38], the evaluation of the correlations between the CPA measurement and the semi-quantitative assessment was carried out according to Laennec, a modified METAVIR scoring system, along with the traditional Ishak and METAVIR grading.

In Banerjee et al. [39], all fibrotic samples were scored with a modified Ishak-based system. For this study, mild fibrosis was defined as Ishak F1–F2, moderate fibrosis as Ishak F3–F4, and severe fibrosis as

Ishak  $\geq$  F5. In Marino et al. [40], expert pathologists blinded to clinical data employed the METAVIR classification system to evaluate the fibrosis progression after LT and for the measurement of specific architectural changes in the liver.

Biopsy specimens in Chen et al. [41] were interpreted by an experienced physician not aware of the LSM results, with the help of the METAVIR scoring system, while NIA was rated as A0 = none, A1 = mild, A2 = moderate, and A3 = severe. In Ding et al. [42], a tissue section fibrosis was staged on a scale from 0 to 4 according to the Scheuer scoring system. In the next step, the semi-quantitative method of Chevallier et al. [14] was chosen to reflect morphometric measurements of fibrosis, including centrilobular veins, perisinusoidal space, portal tract, and the number and width of septa.

Tai et al. [47] examined the Fibro-C-Index diagnostic results by comparing them with those obtained from the Ishak semi-quantitative scoring. Guilbert et al. [48] aimed at good fibrosis level correlations of their proposed method with the METAVIR HAI. Sun et al. [49] estimated the correlation of their fibrosis identification system with the Ishak staging and the METAVIR HAI. On the other hand, the Laennec classification was set to sub-classify cirrhosis based on the type of fibrous tissue regions in each sample.

In Matalka et al. [50], two pathologists performed an Ishak interpretation and evaluation of the biopsy samples. To verify the diagnostic ability of fibrosis in Stanciu et al. [51], an expert clinician used the METAVIR system to assess histologic lesions, using two separate scores, one for necroinflammatory grade and another for the stage of fibrosis.

Xu et al. [52] analyzed the qFibrosis CPA results by employing the METAVIR and Ishak fibrosis scoring systems. The METAVIR staging system was mainly employed to illustrate the histological features of the various collagen patterns caused by CHB. All samples in Sun et al. [53] were assessed by a modified Ishak HAI and by the SHG/TPEF-based qFibrosis software. In particular, the Ishak HAI was employed to assess NIA and fibrosis, while the performance index rating (PIR) score was used to assess the dynamic changes of liver fibrosis pre- and post-treatment. The diagnostic results of Wang et al. [54] were validated by pathologists according to the Ishak scoring system, as well as by conventional collagen staining methods followed by microscopy imaging.

All images in Giannakeas et al. [57] were characterized according to the Ishak HAI, while each image was annotated at a pixel level by a specialist physician. For quantifying the performance of the proposed method, the extracted CPA values were compared to the annotated ones. In Yu et al. [61], each tissue sample was initially submitted to a specialist physician analysis, which determined the fibrosis stage based on the METAVIR scoring system.

### 3.5. Image Analysis Method

#### 3.5.1. Early Methodologies Relying on Morphometric Analysis

Image analysis is a critical step for the quality of collagen measurement and consequently important for fibrosis stage determination. The earliest studies used several morphometric software implementations to count the pixels of the images representing fibrosis, as well as the pixels composed of the whole biopsy specimens. A detailed description of the quantification procedure was only provided by a small portion of the referred works.

More specifically, Manabe et al. [12] and Chevallier et al. [14] employed the histomorphometric analysis software Biocom 200 (Biocom, Les Ulis, France). In Nakayashi et al. [13] and Nakabayashi et al. [15], the histological findings including the total tissue area and the Sirius red-stained fibrosis area were evaluated with the use of the SPICCA II (Olympus, Tokyo, Japan) image analyzer. The degree of fibrosis was expressed as the ratio of the SR-stained area to the total biopsy sample area. On the other hand, Kage et al. [16] used another image analysis system (Nexus Qube, Nexus Inc., Osaka, Japan).

### 3.5.2. DIA Systems Using Manual Thresholding

Commonly, image analysis involves a segmentation procedure (i.e., histogram thresholding) for the detection of collagen regions and the elimination of artifacts and other undesired areas. In most methods for calculating liver fibrosis, the authors processed the images manually.

Pilette et al. [17] analyzed a set of grayscale images obtained from a Leica DMR microscope (Leica, Wetzlar, Germany) to quantify the percentage of fibrosis. The images were digitized at  $\times 100$  magnification. The authors mentioned that a user-defined threshold was applied to convert the images into binary. Then, a semi-automated procedure was used to eliminate the unwanted artifacts in them. Finally, the percentage of the fibrosis (fibrosis index—FI) was computed using the identified areas of fibrosis, divided by the whole area of the hepatic parenchyma, as follows:

$$\text{FI}(\%) = \frac{\text{Total stained fibrosis area}}{\text{Total area of sample}} \times 100. \quad (1)$$

In terms of the image analysis field, the first extended work that presented the FibroQuant software was that by Masseroli et al. [18]. The method consisted of four steps, and it was applied to a biopsy image with a magnification of  $\times 200$ . During the first step, a histogram equalization procedure enhanced the initial image. In the next three steps, different adaptive thresholding techniques have been used to detect specific areas, including percentage areas of perisinusoidal fibrosis (PSF), portal-periportal and septal fibrosis (PPF) and portal vessel and biliary duct lumina (PLA). In all phases, user inference was allowed to improve the inclusion of these areas when the thresholds failed. FibroQuant was also utilized in Caballero et al. [19] in order pre and post-treatment biopsies to be compared and for fibrosis areas isolation, according to their staining and their exterior limits.

Each biopsy image in Tanano et al. [21] was edited with a morphometric program and with an adequate threshold to produce a monochrome image. The staining area was calculated with a computerized digitizer, in which the fibrosis index (FI) was determined according to Equation (1) and calculated in 18–20 randomly selected fields, dependent on the size of the sample. To evaluate the chronological change of FI from the hepatic portoenterostomy to the stoma closure ( $n = 26$ ), the difference rate of FI (FIDR) was defined as follows:

$$\text{FIDR}(\%/mo) = \frac{\text{FI at stoma closure} - \text{FI at HPE}}{\text{Interval from HPE to stoma closure (months)}}. \quad (2)$$

Wright et al. [22] used the Scion (Scion Corporation, Frederick, MD, USA) image software. Microscopy was performed using high levels of illumination, in addition to a green optical filter to control the picro-Sirius red influence without distorting the color values of fibrotic areas. All histological images were initially converted to grayscale and the total tissue area was calculated with a threshold value. The pixel count was then measured and divided by the total tissue area to calculate the proportion of fibrous tissue staining on each slide.

In O'Brien et al. [23], the fibrosis ratio for each liver biopsy was calculated using the Optimas 5.0 (Media Cybernetics, Silver Spring, MD, USA) image analysis software. Images from trichrome-stained sections were captured with a magnification of  $\times 40$  and analyzed as RGB (Red, Green, Blue) 24-bit images. The stained blue section represented the fibrosis, while the red, the parenchyma respectively. After interactive thresholding, the image was converted into binary, where the two-dimensional patterns were measured by direct pixel counting.

In Lazzarini et al. [25], the Image-Pro Plus 4.5 (Media Cybernetics Co., Rockville, MD, USA) software was used to analyze the digitized biopsy specimens as RGB 24-bit color images. Fibrosis (%) measurements for six liver biopsies were compared using both the  $\times 40$  and  $\times 100$  magnifications. The higher magnification did not result in any noticeable difference in fibrosis measurement, therefore the magnification of  $\times 40$  was chosen. More technically, interactive thresholding was applied to each



image sample to highlight the area of fibrosis, followed by binary conversion, so that the CPA could be measured by direct pixel counting.

All digitized images in Arima et al. [26] were edited with the Adobe Photoshop 5.02i (Adobe Systems Inc., San Jose, CA, USA) environment. In particular, fibrosis appeared as blue-colored tissue, and then, National Institute of Health (NIH, Bethesda, MD, USA) image software was used for the quantification of the fibrotic areas. In Xie et al. [27], the image analysis system was based on the Image-Pro plus 6.0 software. According to the research team, the magnification of  $\times 40$  was the ideal choice, where the image size consisted of  $1280 \times 960$  pixels. The light intensity was adjusted according to the magnification level and three non-overlapping visual fields were randomly selected. Subsequently, the areas of red collagen fibers in each visual field were measured according to Equation (1), in which the total area of the biopsy sample indicated the sum of the three visual fields.

Calvaruso et al. [30,31] measured the CPA in liver biopsy images using the Zeiss KS300 (Zeiss, Hertfordshire, UK) image analysis software. Both the threshold estimation and the artifact elimination have been performed manually for each image according to the utilities of the software. The same stands for Tsochatzis et al. [32], where the Zeiss KS300 was the system of choice. Here, the digital image analysis process included a manual RGB threshold for the elimination of unnecessary artifacts, as well as the quantitative measurement of nodule size, septal width, and fibrous tissue expressed in CPA.

Several works including both Huang et al. studies [33,34] operated the Aperio ImageScope (Aperio, Vista, CA, USA) software for manually processing the liver images. The optimal threshold for positive pixel values corresponding to the areas of Sirius red or Masson's trichrome staining was determined by the ImageScope software and by parameterizing the hue and color saturation values. As a result, a binary image was exported, and the CPA was expressed as the percentage of positive pixels to the total number of pixels. Raftopoulos et al. [36] also benefited from using ImageScope (version 10.0), for including large portal tracts and vessels within each scanned slide. Batch analysis was then performed on every image using the Aperio PixelCount (version 9) algorithm, which calculated the prevalence ratio of positive disease pixels (Equation (1)).

Another manual approach for the quantification of the fibrosis percentage was proposed by Campos et al. [37]. This work used techniques from Adobe Photoshop to extract the pixels of the stained tissue, as well as the fibrosis areas. The regions of interest were extracted by thresholding different channels of the image. In Zhou et al. [38], each sample was digitized using a magnification of  $\times 40$ . All images were captured, providing a maximum resolution of  $3748 \times 2736$  pixels as 24-bit RGB Joint Photographic Experts Group (JPEG) images. By applying several commands in Adobe Photoshop, the background and other undesired elements were excluded. Then, the red/magenta area of the biopsy specimen, corresponding to the fibrotic tissue, was selected and the fibrosis index (FI) calculation, expressed as a percentage of red pixels throughout the entire tissue region, was finally applied.

Chen et al. [41] edited all histological images using the Adobe Photoshop CS6 platform to exclude collagenous structures that were mainly structural, such as portal tracts, liver capsules, and others. The threshold was set by consensus between hepatologists and pathologists. The stained collagen and the tissue regions were calculated in pixels by using the Image-Pro 7.0 software (Media Cybernetics Inc., Rockville, MD, USA). As expected, the ratio between the two regions was expressed as a CPA percentage.

The quantitative estimation of collagen in Ding et al. [42] was evaluated as the collagen proportionate area (CPA) using computer-assisted DIA. Image-Pro plus 6.0 software was used by the current research team as the main image analysis system. The magnification was set at  $\times 10$ , the light intensity was adjusted to the same level of magnification, and three non-overlapping visual fields were randomly chosen. Thereafter, areas of red collagen fibers (type I collagen) and green collagen fibers (type III collagen) were measured with

$$CPA_n = \frac{\text{type I collagen} + \text{type III collagen}}{\text{Total tissue area}}, \quad (3)$$



being obtained from five (CPA<sub>1-5</sub>) hot spots in the same histological sample according to the formula:

$$CPA = \frac{CPA_1 + CPA_2 + \dots + CPA_5}{5}. \quad (4)$$

In Thiele et al. [43], the tissue sections for CPA measurement were digitized using a Nanozoomer 2.0 HT (Hamamatsu Photonics, Hamamatsu, Japan) microscope and analyzed using the VISIOmorph 4.3.6 (Visiopharm, Horsholm, Denmark) image analysis tool.

The histological images in Rastellini et al. [44] were acquired with a digital Leica DC 300 F camera, carrying a  $\times 10$  magnification. Collagen proportionate area was subsequently quantified using the Qwin Leica Q550IW (Meyer Instr., Houston, TX, USA) software. The quantitative method allowed the research team to determine the percentage of picro-Sirius red positive area in the biopsy specimen, expressed as (a) 0–5%, (b) 5–10%, (c) 10–20%, and (d) >20%, representing the fibrosis density. Whereas, significant fibrosis was defined as CPA > 5%. In Ali et al. [45], specimen images from three sections per mouse were captured, while fibrosis and total tissue area were measured using Image J (NIH, Bethesda, MD, USA) software.

### 3.5.3. Application of Advanced Image Segmentation Techniques

Sun et al. [46] initially introduced some optimization parameters to  $4096 \times 4096$  SHG/TPEF images. The background in the SHG images was first removed and the corresponding tissue region images were low-pass filtered in the frequency domain. Then, an Otsu threshold segmentation, along with the erosion and dilation operations, were used for noise removal. In Tai et al. [47], image processing was carried out using MATLAB. The optimal threshold levels were determined by identifying the local minima from the pixel intensities of the TPEF image. Once an optimal threshold value was determined, it was multiplied by the SHG. An Otsu segmentation and morphological erosion and dilation were then applied to the SHG image to highlight the areas of collagen.

SHG images in Guilbert et al. [48] were acquired with 12-bit intensity resolution and recorded as TIFF format files using the Fluoview (Olympus, Southborough, MA, USA) software. The team aimed at utilizing a sample as a reference, to obtain a mean SHG intensity in non-fibrotic areas. Practically all SHG images were binarized using an intensity threshold  $t$  and the sample score was obtained by replacing each pixel intensity-value  $I$  by 1 if  $I > t$ , or 0 if  $I < t$ , by calculating the mean number of pixels with a value of 1. That is, the score  $\bar{S}(t)$  could be computed by

$$\bar{S}(t) = \frac{1}{N^2} \sum_{i,j=1}^N H(I_{ij} - t), \quad (5)$$

where  $I_{ij}$  denotes the intensity measured in  $[i, j]$  pixel,  $N^2$  the total number of image pixels,  $t$  the intensity threshold, and  $H(x)$  the Heaviside step function ( $H(x) = 1$  when  $x > 0$  and 0 if  $x \leq 0$ ).

In Sun et al. [49], the total CPA measurement was performed thanks to SHG/TPEF microscopy (Genesis200TM, HistoIndex Pte. Ltd., Singapore). The images were scanned at  $\times 20$  magnification on non-stained biopsy specimens carrying a  $512 \times 512$  pixel resolution. In particular, SHG microscopy was used in identifying the collagen regions, while the corresponding TPEF was used to identify other cellular structures. The percentage of TPEF signals was then applied to normalize the SHG signals so that the normalized total collagen percentage could be determined.

### 3.5.4. Application of Automated Fibrosis Detection Techniques

Each image in the work from Matalka et al. [50] was analyzed through an image analysis process through the MATLAB 6.5 environment. Key steps initially included some grayscale image enhancement techniques followed by a binary morphological opening operation to smooth the fibrosis structures. Due to the presence of many texture types within a single image, K-means clustering was then performed to obtain the segmentation of the image by isolating the fibrosis regions from the

non-fibrosis tissue contents. An MLFFBP-ANN model of one output was then required to classify fibrous samples in one of the six stages of the Ishak system.

In Stanciu et al. [51], before applying a Dense Scale Invariant Feature Transform-Bag of Features (DSIFT-BoF) classification, each image was processed with Wiener filtering to remove any type of detected noise. The DSIFT descriptors came from histogram representations that combined local gradient orientations and magnitudes, in a neighborhood of pixels around a key point indicated by a bin size. The typical BoF strategy required the detected features to be converted into a vector space, taking the form of a normalized histogram. Subsequently, K-means clustering was performed to determine the centroids when grouping these features. Eventually, for every testing image, each feature vector was classified by a weighted k-Nearest Neighbor (k-NN) algorithm.

In Xu et al. [52], a sophisticated method was presented for histology image analysis. According to this work, from 69 good quality biopsy samples (>15 mm), three different types of tissue areas were determined (portal, fibrous, and septa), while a set of feature values was extracted from them. Thereafter, they were used to train a multinomial logistic regression (MLR) model and applying it to the remaining 38 suboptimal (<15 mm) specimens in order to detect changes of collagen patterns in cirrhosis progression or regression. Moreover, a stepwise MLR analysis was performed to find the best combination of non-invasive clinical biomarkers to differentiate Ishak stages 5 and 6.

In Sun et al. [53], all liver biopsy images were acquired at  $\times 20$  magnification and with a  $512 \times 512$  pixel resolution. Correlation analysis between the collagen features and fibrosis stages, along with the Bayesian information criterion (BIC) test, were used to select the most important features for parameterizing the qFibrosis system. In parallel, a nonlinear regression model was selected to combine the collagen features into a single index based on the distribution of fibrosis stages. Eventually, the model revealed 15 features for the best performance to be obtained. Feature selection and model fitting were processed by MATLAB 2015a.

The SHG/TPEF images of Wang et al. [54] were acquired using the Genesis<sup>TM</sup> software, carrying a  $\times 20$  magnification and a  $512 \times 512$  pixel resolution. The collagen strings were detected using a segmentation algorithm based on the binary conversion of scanned images. Multiple features including the total length, width, area and perimeter of collagen strings were measured to build two grading method types. In the first method, a linear combination of n-feature variables ( $X_1, X_2, \dots, X_n$ ) was applied, as shown in the following equation:

$$y = F(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}}, \quad (6)$$

where  $\beta_0, \beta_1, \dots, \beta_n$  denote the coefficients estimated with the maximum likelihood method. The second method employed the supervised SVM classification model with the radial basis function (RBF) kernel, to calculate the distance measure between two feature vectors  $x$  and  $x'$  in an m-dimensional space and with  $m \geq n$ :

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right), \quad (7)$$

where  $\sigma$  was set to control the width of the kernel.

In the Meejaroen et al. [55] image enhancement process, the red channel in the RGB color space was chosen as the determining factor because the red channel histograms of control tissue and fibrosis differ significantly. Thereafter, various low-pass filtering techniques were applied to reduce the amplitude of image variations, with average pixel filtering producing the optimal results. In parallel, ground truth images, in which the fibrosis areas were manually annotated, were used as inputs to train the Bayesian classifier. The classification accuracy was finally calculated by comparing the fibrosis rate obtained from the automated software with the corresponding ground truth images.

The Evolutionary computation (EC) at Thong-on and Watchareeruetai [56] utilized the Genetic Algorithm II (NSGA-II) algorithm to ensure the survival of an elite group of offspring population. Initially, the crossover operation was performed after selecting a random pair of programs as parents.

Following the crossover operation, the mutation step was performed, where one individual's genes were shown to be likely to be inserted, deleted or modified, which increased the diversity of the created programs.

In Giannakeas et al. [57], the application of the K-means algorithm was proven to be more beneficial than thresholding, due to the use of all three channels of the color image. More specifically, the pixel intensity from each channel was used as the feature vector for clustering and overall, the pixel color information was included in the segmentation procedure. Using the above vectors as samples, the K-means algorithm was employed to provide two centroids: (a) 1 = liver tissue and (b) 2 = background. Thus, the number of clusters was set to  $K = 2$  and a square-error criterion for each of the two clusters was calculated, based on the following steps:

$$\text{Error}^2 = \sum_{k=1}^2 e_k^2, \quad (8)$$

with

$$e_k^2 = \sum_{i=1}^{n_k} (x_{ik} - M_k)^2, \quad (9)$$

where  $x_{ik}$  is the  $3 \times 1$  feature vector of the  $i$ th pixel,  $n_k$  the number of pixels which belong to the  $k$ th ( $k = 1, 2$ ) cluster and  $M_k$  the centroids of the  $k$ th cluster, respectively according to

$$M_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_{ik}. \quad (10)$$

As an expansion of the previous work, in Tsiouras et al. [58], after performing a background-tissue separation through the K-means and Fuzzy C-means (FCM) clustering algorithms, the group turned to the identification of fibrosis. This was achieved by training several classification algorithms, including (a) k-NN, (b) C4.5 decision tree (DT), (c) random forest (RF), (d) multilayer perceptron (MLP) neural network, and (e) support vector machines (SVM), to exclude regions of no interest from the final CPA computation.

Advancing with Tsouros et al. [59], by inserting some constraints to the centers' initialization in K-means algorithm and their update during the clustering procedure, each RGB pixel value was able to be segmented into one of three classes: (1) background, (2) liver tissue, and (3) collagen.

In Vicas et al. [60], some common image preprocessing operations were performed. Each image was converted to grayscale and 0.1–99.9% histogram percentiles were then computed. For most of the operations, the image was then converted to HSV (Hue, Saturation, Value) to reduce the contrast/luminosity variance in the data. The optimum hue and saturation channel thresholds were determined using the ROC curve on a validation set consisting of several manually labeled images. In the case of deep neural networks, initially, the visual geometry group (VGG) architecture proved to be unstable and not able to learn basic aspects of the disease findings, with U-net being the second fully automated approach to perform pixel-wise region segmentations.

All images in Yu et al. [61] were at first obtained from SHG microscopy (Olympus IX81, Olympus, Tokyo, Japan) carrying a  $\times 20$  magnification. Grayscale contrast enhancement and binary morphological closing were the initially performed preprocessing steps for smoothing the collagen segments and eliminating the small ones. As a next step, various feature values were extracted from collagen regions of interest, revealed by classic image processing techniques. The final stage included the identification and quantification of collagen distribution with the employment of a conventional artificial neural network (ANN) and the MLR, SVM and RF algorithms. On the contrary, the deep learning algorithm showed the ability to classify images without the need for preprocessing. During this phase, the processed SHG images were first resized to a size of  $224 \times 224 \times 3$  pixels to fit into the

input layer of the pre-trained AlexNet-CNN model. The final layer was set to output five possible prediction labels corresponding to the METAVIR fibrosis stages.

#### 4. Results and Discussion

In this section, a summary is made of how the examined from the literature methods were able to quantitate hepatic fibrosis in microscopic images, obtained from patients with chronic liver conditions. During the PRISMA selection method, when the gathered papers were sorted by methodological approach, a corresponding number of groups were formed in the following tables (Tables 2–5) to summarize all results after the end of the discussion.

**Table 2.** Comparison of works using exclusively semi-quantitative examinations for collagen proportionate area (CPA) calculation and combinatorial systems employing digital image analysis (DIA) with histology activity index (HAI) performance verification.

Author/Year	Dataset/Histological Stain	HAI	Image Analysis Method	Results/Remarks
Jimenez et al. 1985 [11]	38 samples (8 steatosis w/o fibrosis, 8 chronic hepatitis, 7 fibrotic, 15 alcoholic cirrhotic patients) Stains: Sirius red, Fast green	-	-	Histomorphometric—colorimetric CPA correlation ( $r = 0.77$ , $p < 0.001$ )
Manabe et al. 1993 [12]	59 samples (54 HCV antibody patients) Stains: Gordon-Sweet’s silver, Sirius Fast red	Knodell	Histomorphometric analysis (Biocom 200)	Histomorphometric—colorimetric CPA correlation ( $r = 0.778$ )
Nakayashi et al. 1993 [13]	24 samples (autopsy, hepatic transfusion patients) Stains: Sirius red, Fast green, Azan, Van Gieson	-	Biochemical analysis. Planimetric measurement (SPICCA II)	Hydroxyproline—fibrosis content correlation ( $r = 0.887$ )
Chevallier et al. 1994 [14]	200 samples (alcoholic and viral patients) Stain: picro-Sirius red	METAVIR (modified)	Histomorphometric analysis (Biocom 200)	SSS—SDTC correlation ( $r = 0.73$ , $p < 10^{-5}$ )
Nakabayashi et al. 1996 [15]	65 samples (liver disease patients) 5 normal as controls	Knodell (modified)	Biochemical analysis. Planimetric measurement (SPICCA II)	CI—Biochemical analysis correlation ( $r = 0.88$ , $p < 0.0001$ )
Kage et al. 1997 [16]	45 samples (25 CHC patients. 20 CHB patients as controls) Stains: H&E, Azan	Scheuer	Image analysis (Nexus Qube)	Fibrosis stage—area correlations: hepatitis B ( $r = 0.67$ , $p < 0.001$ ), hepatitis C ( $r = 0.75$ , $p < 0.001$ )
Pilette et al. 1998 [17]	243 samples (83 viral liver disease. 160 ALD patients) Stains: H&E-Saffran, MT, picro-Sirius red	METAVIR. Knodell	Manual thresholding. Semi-automated artifact elimination	DIA—semi-quantitative scores correlation ( $r = 0.84$ , $p < 10^{-4}$ )
Masseroli et al. 2000 [18]	48 samples (CHC patients), 11 normal as controls Stain: picro-Sirius red	Knodell. Scheuer	Adaptive thresholding. User interference (FibroQuant)	FibroQuant—HAI correlations ( $0.72 < r < 0.83$ , $p < 0.0001$ )
Caballero et al. 2001 [19]	98 samples (HCV-chronic hepatitis patients) Stain: picro-Sirius red	Knodell (modified). Scheuer (modified)	Automatic segmentation & CPA quantification (FibroQuant)	FibroQuant: porto-periportal—septal fibrosis correlation ( $p \leq 0.05$ )
Colloredo et al. 2003 [20]	161 samples (chronic viral hepatitis patients (B, C or D) Stains: H&E, Periodic Acid Schiff, MT, Perl’s	Ishak	-	Biopsy length Reduction—Ishak correlation ( $p < 0.001$ )
Tanano et al. 2003 [21]	46 samples (BA fibrosis index patients) Stain: Azan Mallory	-	Monochrome image production. Morphometric editing. Adequate thresholding	CV of FI range (%): 2.9–7.7 in $\times 40$ magnification
Wright et al. 2003 [22]	30 sections (from 300 CHC samples) Stain: picro-Sirius red	Ishak (modified)	Grayscale conversion. Optimal thresholding (Scion)	DIA—HAI intra, inter-observer correlations: DIA ( $r = 0.889$ , $r = 0.837$ ), HAI ( $r = 0.878$ , $r = 0.776$ )
O’Brien et al. 2000 [23]	30 samples (CHC patients) Stains: H&E, Mallory trichrome	Ishak	Interactive thresholding (Optimas 5.0)	Fibrosis ratio—Ishak correlation ( $r = 0.58$ , $p < 0.001$ )
Ryder et al. 2004 [24]	214 samples (HCV patients) Stain: Perl’s	Knodell. Ishak	-	Knodell—Ishak correlation ( $p = 0.0001$ )
Lazzarini et al. 2005 [25]	164 samples (CHC patients) Stains: H&E, trichrome	Ishak	Binary interactive thresholding. Direct pixel counting (Image-Pro Plus 4.5)	CPA—Ishak correlation ( $p < 0.001$ ). AUROC = 0.982–1.00
Arima et al. 2004 [26]	25 samples (from HVC, 20 IFN treatment patients) Stains: H&E, Azan-Mallory	-	Image editing (Adobe Photoshop 5.02i). CPA quantification (NIH)	Significant fibrosis regression (%): viral eradication ( $7.2 \pm 1.5$ ), IFN treatment ( $2.7 \pm 0.5$ )

Table 2. Cont.

Author/Year	Dataset/Histological Stain	HAI	Image Analysis Method	Results/Remarks
Xie et al. 2011 [27]	53 samples (CHB cirrhosis patients) Stains: H&E, Masson's trichrome	-	Intensity adjustment. Random selection of non-overlapping visual fields (Image-Pro plus 6.0)	CPA: MELD score—serum TBIL—prothrombin B ISR correlation ( $p < 0.001$ )
Manousou et al. 2011 [28]	155 samples (HCV patients after LT) Stain: picro-Sirius red	Ishak	Direct CPA determination, quantification	MVA in CPA decompensation ( $p = 0.010$ )
Manousou et al. 2013 [29]	155 samples (HCV patients after LT) Stain: picro-Sirius red	Ishak	Direct CPA determination, quantification	MVA in CPA decompensation ( $p = 0.001$ )
Calvaruso et al. 2009 [30]	250 samples (from HCV 115 patients) Stain: picro-Sirius red	Ishak	Manual thresholding, artifact elimination (Zeiss KS300)	Median CPA—Ishak 6—HVPg ( $r = 0.62, p < 0.001$ )
Calvaruso et al. 2012 [31]	62 samples (HCV patients after LT) Stain: picro-Sirius red	Ishak	Manual thresholding, artifact elimination (Zeiss KS300)	CPA—HVPg $\geq 10$ mmHg correlation ( $p = 0.04$ )
Tsochatzis et al. 2014 [32]	69 samples (post-LT CLD patients) Stains: H&E, DPAS, Orcein, Victoria Blue, Perl's	Laennec, Kumar & Nagula	RGB thresholding. Manual artifacts elimination. CPA measurement (Zeiss KS300)	CPA AUROC = 0.909. CPA correlation: Laennec ( $p < 0.001$ ), Kumar ( $p = 0.003$ )
Huang et al. 2013 [33]	201 samples (CHC patients) Stains: picro-Sirius red, MT	METAVIR	Threshold determination. Fibrotic areas manual extraction.	CPAs > CPAt—METAVIR correlation ( $p < 0.001$ )
Huang et al. 2014 [34]	533 samples (CHC patients) Stain: Masson's trichrome	METAVIR	Manual area exclusion. Optimus threshold determination (Aperio)	CPA—METAVIR correlation ( $p < 0.001$ )

Table 3. Studies comparing histopathological imaging with non-invasive imaging modalities.

Author/Year	Dataset/Histological Stain	HAI	Image Analysis Method	Results/Remarks
Naveau et al. 2009 [35]	218 samples (ALD patients) Stains: H&E-Saffran, MT, picro-Sirius red	METAVIR	-	FibrometerA, Hepascore vs. FibroTest fibrosis (AUROC = $0.83 \pm 0.03$ ), vs. APRI—Forns—FIB4 correlation ( $p < 0.01$ )
Raftopoulos et al. 2012 [36]	179 samples (CHB patients) Stain: Masson's trichrome	METAVIR	Manual areas exclusion (Aperio ImageScope). Batch analysis (Pixelcount 9.0)	Morphometry correlation: Hepascore ( $r = 0.603, p < 0.001$ ), FibroTest ( $r = 0.392, p = 0.03$ ), HBeAg-positive ( $r = 0.492, p = 0.001$ )
Campos et al. 2014 [37]	282 samples (CHC patients) Stain: Masson's trichrome	METAVIR. Ishak	Image enhancement. Background removal. Fibrotic area extraction (Adobe Photoshop)	DIA—Ishak, METAVIR correlation ( $r = 0.95, 0.92, p < 0.001$ )
Zhou et al. 2017 [38]	142 samples (CHB patients) Stain: Sirius red, Masson's trichrome	Laennec. METAVIR (modified). Ishak. METAVIR	Background, undesired areas exclusion (Adobe Photoshop)	DIA—Ishak, METAVIR, Laennec correlation ( $r = 0.83, 0.86, 0.84, p < 0.001$ )
Banerjee et al. 2014 [39]	79 samples, 7 normal as controls	Ishak (modified)	ROI volume assessment	MR—Histology fibrosis correlation ( $r = 0.68, p < 0.0001$ ). Fibrosis AUROC (0.94)
Marino et al. 2014 [40]	101 samples (HCV patients after LT) Stain: H&E, MT, Sirius red	METAVIR	-	MT—SR stains SF correlation ( $p < 0.01$ )
Chen et al. 2015 [41]	137 samples (CHC patients) Stain: Masson's trichrome	METAVIR	Threshold setting (Adobe Photoshop CS6). CPA calculation (Image-Pro 7.0)	CPA—SWV correlation ( $p < 0.2585$ )
Ding et al. 2015 [42]	78 samples (CHB patients)	Scheuer	Intensity adjustment. Random selection of non-overlapping visual fields. CPA assessment (Image-Pro plus 6.0)	PSWE correlation: CPA ( $r = 0.628, p < 0.05$ ), Scheuer ( $r = 0.473, p < 0.05$ ), [13] system ( $r = 0.487, p < 0.05$ )
Thiele et al. 2016 [43]	99 samples (chronic ALD patients) Stain: Sirius red	Ishak	Samples digitization (Nanozoomer 2.0 HT). Image analyzing (VISIOMorph 4.3.6)	TE—2D-SWE AUC $\geq 0.92$
Rastellini et al. 2018 [44]	61 samples (cALD patients), 9 normal as controls Stain: picro-Sirius red	-	CPA quantification (Qwin Leica Q550IW)	Mean CPA in ALD (7.1%, $p = 0.17$ ). Fibrosis—HVPg correlation ( $p = 0.02$ )
Ali et al. 2012 [45]	11 samples (SNP-pretreated mice) Stain: Sirius red	-	Fibrosis areas quantification (Image J)	Fibrosis: SNP-MSc ( $0.4 \pm 0.3$ ), CCl <sub>4</sub> ( $4.4 \pm 2.4$ ), MSc ( $1.9 \pm 1.5$ ), SNP ( $1.3 \pm 0.4$ )



**Table 4.** Modern methodologies based on nonlinear SHG/TPEF imaging microscopy.

Author/Year	Dataset/Histological Stain	HAI	Image Analysis Method	Results/Remarks
Sun et al. 2008 [46]	3 samples (CLD Wistar rats) Stains: Sirius red, Masson's trichrome	-	SHG background removal. Frequency domain filtering. Otsu segmentation. Morphological erosion & dilation	CPA (%): SHG > MT stain, SHG = SR stain
Tai et al. 2009 [47]	3 human samples, 20 animal samples (15 viral Wistar rats, 5 normal as controls) Stain: Masson's trichrome	Ishak (modified). Ishak	Threshold levels determination (MATLAB)	Human CPA (4.21 ± 0.67%) vs. Fibro-C-Index (2.78% ± 0.44%)
Guilbert et al. 2014 [48]	12 samples (CLD, ALD, hepatitis B or C patients) Stain: with Sirius red	METAVIR	Mean SHG intensity. Intensity threshold determination (Fluoview)	CPA—METAVIR (F3, F4) linear correlations ( $R^2 = 0.997, 0.9944$ )
Sun et al. 2017 [49]	71 samples (CHB patients) Stain: H&E, reticulin and MT (at a later stage)	Ishak. METAVIR. Laennec	SHG image collagen area extraction and normalization with TPEF signals	CPA—HAIs correlation ( $p = 0.001$ )

**Table 5.** Recent publications building intelligent systems for automated CPA detection.

Author/Year	Dataset/Histological Stain	HAI	Image Analysis Method	Results/Remarks
Matalka et al. 2006 [50]	210 samples (CLD patients), 50 normal samples Stain: Van Geison	Ishak	Contrast enhancement. Morphological opening. K-means clustering. ANN classification (MATLAB 6.5)	AQS testing accuracy: 94.69%. AQS—Ishak correlations (Ph1: 0.889, Ph2: 0.815)
Stanciu et al. 2014 [51]	200 samples (TAA-induced rats) Stain: MT (at a later stage)	METAVIR	SHG/TPEF image preprocessing. K-means clustering. BoF, k-NN classification	Mean PR-area ≤ 0.9719
Xu et al. 2014 [52]	162 samples (CHB patients), 25 TAA-induced rats) Stain: MT (at a later stage)	METAVIR. Ishak	Automated SHG/TPEF image preprocessing. MLR classification (qFibrosis)	qFibrosis—Ishak 5, 6 (AUC = 0.73, $p = 0.008$ ). qFibrosis—METAVIR: animals (AUC ≤ 0.99, $p < 0.001$ ), humans (AUC ≤ 0.97, $p < 0.05$ )
Sun et al. 2018 [53]	162 samples (CHB patients) Stain: H&E, MT, Reticulin	Ishak (modified)	Automated SHG/TPEF image preprocessing. (qFibrosis). Feature selection, model fitting (MATLAB 2015a)	qFibrosis—Ishak correlation: LS ( $r = 0.6$ , $p < 0.001$ ), LS after treatment ( $r = 0.3$ , $p < 0.001$ )
Wang et al. 2015 [54]	175 samples (cHBV patients) Stain: MT (at a later stage)	Ishak	Binary image conversion. Collagen area segmentation GLM, SVM classification	GLM—Ishak 1–4 vs. 5–6 (AUC = 0.829). SVM—Ishak AUC (1 vs. 2–6 = 0.745, 1–2 vs. 3–6 = 0.774, 1–3 vs. 4–6 = 0.797)
Meejaroen et al. 2015 [55]	34 samples (fibrosis patients)	-	Image enhancement. Low-pass filtering. Bayesian classification	Classification accuracy = 91.42%
Thong-on and Watchareeruetai 2017 [56]	32 samples (fibrosis patients)	-	MOGP Segmentation (NSGA-II). Bayesian classification	NSGA-II segmentation accuracy = 93.05%. Bayesian classification accuracy = 92.33%
Giannakeas et al. 2015 [57]	25 samples (CHC patients) Stain: picro-Sirius Red	Ishak	Manual fibrotic area annotation. K-means color image clustering	CPA segmentation error rates ≤ 2.93%
Tsipouras et al. 2017 [58]	79 samples (HCV patients) Stain: picro-Sirius red	-	K-means, Fuzzy C-means clustering. KNN, C4.5, RF, MLP, SVM fibrosis regions classification	Mean CPA classification error (abs) = 1.31%. CCC = 0.923
Tsouros et al. 2017 [59]	8 samples (HCV patients)	-	K-means algorithm (parameterized) image clustering. Bayesian classification	Mean classification accuracy = 97.79%, 1/2 of image dataset error < 1%
Vicas et al. 2017 [60]	107 samples (fibrosis and steatosis prevalence)	-	Image preprocessing. GBT, SVM, LR, RF, CNN classification. U-Net Segmentation	$R^2 \leq 89.3^1$
Yu et al. 2018 [61]	25 samples (21 TAA-induced rats and 4 as controls) Stain: picro-Sirius red (at a later stage)	METAVIR	SHG contrast enhancement. Morphological closing. ANN, MLR, SVM, RF, AlexNet-CNN classification	AUROC: CNN ≤ 0.95, ANN ≤ 1.00, MLR ≤ 1.00, SVM ≤ 0.99, RF ≤ 0.99

<sup>1</sup> Correlation coefficient with human expert quantification.

#### 4.1. Results from Earlier Works Employing DIA and HAI Systems

In Jimenez et al. [11], the morphometric method indicated that the proportion of fibrosis was significantly higher in biopsies with cirrhosis than in those with steatosis ( $p < 0.001$ ) or chronic hepatitis ( $p < 0.001$ ) only. The degree of fibrosis was higher in biopsies with alcoholic fibrosis compared to those with steatosis ( $p < 0.01$ ) and chronic hepatitis ( $p < 0.05$ ) only. A significant correlation was

observed between the percentage of fibrosis estimated histomorphometrically and the amount of collagen measured colorimetrically ( $r = 0.77, p < 0.001$ ).

Manabe et al. [12] observed a regression of liver collagen in patients treated with interferon, whereas conventional histological evaluation (Knodell's HAI) did not detect any significant improvement in fibrosis. Moreover, the colorimetric and morphometric measurements of total collagen and the collagen content in the Disse space correlated significantly ( $r = 0.774, p < 0.005$  and  $r = 0.709, p < 0.01$ ). A highly significant correlation ( $r = 0.778$ ) was also observed between the collagen content measured colorimetrically and the percentage of fibrosis assessed histomorphometrically.

Nakayashi et al. [13] noted that the collagen index of the liver tissue (ratio of the stained fibrotic area to the total tissue area) correlated well with the biochemically measured hydroxyproline content in biopsy specimens as well as with the fibrosis percentage retrieved from samples colored with SR and Fast green double stain ( $r = 0.887$ ).

Chevallier et al. [14] proved that the microscopy histological image analysis was better related to morphometric measurement of fibrosis as well as with four main fibrotic deposit sites embedded in the developed by the researchers' semi-quantitative scoring system (SSS). In particular, the estimation of the total fibrosis score was expressed by the following formula:

$$\text{SSS} = \text{CLV} + \text{PS} + \text{PT} + 2(\text{WS} + \text{NS}), \quad (11)$$

where CLV refers to centrilobular vein, PS to the perisinusoidal space, PT to the portal tract, while WS and NS stand for the width and the number of the septa, respectively. The results showed that the correlations between the novel SSS and the surface density of total collagen (SDTC) were ( $r = 0.73, p \leq 10^{-5}$ ). Each one of the above five individual histological variables, except fibrosis in the CLV, presented an excellent intra-observer and inter-observer agreement.

Several studies were based on microscopy image analysis to examine differences in liver collagen content at several stages of chronic hepatitis. In Nakabayashi et al. [15], the total collagen index (CI) was correlated well with the one quantified by biochemical analysis ( $r = 0.88, p < 0.0001$ ). Kage et al. [16] observed a strong and significant Scheuer correlation between the stage and the area of fibrosis in both hepatitis B ( $r = 0.67, p < 0.001$ ) and hepatitis C patients ( $r = 0.75, p < 0.001$ ). Specifically, the area of the collagen fiber in the initial biopsy was related significantly to the period of evolution from chronic hepatitis C to cirrhosis. Also, the results proved to be well correlated and the increase in the rate of fibrosis in hepatitis C patients was noticeably higher than in hepatitis B patients.

Pilette et al. [17] showed that the area of fibrosis determined by microscopy image analysis and the semi-quantitative scores were well correlated ( $r = 0.84, p < 10^{-4}$ ). When serum markers of fibrosis were taken into account, the correlation was higher when the area of fibrosis was used as reference compared to a semi-quantitative score. In conclusion, most markers were not correlated with major fibrotic histological findings, meaning that they lack accuracy over the quantitative measurement of liver fibrosis. Thus, the merits of microscopy image analysis motivated the researchers to focus on designing and developing image analysis tools, such as the FibroQuant application, to quantify the percentage of the glomerular porto-periportal and septal fibrosis areas.

Masseroli et al. [18] showed significant correlations ( $0.72 < r < 0.83, p < 0.0001$ ) between the microscopy image analysis based on FibroQuant application and the Knodell, Scheuer histologic evaluations of liver fibrosis. Caballero et al. [19], using the FibroQuant application, showed a significant reduction in porto-periportal and septal fibrosis areas among IFN responders ( $p < 0.001$ ) and non-responders ( $p < 0.05$ ). Finally, an increase in the fibrosis percentage in non-responders was observed ( $p < 0.001$ ).

In Colloredo et al. [20], the reduction of the biopsy length led to an increase of values in cases of mild fibrosis ( $p < 0.001$ ). Similarly, the number of complete and incomplete portal tracts decreased significantly when the specimen was shortened from the original  $\geq 3$  cm to 1 cm ( $p < 0.001$ ). The values in cases of severe fibrosis were also decreased among the shorter specimens. In terms of width,

both fibrosis grade and stage were significantly underscored in the 1 mm samples, regardless of their length.

In Tanano et al. [21], due to experimentation with different magnification factors, the coefficient of variation (CV) of the fibrosis index (FI) ranged from (a) 3.2% to 8.2% at the magnification of  $\times 20$ , (b) 2.9% to 7.7% at  $\times 40$ , and (c) 5.4% to 24.2% at  $\times 100$ . It was concluded that FI at stoma closure or FIDR was significantly lower in living patients than in patients who eventually died or were subjected to liver transplantation. In Wright et al. [22], at lower  $\times 40$  magnification, the peak proportion area change (PPAC) of fibrosis method was highly correlated with the modified HAI semi-quantitative fibrosis score ( $r = 0.621$ ,  $p = 0.0002$ ). During the comparison of semi-quantitative and DIA systems, both methods showed a good intra and inter-observer correlation: (a) DIA:  $r = 0.889$ ,  $r = 0.837$  and (b) modified HAI:  $r = 0.878$ ,  $r = 0.776$ , respectively.

O'Brien et al. [23] found a significant Ishak correlation between the fibrosis ratio and an ordinal score change ( $r = 0.58$ ,  $p < 0.001$ ). However, a subset analysis showed that this correlation was restricted to biopsy specimens with high scores ( $r = 0.63$ ,  $p < 0.001$ ), and there was no correlation or difference between category means found among biopsy specimens with low scores (0–3: normal to early bridging fibrosis). Moreover, an agreement between the two estimates of fibrosis progression was found in only 11 (30%) of the 37 pairs compared.

Ryder et al. [24] demonstrated that there were no significant differences relating to the risk of fibrosis progression, either using the Knodell or the Ishak scoring system. This was supported by the fact that there was a very strong correlation between the two staging systems ( $p = 0.0001$ ). In Lazzarini et al. [25], the digital analysis of fibrosis percentage was highly correlated with the Ishak scores ( $p < 0.001$ ). The area under receiver operating characteristic (AUROC) curves showed a reliable discriminative method capability when compared with semi-quantitative assessments of fibrosis and also an excellent inter-observer reliability (AUROC = 0.982–1.00).

In Arima et al. [26], patients with viral eradication presented a  $7.2\% \pm 1.5\%$  fibrosis rate before therapy and  $2.7\% \pm 0.5\%$  after therapy, showing a significant regression. Instead, in the conservative therapy group, the rate was  $8.4\% \pm 4.3\%$  in the first biopsy and  $15.9\% \pm 7.7\%$  in the second, respectively. Regression was overall confirmed to be a virologic response to IFN even in patients with liver cirrhosis.

In Xie et al. [27], the mean CPA value of the decompensated stage of CHB-related cirrhosis was  $35.93\% \pm 14.42\%$ . The correlation coefficients of the CPA with a model for end-stage liver disease (MELD) score, serum total bilirubin (TBIL), and the international standard ratio (ISR) of prothrombin B were 0.553, 0.519 and 0.533, respectively ( $p < 0.001$ ).

In Manousou et al. [28], the multivariate analysis (MVA) showed that only CPA was associated with clinical liver decompensation ( $p = 0.010$ ). The current group continued their research in reference [29] and concluded that in the univariate analysis the first clinical decompensation was associated with a fibrosis rate of increase based on CPA ( $p < 0.001$ ), with an Ishak stage fibrosis rate of increase ( $p = 0.001$ ), with advanced donor age ( $p = 0.007$ ) and with histological hepatitis C ( $p = 0.013$ ). Moreover, the only factor associated with clinical decompensation was the rate of fibrosis increase according to CPA ( $p = 0.001$ ).

New studies evaluated the relationship between the CPA measured by microscopy image analysis, Ishak score and HVPG. Calvaruso et al. [30] showed that the CPA was a better histological correlate with HVPG ( $r = 0.62$ ,  $p < 0.001$ ) than the Ishak stage, it had a greater numerical change when HVPG was low and resulted in further quantitation of fibrosis in cirrhosis. Respectively in a later work [31], Ishak 5 and 6 stages of CPA correlated with HVPG ( $p = 0.04$ ), but due to the wide range of output values, a greater sensitivity for distinguishing “early” from “late” severe fibrosis/cirrhosis was indicated. Besides, CPA was a unique and independent predictor of HVPG  $\geq 10$  mm of mercury (mmHg), with the endpoint being the clinical liver decompensation and CPA providing better discrimination than HVPG or Ishak stage.

Tsochatzis et al. [32] found that the CPA was associated with future clinical decompensation ( $p = 0.017$ ). Notably, CPA increased significantly across the Laennec scoring system, with mean values in

cirrhosis stages carrying a correlation value of  $p < 0.001$ . This difference was also present in the Kumar classification with a correlation of  $p = 0.003$ . To conclude, CPA positively correlated with the septal width ( $r = 0.622$ ,  $p < 0.001$ ) and negatively with the size of nodules ( $r = -0.620$ ,  $p < 0.001$ ), while there was no correlation with the number of nodules. More importantly, CPA's AUROC when predicting clinical decompression was at 0.909.

In Huang et al. [33], CPA Sirius red (CPAs) along with CPA trichrome (CPAt) were well correlated with the METAVIR stage, with CPAs showing a better performance and correlating with seven of the eight total serum markers ( $p < 0.001$ ). Later, Huang et al. [34] found a good correlation between the CPA stage values and the METAVIR fibrosis grading system ( $p < 0.001$ ). In conclusion, both research works presented similar sensitivity values, with CPA presenting a significantly higher predictive ability for the HCC development than the METAVIR stage.

#### 4.2. Correlation of Histology Imaging with Non-Invasive Markers of Fibrosis

The general summary in Naveau et al. [35] was that the FibrometerA and Hepascore diagnostic values did not differ from that of FibroTest for advanced fibrosis (AUROC =  $0.83 \pm 0.03$ ) and cirrhosis, and they were significantly greater than those of APRI, Forns, and FIB4 biomarkers ( $p < 0.01$ ). Multivariate analysis (MVA) indicated FibroTest as the most informative biomarker, as in patients with chronic viral hepatitis B and C, its prognostic value was similar to that of biopsy samples.

Raftopoulos et al. [36] found that the total area of fibrosis assessed morphometrically, correlated with Hepascore ( $r = 0.603$ ,  $p < 0.001$ ), FibroTest ( $r = 0.392$ ,  $p = 0.03$ ), and HBeAg-positive ( $r = 0.492$ ,  $p = 0.001$ ) scores only, whereas Hepascore proved to be more accurate compared to the FibroTest and HBeAg-positive biomarkers.

Campos et al. [37] showed that the incorporation of microscopy image analysis provided a more complete evaluation of fibrosis. Indeed, a high correlation was observed between the fibrosis index calculated by the DIA quantification technique and the Ishak and METAVIR scores ( $r = 0.95$  and  $0.92$ , respectively,  $p < 0.001$ ). An excellent intra-observer reproducibility was observed with an intraclass correlation index of 0.99.

In Zhou et al. [38], a strong correlation was observed between the fibrosis index and the Ishak ( $r = 0.83$ ), METAVIR ( $r = 0.86$ ), and Laennec ( $r = 0.84$ ) HAIs ( $p < 0.001$ ). The cut-off value associated with cirrhosis was 7.7% and with an AUROC curve of 0.95 ( $p < 0.001$ ). In addition, the fibrosis index yielded a cut-off value of: (a) 8.9% (AUROC = 0.74), (b) 12% (AUROC = 0.84), and (c) 14% (AUROC = 0.97) for the diagnosis of cirrhotic stages 4a, 4b and 4c, respectively.

The Banerjee et al. [39] MR measures were strongly correlated with those of histology, indicating: (a)  $r = 0.68$ ,  $p < 0.0001$  for fibrosis, (b)  $r = 0.89$ ,  $p < 0.001$  for steatosis, and (c)  $r = -0.69$ ,  $p < 0.0001$  for hemosiderosis. The area under the ROC curve was 0.94, 0.93, and 0.94 for the diagnosis of any degree of fibrosis, steatosis, and hemosiderosis, respectively.

In Marino et al. [40], the presence of early significant sinusoidal fibrosis (SF) allowed the identification of 78.9% and 90.6% of patients with fibrosis (F)  $\geq 2$  and HVPG  $\geq 6$  mmHg, respectively, while patients with early SF had older donor age and higher necroinflammatory activity (NIA). More importantly, it was concluded that Masson's trichrome (MT) staining during the SF classification showed a fair correlation with that of Sirius red (SR) staining ( $p < 0.01$ ).

In Chen et al. [41], the METAVIR fibrosis stages were dichotomized using the CPA and shear wave velocity (SWV). To dichotomize F1 versus F2–F4, the AUROC curves for the CPA was 0.9349 and for SWV equal to 0.8434. This comparison between CPA and SWV showed a correlation of  $p = 0.0063$ . Respectively for F1, F2 versus F3, F4, the CPA AUROC was 0.9436 and the SWV AUROC was 0.8997 ( $p = 0.1587$ ). For F1–F3 versus F4, the CPA AUROC was 0.8647, and the SWV AUROC was 0.9036 ( $p = 0.2585$ ).

In Ding et al. [42], the point shear wave elastography (PSWE) values increased according to the severity of hepatic fibrosis, while there were significant differences among fibrosis stages according to the Scheuer scoring system ( $p < 0.05$ ). PSWE correlated positively with CPA ( $r = 0.628$ ,  $p < 0.05$ ) and

showed a more logical relationship with CPA than with the stages evaluated by the Scheuer HAI ( $r = 0.473$ ,  $p < 0.05$ ) or the Nakayashi et al. [13] semi-quantitative scoring system ( $r = 0.487$ ,  $p < 0.05$ ).

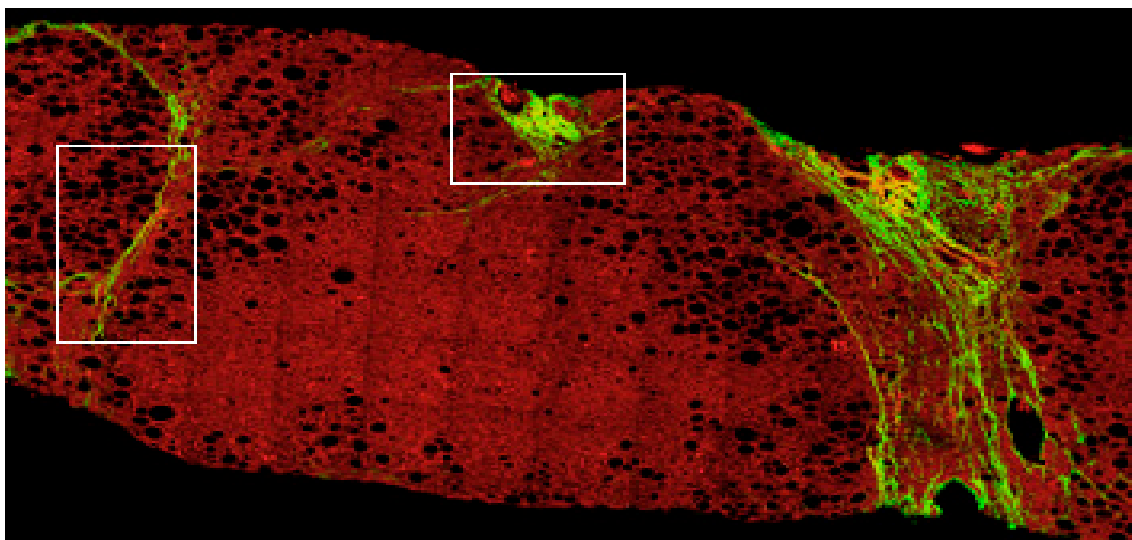
Thiele et al.'s [43] diagnostic results showed that transient elastography and two-dimensional shear wave elastography (2D-SHE) identified subjects with significant fibrosis (Ishak score  $\geq 3$ ) and cirrhosis (Ishak score  $\geq 5$ ) with a high-performance level ( $AUC \geq 0.92$ ). An important fact is that there was no difference in the diagnostic accuracy between the two techniques.

In Rastellini et al. [44], the mean CPA in patients with ALD was 7.1%, with no difference between active drinkers and abstinent patients ( $p = 0.17$ ). Using a fibrosis density cut-off of 5%, a positive correlation between high fibrosis density and the HVPG was observed only in the active drinkers ( $p = 0.02$ ).

Ali et al. [45], after performing the quantitative determination of collagen fibers with an image processing software, indicated that the mesenchymal stem cells (MSc) therapy resulted in a significant reduction of liver fibrosis. The percentage (%) of fibrotic area was significantly reduced in sodium nitroprusside (SNP)-MSc combination treatment group ( $0.4 \pm 0.3$ ) compared to other groups (CCl<sub>4</sub> induced fibrosis =  $4.4 \pm 2.4$ , MSc =  $1.9 \pm 1.5$  and SNP =  $1.3 \pm 0.4$ ).

#### 4.3. Results from SHG/TPEF Microscopy Methodologies

Sun et al. [46] compared the obtained diagnostic results with those of Sirius red and Masson's trichrome staining images. The outcomes indicated that features observed in the SHG/TPEF images also appeared in the SR biopsy specimens, confirming that the SHG signals reflected the fibrillar collagen deposition faithfully (Figure 7). In comparison with MT staining, SHG microscopy was able to provide more sensitive and quantitative information about collagen in the liver tissues, having the ability to detect collagen fibers inside the sinusoids.



**Figure 7.** Localization of 2D collagen patterns in second-harmonic generation/two-photon excitation fluorescence (SHG/TPEF) microscopy imaging. The TPEF image provides red color values for the scanned tissue sample, while SGH signals indicate the areas of collagen deposition in green.

In Tai et al. [47], the calculated CPA in a human tissue sample was  $4.21\% \pm 0.67\%$  and equivalent to a Fibro-C-Index value of  $2.78\% \pm 0.44\%$ . The semi-quantitative Ishak assessment performed by a histopathologist produced a good agreement, confirming the use of Fibro-C-Index as accurate, if not better than, the conventional staining methods for collagen content quantification.

When SHG was used as reference, Guilbert et al. [48] showed fibrosis areas saturated with high collagen content, whereas non-saturated areas corresponded to the non-fibrotic regions of the human



liver. After the image reconstruction phase, the diagnostic results showed a good linear correlation between the sophisticated method's scoring and the METAVIR (F3, F4) assessment ( $R^2 = 0.997, 0.9944$ ).

According to the Beijing classification by Sun et al. [49], progressive, indeterminate, and regressive types of fibrosis were observed in 58%, 29%, and 13% of total patients before Entecavir treatment, versus 11%, 11%, and 78% after treatment. In general, the CPA showed a correlation equal to  $p = 0.001$  with the semi-quantitative HAI systems. Of the 55 patients who showed regressive changes on post-treatment phase, 29 cases (53%) showed fibrosis recurrence at least one Ishak stage, whereas 25 cases (45%) had significant fibrosis improvements in terms of Laennec sub-stage, CPA quantification and liver stiffness, despite remaining in the same Ishak stage.

#### 4.4. CPA Results Produced by Automated Intelligent Systems

The overall testing accuracy in Matalka et al. [50] was 94.69%. Also, its correlation with the semi-quantitative Ishak scoring of the two physicians was acceptable (Ph1: 0.889, Ph2: 0.815). In Stanciu et al. [51], the classification of DSIFT-BoF was evaluated by calculating the under Precision-Recall (PR) curves area. When various bin sizes were applied, an optimal PR value in METAVIR stage 0\_S images was observed, with a mean PR-area of 0.9719. It was made clear that a 6-pixel bin size provided satisfactory mean PR-area values, especially in the case of METAVIR S012\_vs.\_S34 (0.7948) and METAVIR S0123\_vs.\_S4 (0.9451) images.

The Xu et al. [52] fully automated method identified differences between all METAVIR stages in both animal samples ( $p < 0.001$ ) and human biopsies ( $p < 0.05$ ). qFibrosis was able to differentiate Ishak stages 5 and 6 (AUC = 0.73,  $p = 0.008$ ) and was able to predict staging underestimation in suboptimal biopsies (<15 mm) and under/over-scoring by different pathologists ( $p < 0.001$ ). In Sun et al. [53], the qFibrosis results correlated well with liver stiffness ( $r = 0.6, p < 0.001$ ). Following antiviral treatment, a decline in inflammation prevalence was observed, with LSM significantly decreasing and significantly correlating with the qFibrosis diagnosis ( $r = 0.3, p < 0.001$ ).

A receiver operating characteristic (ROC) curve was used in Wang et al. [54] for the evaluation of SHG/TPEF fibrosis scoring performance. Overall, the non-linear SVM model was superior to the generalized linear model (GLM). In more detail, the multivariate generalized linear model was able to perform an optimal separation between non-cirrhotic (Ishak stages 1–4) and cirrhotic (Ishak 5–6) patients with AUC = 0.829. However, the SVM model was optimal in classifying early fibrotic stages: (a) Ishak 1 vs. 2–6: AUC = 0.745, (b) Ishak 1–2 vs. 3–6: AUC = 0.774, and (c) Ishak 1–3 vs. 4–6: AUC = 0.797.

The experimental results of Meejaroen et al. [55] showed that an  $8^3$  bin value combined with a  $7 \times 7$  average filtering gave the best performance, yielding a fibrosis classification accuracy of 91.42%. In addition, the average differences between the fibrosis percentage obtained from the proposed method and that of ground truth images were 2.29 points.

The evolutionary method of Thong-on and Watchareeruetai [56] was executed 10 times with the same parameter setting. The fibrosis estimation error was measured as the absolute difference between the percentage of segmented fibrosis and the corresponding ground truth, determined by other methods. Of all generated individuals, the optimum yield had a fibrosis segmentation accuracy of 93.05 and a corresponding error of 2.09, which was less than the handcraft method (error = 2.63). Concerning the rest of the individuals, about half of them were able to achieve better results, while the rest had similar error range and compared to a Bayesian classification approach, which showed accuracy and error rates of 92.33 and 2.63, respectively.

After the study by Giannakeas et al. [57], computed CPA errors were found to be less than 1% in most images. However, there were three cases (from 25 in total) where the calculated CPA from the proposed method greatly differed from the value provided by the physician. In all other cases, error rates ranged from 0.05% to 2.93%. In Tsipouras et al. [58], all CPA results derived from

the proposed methodology were compared to those defined from the experts' annotations for each clustering/classification algorithm, with the calculation of the absolute CPA error, defined as

$$AE_{CPA}(\%) = |CPA_{annot} - CPA_{method}|, \quad (12)$$

where the mean absolute CPA error for all biopsy images was finally calculated, as well as the absolute CPA error standard deviation. To test the agreement between the CPA provided by the experts and the CPA that was automatically evaluated, the concordance correlation coefficient (CCC)  $[-1, 1]$  was used, with value 1 corresponding to perfect agreement, value  $-1$  to perfect disagreement and value of 0 to no agreement. The final results produced a 1.31% mean absolute CPA error and a CCC of 0.923.

In Tsouros et al. [59], all results were obtained based on the sensitivity and precision values for each class, as well as the classification accuracy for each image. The methodology produced high classification results, with accuracy ranging from 93.69% to 99.50% and the average accuracy for all images being 97.79%. CPA results were encouraging, as four out of eight images showed an error of  $< 1\%$ .

In Vicas et al. [60], the validation of the automated quantitative analysis was based on the qualitative scores set by a physician. This was done once per approach and could not be used for hyper-parameter optimization. In conclusion, the correlation coefficient  $R^2$  was 0.748 for the classical computer vision approach and 0.893 for the CNN approach.

In Yu et al. [61], all comparisons between the five fibrosis stages produced a high area under the receiver operating characteristic values (AUROC  $> 0.85$ ) in the majority of the classification models. The deep AlexNet convolutional neural network presented balanced AUROC values (0.85–0.95) along with the conventional ANN (0.87–1.00) and RF classifier (0.94–0.99). For the MLR and SVM algorithms, the corresponding values were much lower (0.73–1.00 and 0.69–0.99, respectively), as the system performance was affected in both cases by less critical features.

## 5. Conclusions

The present study focused on the analysis of research works on the automated detection of liver fibrosis in microscopic images derived from the needle biopsy gold standard. The purpose of this review was to identify the benefits and limitations of the digital image analysis (DIA) techniques, used as diagnostic tools, but also to identify the "gray-areas" including the obstacles in the collagen proportional area (CPA) quantification process. Over the last few decades, research efforts have been based on this principle, which aims to eliminate the diagnostic limitations coming from subjective human interpretations by building computational methods that can obtain accurate clinical determinations for chronic liver diseases.

Most of the initial image analysis methods in the literature required the user's interference to manually estimate the percentage of fibrosis. This proved to be a difficult task, as the quantitation of fibrosis in a large set of images refers to a time-consuming procedure. FibroQuant was at first the only approach that provided both automated threshold estimation and artifact elimination, using several thresholding techniques. Thresholding remained the simplest technique for image segmentation so that more sophisticated techniques could be employed for more accurate detection, either in the hepatic parenchyma or in the fibrosis areas. However, sometimes the precise inclusion of the fibrosis area had to be determined by the user.

On the other hand,  $\times 20$  to  $\times 200$  magnifications refer to an important limitation of the proposed image analysis techniques. As for higher magnifications, these resulted in several image fragments, which needed to be merged into the final image for image processing. For this reason, the magnification of  $\times 40$  proved to be the ideal choice for equalizing the quality of a single image with computational processing resources. It is important to note that biopsy specimens with a length of 3 cm and a thickness of 2–5  $\mu\text{m}$  remain the ideal choice for microscopy image analysis. Thanks to these dimensions, 4–6 complete portal tracks can be included and lead to a more reliable fibrosis prevalence quantification.

In other cases, reducing the length of the needle biopsy has been shown to lead to an increase in mild fibrosis.

In recent years, unsupervised machine learning techniques have been proposed for the segmentation of the biopsy images through clustering, therefore providing an alternative to the manual thresholding automated solution. Additionally, the detection of areas of exclusion is another issue of image analysis. These areas could be a capsule, portal vessel, or biliary duct. In most of the previous works, the authors removed such areas manually. To overcome this challenge, supervised classification-based methods have successfully been developed to characterize each segmented stained area and exclude the undesired ones. The newest research efforts have increasingly employed deep convolutional neural network topologies in state-of-the-art diagnostic methods. This new generation of algorithms has proven capable to provide fully automated and highly accurate solutions to the segmentation and classification of fibrous specimens without the need for manual annotations.

**Author Contributions:** A.A., R.F., and N.G. conceived of the idea and apply the PRISMA methodology along with the inclusion and exclusion criteria. M.G.T. and A.T.T. explored and categorized the proposed methodologies according to their technical contribution. R.F. and P.M. review the presented methodologies regarding the medical question and provide the pathological point of view of the survey. A.A., R.F., and all other authors prepared the manuscript. M.G.T. and A.T.T. organized the research team, and P.M. and N.G. supervised the survey. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is partly funded by the project entitled xBalloon, co-financed by the European Union and 468 Greek national funds through the Operational Program for Research and Innovation Smart Specialization 469 Strategy (RIS3) of Ipeiros (Project Code: 5033187).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Koumbi, L. Gut microbiota alterations in liver diseases. *EC Gastroenterol. Dig. Syst.* **2018**, *5*, 191–198.
2. Lo, R.C. Histopathological evaluation of liver fibrosis and cirrhosis regression. *Clin. Mol. Hepatol.* **2017**, *4*, 302–307. [[CrossRef](#)] [[PubMed](#)]
3. Asselah, T.; Marcellin, P.; Bedossa, P. Improving performance of liver biopsy in fibrosis assessment. *J. Hepatol.* **2014**, *61*, 193–195. [[CrossRef](#)] [[PubMed](#)]
4. Desmet, V.J.; Gerber, M.; Hoofnagle, J.H.; Manns, M.; Scheuer, P.J. Classification of chronic hepatitis: Diagnosis, grading and staging. *Hepatology* **1994**, *19*, 1513–1520. [[CrossRef](#)] [[PubMed](#)]
5. Knodell, R.G.; Ishak, K.G.; Black, W.C.; Chen, T.S.; Craig, R.; Kaplowitz, N.; Kiernan, T.W.; Wollman, J. Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. *Hepatology* **1981**, *1*, 431–435. [[CrossRef](#)]
6. Scheuer, P.J. Classification of chronic viral hepatitis: A need for reassessment. *J. Hepatol.* **1991**, *13*, 372–374. [[CrossRef](#)]
7. Bedossa, P.; Poynard, T. An algorithm for grading activity in chronic hepatitis C. *Hepatology* **1996**, *24*, 289–293. [[CrossRef](#)]
8. Ishak, K.; Baptista, A.; Bianchi, L.; Callea, F.; De Groote, J.; Gudat, F.; Denk, H.; Desmet, V.; Korb, G.; MacSween, R.N. Histological grading and staging of chronic hepatitis. *J. Hepatol.* **1995**, *22*, 696–699. [[CrossRef](#)]
9. Brunt, E.M. Grading and staging the histopathological lesions of chronic hepatitis: The Knodell histology activity index and beyond. *Hepatology* **2000**, *31*, 241–246. [[CrossRef](#)]
10. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* **2009**, *6*. [[CrossRef](#)]
11. Jimenez, W.; Pares, A.; Caballeria, J.; Heredia, D.; Bruguera, M.; Torres, M.; Rojkind, M.; Rodes, J. Measurement of fibrosis in needle liver biopsies: Evaluation of a colorimetric method. *Hepatology* **1985**, *5*, 815–818. [[CrossRef](#)] [[PubMed](#)]
12. Manabe, N.; Chevallier, M.; Chossegros, P.; Causse, X.; Guerret, S.; Trepo, C.; Grimaud, J.A. Interferon-alpha 2b therapy reduces liver fibrosis in chronic non-A, non-B hepatitis: A quantitative histological evaluation. *Hepatology* **1993**, *18*, 1344–1349. [[CrossRef](#)] [[PubMed](#)]

13. Nakayashi, H.; Takamatsu, S.; Nakano, H. A simple estimation of liver collagen content by the computed color image analyzer using sirius red stained sections. *Kanzo* **1993**, *34*, 212–218. [[CrossRef](#)]
14. Chevallier, M.; Guerret, S.; Chossegros, P.; Gerard, F.; Grimaud, J.A. A histological semiquantitative scoring system for evaluation of hepatic fibrosis in needle liver biopsy specimens: Comparison with morphometric studies. *Hepatology* **1994**, *20*, 349–355. [[CrossRef](#)]
15. Nakabayashi, H.; Takamatsu, S.; Tsujii, H.; Okamoto, Y.; Nakano, H.; Yamada, E.; Ichijima, K. Collagen content of liver biopsy specimens in patients with chronic hepatitis. *Int. Hepatol. Commun.* **1996**, *4*, 311–315. [[CrossRef](#)]
16. Kage, M.; Shimamatu, K.; Nakashima, E.; Kojiro, M.; Inoue, O.; Yano, M. Long-term evolution of fibrosis from chronic hepatitis to cirrhosis in patients with hepatitis C: Morphometric analysis of repeated biopsies. *Hepatology* **1997**, *25*, 1028–1031. [[CrossRef](#)]
17. Pilette, C.; Rousselet, M.C.; Bedossa, P.; Chappard, D.; Oberti, F.; Rifflet, H.; Maiga, M.Y.; Gallois, Y.; Cales, P. Histopathological evaluation of liver fibrosis: Quantitative image analysis vs semi-quantitative scores. *J. Hepatol.* **1998**, *28*, 439–446. [[CrossRef](#)]
18. Masseroli, M.; Caballero, T.; O'Valle, F.; Del Moral, R.M.; Perez-Milena, A.; Del Moral, R.G. Automatic quantification of liver fibrosis: Design and validation of a new image analysis method: Comparison with semi-quantitative indexes of fibrosis. *J. Hepatol.* **2000**, *32*, 453–464. [[CrossRef](#)]
19. Caballero, T.; Perez-Milena, A.; Masseroli, M.; O'Valle, F.; Salmeron, F.J.; Del Moral, R.M.; Sanchez-Salgado, G. Liver fibrosis assessment with semiquantitative indexes and image analysis quantification in sustained-responder and non-responder interferon-treated patients with chronic hepatitis C. *J. Hepatol.* **2001**, *34*, 740–747. [[CrossRef](#)]
20. Colloredo, G.; Guido, M.; Sonzogni, A.; Leandro, G. Impact of liver biopsy size on histological evaluation of chronic viral hepatitis: The smaller the sample, the milder the disease. *J. Hepatol.* **2003**, *39*, 239–244. [[CrossRef](#)]
21. Tanano, H.; Hasegawa, T.; Kimura, T.; Sasaki, T.; Kawahara, H.; Kubota, A.; Okada, A. Proposal of fibrosis index using image analyzer as a quantitative histological evaluation of liver fibrosis in biliary atresia. *Pediatric Surg. Int.* **2003**, *19*, 52–56. [[CrossRef](#)] [[PubMed](#)]
22. Wright, M.; Thursz, M.; Pullen, R.; Thomas, H.; Goldin, R. Quantitative versus morphological assessment of liver fibrosis: Semi-quantitative scores are more robust than digital image fibrosis area estimation. *Liver Int.* **2003**, *23*, 28–34. [[CrossRef](#)] [[PubMed](#)]
23. O'Brien, M.J.; Keating, N.M.; Elderiny, S.; Cerda, S.; Keaveny, A.P.; Afdhal, N.H.; Nunes, D.P. An assessment of digital image analysis to measure fibrosis in liver biopsy specimens of patients with chronic hepatitis C. *Am. Soc. Clin. Pathol.* **2000**, *114*, 712–718. [[CrossRef](#)] [[PubMed](#)]
24. Ryder, S.D.; Irving, W.L.; Jones, D.A.; Neal, K.R.; Underwood, J.C. Progression of hepatic fibrosis in patients with hepatitis C: A prospective repeat liver biopsy study. *Gut* **2004**, *53*, 451–455. [[CrossRef](#)]
25. Lazzarini, A.L.; Levine, R.A.; Ploutz-Snyder, R.J.; Sanderson, S.O. Advances in digital quantification technique enhance discrimination between mild and advanced liver fibrosis in chronic hepatitis C. *Liver Int.* **2005**, *25*, 1142–1149. [[CrossRef](#)]
26. Arima, M.; Terao, H.; Kashima, K.; Arita, T.; Nasu, M.; Nishizono, A. Regression of liver fibrosis in cases of chronic liver disease type C: Quantitative evaluation by using computed image analysis. *Intern. Med.* **2004**, *43*, 902–910. [[CrossRef](#)]
27. Xie, S.B.; Ma, C.; Lin, C.S.; Zhang, Y.; Zhu, J.Y.; Ke, W.M. Collagen proportionate area of liver tissue determined by digital image analysis in patients with HBV-related decompensated cirrhosis. *Hepatobiliary Pancreat. Dis. Int.* **2011**, *10*, 497–501. [[CrossRef](#)]
28. Manousou, P.; Dhillon, A.P.; Isgro, G.; Calvaruso, V.; Luong, T.V.; Tsochatzis, E.; Xirouchakis, E.; Kalambokis, G.; Cross, T.J.; Rolando, N.; et al. Digital image analysis of liver collagen predicts clinical outcome of recurrent hepatitis C virus 1 year after liver transplantation. *Liver Transplant.* **2011**, *17*, 178–188. [[CrossRef](#)]
29. Manousou, P.; Burroughs, A.K.; Tsochatzis, E.; Isgro, G.; Hall, A.; Green, A.; Calvaruso, V.; Ma, G.L.; Gale, J.; Burgess, G.; et al. Digital image analysis of collagen assessment of progression of fibrosis in recurrent HCV after liver transplantation. *J. Hepatol.* **2013**, *58*, 962–968. [[CrossRef](#)]

30. Calvaruso, V.; Burroughs, A.K.; Standish, R.; Manousou, P.; Grillo, F.; Leandro, G.; Maimone, S.; Pleguezuelo, M.; Xirouchakis, I.; Guerrini, G.P.; et al. Computer-assisted image analysis of liver collagen: Relationship to Ishak scoring and hepatic venous pressure gradient. *Hepatology* **2009**, *49*, 1236–1244. [[CrossRef](#)]
31. Calvaruso, V.; Dhillon, A.P.; Tsochatzis, E.; Manousou, P.; Grillo, F.; Germani, G.; Patch, D.; O’Beirne, J.; Burroughs, A.K. Liver collagen proportionate area predicts decompensation in patients with recurrent hepatitis C virus cirrhosis after liver transplantation. *J. Gastroenterol. Hepatol.* **2012**, *27*, 1227–1232. [[CrossRef](#)] [[PubMed](#)]
32. Tsochatzis, E.; Bruno, S.; Isgro, G.; Hall, A.; Theocharidou, E.; Manousou, P.; Dhillon, A.P.; Burroughs, A.K.; Luong, T.V. Collagen proportionate area is superior to other histological methods for sub-classifying cirrhosis and determining prognosis. *J. Hepatol.* **2014**, *60*, 948–954. [[CrossRef](#)] [[PubMed](#)]
33. Huang, Y.; de Boer, W.B.; Adams, L.A.; MacQuillan, G.; Rossi, E.; Rigby, P.; Raftopoulos, S.C.; Bulsara, M.; Jeffrey, G.P. Image analysis of liver collagen using sirius red is more accurate and correlates better with serum fibrosis markers than trichrome. *Liver Int.* **2013**, *33*, 1249–1256. [[CrossRef](#)]
34. Huang, Y.; de Boer, W.B.; Adams, L.A.; MacQuillan, G.; Bulsara, M.K.; Jeffrey, G.P. Image analysis of liver biopsy samples measures fibrosis and predicts clinical outcome. *J. Hepatol.* **2014**, *61*, 22–27. [[CrossRef](#)]
35. Naveau, S.; Gaude, G.; Asnacios, A.; Agostini, H.; Abella, A.; Barri-Ova, N.; Dauvois, B.; Prevot, S.; Ngo, Y.; Munteanu, M.; et al. Diagnostic and prognostic values of noninvasive biomarkers of fibrosis in patients with alcoholic liver disease. *Hepatology* **2009**, *49*, 97–105. [[CrossRef](#)]
36. Raftopoulos, S.C.; George, J.; Bourliere, M.; Rossi, E.; de Boer, W.P.; Jeffrey, G.P.; Bulsara, M.; Speers, D.J.; MacQuillan, G.; Ching, H.L.; et al. Comparison of noninvasive models of fibrosis in chronic hepatitis B. *Hepatol. Int.* **2012**, *6*, 457–467. [[CrossRef](#)]
37. Campos, C.F.; Paiva, D.D.; Perazzo, H.; Moreira, P.S.; Areco, L.F.; Terra, C.; Perez, R.; Figueiredo, F.A. An inexpensive and worldwide available digital image analysis technique for histological fibrosis quantification in chronic hepatitis C. *J. Viral Hepat.* **2014**, *21*, 216–222. [[CrossRef](#)]
38. Zhou, Y.; Ru, G.Q.; Yan, R.; Wang, M.S.; Chen, M.J.; Yu, L.L.; Wang, H. An inexpensive digital image analysis technique for liver fibrosis quantification in chronic hepatitis B patients. *Ann. Hepatol.* **2017**, *16*, 881–887. [[CrossRef](#)]
39. Banerjee, R.; Pavlides, M.; Tunnicliffe, E.M.; Piechnik, S.K.; Sarania, N.; Philips, R.; Collier, J.D.; Booth, J.C.; Schneider, J.E.; Wang, L.M.; et al. Multiparametric magnetic resonance for the non-invasive diagnosis of liver disease. *J. Hepatol.* **2014**, *60*, 69–77. [[CrossRef](#)]
40. Marino, Z.; Mensa, L.; Crespo, G.; Miquel, R.; Bruguera, M.; Perez-Del-Pulgar, S.; Bosch, J.; Forns, X.; Navasa, M. Early periportal sinusoidal fibrosis is an accurate marker of accelerated HCV recurrence after liver transplantation. *J. Hepatol.* **2014**, *61*, 270–277. [[CrossRef](#)]
41. Chen, S.H.; Peng, C.Y.; Lai, H.C.; Chang, I.P.; Lee, C.J.; Su, W.P.; Lin, C.H.; Kao, J.T.; Chuang, P.H. Head-to-head comparison between collagen proportionate area and acoustic radiation force impulse elastography in liver fibrosis quantification in chronic hepatitis C. *PLoS ONE* **2015**, *10*, e0140554. [[CrossRef](#)] [[PubMed](#)]
42. Ding, H.; Ma, J.J.; Wang, W.P.; Zeng, W.J.; Jiang, T.; Huang, B.J.; Chen, S.Y. Assessment of liver fibrosis: The relationship between point shear wave elastography and quantitative histological analysis. *J. Gastroenterol. Hepatol.* **2015**, *30*, 553–558. [[CrossRef](#)] [[PubMed](#)]
43. Thiele, M.; Detlefsen, S.; Sevelsted, M.L.; Madsen, B.S.; Fuglsang, H.J.; Fialla, A.D.; Trebicka, J.; Krag, A. Transient and 2-dimensional shear-wave elastography provide comparable assessment of alcoholic liver fibrosis and cirrhosis. *Gastroenterology* **2016**, *150*, 123–133. [[CrossRef](#)] [[PubMed](#)]
44. Rastellini, S.; Goossens, N.; Clement, S.; Lanthier, N.; Negro, F.; Rubbia-Brandt, L.; Spahr, L. Collagen proportionate area correlates to hepatic venous pressure gradient in non-abstinent cirrhotic patients with alcoholic liver disease. *World J. Hepatol.* **2018**, *10*, 73–81. [[CrossRef](#)] [[PubMed](#)]
45. Ali, G.; Mohsin, S.; Khan, M.; Nasir, G.A.; Shams, S.; Khan, S.N.; Riazuddin, S. Nitric oxide augments mesenchymal stem cell ability to repair liver fibrosis. *J. Transl. Med.* **2012**, *10*, 75. [[CrossRef](#)] [[PubMed](#)]
46. Sun, W.; Chang, S.; Tai, D.C.; Tan, N.; Xiao, G.; Tang, H.; Yu, H. Nonlinear optical microscopy: Use of second harmonic generation and two-photon microscopy for automated quantitative liver fibrosis studies. *J. Biomed. Opt.* **2008**, *13*, 064010. [[CrossRef](#)]



47. Tai, D.C.; Tan, N.; Xu, S.; Kang, C.H.; Chia, S.M.; Cheng, C.L.; Wee, A.; Wei, C.L.; Raja, A.M.; Xiao, G.; et al. Fibro-C-Index: Comprehensive, morphology-based quantification of liver fibrosis using second harmonic generation and two-photon microscopy. *J. Biomed. Opt.* **2009**, *14*, 044013. [[CrossRef](#)]
48. Guilbert, T.; Odin, C.; Le Grand, Y.; Gailhouste, L.; Turlin, B.; Ezan, F.; Desille, Y.; Baffet, G.; Guyader, D. A robust collagen scoring method for human liver fibrosis by second harmonic microscopy. *Opt. Express* **2010**, *18*, 25794–25807. [[CrossRef](#)]
49. Sun, Y.; Zhou, J.; Wang, L.; Wu, X.; Chen, Y.; Piao, H.; Lu, L.; Jiang, W.; Xu, Y.; Feng, B.; et al. New classification of liver biopsy assessment for fibrosis in chronic hepatitis B patients before and after treatment. *Hepatology* **2017**, *65*, 1438–1450. [[CrossRef](#)]
50. Matalka, I.I.; Al-Jarrah, O.M.; Manasrah, T.M. Quantitative assessment of liver fibrosis: A novel automated image analysis method. *Liver Int.* **2006**, *26*, 1054–1064. [[CrossRef](#)]
51. Stanciu, S.G.; Xu, S.; Peng, Q.; Yan, J.; Stanciu, G.A.; Welsch, R.E.; So, P.T.; Csucs, G.; Yu, H. Experimenting liver fibrosis diagnostic by two photon excitation microscopy and Bag-of-Features image classification. *Sci. Rep.* **2014**, *4*. [[CrossRef](#)] [[PubMed](#)]
52. Xu, S.; Wang, Y.; Tai, D.C.S.; Wang, S.; Cheng, C.L.; Peng, Q.; Jan, J.; Chen, Y.; Sun, J.; Liang, X.; et al. qFibrosis: A fully-quantitative innovative method incorporating histological features to facilitate accurate fibrosis scoring in animal model and chronic hepatitis B patients. *J. Hepatol.* **2014**, *61*, 260–269. [[CrossRef](#)]
53. Sun, Y.; Zhou, J.; Wu, X.; Chen, Y.; Piao, H.; Lu, L.; Ding, H.; Nan, Y.; Jiang, W.; Wang, T.; et al. Quantitative assessment of liver fibrosis (qFibrosis) reveals precise outcomes in Ishak "stable" patients on anti-HBV therapy. *Sci. Rep.* **2018**, *8*, 2989. [[CrossRef](#)] [[PubMed](#)]
54. Wang, T.H.; Chen, T.C.; Teng, X.; Liang, K.H.; Yeh, C.T. Automated biphasic morphological assessment of hepatitis B-related liver fibrosis using second harmonic generation microscopy. *Sci. Rep.* **2015**, *11*, 12962. [[CrossRef](#)] [[PubMed](#)]
55. Meejaroen, K.; Chaweechan, C.; Khodsiri, W.; Khu-smith, V.; Watchareeruetai, U.; Sornmagura, P.; Kittiyakara, T. Detection of fibrosis in liver biopsy images by using Bayesian classifier. In Proceedings of the 7th International Conference on Knowledge and Smart Technology (KST), Chon Buri, Thailand, 28–31 January 2015; Institute of Electrical and Electronics (IEEE): Piscataway, NJ, USA, 2015. [[CrossRef](#)]
56. Thong-on, P.; Watchareeruetai, U. Detection of fibrosis in liver biopsy images using multi-objective genetic programming. In Proceedings of the 9th International Conference on Information Technology and Electrical Engineering (ICITEE), Phuket, Thailand, 12–13 October 2017; Institute of Electrical and Electronics (IEEE): Piscataway, NJ, USA, 2017. [[CrossRef](#)]
57. Giannakeas, N.; Tsiouras, M.G.; Tzallas, A.T.; Kyriakidi, K.; Tsianou, Z.E.; Manousou, P.; Hall, A.; Karvounis, E.C.; Tsianos, V.; Tsianos, E. A clustering based method for collagen proportional area extraction in liver biopsy images. *Proc. IEEE Eng. Med. Biol. Soc. Annu. Conf.* **2015**, *2015*, 3097–30100. [[CrossRef](#)]
58. Tsiouras, M.G.; Giannakeas, N.; Tzallas, A.T.; Tsianou, Z.E.; Manousou, P.; Hall, A.; Tsoulos, I.; Tsianos, E. A methodology for automated CPA extraction using liver biopsy image analysis and machine learning techniques. *Comput. Methods Progr. Biomed.* **2017**, *140*, 31–68. [[CrossRef](#)] [[PubMed](#)]
59. Tsouros, D.C.; Smyrlis, P.N.; Tsiouras, M.G.; Tsalikakis, D.G.; Giannakeas, N.; Tzallas, A.T.; Manousou, P. Automated collagen proportional area extraction in liver biopsy images using a novel classification via clustering algorithm. In Proceedings of the 30th International Symposium on Computer-Based Medical Systems, Thessaloniki, Greece, 22–24 June 2017; Institute of Electrical and Electronics (IEEE): Piscataway, NJ, USA, 2017; pp. 30–34. [[CrossRef](#)]
60. Vicas, C.; Rusu, I.; Al Hajjar, N.; Lupsor-Platon, M. Deep convolutional neural nets for objective steatosis detection from liver samples. In Proceedings of the 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, Romania, 7–9 September 2017; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2017; pp. 385–390. [[CrossRef](#)]
61. Yu, Y.; Wang, J.; Ng, C.W.; Ma, Y.; Mo, S.; Fong, E.L.S.; Xing, J.; Song, Z.; Xie, Y.; Si, K.; et al. Deep learning enables automated scoring of liver fibrosis stages. *Sci. Rep.* **2018**, *30*, 16016. [[CrossRef](#)]

