澳門大學

UNIVERSIDADE DE MACAU

UNIVERSITY OF MACAU

UNIVERSITY OF MACAU

FACULTY OF SCIENCE AND TECHNOLOGY

CISC3023 Course Project Report

Wound area location in animal model images

Students' Number & Name

DC127235, YANG ZHEYU

DC228002, GU SHUNSHUN

Date of Submission: 15/12/2024

Table of Contents

1. Two Types of Machine Learning Models Used

1.1 Random Forest: The Random Forest (RF) model exhibits remarkable robustness and outstanding tolerance to noise and outliers in the data. In processing animal wound image data, factors such as the shooting environment and individual differences among animals introduce inevitable noise. The RF model effectively resists such interference. Through mechanisms like bootstrapping and random feature selection, RF constructs multiple decision trees, preventing overfitting, especially with a limited sample size (only 150 images in this project).

1.2 XGBoost: XGBoost is adept at handling complex nonlinear relationships, crucial for wound area localization in animal model images. Its gradient boosting framework iteratively enhances prediction outcomes by minimizing residuals, uncovering nuanced patterns between image features and wound coordinates. XGBoost's regularization mechanisms (L1 and L2) prevent overfitting by penalizing excessive weights. Additionally, XGBoost's computational efficiency, through parallel computing and cache-aware algorithms, ensures rapid training even with resource-intensive image data. Both RF and XGBoost offer extensive hyperparameter tuning options.

2. Handling the Prediction of 4 Outputs

We adopted the multi-output regression method, treating each output as an interdependent variable within a unified regression framework. During model training, the complex relationships and correlations among these outputs were learned simultaneously.

3. Data Preparation for Better Results

3.1 Image Preprocessing and Normalization:

- The original image size was uniformly adjusted to 32x32 pixels using Bicubic interpolation, significantly reducing computational cost.
- Pixel values were normalized to the range of 0-1, crucial for accelerating model convergence and stabilizing training.

3.2 Data Augmentation: We applied the following strategies, performing corresponding transformations on the original data and labels:

- **Gaussian Blur:** Randomly (50%) applied.

- **Horizontal Flip:** Randomly (50%) flipped.
- **Vertical Flip:** Randomly (50%) flipped.
- **Rotation:** Randomly (50%) rotated.
- **Color Jitter:** Randomly (50%) adjusted brightness, contrast, saturation, and hue.

Given the limited dataset (150 images), we expanded it with augmented data, enhancing data diversity and improving generalization ability, reducing overfitting risks.

3.3 Random Seeds: To ensure reproducibility and avoid random factor interference, we set specific seed values for different models.

4. Parameter Selection for Different Models

4.1 Method: We utilized GridSearchCV for parameter selection, training the model with a custom parameter grid to identify optimal parameters. (The training process was not included in the submitted code due to its time-consuming nature.)

Here is an example for XGBoost model:

```
# Define the parameter grid
param_grid = {
    'max_depth': [3, 6, 9],
    'eta': [0.01, 0.1, 0.3],
    'subsample': [0.6, 0.8, 1.0],
    'colsample_bytree': [0.6, 0.8, 1.0]
}

# Initialize the XGBRegressor
xgb_model = XGBRegressor(objective='reg:squarederror', seed=7)

# Initialize GridSearchCV
grid_search = GridSearchCV(estimator=xgb_model, param_grid=param_grid, scoring='neg_mean_squared_error',

# Fit GridSearchCV
grid_search.fit(X, Y)

# Get the best parameters
best_params = grid_search.best_params_
print(f"Best parameters found: {best_params}")
```

4.2 Best Model Parameters Selection:

The parameters of both models were meticulously determined via experimental verification:

- **Random Forest:**

```
reg = RandomForestRegressor(n_estimators=100,
                            random_state=2,
                            n_jobs=-1)
```

Key parameter choices:

- n_estimators=100: Strikes a balance between model performance and computational expense.
- random_state=2: Guarantees reproducible outcomes
- n_jobs=-1: Utilizes all CPU cores for parallel processing

- **XGBoost:**

```
params = {
    'objective': 'reg:squarederror',
    'max_depth': 6,
    'eta': 0.1,
    'subsample': 0.8,
    'colsample_bytree': 0.8,
    'eval_metric': 'rmse',
    'seed': 7  # Set random seed
}
num_boost_round = 100
```

Key parameter choices:

- objective='reg:squarederror': Selected for regression task
- max_depth=6: Regulate tree complexity and avert overfitting.
- eta=0.1: Provide a moderate learning rate for stable training
- subsample=0.8: 80% of the data was utilized per tree to mitigate overfitting
- colsample_bytree=0.8: Employ 80% of the features per tree to enhance robustness.
- num_boost_round=100: Ensure an adequate number of iterations for model convergence

5. Best Model Selection

We employed two performance metrics:

1. **5.1 Mean Squared Error (MSE):** Measures the average squared difference between predicted and actual values of the four output variables (x, y, x*width,* *y*width). Lower MSE indicates higher prediction accuracy.
2. **5.2 Overlap Ratio:** Assesses the spatial overlap between the predicted wound area (represented by an ellipse) and the actual wound area. Higher overlap ratio implies better fit and accurate wound localization.

6. Performance of the Models and Improvement Strategies

6.1 Random Forest:

- **Training Performance:** MSE = 327.89615, Average Overlap Ratio = 0.9623662678848166
- **Testing Performance:** MSE = 1860.971435546875, Average Overlap Ratio = 0.9299558192384737

6.2 XGBoost:

- **Training Performance:** MSE = 12.06512953726848, Average Overlap Ratio = 0.9923394748890079
- **Testing Performance:** MSE = 1389.17128053552, Average Overlap Ratio = 0.9143395754089013

6.3 Comparation: Both models managed to fulfill the project's main objective, as evidenced by their testing overlap ratios exceeding 0.90, which implies accurate and dependable wound area localization. In terms of training, the XGBoost model exhibited a better Mean Squared Error (MSE) metric (12.07 compared to 327.89). However, it demanded more meticulous parameter adjustment and also had a more significant performance disparity in testing (1389.17 against 1860.97). The Random Forest model, on the other hand, demonstrated more consistent performance between the training and testing phases. Its overlap ratio ranged from 0.962 in training to 0.930 in testing, while the XGBoost model had a more pronounced drop from 0.992 in training to 0.914 in testing. This consistency of the Random Forest model indicates greater reliability and stability when applied in practical scenarios.

6.4 Improvement Strategy: To improve performance, we used data augmentation to expand the training set, enhancing generalization. We also optimized parameters through a parameter grid and set optimal seed values to reduce random factor interference.

7. Difficult Samples and Explanations

We identified the top 10% error samples as difficult. Characteristics of these samples include high brightness and low contrast, indicating overexposure and subtle differences between wound and surrounding tissue.

7.1 Random Forest:

- **Training:** 15 difficult samples (error >= 58.73), Average brightness = 152.56, Average contrast = 53.61
- **Testing:** 4 difficult samples (error >= 135.68), Average brightness = 159.20, Average contrast = 60.86

7.2 XGBoost:

- **Training:** 75 difficult samples (error >= 10.87), Average brightness = 154.36, Average contrast = 60.26
- **Testing:** 4 difficult samples (error >= 108.21), Average brightness = 155.36, Average contrast = 58.88

7.3 visualization

The following visualizations focus on the most challenging cases for each model, showing examples where the prediction error was highest. These cases help identify potential limitations and areas for improvement:

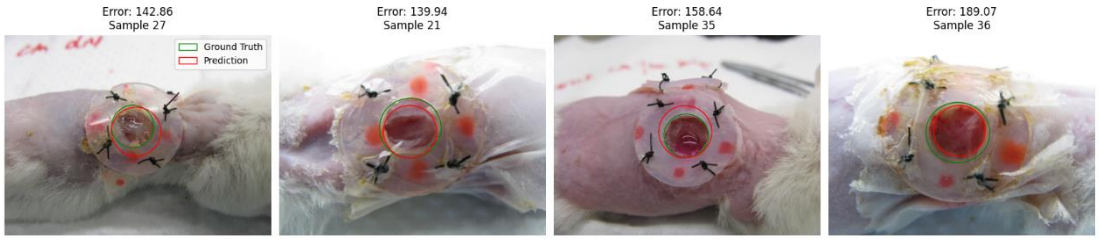


Figure 1 Test Difficult Samples Visualization - Random Forest

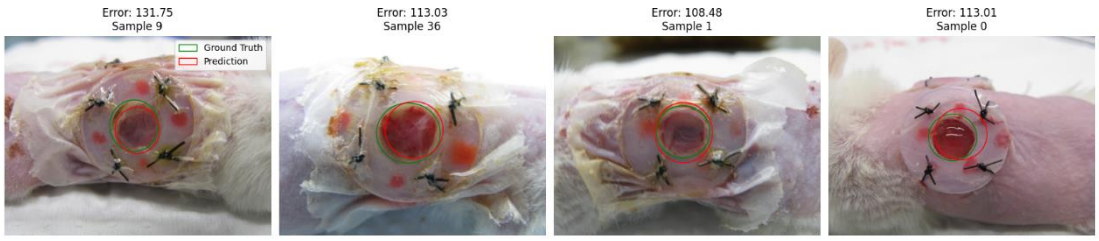


Figure 2 Test Difficult Samples Visualization – XGBoost

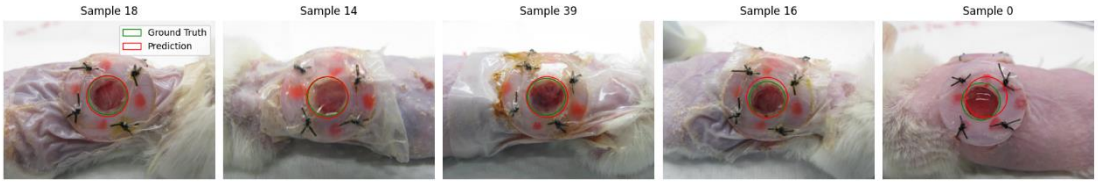
7.4 Explanation: Difficult samples often show translucent, pale skin with significant glare or reflection, high brightness, and low contrast. Moreover, the shapes of the wounds in such samples are irregular.

8. Visualization of Prediction Results

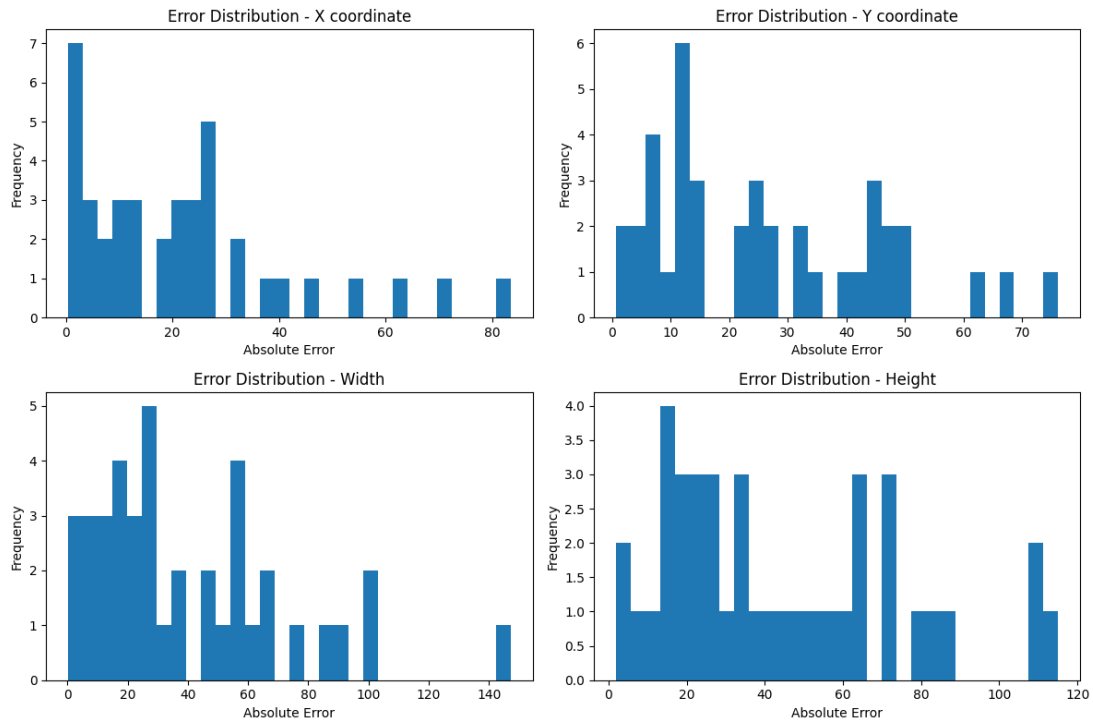
In order to conduct a thorough assessment of the performance of our models, we have created a number of visualization sets that emphasize diverse aspects of the prediction outcomes. The subsequent figures present a comparison between the ground truth, which is indicated in green, and the predicted wound areas, shown in red, as well as the distribution of prediction errors with respect to all four parameters, namely x, y, width, and height.

8.1 Random Forest:

- **Prediction Results:**

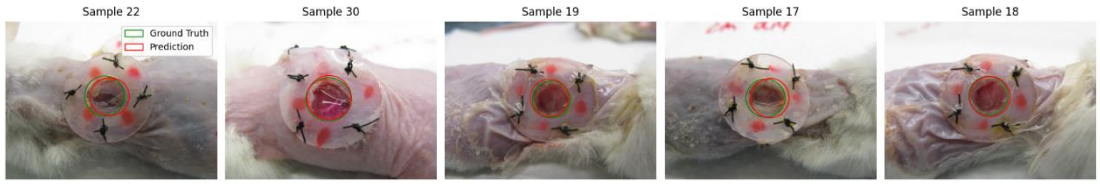


- **Error distribution:**

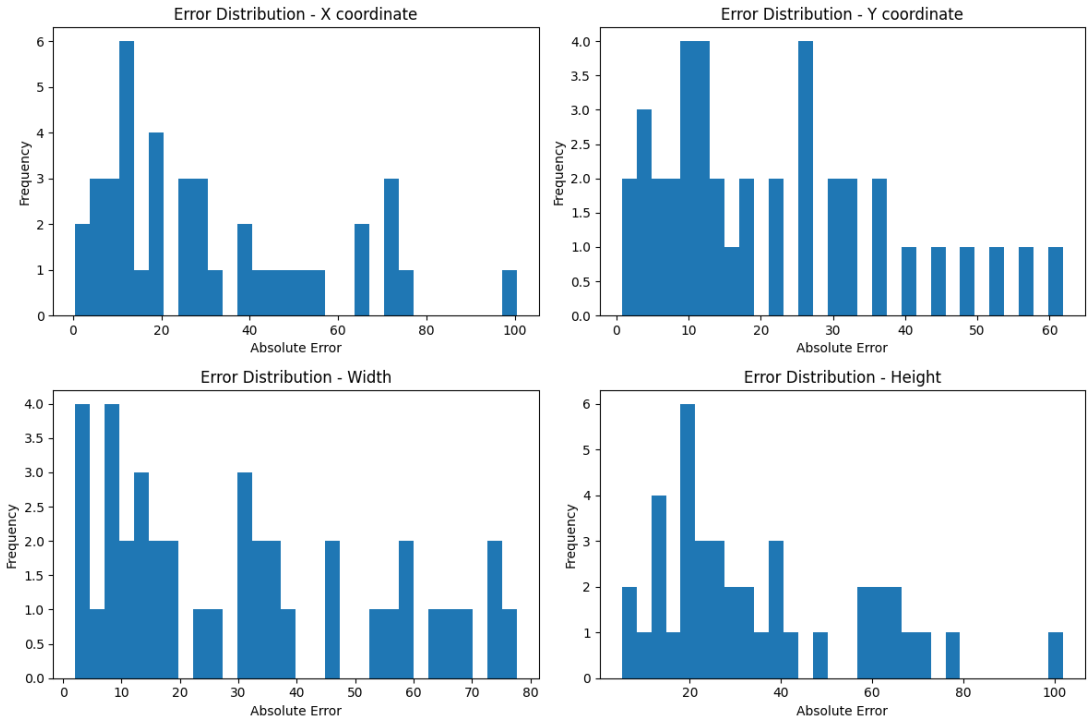


8.2 XGBoost:

- **Prediction Results:**



- **Error Distribution:**



9. Conclusion

The main difficulty of the entire project lies in handling the original data set. The original data has a large number of features (3264x2448 pixels), but the total amount of data is extremely small (150 images). Therefore, extracting key features from the original data and effectively improving the generalization ability of the model are the primary challenges. In this project, we focused on enhancing the generalization ability of the model. Mainly through the use of data augmentation techniques to expand the training data set, increase the diversity of data, and adjust the parameters of the model to improve the learning performance. Future improvements may further include using PCA technology to reduce the dimensionality of the original data to retain key features, or using neural networks to extract key features from the original data, or mapping the original data from the spatial domain to the frequency domain for further training. Certainly, there are more training models worthy of exploration.