Liver Fat Assessment with Body Shape

Yijiang Zheng¹, Qiyue Wang², James K. Hahn³

Abstract— Hepatic steatosis has become a serious health concern among the general population, but especially for those who are obese. Liver fat can increase the risk of cirrhosis and even liver cancer. Current standard methods to assess hepatic steatosis, such as liver biopsy and CT/MR imaging techniques, are expensive and/or may have associated risks to health. In this paper, we use body shapes to assess hepatic steatosis using both traditional linear regression models and a deep neural network. We apply our models to a medical dataset and evaluate the approaches for both regression and classification. We compare the performance of several models via popular evaluation metrics. The experimental results indicate that our proposed neural network outperforms the vanilla linear regression model by 22.37% in RMSE and the accuracy by 18%. The R-squared value of the neural model is more than 0.72 and the accuracy reaches 78%. Hence, the body shape features can provide an additional accurate and affordable choice to monitor the degree of the patient’s liver fat.

Clinical Relevance— This paper presents a low cost and convenient approach to predict liver fat percentage using body shapes. This approach will not replace the gold standard for assessing hepatic steatosis. However, with the wide availability for depth cameras (including on smartphones), the approach promises to provide another modality that can be deployed widely in clinical setting as well for home use for telehealth.

I. INTRODUCTION

Hepatic steatosis occurs when there is extra fat in the liver. Though the simple accumulation of liver fat could be hepatoprotective, excess storage may cause some common diseases e.g., Non-Alcoholic Fatty Liver Disease (NAFLD) [1]. NAFLD is associated with obesity, type 2 diabetes, insulin resistance, and hyperlipidemia [1-3]. With the prevalence of obesity in the United States, NAFLD affects about 30% of the general population and over 75% of patients with type 2 diabetes [4]. About 20% of people with NAFLD have Nonalcoholic steatohepatitis (NASH), which can cause more severe problems such as liver cancer and cirrhosis [5]. Liver fat can be easily reversed and even cured when the patient has only simple fatty liver via proper treatments e.g., weight loss and healthy diet [6]. Hence, detecting the presence of steatosis and assessing the severity can help patients start treatment as early as possible.

There are no specific biochemical or serological tests for the detection of extra liver fat. Liver biopsy is currently the most accurate means to diagnose hepatic steatosis [7]. The procedure removes a small tissue sample from the patient’s liver and the sample undergoes histological examination to analyze liver conditions. However, it is not recommended when patients have severe ascites or clotting factor abnormalities [8]. In addition, there can be sampling variations in liver biopsy which can result in misdiagnosis [9].

Among non-invasive methods in clinic practice, imaging techniques are popular in the diagnosis of hepatic steatosis such as ultrasonography (US), computed tomography (CT), and magnetic resonance imaging (MRI) [10]. Ultrasonography applies high-frequency sound waves to generate images of the patient’s abdomen. The images can be used to evaluate the condition of the liver. The cost of US is relatively low. However, ultrasound does not perform well in quantifying hepatic fat content compared to the other two methods, especially in morbid obesity. CT scanning is a technique that combines special x-ray equipment with computers to diagnose fatty liver using Hounsfield Units (HU) [11]. One of the main limitations of diagnosing steatosis using CT is the inability to determine if the steatosis is reactive to infectious or inflammatory conditions, such as hepatitis or alcohol-induced or secondary to metabolic syndrome [12]. Additionally, CT scans could produce ionizing radiation which may cause damage to health and has limited use in children. The MRI uses a magnetic field to detect the fat signal from the liver based on the frequency difference between water and fat. MRI is the most sensitive imaging test for steatosis and is highly accurate even in mild steatosis. Although accurate, CT and MRI are expensive for in-clinic diagnoses. Therefore, a reliable, inexpensive, and easily accessible method is desired detect and monitor hepatic steatosis.

Body shape is associated with body fat distribution [13]. Some anthropomorphic features have shown a strong correlation with hepatic steatosis such as BMI, waist circumference (WC), and waist-to-hip ratio (WHR) [14]. However, despite the simplicity of the methods, they suffer from measurement errors and can only provide a rough assessment [15]. These approaches do not account for the whole abdominal region. Level circumferences, on the contrary, include more circumferences around the abdomen than WC. It has been used as a strong predictor for visceral abdominal tissue (VAT) [16]. It can capture more information about the abdominal shape. With the popularity of optical scan technologies (e.g., computer vision, structured light, time-of-flight) and the accessibility of such sensors on commodity devices such as smartphones, 3D body shape data can be easily obtained. It has the potential to become an important predictor in many health domain [17]. The paper will examine the

* Research supported by NIH R01DK129809.
¹Y. Zheng is with the Department of Computer Science, George Washington University, Washington, DC 20052 USA, email: yijiangzheng@gwu.edu
²Q. Wang is from the Department of Computer Science, George Washington University, Washington, DC 20052 USA, email: wangqiyue@gwu.edu
³J. K. Hahn is from the Department of Computer Science, George Washington University, Washington, DC 20052 USA, email: hahn@gwu.edu
relationship between body shape and liver fat percentage. This paper has two main contributions:

- We introduce an affordable and accessible body feature to predict the liver fat percentage which offers a new non-invasive way to assess the liver condition.
- We propose a deep neural network to predict liver fat with body shape descriptors using level circumference. This model provides much better performance compared to traditional regression models in both regression and classification evaluations.

II. DATASET AND PREPROCESSING

A. CT Dataset

The study aims to evaluate body shape scan features to predict the degree of hepatic steatosis using both regression and classification approaches. We currently do not have access to the optical body scan along with a gold standard measurement of hepatic steatosis. CT scan is an alternate way to generate accurate body shape data. The precision of CT reconstruction can be very high (0.5mm - 2mm). Therefore, we reconstruct the iso-surface from the CT scans as the feature of body shape. We also calculate the liver fat percentage directly from the CT images. The dataset includes 100 subjects’ CT scans which include their axial abdominal images. The CT scan dataset is from the George Washington University Hospital. The dataset and study are approved by the Institutional Review Board (IRB) of the George Washington University.

B. CT Processing

Fig. 1 demonstrates the process of extracting level circumferences and liver fat percentages. Since the CT images may follow different scanning protocols, we manually align the CT scans before the label processing. To ensure the consistency of the level circumference number for each subject, we re-sample all the data into 64 slices. We extract the circumferences from the 64 slices in a top to down order. Since the attenuation value of fat (about -100 HU) is much lower than soft tissue (30-40 HU) [18], we can easily calculate the liver fat percentage from the images. The data is processed by the segmentation tool ITK-SNAP [19]. We reviewed the processed data at least twice to ensure the validity of training features and ground truth.

![Data Processing and Label Processing Diagram](image)

**Figure 1.** The extraction pipeline of the level circumferences and liver fat percentage. In data processing, the raw CT scans are manually aligned and re-sampled to ensure 64 slices for each subject. In label processing, level circumferences and liver fat percentage are extracted directly from the processed slices.

III. METHODS

A. Ridge Regression

To evaluate the validity of the body shape, we apply a regression model to analyze the dataset. In the multiple linear regression model, the vanilla approach assumes independence between explanatory variables, when in actuality strong relationships could occur. Neglecting the dependency might cause the multicollinearity problem. Multicollinearity occurs when two or more features are correlated. This will increase the standard error of the coefficient and the inaccurate coefficients could impact the model precision [20]. Each subject includes 64 consecutive circumference features of the abdominal region which suggests the likelihood of dependency among features. Ridge Regression (RR) is a variant of linear regression and a popular method to deal with multicollinearity problem [21]. RR uses L2 regularization by adding a small positive constant number λ (λ > 0) to the ordinary least square estimator (OLS). The estimated coefficients \( \hat{\beta} \) are calculated by Eq. (1):

\[
\hat{\beta} = (X'X + \lambda I_p)^{-1}X'Y
\]

Where \( X \) represents the independent variables, \( Y \) is the output variable and \( I_p \) is the identity matrix. By providing the “ridge” parameter, \( \lambda \), to the diagonal elements of the input correlation matrix, the regression model can have a smaller mean squared error (MSE) compared to OLS models.

B. Deep Neural Network

It is not necessary for the relationship between body shape features and liver fat percentage to remain linear. Deep Neural network (DNN) is not only able to capture the linear separable regions inside the data, but its inherent nonlinearity can provide a more comprehensive capability to interpret the data [22]. The non-linear activation function allows DNN to learn more complex relationships.

Instead of applying a complicated neural network, we develop a 3-layer, fully connected (FC) network to test the performance in both regression and classification. The network structure and data preprocessing pipeline are shown in Fig. 2. At preprocessing phase, the standardization of training data is followed by the application of Principal Component Analysis (PCA). PCA is another method used to handle multicollinearity via maximizing the variance of the training data [23]. Hence, we process our level circumferences by PCA and choose the top 8 principal components of body shape maps which account for at least 98% of the variation in body shape. In the first two layers, we apply rectified linear activation function (ReLU). The ReLU activation function is computationally less expensive and reduces the likelihood of vanishing gradient. The final layer only contains one output as our prediction value. Due to the size of the dataset, the model contains only about 1,000 parameters to alleviate the overfitting problem. For classification, we directly compare the output value against the threshold to gain the final label.
We select Mean Squared Error (MSE) as our loss function. MSE is the most used loss function to train a regression model by narrowing down the differences between the observed and predicted values.

IV. RESULTS

We firstly review the distribution of the abdominal circumferences from CT slices. We calculate the mean value of circumference (expressed as number of pixels) per feature. The mean circumference varies by the position on the abdomen with the highest value (1072 pixels) at index 63 and lowest value (746 pixels) at index 29 with average 900 pixels and standard deviation of 96. The relative standard deviation is 10.69%, which indicates the instability of a single waist measurement. The high variance of the level circumferences could lead to the large difference when conducting waist measurements.

To alleviate the impact of multicollinearity between features, we apply ridged regression to the whole dataset to evaluate the degree of fitness and the model coefficients. We consider p-value < 0.05 as statistically significant. In Fig. 3, the R-squared value of the model is about 0.86, and the best feature p-value < 0.02 which indicates the observed features can be strongly explained by the model. We also evenly divided the features into four parts by the feature index with each part consisting of 16 features. We can see the mean absolute coefficient value in the fourth section is higher than in other locations. Its average value is around 0.32 while other parts are between 0.23 and 0.24. This suggests the circumferences from the lowest part play more important roles in the regression model.

For the regression, we analyze the performance of some simple and popular models over the dataset (e.g., traditional regression models, neural network models). The models include vanilla Linear Regression (LR), Support Vector Regression (SVR), Gaussian Process Regression (GPR), Lasso Regression (LSR), Ridge Regression (RR), and Full Connected Neural Network (FC). To better evaluate the generalizability of the models, the experiment applies 5-fold cross-validation. In each fold, the testing size ratio remains 0.2. For regression evaluation metrics, we apply Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R-Squared. In addition to the regression, classification is another important way to assess the performance of the model. The median value (10.25%) of liver fat percentage from the whole dataset is selected as the classification threshold. Class-0 has a value smaller than the threshold, while Class-1 has a value larger than or equal to the threshold.

Table I compares the results of the methods for both regression and classification evaluation. Among traditional regression models, RR performs best in all regression metrics with 0.60 R-squared value. The accuracy of RR reaches 0.68 in classification. It is worth noting that the vanilla linear regression model cannot fit the data well and has the worst RMSE as well as MAE. The performance of the neural network is better than other methods in all measurement metrics. The R-squared of FC is more than 0.7 which shows a better model fitting. FC has 0.78 accuracy in classification. The confusion matrices from Fig. 4 present the details in classification. We can see the FC has much higher accuracy in Class-0 (88%) while other methods do not perform well (between 50% - 60%).

### TABLE I. COMPARISON OF EXPERIMENTAL RESULTS

<table>
<thead>
<tr>
<th>Methods</th>
<th>Evaluation Metrics</th>
<th>RMSE</th>
<th>MAE</th>
<th>R-Squared</th>
<th>Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR</td>
<td></td>
<td>9.03</td>
<td>7.28</td>
<td>0.55</td>
<td>66</td>
</tr>
<tr>
<td>SVR</td>
<td></td>
<td>8.82</td>
<td>7.05</td>
<td>0.57</td>
<td>63</td>
</tr>
<tr>
<td>GPR</td>
<td></td>
<td>8.66</td>
<td>6.90</td>
<td>0.59</td>
<td>64</td>
</tr>
<tr>
<td>LSR</td>
<td></td>
<td>8.63</td>
<td>6.83</td>
<td>0.59</td>
<td>65</td>
</tr>
<tr>
<td>RR</td>
<td></td>
<td>8.50</td>
<td>6.68</td>
<td>0.60</td>
<td>68</td>
</tr>
<tr>
<td>DNN</td>
<td></td>
<td>7.10</td>
<td>5.38</td>
<td>0.72</td>
<td>78</td>
</tr>
</tbody>
</table>

Figure 4. The confusion matrices for the different models. Class-0 and Class-1 include 50 instances respectively.
From the above experiments, we can observe the drawbacks of simple anthropomorphic measurements such as waist circumference. The irregular distribution of abdominal circumferences and the difference between them could lead to inaccurate assessment. The lower part of the abdomen is more statistically significant for liver fat. Ridge regression can be a good fit model for the level circumference since the R-squared value in the whole dataset reaches more than 0.85. The vanilla linear regression suffers from the multicollinearity problem while ridge regression offers a better solution to mitigate the dependency among features. The activation function of the neural network brings the nonlinearity into the model fitting which increases the accuracy in both classification and regression. The proposed simple neural model outperforms the baseline linear regression model by 18% in classification accuracy and reduces 22.37% RMSE in regression. It also reaches the highest accuracy 0.78 compared to all other regression methods. In summary, we conclude that the body shape can be an important feature for hepatic steatosis assessment. In this study, the shape feature is extracted from the CT slices. We are in the process of collecting the body shapes using an optical 3D body scanner (e.g., Fit3D, San Mateo, CA). This will allow the approaches developed here to be used with a relatively inexpensive commodity hardware.

ACKNOWLEDGMENT

This work supported by the National Institutes of Health under grants R01DK129809 and the George Washington University Cross-Disciplinary Research Fund.

REFERENCES


