Multi-class SVM model for fMRI-based classification and grading of liver fibrosis

M. Freiman, Y. Sela, Y. Edrei, O. Pappo, L. Joskowicz and R. Abramovitch

School of Engineering and Computer Science, The Hebrew University of Jerusalem, Israel.
The G. Savad Inst. for Gene Therapy, and MRI/MRS Lab HBRC, Hadassah Hebrew Univ. Medical Center, Jerusalem, Israel.
Dept. of Pathology, Hadassah Hebrew Univ. Medical Center, Jerusalem, Israel.

ABSTRACT

We present a novel non-invasive automatic method for the classification and grading of liver fibrosis from fMRI maps based on hepatic hemodynamic changes. This method automatically creates a model for liver fibrosis grading based on training datasets. Our supervised learning method evaluates hepatic hemodynamics from an anatomical MRI image and three T2*-W fMRI signal intensity time-course scans acquired during the breathing of air, air-carbon dioxide, and carbogen. It constructs a statistical model of liver fibrosis from these fMRI scans using a binary-based one-against-all multi class Support Vector Machine (SVM) classifier. We evaluated the resulting classification model with the leave-one out technique and compared it to both full multi-class SVM and K-Nearest Neighbor (KNN) classifications. Our experimental study analyzed 57 slice sets from 13 mice, and yielded a 98.2% separation accuracy between healthy and low grade fibrotic subjects, and an overall accuracy of 84.2% for fibrosis grading. These results are better than the existing image-based methods which can only discriminate between healthy and high grade fibrosis subjects. With appropriate extensions, our method may be used for non-invasive classification and progression monitoring of liver fibrosis in human patients instead of more invasive approaches, such as biopsy or contrast-enhanced imaging.

Keywords: Abdominal, Characterization, Machine Learning

1. INTRODUCTION

The degree and rate of liver fibrosis progression are important prognostic factors in patients with chronic liver disease. The degree of liver fibrosis is currently assessed by a biopsy. Liver biopsy is an invasive procedure that carries a risk of serious complications with a procedure-related mortality rate of 1:10,000. Other limitations of liver biopsy include: 1) lack of functional assessment of the whole organ and disease development grading due to the local nature of the information; 2) biopsy sampling error; 3) complications and patient discomfort, and; 4) morbidity and ethical issues of repeated liver biopsies.

Existing non-invasive imaging methods for liver fibrosis grading are based on elastography, texture, diffusion, and perfusion. For a detailed review of current liver fibrosis imaging methods, see. Elastography-based methods require modeling the mechanical properties of the liver parenchyma. Current models of the liver mechanical behavior are not patient-specific, and therefore have limited accuracy. In addition, elastography-based imaging is cumbersome, as it requires attaching specialized mechanical equipment to the patient that must also fit into the imaging device. Texture and perfusion-based methods require the intravenous injection of contrast agent, which may lead renal toxicity and is a major cause of acute renal failure. The main limitations of other imaging methods include poor spatial resolution in radio nuclear studies and lack of reproducibility in Doppler ultrasound and radiation exposure in CT. In hemodynamic perfusion-based CT and MRI images, good separation of the arterial from the portal phase requires a high temporal resolution, which comes at the cost of reduced spatial resolution. Current experience shows that these methods can differentiate between healthy and high-grade fibrosis patients (f0 and f4-5 stages, respectively). However, these imaging methods cannot differentiate the earlier stages (f2-3) based on the Batts and Ludwig scoring system.
Recently, Abramovitch et al.\textsuperscript{16,17} demonstrated the feasibility of functional MRI (fMRI) combined with hypercapnia and hyperoxia for monitoring changes in liver perfusion and hemodynamics without contrast agent administration in liver tumors, cirrhosis and fibrosis in animal models. Unlike contrast agent based methods, this method can detect steady state levels without compromising between the spatial and temporal resolutions. Since hemodynamical changes are subtle, relative, and spatially distributed, direct observation of these changes is difficult, unreliable, and time-consuming. Previously, only two parameters were extracted from the entire time-course scans based on the differences between the mean fMRI signal intensity values during the inhalation of each gas.\textsuperscript{16} However, basing the entire analysis on mean values may lead to erroneous conclusions, as it ignores the entire kinetics information. The hemodynamics fMRI technique has shown great promise on mice and is currently being modified and tested on humans.

In this paper, we present a supervised learning method for the automatic creation of a classification and grading model for mice liver fibrosis from fMRI signal intensity time-course scans. Our method uses the entire time-course scans as input to a multi-class classifier in order to perform multi-class classification according to Batts and Ludwig fibrosis grades. Since it does not require a complex mechanical model and uses all the time-course scans information, it is potentially more accurate than the existing methods and provides a quantitative evaluation of the entire time course information. This provides the radiologist with a better grading of the fibrosis level.

In order to evaluate our method, we compared the performance of three multi-class classifiers: 1) One Against All Support Vector Machine (OAA-SVM),\textsuperscript{18–20} 2) Multi-class Support Vector Machine (MSVM),\textsuperscript{21, 22} and; 3) K-Nearest Neighbor (KNN) classifier\textsuperscript{23} to ground-truth classification based on histology analysis performed by an expert radiologist. Experimental results for 57 slices sets from 13 mice show that fMRI based multi-class classification can for the first time accurately differentiate between healthy and low grade fibrosis subjects based on non-invasive imaging only. This is in contrast with existing imaging-based methods, which can only distinguish between healthy and high grade fibrosis subjects.\textsuperscript{13} These results suggest that the fMRI based approach can be useful for human datasets, with appropriate pre-processing steps such as respiratory motion correction\textsuperscript{24} and liver segmentation.\textsuperscript{25}

## 2. METHOD

Our method follows the supervised learning paradigm. First, we build a multi-class classification model from a set of image datasets from healthy subjects and subjects with different grades of fibrosis: healthy $f_0$, periportal fibrosis $f_2$, and septal fibrosis $f_3$.\textsuperscript{15} The model is then used to classify and grade datasets of new subjects. Each dataset is processed with the same analysis method.

The input dataset consists of an anatomical MRI image and a set of fMRI-based time-courses representing hepatic hemodynamics for each pixel. The fMRI images are acquired during the breathing of air, air-carbon dioxide (5% $CO_2$), and carbogen (95% oxygen; 5% $CO_2$) with the protocol described in.\textsuperscript{16} There is no need to co-register the images, since the mice were under anesthesia and immobilized to the table. For each inhaled gas, eight time points were taken.

### 2.1 Features normalization

Normalization of the hemodynamic activity vectors is necessary due to the high variability of the gradient echo signal from different subjects. Since the fibrosis does not affect the signal intensity of the fMRI images during normal air breathing, we use this part of the time-courses to normalize the fMRI intensity values of the entire vector.

The normalization step proceeds as follows. First, the liver borders are interactively segmented in the anatomical MRI images and the remaining tissues are eliminated from the fMRI images. Next, the first two time points for each inhaled gas are removed to ensure the stability of the fMRI signal during the inhalations. Then, the hemodynamic activity vectors are normalized to have a mean value of 1:

$$\Delta I(\vec{x}) = \frac{I(\vec{x})}{\mu_{I_{air}}}$$  \hspace{1cm} (1)
Figure 1. (a)-(c) Anatomical MRI slices of three mice specimens with different fibrosis grades and their corresponding (d)-(i) fMRI-based hemodynamic maps and (j)-(l) histological specimens stained with Masson’s trichrome stain (the connective tissue is stained in blue).

where $\vec{x}$ is the pixel coordinates and $\mu_{I_{air}}$ is the mean intensity value during normal air breathing. Finally, the median hemodynamic activity vector from the entire liver is computed for each slice.
2 Liver fibrosis hemodynamics fMRI time-course

Figure 2. Representative median fMRI signal intensity time-course from three mice: healthy $f_0$ (blue), periportal fibrosis $f_2$ (red), and septal fibrosis $f_3$ (green). The dashed lines denote the changes in the inhaled gas, starting with normal air, followed by hypercapania and by hyperoxia.

Fig. 1 shows the anatomical MRI images of three representative mice along with their corresponding fMRI-based hemodynamic maps$^{16}$ and their corresponding histological specimens. Fig. 2 shows the median feature-vectors from three representative livers that were used for the classification.

2.2 Multi-class classification

When a new set of hemodynamic activity vectors is given, our goal is to classify each vector as either no fibrosis ($f_0$), periportal fibrosis ($f_2$), or septal fibrosis ($f_3$). For training and testing purposes, each liver was individually diagnosed by an expert pathologist according to the actual liver fibrosis degree in the histological specimen, stained with Masson’s trichrome staining.

We tested the performance of three multi class classifiers: 1) One Against All (OAA-SVM),$^{19,20}$ 2) single optimization Multi class SVM (MSVM),$^{21,22}$ and 3) K nearest neighbor (KNN) classifier$^{23}$

The OAA-SVM classifier$^{19}$ uses one SVM binary model to separate each of the $k$ classes from the rest. Formally, given a training set $(\vec{x}_1, y_1)\ldots(\vec{x}_n, y_n)$ with labels $y_i$ in $[1..k]$, the $i$th SVM is defined as:

$$\min_{\vec{w}, b, \xi} \frac{1}{2}||\vec{w}||^2 + C \sum_{i=1}^{k} \xi_i$$

s.t.

$$\forall i \in [1..n] : y_i (\vec{w} \cdot \vec{x}_i - b) \geq 1 - \xi_i$$

$$\xi_i \geq 0$$
where \( \vec{w} \) is the normal vector perpendicular to the hyper-surface that separates class \( y_i \) from the others, \( C \) is a regularization parameter that trades off margin size and training error, \( \xi_i \) are the slack variables that measure the degree of misclassification of the datum \( x_i \), and \( b \) is an offset scalar.

This \( i \)th classifier classifies each instance as belonging to the class \( i \), or to the rest. The final classification of each instance is obtained by applying the following rule:

\[
y_i = \arg\max_{i \in \{1, \ldots, k\}} \left( \vec{w}_i^\top y_i(x_i) + b_i \right)
\]

(3)

In the training phase, the \( i \)th SVM is trained with all of the examples in the \( i \)th class with positive labels, and all other examples with negative labels. In the classification step, the \( i \)th SVM classifier classifies the new instance as belonging to the \( i \)th class or belonging to one of the other classes. Then, for each new instance, its \( k \) values are the likelihoods that the instance belongs to each class. The final classification is obtained by choosing the class with the highest likelihood (Eq. 3).

The MSVM classifier uses a multi-class classification function that is computed by solving a single optimization problem. Given a training set as above, the multi-class classification is defined as:

\[
\min \frac{1}{2} \sum_{i=1}^{k} \| \vec{w}_i \|^2 + \frac{C}{n} \sum_{i=1}^{n} \xi_i
\]

s.t.

\[
\forall y \in \{1..k\} : \langle \vec{x}_1, \vec{w}_1 \rangle \geq \langle \vec{x}_1, \vec{w}_i \rangle + \Delta(y_1, y) - \xi_1
\]

...

\[
\forall y \in \{1..k\} : \langle \vec{x}_n, \vec{w}_n \rangle \geq \langle \vec{x}_n, \vec{w}_n \rangle + \Delta(y_n, y) - \xi_n
\]

(4)

where \( \vec{w}_i \) is the normal vector perpendicular to the hyper-surface that separates class \( y_i \) from the others, \( \langle \vec{x}_i, \vec{w}_n \rangle \) is the vector dot product, and \( \Delta(y_n, y) \) is the loss function that is 0 if \( y_n \) equals \( y \), and 1 otherwise.

For both classifiers, we use a Gaussian Radial Basis Function (GRBF) kernel defined as:

\[
k(\vec{x}_i, \vec{x}_j) = \exp \left( -\frac{\| \vec{x}_i - \vec{x}_j \|^2}{2\sigma^2} \right)
\]

(5)

The KNN classifier is a non-parametric classifier. Given a training set \((\vec{x}_1, y_1)\ldots(\vec{x}_n, y_n)\) with labels \( y_i \) in \([1..k]\), it classify a new sample \( \vec{x} \) as follows. First, compute the distance between the new sample \( \vec{x} \) and each sample of the training data \( \vec{x}_1 \ldots \vec{x}_n \). Next, choose the \( K \) closest samples. Based on this information, set the new label \( y \) as the label with the largest number of instances among the \( K \) closest samples.

3. EXPERIMENTAL RESULTS

In order to evaluate our computerized approach, we conducted an in-vivo mice study and performed two modeling experiments on the data.

The mice study consists of thirteen 2-month-old and 4-month-old multi-drug resistance protein 2 gene knock-out mice (Mdr2\(^{−/−}\)) and equivalent aged matched (FVB/NJ) mice. The Mdr2\(^{−/−}\) mice are deficient in the canalicular phospholipid flippase and spontaneously develop liver injury and chronic inflammation due to the absence of phospholipid from bile, which leads to the formation of periportal biliary fibrosis at an early age.\(^{26}\)

MRI scans of the mice, fixed and immobilized under anesthesia, were acquired in a horizontal 4.7-T Biospec spectrometer (Bruker Medical, Ettlingen, Germany) using a 3.5-cm bird-cage coil. Coronal and axial T1-weighted spin echo images of the whole liver were acquired for alignment and to determine liver borders (repetition time = 360 ms; echo time = 18 ms). Hepatic perfusion and hemodynamics were evaluated from T2*-weighted gradient echo images (repetition time=147 ms; echo time=10 ms; flip angle=30; field of view=3cm; 256x128 pixels; in-plane resolution=117\(\mu \)m; five slices with slice thickness=1mm; spectral width of 25,000 Hz; two averages; 37
<table>
<thead>
<tr>
<th></th>
<th>MSVM</th>
<th>OAA-SVM</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>f0</td>
<td>14</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>f2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>f3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

| Sensitivity | 100% | 44.4% | 76% |

Table 1. Confusion matrices for the MSVM, OAA-SVM and KNN classifiers. The columns indicate the true class and the rows indicate the hypothesized class.

...secs/image) acquired during breathing of air, air-CO$_2$ (95% air and 5% CO$_2$), and oxygen-CO$_2$ (95% oxygen and 5%CO$_2$). Eight repeats were acquired at each gas mixture. Zero filling of k-space data was applied to obtain 256 × 256 pixels matrices.

To assess the fibrosis grades, the mice were sacrificed at parallel time points to the MRI scans and histological slides were stained with Masson’s trichrome for the identification of connective tissue. Histological slides were reviewed by an expert pathologist who was blind to the MRI scans. The fibrosis grade of the 57 slices was as follows: 14 were graded as no fibrosis (f0), 18 as periportal fibrosis (f2), and 25 as septal fibrosis (f3).

We conducted two modeling experiments. In the first experiment, we separated between healthy mice, and mice with f2 and f3 grade fibrosis using the binary SVM classification model. The kernel σ parameter was optimized according to the Leave One Out (LOO) results using binary search and set to 300. The accuracy was evaluated with LOO, where the entire current mouse slices were removed and classified to avoid the relations between same subject slices. The classifier accuracy was 100% with 100% precision and recall.

In the second experiment, we graded each slice from the dataset as f0, f2 or f3 using two multi-class SVM based classification models and a KNN classifier. A OAA-SVM multi-class SVM model, a MSVM and a KNN classification models were built. Both models were evaluated using LOO. As in the the first experiment, we set σ = 300 for the SVM models kernels. The KNN classifier was then evaluated for $K = 4$ with the Euclidean distance function. All parameters were optimized experimentally.

Table 1 presents the confusion matrices obtained from the three classifiers. The sensitivity of our model was computed for each class against the others separately. The MSVM and KNN classifiers separated with 100% accuracy between non-fibrotic and fibrotic subjects, while the OAA-SVM mis-classified one slice with f2 grade as f0. However, the OAA-SVM classifier performed better for the classification between subjects with f3 fibrosis grade and subjects with f2 fibrosis grade, compared to both MSVM and KNN classifiers. The overall accuracy of the OAA-SVM was 84.2% which is better than the MSVM accuracy of 71.9% and the KNN classifier accuracy of 75.5%.

### 4. CONCLUSION

In this work we shows a new method for liver fibrosis grading based on the fMRI hemodynamic response maps to hypercapnia and hyperoxia combined with multi-class SVM\ KNN classification. Our combined method enables the automatic non-invasive grading of liver fibrosis even for low grade fibrosis. Leave-One-Out evaluation of our data yields nearly 100% accuracy for the classification between no-fibrosis and low grade fibrosis subjects, and the overall accuracy was 84.2% for fibrosis grading.

These results indicates that our method can, for the first time, differentiate between healthy and low grade fibrosis subjects based on MRI images with comparable accuracy to histology-based human grading. Our method may be used for non-invasive classification and progression monitoring of liver fibrosis patients instead of more invasive approaches, such as biopsy, or contrast-enhanced imaging.

We are currently expanding our database with an additional animal model of chemically induced fibrosis (CCLU) and plan to perform a comparative study to evaluate the performance of different classification methods on it. In these days we are trying to implement this fMRI method in healthy human subjects. We modify the fMRI imaging protocol for human liver hemodynamics. We are also developing an automatic liver contour
segmentation algorithm and a method to co-register datasets to compensate for respiratory motions. Once the liver regions are identified and co-registered, the algorithm described in this paper would hopefully can be used for liver fibrosis grading.

ACKNOWLEDGMENTS

Moti Freiman is partially supported by the Hebrew University Hoffman academic and social leadership program scholarship.

REFERENCES


