Support vector machine (SVM) based liver classification: fibrosis, steatosis, and inflammation

Jihye Baek
Department of Electrical and Comupter Engineering
University of Rochester
Rochester, NY, USA
jbaek7@ur.rochester.edu

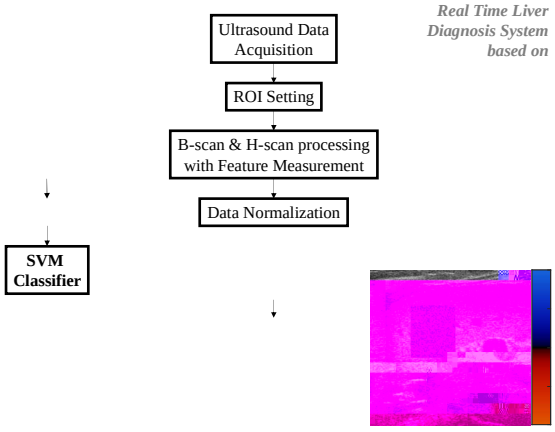
Terri A. Swanson
Pfizer Inc.
Cambridge, MA, USA
terri.a.swanson@pfizer.com

Theresa Tuthill
Pfizer Inc.
Cambridge, MA, USA
theresa.tuthill@pfizer.com

Kevin J. Parker
Department of Electrical and Comupter Engineering
University of Rochester
Rochester, NY, USA
kevin.parker@rochester.edu

***Abstract*— An SVM based liver classifier was developed to differentiate liver conditions, including normal, fibrosis with low fat, fibrosis with high fat, and inflammation.**

An in-vivo study was performed with 35 rats under normal conditions or after carbon tetrachloride (CCl_4) or concanavalin A (ConA) dosing to induce fibrosis with varying degrees of steatosis, and inflammation, respectively. These livers were imaged in-vivo by an ultrasound, and approximately 30 frames for each rat were acquired. Therefore, a total of 998 ultrasound images were analyzed and used for training a SVM classifier. Each image has three measured parameters: H-scan scattering A(EB



Real Time Liver Diagnosis System based on

fat (> 9%), 11 for fibrosis with low fat (< 6.5%), and 9 for inflammation.

Ultrasound scans were performed at 8 weeks after starting the protocol. A Vevo 2100 (Visual Sonics, Toronto, Canada)

positions of diseased cases compared to normals. Although no single parameter can differentiate among the 4 classes without overlap, each feature tends to discriminate specific groups. The H-scan provides the best separation between normal and inflammation, and attenuation estimation is also likely to distinguish inflammation from normal. The B-scan intensity separates fibrosis from the others. For the separation between normal and low fatty group, all the three measurements provide discrimination. The H-scan and attenuation tend to differentiate low and high fatty groups. The B-scan can separate low fatty fibrosis and inflammation. Multi-parametric analysis combines all these results for an improved classification.

Figure 3 (a) and (b) represent two-dimensional (2D) and three-dimensional (3D) view of clusters, respectively. For 2D view, principal component analysis (PCA) reduced the three parameters into two, and the retained variances are 84.9%, 13.0%, and 2.1% for the first, second, and third principal components (PC), respectively. **Figure 3(a)** exclusively used the first two PCs with 97.9% of retained variances. Combining the three features in 2D space provides better separation between the four liver groups compared to each feature results in **Figure 2**, but there are still overlaps between the groups.

Figure 3 (b, c) illustrates two views of clusters in 3D space by using all the three measured features with normalized scales. The 3D views provide better separation