Report 2

Measurement of fetal head circumferent using ultrasound

By

BI12-423 Nguyen Thi Thao

A report presented for the course

BI12— DS

ICT department

Ha Noi university of science and technology

March 20, 2024

Contents

1 Introduction

One important biological sign used in prenatal ultrasound screening is fetal head circumference (HC). Historically, sonographers have measured HC by hand, a method that is laborious and time-consuming and is subject to variance between and among observers. Despite the many attempts to use artificial intelligence (AI) techniques to automate HC assessment, prior solutions had trouble generalizing, particularly when dealing with missing or unclear skull borders.

In this study, we propose a novel approach for automatic assessment of fetal HC, offering a rapid and accurate solution. Utilizing two-dimensional (2D) ultrasound imaging, our method capitalizes on the benefits of low cost, non-invasiveness, real-time acquisition, and absence of radiation hazards, making it the preferred choice for prenatal screening and functional monitoring in fetuses. The significance of fetal HC extends beyond mere measurement; it serves as a pivotal parameter for monitoring fetal growth, estimating gestational age, and informing decisions regarding delivery mode for pregnant women. As such, accurate and efficient HC assessment holds paramount importance in prenatal ultrasound examinations.

In practical practice, the obstetric ultrasound guidelines [1], [2] dictate that the current fetal HC measuring methods are carried out on the standard plane of the fetal thalamus in 2D ultrasound pictures. Because the shape of the skull resembles an ellipse, the circumference of the ellipse may be recognized as the fetal HC. Sonographers often have to hand sketch an entire ellipse using the "outer to outer" approach, beginning and ending at the outside edge of the skull. Post-processing software is then used to compute the HC. Sonographers find manual measurement to be cumbersome and time-consuming, and the accuracy of the measurement is mostly reliant on the operator's ability and experience [3]. Consequently, one of the most important steps in HC measurement is identifying the skull border on the fetal thalamus standard plane. The standard plane of the fetal thalamus in 2D ultrasound pictures is hampered by a number of issues, including low signal-to-noise ratio, unclear texture or structure generated by the uterine wall and amniotic fluid, missing skull boundary caused by acoustic shadows, and fuzzy boundary induced by motion. These elements would make it difficult or impossible to determine the fetal skull border and present significant hurdles.

In this work, we used random forest to identify measurements of the fetal head circumference. In this report i was organized as follows: In Section 2, i present the background. Method part, which i put in the section 3. Section 4 and 5 are evaluation and conclusion repectively.

2 Background

The Random Forest machine learning algorithm comprises a group of decision trees. This ensemble approach is applied in the context of fetal head circumference measurement using ultrasound, where multiple decision trees are trained on different subsets of training data and attributes. This diversity is crucial to prevent overfitting and enhance the model's ability to generalize to new inputs. Random Forest adopts a technique akin to bootstrap aggregation, or bagging, to construct varied trees by sampling training data with replacement to create subsets for each decision tree.

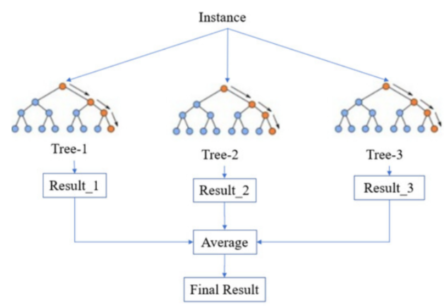

Figure 1: Random forest model

Ultrasound imaging technology plays a pivotal role in determining fetal head circumference during prenatal screening, providing insights into the size of the unborn head. This process is fundamental in obstetrics for monitoring fetal growth, estimating gestational age, and detecting anomalies. During the examination, sonographers capture two-dimensional (2D) ultrasound images of the fetal head. Subsequently, specialized software is employed to measure the circumference of the fetal skull, either manually or automatically, typically following established protocols at the level of the fetal thalamus. The fetal skull circumference serves as a critical biometric measure to assess fetal health and development.

The Random Forest ensemble learning technique involves a collection of decision trees, each trained on a distinct subset of data and features, enhancing generalization capacity and reducing overfitting. Similarly, in fetal skull circumference measurement using ultrasound, whether conducted manually or automatically, the procedure entails capturing ultrasound images to determine fetal skull circumference. Analogous to Random Forest's ensemble technique, which aggregates predictions from multiple decision trees to enhance accuracy, this process is essential for monitoring fetal growth and development throughout prenatal care. Both methodologies leverage ensemble methods and randomness to yield more consistent and reliable results, whether evaluating fetal biometrics or making predictions about outcomes.

3 Methodology

I decided to use the random forest method to assess the fetal head circumference using ultrasonography in order to build a regression model. First, the information was acquired. The HC18 dataset is the one that was used. The head circumference (HC) is a useful marker for tracking fetal growth and determining the gestational age. It is measured in the standard plane, which is a certain cross-section of the fetal skull. Thirteen hundred thirty-four two-dimensional (2D) ultrasound pictures in the standard plane that are appropriate for HC measurement make up the dataset that is made available for this work. This challenge makes it easier to compare algorithms that have been developed for automated fetal head circumference assessment in two-dimensional ultrasound pictures. The dataset for this challenge is available at DOI 10.5281zenodo.1322001 on Zenodo. The dataset is split into two parts: a test set with 335 photos and a training set with 999 images. The dimensions of each 2D ultrasound picture are 800 by 540 pixels, with a pixel size that varies from 0.052 to 0.326 mm. The CSV files 'training_set_pixel_size_and_HC.csv' and 'test_set_pixel_size.csv' contain the pixel sizes for every picture. The training set also consists of pictures that have been carefully annotated by qualified sonographers with regard to head circumference. For every annotated HC in the training set, the head circumference measurements (in millimeters) are included in the CSV file 'training_set_pixel_size_and_HC.csv'. Every filename starts with a digit. The training set has 999 photos, however the filenames only go up to 805 of them. Certain ultrasonography pictures were taken at the same echoscopic assessment, which explains why they seem so similar. These images are distinguished by an additional number in the filename between "_" and "HC" (for instance, 010_HC.png and 01_2HC.png).

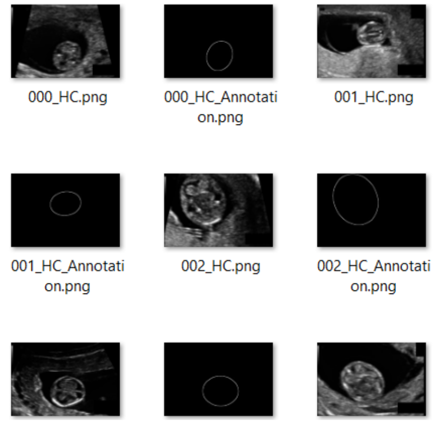

Figure 3: Some sample about dataset

Next, I moved on to the preparation of the data, eliminating any unnecessary information (for example, deleting the "filename" column) and employing "pixel size (mm)" and "head circumference (mm)" as the target variables. For validation, we divided the data into testing and training sets. After that, we built the random forest model. First, we used the RandomForestRegressor with 100 decision trees and a random_state of 42 from the sklearn.ensemble package. The model was then trained on the training set after the data was divided into training and testing sets using a 75/25 ratio.

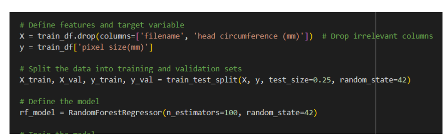

Figure 4: The example for building model

4 Evaluation

Mean_squared_error (MSE) and mean_absolute_error (MAE) are two metrics I used to evaluate the model's performance on the training and test sets. The difference between expected and actual values may be quantified using two popular statistical measures: mean square error (MSE) and mean average deviation (MAE). Whereas MAE computes the average of the absolute errors, MSE computes the average of the squared errors.

```python
# Evaluate the model
train_preds = rf_model.predict(X_train)
val_preds = rf_model.predict(X_val)

train_mse = mean_squared_error(y_train, train_preds)
train_mae = mean_absolute_error(y_train, train_preds)
val_mse = mean_squared_error(y_val, val_preds)
val_mae = mean_absolute_error(y_val, val_preds)
```

Figure 5: Code perform evaluation

The model performs well in prediction since the MSE and MAE values are relatively low in both datasets. Nonetheless, the validation set's MSE value is somewhat greater than the training set's, indicating that the model could be overfitting in certain areas. However, the model works well on both datasets, as seen by the lack of significance in the difference between MSE and MAE. This suggests that the model is not unduly dependent on the training set and can generalize well to fresh data. All things considered, these findings show that the model performs well when it comes to estimating the fetal head circumference from ultrasound pictures.

```
Training MSE: 1.1888433104172739e-07
Training MAE: 5.988672022007262e-05
Validation MSE: 1.165289937957883e-05
Validation MAE: 0.0003579804046050854
```

Figure 6: The evaluation's outcome

I will use the trained Random Forest model to make predictions on a fresh dataset or the test dataset. In order to make predictions, we must first load the fresh data into the model. Next, using the characteristics that have been provided, we will use the trained model to forecast the predicted values of the target variable (in this example, the fetal head diameter pixel size). To facilitate administration and future usage, the expected results will be stored as a CSV file or as a DataFrame.

```python
# Load the test data
test_df = pd.read_csv('/content/drive/MyDrive/Data/practice2/test_set_pixel_size.csv')

# Prepare the test data
X_test = test_df.drop(columns=['filename'])

# Predict on the test set
test_preds = rf_model.predict(X_test)

# Optionally, save the predictions to a CSV file
test_df['pixel size(mm)_predicted'] = test_preds
test_df.to_csv('/content/drive/MyDrive/Data/practice2/test_set_pixel_size_with_predictions.csv', index=False)
```

Figure 7: The code to predict result

In this method, the test or new data is included in a DataFrame called test_df, the dataset containing its features is called X_test, and the predicted result is called test_preds, which is produced by using the previously trained model to forecast the test or new data. Now that we have the anticipated outcomes, we can use the matplotlib.pyplot module to visualize them. We are going to make a scatter plot, with the actual values shown on the yaxis and the expected values on the xaxis. This aids in our evaluation of how well the model fits real data and projections. Afterwards, I compute the Mean Absolute Error to assess the model's general effectiveness, regardless of the particular dataset used for training. It calculates the mean of the absolute differences between all data points' actual and anticipated values.

The overall statistic in this instance, the Mean Absolute Error, is independent of the dataset used to train the model. This implies that while the Mean Absolute Error may vary between the training and validation sets, it still accurately represents the average difference between the values that were predicted and those that were observed for each and every data point. The MAE in this instance is 0.00034839596941663463, which shows that there is often very little variation between the expected and actual numbers.

A reduced MAE often indicates that the model is more capable of making predictions, with less variation between predicted and actual values. The fact that the MAE is so low in this case suggests that the model can reliably infer the baby head circumference from ultrasound data.

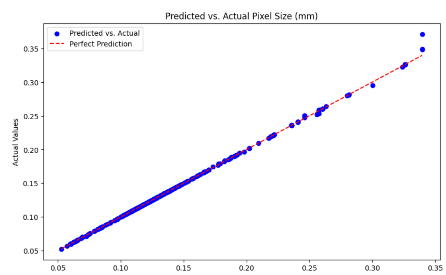

Figure 8: Performance of prediction result

5 Conclusion

Strong performance is shown by the Random Forest model that was trained to estimate fetal head circumference using ultrasound data. The model's capacity to produce accurate predictions on both training and validation datasets is demonstrated by its small mean squared error and mean absolute error. Furthermore, a tight alignment between expected and actual values is shown by the model's low mean absolute error, confirming its accuracy in forecasting fetal head size from ultrasound pictures. All things considered, our results highlight how well the Random Forest technique works in this situation and present encouraging opportunities for precise prenatal screening and fetal biometric evaluation.

References

[1] T. Chudleigh, "The 18+0–20+6 weeks fetal anomaly scan national standards," *Ultrasound*, vol. 18, pp. 92 – 98, 2010. [Online]. Available: https://api.semanticscholar.org/CorpusID:70988072

[2] A. C. Fleischer, "The principles and practice of ultrasonography in obstetrics and gynecology," 1990. [Online]. Available: https://api.semanticscholar.org/CorpusID:68974414

[3] I. Sarris, C. Ioannou, P. F. Chamberlain, E. O. Ohuma, E. O. Ohuma, F. Roseman, L. Hoch, D. G. Altman, and A. T. Papageorghiou, "Intra- and interobserver variability in fetal ultrasound measurements," *Ultrasound in Obstetrics & Gynecology*, vol. 39, 2012. [Online]. Available: https://api.semanticscholar.org/CorpusID:25594744