International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0130
e ISSN: 2584-2854
Volume: 03
Issue:03 March 2025
Page No: 811-815

Evaluating Machine Learning Models for Gender Detection Using Palm Images: A Comparative Approach

Jahnvi D Rajyaguru[1], *Dr. Falguni Parsana*[2]
[1,2]*Department of Computer Science, Atmiya University, Rajkot, India.*
Email ID: *jhanvi.rajyaguru@atmiyauni.ac.in*[1], *falguni.parsana@atmiyauni.ac.in*[2]

Abstract

This study examines how machine learning can identify a person's gender based on palm images. It compares two widely used models, Random Forest and Gradient Boosting, to evaluate their effectiveness. The dataset, obtained from Kaggle, was selected for its reliability and diversity. The research goes through essential steps like data preparation, feature selection, model optimization, and performance comparison. It also highlights the advantages and limitations of each model, offering valuable insights for those interested in this field.

Keywords: *Gender Detection, Palm Images, Random Forest, Gradient Boosting, Machine Learning, Comparative Analysis Introduction*

1. Introduction

For a long time, people have believed that the lines and shapes on a person's palm can reveal things about their future, health, and personality—this is known as palmistry. [1] Today, palm images are also used in technology, especially in biometric systems, because every person's palm has unique features like fingerprints and palm prints, making them useful for identity verification. [2] Using palm images to determine gender is an interesting and modern approach. [3] Unlike face or voice recognition, it is less intrusive and can still provide valuable information. This method could be helpful in many areas, such as healthcare, security, and forensic investigations. For example, identifying gender from a palm image could support criminal investigations or help doctors offer more personalized treatments. [4] However, it's not always easy to tell the difference between male and female palms since the variations can be subtle. [5] My research aims to solve this problem by developing better techniques to improve accuracy in gender classification using palm Images. [6]

2. Related Work

Researchers have been exploring different ways to determine a person's gender using biometric data, focusing on features like the palm, back of the hand, and fingers. These studies show that these features can help not only in identifying gender but also in estimating age. [7] Today, many applications use computer vision and machine learning to analyse biometric images. These advanced technologies help automatically recognize important details, making gender detection more efficient and accurate. [8] Among various machine learning models, Random Forest and Gradient Boosting have been particularly successful in similar tasks. Random Forest works well with large and complex datasets, making it ideal for biometric analysis. [9] Gradient Boosting is especially useful when dealing with imbalanced data, helping to improve accuracy by balancing precision and recall. Despite these advancements, not much research has been done specifically on using palm images for gender detection. [10] Most biometric studies focus on face or eye recognition, leaving palm-based analysis relatively unexplored. [11] One of the biggest challenges in using palm images is the natural variation in skin texture, lighting conditions, and age-related changes. [12] These factors can affect the accuracy of gender detection models. However, overcoming these challenges could make palm-based analysis a valuable tool for biometric research in the future. [13]

3. Literature Survey

I studied several works related to palm image recognition techniques like palm vein recognition using CNN (Convolutional Neural Newt wok), simplify features extraction and achieving high accuracy. [14] Palm print analysis focuses on tiny

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0130
e ISSN: 2584-2854
Volume: 03
Issue:03 March 2025
Page No: 811-815

religion of interest to extract key features. These studies guide my approach to designing an effective model for gender detection. [15]

4. Dataset & Pre-Processing

The dataset used in this study is a mix of numbers and text, making it rich in features for machine learning. It includes details like whether the hand is shown from the palm side or the back, if there's nail polish, rings, or bracelets, and any unique features like scars or marks. This dataset is designed for machine learning and includes key details like skin color, image names, and gender labels. Palm images were taken in controlled environments to ensure high quality and consistency. The dataset is evenly distributed, containing an equal number of male and female participants with different ages and skin tones. The images are saved in standard formats like JPEG or PNG, all with the same resolution. Before being used, the images go through preprocessing steps such as resizing and adjusting brightness and contrast. To introduce variety, some images are rotated, flipped, or cropped. If one gender has more images than the other, techniques like SMOTE help balance the dataset by generating synthetic images. Finally, the images and their details are converted into numerical data to be effectively processed by machine learning models.

5. Methodology

This research explores how Random Forest and Gradient Boosting can be used to predict gender. To ensure fair and accurate results, the dataset is divided into three parts: training, validation, and testing. Both models are adjusted to improve their accuracy. Their performance is measured using key factors like accuracy, precision, recall, and F1-score. The study also looks at how fast and efficient these models are to determine if they would be useful in real-world applications.

6. Random Forest Algorithm

Random Forest is a powerful technique that helps computers make predictions by using multiple decision trees. Instead of relying on a single tree, it creates many and combines their results to improve accuracy and consistency. This approach is useful for classifying data, such as determining gender, and also for making numerical predictions. By using multiple trees, it reduces errors and enhances overall performance. This method is well-suited for gender prediction because it can handle different types of data and identify the key factors influencing the outcome. Since it builds several decision trees, it prevents overfitting, ensuring the model doesn't become too dependent on any one tree. It also works effectively with large datasets. With its balance of accuracy, efficiency, and simplicity, Random Forest is a strong option for research in this area. (Figure 1)

6.1. How Random Forest Work

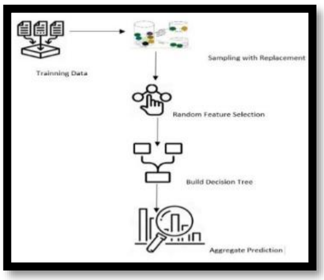

Figure 1 Random Forest Work Flow

6.2. Performance Evaluation Metrics of the Model and Gender Based Model Outcomes Comparison

The graph and classification report clearly show how well the model is performing. The bar chart reveals that the model has an accuracy of approximately 80.37%, meaning it gets most predictions right but still has a 19.63% margin of error. While the overall performance is strong, there is still room for improvement. Based on the classification report, the model performs exceptionally well for Class 0, achieving 83% precision, 88% recall, and an F1-score of 85%. However, for Class 1, its performance is slightly weaker, with 75% precision, 67% recall, and an F1-score of 71%. On average, the model maintains an 80% accuracy across both categories. A bar chart created using matplotlib visually represents the model's gender predictions. It displays the number of "Male" and "Female" predictions, using cyan and pink bars for easy differentiation. Titles and labels are included to improve readability and interpretation. This visual representation helps assess how well the model distributes gender predictions and highlights areas where improvements may be needed. (Figure 2)

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0130

e ISSN: 2584-2854
Volume: 03
Issue:03 March 2025
Page No: 811-815

6.3. Tables

Table 1 Models and Performance Metrics

Steps	Description
1. Load the Data	Read the dataset from the given file into a table (Data Frame) so we can work with it easily.
2. Clean the Data	Remove columns like id and image Name that aren't useful for training. Change text columns (e.g., gender, skin-color, aspectOfHand) into numbers using encoding.
3. Separate Inputs and Target	Take the useful data (features) for training into one table (X). Keep the column you want to predict (gender) in another table (y).
4. Split the Data	Divide the data into two parts: one for training the model (X_train, y_train) and another for testing it (X_test, y_test) with an 80:20 ratio.
5. Train the Model	Create a Random Forest model and teach it to predict gender using the training data.
6. Test the Model	Ask the trained model to predict gender for the test data and see how well it works.
7. Check Results	Calculate how accurate the model is. Also, create a detailed report showing how well it predicts each category (e.g., male, female).

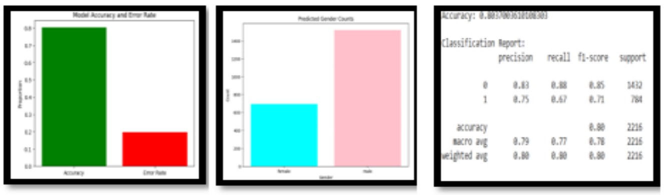

Figure 2 Model Evaluation Output

7. Gradient Boosting Classifier

7.1. Performance Evaluation Metrics of the Model

The model performs exceptionally well, achieving an impressive accuracy of 98.87%. It does a great job at recognizing the majority class (Class 0) with 99% precision and almost perfect recall. However, it struggles with the minority class (Class 1), correctly identifying only 19% of actual cases, despite having perfect precision. This means that while the model is highly accurate overall, it tends to overlook many instances of the less common class. A gender classification bar chart visually represents how the model predicts each gender. The X-axis represents gender categories (Male and Female), while the Y-axis shows the number of predictions. In an ideal scenario, both bars would be of similar height, indicating a balanced dataset and fair predictions. However, if one gender's bar is significantly taller, it suggests the model may be biased toward predicting that gender more often. This imbalance could stem from the dataset itself or the way the model was trained. The chart serves as a useful tool to assess gender prediction patterns, highlighting potential fairness and accuracy concerns. (Figure 3)

7.2. Workflow of Gradient Boosting Classifier

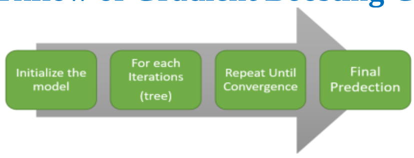

Figure 3 Working Process of Gradient Boosting

Table 2 of Palm Image Classification Models and Performance Metrics

Steps	Description
1. Start with Initial Prediction	Begin by making an initial guess (for regression, it could be the average value)
2. Set Parameters	Decide on settings like how much each model should influence the final result (learning rate) and how many models (trees) to use.
3. Calculate Errors (Residuals)	Find out where the current predictions went wrong by calculating the errors (residuals).
4. Build a New Model	Create a new model (usually a decision tree) that focuses on fixing these errors from the previous model.
5. Get New Model's Predictions	Get the predictions from the newly trained model (tree)
6. Update Predictions	Adjust the current predictions by adding the new model's predictions, making sure it doesn't dominate too much by using a learning rate
7. Repeat	Do this process multiple times (for the set number of trees), each time improving the predictions.
8. Final Prediction	Once all the models (trees) are trained, the final prediction is the result of combining all the models.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0130
e ISSN: 2584-2854
Volume: 03
Issue:03 March 2025
Page No: 811-815

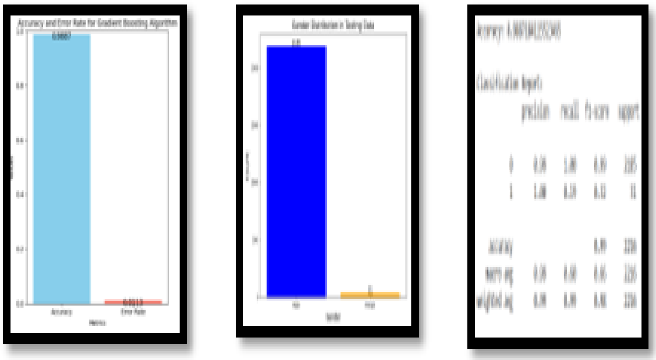

Figure 4 **Model Evaluation Output**

8. Comparative Analysis of Model

Random Forest and Gradient Boosting use different strategies to make predictions. Random Forest creates multiple decision trees separately, with each tree trained on a different part of the data. The final result is determined by averaging (for numbers) or voting (for categories). This makes it a reliable model, but it may not always capture complex patterns unless properly fine-tuned. On the other hand, Gradient Boosting builds trees in a sequence, with each new tree learning from the mistakes of the previous ones. This helps the model improve step by step, making it better at identifying complex patterns and achieving higher accuracy. However, if not carefully adjusted, it can become too focused on the training data and over fit. While Random Forest is great for general tasks, Gradient Boosting is more effective for solving complex problems when properly optimized.

Conclusion

Both Random Forest and Gradient Boosting are powerful machine learning models, but they approach learning in different ways. Random Forest is reliable and works well with various types of data, making it a great choice for general use. However, it might not always capture complex patterns unless fine-tuned carefully. On the other hand, Gradient Boosting continuously learns from its mistakes, making it more accurate in many cases—but it requires more tuning and can over fit if not properly managed. If you're looking for a simple and effective model, Random Forest is a solid choice. But if accuracy is your top priority and you're willing to fine-tune the model, Gradient Boosting offers more precision and adaptability.

Refrences

[1]. Saha, D., De, D., Ghosh, P., Sengupta, S., & Majumdar, T., "An Extensive Analysis of Human Gender Prophecy using Computer Vision," International Research Journal of Engineering and Technology (IRJET), 7(5), 2020, 1223–1229. https://www.irjet.net/archives/V7/i5/IRJET-V7I5412.pdf

[2]. M. Wu and Y. Yuan, "Gender Classification Based on Geometry Features of Palm Image," The Scientific World Journal, vol. 2014, pp. 1-7, 2014. Available: 10.1155/2014/734564

[3]. R. Wadhwa and M. Kaur, "Age and Gender Determination from Finger Prints using RVA and dct Coefficients", Semanticscholar.org, 2020. [Online]. Available: https://www.semanticscholar.org/paper/Age-andGender-Determination-from-Finger-Prints-RVA-WadhwaKaur/2a6094fc7fc7cb8092a50218768ef2e416599198. [Accessed: 23- Jun- 2020]

[4]. G. Amayeh, G. Bebis and M. Nicolescu, "Gender classification from hand shape," 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Anchorage, AK, 2008, pp. 1-7, doi: 10.1109/CVPRW.2008.4563122

[5]. Sudha S Ponnarasi and M Rajaram. Article: Gender Classification System Derived from Fingerprint Minutiae Extraction. IJCA Proceedings on International Conference in Recent trends in Computational Methods, Communication and Controls (ICON3C 2012) ICON3C (2):1-6, April 2012

[6]. Y. Yuan and T. Huang, "A polynomial smooth support vector machine for classification," in Advanced Data Mining and Applications, vol.3584 of Lecture Notes in Computer Science, pp. 157–164,2005

[7]. J. Liang and D. Wu, "A new smooth support vector machine," in Artificial Intelligence and Computational Intelligence, vol. 3801 of Lecture Notes in Artificial Intelligence,

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0130
e ISSN: 2584-2854
Volume: 03
Issue:03 March 2025
Page No: 811-815

pp.392–397, Springer, 2005
[8]. Y. B. Yuan, "Canonical duality solution for alternating support vector machine," Journal of Industrial and Management Optimization, vol.8, no.3, pp.611–621, 2012
[9]. Y. Yuan, W. Fan, and D. Pu, "Spline function smooth support vector machine for classification," Journal of Industrial and Management Optimization, vol.3, no.3, pp.529–542, 2007
[10]. Gnanasivam P1, and Dr. Muttan S2, "Fingerprint Gender Classification using Wavelet Transform and Singular Value Decomposition," IEEE2012
[11]. Afifi M, Abdelhamed A (2017) Afif4: Deep gender classification based on adaboost-based fusion of isolated facial features and foggy faces. arXiv:170604277
[12]. J. E. Tapia, C. A. Perez, and K. W. Bowyer, "Gender Classification From the Same Iris Code Used for Recognition," IEEE TIFS, vol. 11, no. 8, pp. 1760–1770, Aug. 2016.
[13]. R. A. Lazenby, "Identification of Sex from Metacarpals: Effect of Side Asymmetry," J. Forensic Sci., vol. 39, no. 5, pp. 1188–1194, 1994.
[14]. A. K. Jain, K. Nandakumar, and A. Ross, "50 years of biometric research: Accomplishments, challenges, and opportunities," PRL, vol. 79, pp. 80–105, 2016
[15]. A. Dantcheva, F. Bremond, and P. Bilinski, "Show me your face and i will tell you your height, weight and body mass index," in Proceedings - International Conference on Pattern Recognition, 2018, vol. 2018-Augus, pp. 3555– 3560.