Fine-Grained Palmar Morphology Detection: A Comparative Analysis of YOLOv5 and YOLOv8 Architectures for Palmistry Knowledge Representation

Ni Kadek Dwi Rusjayanthi[1]*, I Ketut Gede Darma Putra[2], Ngoc Le[3], Made Sudarma[4], Anak Agung Kompiang Oka Sudana[5]

[1]Department of Information Technology, Udayana University, Bali, Indonesia. Email: dwi.rusjayanthi@unud.ac.id

[2]Department of Information Technology, Udayana University, Bali, Indonesia. Email: ikgdarmaputra@unud.ac.id

[3]Department of Communication Network, Electric Power University, Hanoi, Vietnam. Email: ngocla2@fe.edu.vn

[4]Department of Electrical Engineering, Udayana University, Bali, Indonesia. Email: msudarma@unud.ac.id

[5]Department of Information Technology, Udayana University, Bali, Indonesia. Email: agungokas@unud.ac.id

*Corresponding Author: Ni Kadek Dwi Rusjayanthi, dwi.rusjayanthi@unud.ac.id

Abstract

Palmar morphology offers a unique and complex taxonomy of features; however, its analysis in traditional practices is still hindered by human subjectivity and inconsistent interpretation. While biometrics have advanced, the automated detection of 12 diverse and fine-grained palmar elements remains a significant challenge due to high spatial complexity. This study aims to establish a standardized digital framework for identifying 12 heterogeneous palmar morphologies, bridging the gap between traditional taxonomy and modern computer vision while supporting palmistry knowledge representation through structured and computationally reproducible morphological analysis. We propose a comparative benchmark between anchor-based (YOLOv5) and anchor-free (YOLOv8) architectures. Optimization strategies were implemented through transfer learning using MS COCO pre-trained weights to overcome small-scale dataset constraints and enhance multi-scale feature extraction on subtle biological markers. Experimental results demonstrate that YOLOv8-L is the superior model, achieving an mAP@0.5 of 0.9244, significantly outperforming YOLOv5 in localizing fine-grained features. The integration of C2f structures and the anchor-free paradigm proved effective in handling intricate spatial relationships, achieving a high inference speed of 97.2 FPS. This research presents an automated and standardized framework for biological feature extraction, enabling objective digital documentation and consistent interpretation of palmar morphological features. The findings also indicate that anchor-free models are more robust for detecting complex biological morphologies.

Keywords: *Palmar morphology; YOLOv8, Anchor-free, Fine-grained object detection, Comparative analysis*

I. INTRODUCTION

The human hand is a fundamental organ for environmental interaction, manifesting significant anatomical and functional complexity (Hanger & Higgins, 2025). Its capacity for fine motor tasks, ranging from power grips to precision dexterity, is essential for both daily activities and high-level cognitive functions. Furthermore, specific epidermal configurations on the palm serve as unique biometric identifiers for individual recognition (Cummins & Midlo, 1943), while also functioning as biological markers capable of indicating genetic anomalies through surface morphological analysis (J. Li et al., 2022).

Evidenced by these diverse biophysical utilities, hand pattern analysis has become a cornerstone in computational intelligence, driving innovations in biometric authentication, human-computer interaction (HCI), and fine-grained object recognition frameworks.Beyond its established scientific value, the hand is a central object in the traditional practice of palmistry. In the context of this research, palmistry is positioned as a source of a unique, heterogeneous, and systematic feature taxonomy. This taxonomy provides a comprehensive categorization of features, including hand-type geometric configurations, the morphology of palmar mounts, line frequency density, spatial relationships between digits, and the intrinsic characterization of palmar creases (Goldberg & Bergen, 2016). However, feature detection within this taxonomy often yields inconsistent interpretations, as the identification process remains heavily reliant on human subjectivity. Dependence on a reader's intuition leads to instability in hand pattern identification; therefore, a transformation from manual methods toward a computational approach is required to automate palmar morphology detection objectively, ensuring accuracy and analytical uniformity. In addition, the transition toward computational analysis facilitates palmistry knowledge representation by encoding traditional morphological observations into a standardized digital framework that can be consistently reproduced across different observers and applications.

Recent research in hand pattern recognition reveals a continued reliance on hand-crafted feature extraction methods alongside the emerging adoption of deep learning architectures. Prior studies deployed algorithms such as Random Forests (RF) and Linear Support Vector Machines (SVM) to predict individual characteristics based on the four Major Lines (Patil, 2025). To facilitate this, these methodologies utilize multi-stage pipelines incorporating Gaussian smoothing and Canny edge detection to capture boundaries, from which geometric descriptors such as contour arc length and line depth are extracted. However, these approaches are constrained by their dependency on deterministic feature engineering and an analytical scope that is limited to predefined linear and contour structures. While Deep Neural Network (DNN) approaches, such as Convolutional Neural Networks (CNN) and U-Net, have been introduced to reduce manual intervention, a research gap persists regarding multi-class palmar feature detection (Ariyanto et al., 2018; Van et al., 2020). These frameworks primarily focus on either image-level global classification or pixel-level segmentation masks, which do not inherently provide the direct, multi-class bounding-box localization required to analyze heterogeneous palmar structures alongside macro-geometric finger proportions.

Initial efforts to address these limitations were conducted using the YOLOv5 model. However, this model faces significant challenges in detecting micro-objects, such as the Lower Mars Mount, or features with indistinct visual boundaries (Rusjayanthi et al., 2025). This limitation stems from YOLOv5's reliance on rigid anchor-based mechanisms, which struggle to adapt to the biological geometry of objects with high aspect ratio variability. To bridge this gap, this study presents an innovation in palmar morphology-based feature detection using the YOLOv8 architecture. Unlike YOLOv5, YOLOv8 implements an anchor-free approach, which is theoretically more adaptive in capturing irregular biological features. Furthermore, the integration of the C2f structure in YOLOv8 enables superior feature extraction (G. Wang et al., 2023), particularly in identifying fine-grained palmar morphologies with low inter-class variance.

This implementation serves as a strategic approach to overcome small-scale dataset constraints through a transfer learning mechanism utilizing the Microsoft Common Objects in Context (MS COCO) dataset. By adopting pre-trained weights, the model effectively performs domain adaptation from general objects (Reddy & Basha, 2025; Reis et al., 2023) to complex palmar textures and geometries. This research conducts an in-depth comparative analysis between YOLOv5 and YOLOv8 (Small and Large variants) to evaluate the effectiveness of this architectural transition. The primary contributions of this study involve providing a comparative benchmark between anchor-based and anchor-free architectures specifically optimized for fine-grained biological feature detection and establishing a standardized digital framework for the identification of 12 heterogeneous palmar morphologies. This framework bridges the gap between traditional taxonomy and modern computer vision by enabling the automated extraction of diverse and complex features and supporting palmistry knowledge representation through machine-interpretable morphological classifications that preserve their semantic relationships.

II. Method

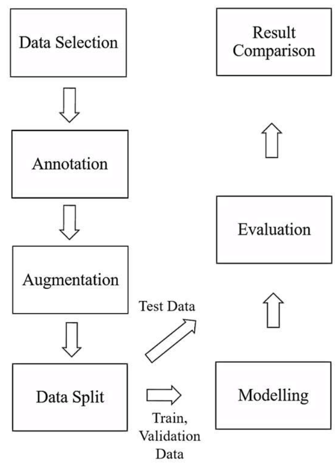


Figure 1: Research Methodology Workflow for Palmar Element Detection

As illustrated in Figure 1, the workflow for detecting palmar morphological elements in this study was developed through a series of systematic stages, beginning with data selection. In this phase, the dataset was constructed by focusing on intricate hand features based on a palmar morphological taxonomy, which encompasses variations in geometry, linear structures, and palmar mounts. The subsequent stage involves image annotation, where class labels were precisely assigned to the selected raw images by determining pixel coordinates. This structured annotation process also serves as the basis for palmistry knowledge representation by converting qualitative morphological descriptions into quantitatively defined computational entities. Each object was localized using bounding boxes corresponding to the highly detailed hand features across the 12 palmar-based element classes.

Following annotation, both geometric and photometric data augmentation techniques were employed to expand the annotated sample size. This procedure is crucial for preventing overfitting and ensuring the model remains robust in recognizing features under heterogeneous color intensities and lighting conditions. To prepare the data for the subsequent network architecture, a stochastic split scheme was performed, allocating 70% of the data for the training set, 15% for the validation set to monitor performance during the learning process, and the remaining 15% as an independent test set to evaluate the model's generalization capability.

The training and validation processes were conducted to develop a model capable of extracting fine-grained features across multiple scales. In this study, a transfer learning strategy was implemented by utilizing pre-trained weights from the MS COCO dataset to accelerate model convergence on a smaller-scale biological dataset. Finally, a comparative analysis was performed between the YOLOv5 and YOLOv8 architectures. This evaluation focused on benchmarking accuracy and efficiency metrics to determine the effectiveness of the anchor-free approach in

YOLOv8 versus the anchor-based mechanism in YOLOv5 in handling complex and irregular palmar geometric variability.

The dataset for this study was comprehensively structured based on a palmar element taxonomy, encompassing hand types, the morphology of the lower Mars mount, pinky finger proportions, and palmar line characteristics (line frequency and the Saturn line). The resulting taxonomy provides a consistent semantic structure for palmistry knowledge representation, enabling each morphological characteristic to be encoded as an identifiable computational category. Primary data were selected from three widely recognized public databases: the 11k Hands Dataset (Afifi, 2019), the Birjand University Mobile Palmprint Database (BMPD) (Mahdie Izadpanah, 2019), and the COEP Palm Print Database (Autonomous Institute of Government of Maharashtra, 2024).

The dataset is categorized into twelve classes representing heterogeneous features, which include hand types, the lower Mars mount, pinky finger length, line frequency, and the Saturn line. The hand type category comprises Water, Fire, Earth, and Air classes, while the lower Mars mount is categorized into Low and High classes. The pinky finger is differentiated into Short and Long classes based on its distal position. Furthermore, line frequency is classified as High, Medium, and Low depending on the density of minor lines, and the Saturn line is identified by the presence of a specific vertical line in the central palmar region. The total dataset, including augmented samples generated through saturation adjustment, vertical flipping, and horizontal flipping, consists of 1,000 images. In accordance with the defined distribution, the dataset was partitioned into training, validation, and testing sets, with the detailed data distribution per class presented in Figure 2.

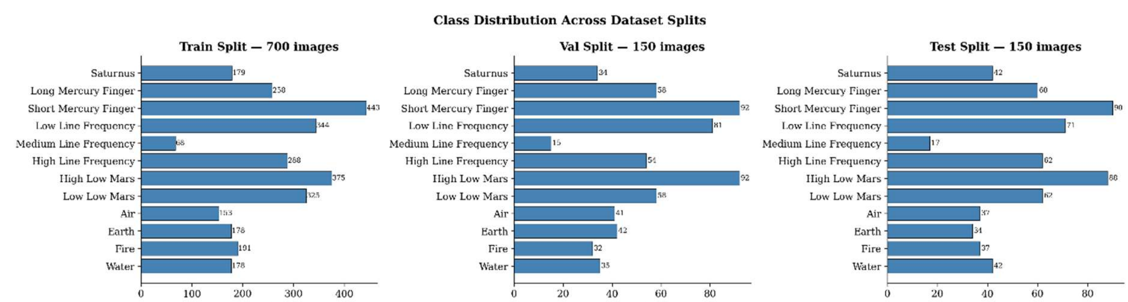
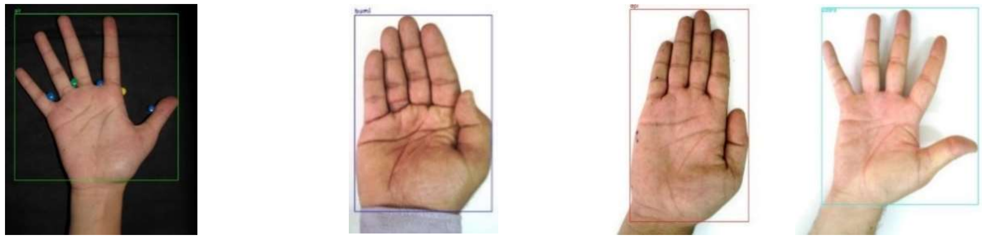


Figure 2: Class Distribution

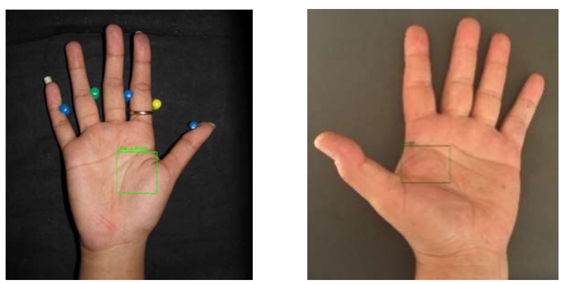
The annotation criteria for each element and its corresponding classes were defined based on specific morphological characteristics (Goldberg & Bergen, 2016). For hand type, annotations encompassed the entire palm and finger regions as illustrated in Figure 3.



(a) (b) (c) (d)

Figure. 3 Annotation results for hand type: (a) Water, (b) Earth, (c) Fire, and (d) Air

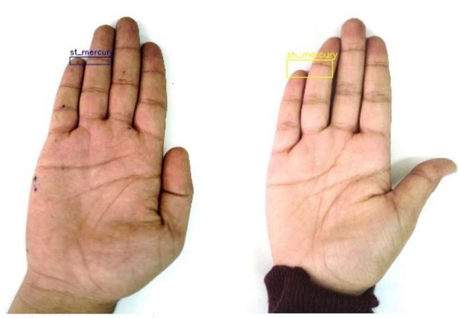
The Water class is characterized by a rectangular palm and finger structure, while the Earth class tends toward a square shape. In contrast, the Fire class features rectangular palms combined with square fingers, and the Air class typically exhibits a square palm with rectangular fingers. Furthermore, the lower Mars mount feature was localized on the palm between the base of the index finger and the thumb's basal joint, as shown in Figure 4.



(a) (b)

Figure 4: Annotation Results for the Lower Mars mount: (a) Low, and (b) High

The High class was annotated for prominent or expansive mount areas, whereas the Low class was applied to regions that appeared flatter or where fine lines reduced the perceived height of the mount. Annotations for the pinky finger focused on the relationship between the little finger's tip and the ring finger’s distal knuckle, as depicted in Figure 5.



(a) (b)

Figure 5. Annotation Results for the Pinky Finger: (a) Long, and (b) Short

The Long class was assigned when the fingertip exceeded the knuckle, while the Short class was used when it did not. Regarding line frequency, this element was annotated based on the density of minor or fine lines on the palm, excluding the principal lines, as illustrated in Figure 6.

(a) (b) (c)

Figure 6: Annotation Results for Line Frequency: (a) High, (b) Medium, and (c) Low

The High class represents a dense concentration of numerous, less distinct lines, while the Medium class was used for minor lines that were numerous but more clearly visible, and the Low class was assigned where fine lines were sparse. Finally, the Saturn line was localized as a vertical line situated between the middle finger and the base of the palm, as illustrated in Figure 7.

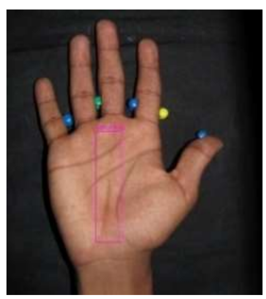


Figure 7: Annotation Results for the Saturn line

The selection of YOLOv5 and YOLOv8 for this comparative study is based on the fundamental architectural differences in how both methods process spatial information. YOLOv5 employs an anchor-based mechanism that relies heavily on predefined bounding box priors (Azevedo & Santos, 2024; Khanam & Hussain, 2024). Theoretically, this dependency poses a significant constraint in recognizing palmar morphology; unlike rigid industrial objects, hand features exhibit organic variability in scale, texture, topography, and aspect ratios that often exceed the capacity of fixed-ratio anchors. Consequently, the transition to the anchor-free approach in YOLOv8 becomes crucial. By eliminating anchor constraints, the model gains greater flexibility to localize features based on the object's center of mass and direct distance regression (Casas et al., 2024; Shukhratov et al., 2024; Yaseen, 2024). This enables more precise detection of 12 heterogeneous palmar morphologies, ranging from thin linear lines to amorphous dermal texture areas.

The architectural superiority of YOLOv8 over YOLOv5 is supported by four key technical evolutionary components. First, in the backbone network, YOLOv8 introduces the C2f module (Cross-Stage Partial with 2 Convolutions), replacing the C3 module found in YOLOv5. Through a gradient shunt mechanism, the C2f module bifurcates input feature maps to enrich gradient flow while preserving contextual information (G. Wang et al., 2023), which is instrumental in maintaining fine-grained textural details. Second, YOLOv8 implements a decoupled head structure that separates the classification and regression processes into distinct network branches with independent weight parameters (G. Wang et al., 2023; J. Wang et al., 2023). This separation ensures that fine-grained spatial details are not distorted by the abstract textural features required for classification—a contrast to the coupled detection head in YOLOv5, which often triggers task conflicts.

Third, regarding localization optimization, YOLOv8 integrates Distribution Focal Loss (DFL) alongside Complete Intersection over Union (CIoU), calculated using Eq. 1 and Eq. 2. Unlike the rigid coordinate estimation in YOLOv5, DFL models boundary locations as a continuous probability distribution, effectively handling positional uncertainty in micro-features with soft edges (Guo et al., 2025; Zhou et al., 2025). The DFL is formulated as shown in Eq. (1):

$$L_{DFL}(P_i, P_{i+1}) = -\big((y_{i+1} - y)\,log(P_i) + (y - y_i)\,log(P_{i+1})\big) \quad (1)$$

In the DFL equation, y represents the ground truth label or target value, while P_i and P_(i+1) represent the predicted probabilities of the discrete bins flanking the target value y. These probabilities are generated through a Softmax activation function to ensure a valid probability distribution. Furthermore, y_i and y_(i+1) denote the values of the nearest discrete bins surrounding the ground truth label. This mechanism forces the model to concentrate the probability distribution on the bins closest to the actual value, which is particularly beneficial for localizing palmar features with indistinct visual boundaries.

The localization accuracy is further optimized using the CIoU loss function. Unlike the standard IoU, CIoU accounts for three critical geometric factors, namely overlap area, central point distance, and aspect ratio, as defined in Eq. (2):

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (2)$$

In the CIoU loss calculation, IoU is the intersection area divided by the union area of the predicted bounding box and the ground truth. The termϱ^2 (b,b^gt)represents the squared Euclidean distance between the central points of the predicted box and the ground truth box, while c is the diagonal length of the smallest enclosing box. Additionally, α serves as a positive trade-off parameter used to measure aspect ratio consistency, and v is the parameter that quantifies the consistency of the aspect ratio between the two boxes. The calculation of α dan v is detailed in Eq. (3) and (4):

$$\alpha = \frac{v}{(1-I \quad)+v} \tag{3}$$

$$v = \frac{4}{\pi^2}\left(arctan\frac{w^{gt}}{h^{gt}} - arctan\frac{w}{h}\right)^2 \tag{4}$$

In the aspect ratio parameter calculation, w^gt and h^gt represent the width and height of the ground truth bounding box, while w and h denote the width and height of the predicted bounding box. Finally, in terms of sample management, YOLOv8 adopts a Task-Aligned Assigner and utilizes Binary Cross-Entropy (BCE) for classification loss (Wu et al., 2024). The Task-Aligned Assigner ensures a more accurate and task-specific assignment of object proposals by dynamically calculating an alignment score t for each anchor, as defined in Eq. (5) (X. Li et al., 2024):

$$t = s^{\alpha} \cdot IoU^{\beta} \tag{5}$$

Within the alignment formula, t is the calculated alignment score, s is the predicted classification score, and IoU is the Intersection over Union value. The variables α and β are weighting hyper-parameters that control the relative influence of the classification score and the box regression accuracy, respectively. Furthermore, the classification task is optimized using BCE Loss, as shown in Eq. (6):

$$L_{BCE}(y,\hat{y}) = -\frac{1}{N}\sum_{i=1}^{N}[y_i \, log(\hat{y_i}) + (1-y_i)\, log(1-\hat{y_i})] \tag{6}$$

In the classification loss equation, y_i is the ground truth label for the i-th sample, which is assigned a value of 1 if the object is present and 0 otherwise. The term (y_i)̂represents the predicted probability for the class after passing through the Sigmoid activation function for the i-th sample, and N denotes the total number of samples included in the calculation. The removal of the Objectness Loss branch present in YOLOv5 theoretically reduces false negatives in indistinct features, as the model prioritizes dynamic alignment between classification and localization (Gao & Li, 2025; Huang et al., 2025).

The model training and evaluation experiments were conducted using a cloud-based infrastructure via Kaggle Notebooks, accelerated by dual NVIDIA Tesla T4 GPUs with 16GB GDDR6 per unit. The training process was configured for 200 epochs utilizing the Adam with Decoupled Weight Decay (AdamW) optimization algorithm and an input resolution of 320 x 320 pixels. This specific resolution was selected to maintain an optimal balance between computational efficiency and feature fidelity across the twelve detected palmar morphologies, ensuring that fine-grained details are preserved during the feature extraction process.

To mitigate the risk of overfitting on the limited dataset, a transfer learning mechanism using MS COCO pre-trained weights was implemented, complemented by mosaic augmentation techniques. These techniques were crucial for enhancing the model's robustness against organic variability in palmar textures. The entire technical implementation was developed using the Python programming language and the Ultralytics framework to ensure a consistent and rigorous evaluation environment between the YOLOv5 and YOLOv8 architectures.

Performance evaluation was measured using standard multi-class object detection metrics, including mean Average Precision at a 0.5 Intersection over Union (mAP@0.5) and the average mAP across the 0.5 to 0.95 range (mAP@0.5:0.95). Furthermore, Precision, Recall, and F1-Score were utilized to provide a comprehensive overview of the model's capability to balance relevant detection with the minimization of false positives. These metrics specifically address the visual challenges inherent in micro-object classes, such as the Lower Mars mount and specific digit landmarks, providing an objective benchmark for architectural comparison. Beyond evaluating detection performance, the experimental protocol also examines the feasibility of employing deep learning models as reliable computational instruments for palmistry knowledge representation.

III. Results and Discussions

Baseline Performance and Local Feature Characteristic Analysis (YOLOv5)

The performance analysis initiates by evaluating the YOLOv5 architecture as the baseline model to dissect how anchor-based detection mechanisms respond to the morphological variability of the palm. As a model reliant on static reference boxes or pre-defined anchors, YOLOv5 provides crucial insights into the constraints of mapping biometric features linearly through rigid coordinate regression. Experimental data indicates that the YOLOv5-L variant demonstrates superiority with a mean Average Precision at 0.5 (mAP@0.5) of 0.8896, outperforming the YOLOv5-S variant, which is constrained at 0.8507, as illustrated in Figure 8. However, an in-depth class-level analysis reveals significant performance disparities rooted in the inability of anchor-based architectures to handle objects with high organic variability.

Regarding the efficacy in linear textural features, the data demonstrates that YOLOv5, particularly the Small (S) variant, exhibits high efficacy in extracting low-level features. This is evidenced by an Average Precision (AP) of 0.9900 for the Low Line Frequency class. Technically, this feature is dominated by distinct lines with high visual contrast against the dermis, allowing the convolution kernels in the C3 blocks of the CSP-Darknet53 architecture to recognize repetitive patterns without requiring extreme network depth. This success indicates that for biometric features with low spatial complexity and high-frequency edge information, models with limited parameters already reach an accuracy saturation point. Consequently, increasing parameters in the Large variant yielded no significant marginal improvement, as these features do not require a deep hierarchy of abstract features for precise recognition.

Fundamental limitations of the anchor-based mechanism emerge when confronted with objects defined by global spatial proportions, such as Hand Types. For the Fire class, YOLOv5-S recorded its lowest score of 0.7397. Morphologically, the Fire hand type is determined by specific ratios between short finger lengths and wide palms, and this low accuracy stems from the inability of static anchors to generalize across objects with non-rigid organic shape variability. Although the Large (L) variant increased accuracy to 0.8573, this disparity underscores that detecting heterogeneous hand morphologies requires a broader receptive field and larger parameter capacity to simultaneously process the relational features between fingers and the palm.

Furthermore, a performance anomaly was observed in the detection of palmar mounts, specifically the High and Low Lower Mars classes, which possess abstract boundaries. While YOLOv5-L achieved maximum precision of 0.9900, YOLOv5-S suffered a performance degradation to 0.9093. From a computer vision perspective, palmar mounts are objects with low-contrast boundaries relative to the surrounding skin texture, demanding sharp gradient sensitivity to distinguish topographical elevations from background noise. This limitation is associated with the C3 block architecture in YOLOv5, which utilizes a more linear gradient flow scheme compared to the C2f modules in its successor. The relative failure of the S variant suggests constraints in receptive field capacity, where shallower layers fail to map functional features that require a broader spatial context, causing essential biometric features to be misclassified as irrelevant background dermis. Operationally, although YOLOv5-L provides higher detection accuracy, it recorded the lowest inference speed at 64.6 frames per second (FPS), as illustrated in Figure 8.

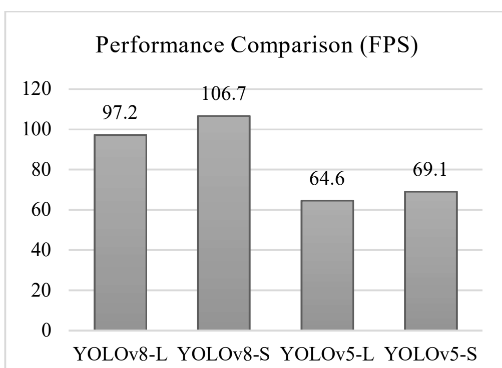


Figure 8: Inference Speed (Frames per Second) Across All Tested Model Configurations

The Small variant led with 69.1 FPS; however, this marginal speed difference of approximately 4.5 FPS does not justify the drastic accuracy drop in critical classes such as Hand Types and Lower Mars Mounts. These findings suggest that reliance on anchor mechanisms creates an inefficient computational load relative to accuracy gains, as the inability of anchors to adapt to fine-grained objects forces the network to perform redundant calculations on misaligned bounding boxes.

The systematic failure of the anchor-based mechanism to adapt to organic variability provides a strong justification for evaluating more flexible architectures. The performance disparity between the Small and Large variants of YOLOv5 indicates that increasing parameter capacity alone is insufficient to overcome spatial ambiguity in subtle biometric features. Therefore, the next section will analyze how the transformation toward the anchor-free paradigm and the integration of C2f modules in YOLOv8 reduce computational redundancy while enhancing localization precision for features that previously challenged the baseline model.

Overall Performance

The comprehensive presentation of experimental results aims to provide an objective comparative overview between the YOLOv5 baseline models and the proposed YOLOv8 models. All model variants were evaluated using the same palmar morphology dataset to ensure the validity of the comparison. The multi-metric evaluation results, encompassing precision, recall, and mean Average Precision, are summarized in Figure 9.

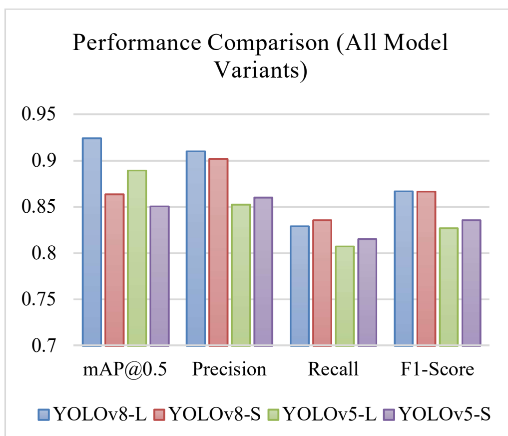


Figure 9: Comparative Analysis of Detection Metrics (mAP@0.5, mAP@0.5:0.95, Precision, and Recall)

Based on the data presented in Figure 9, it is generally observed that the YOLOv8 architecture provides a consistent performance enhancement across all metrics compared to YOLOv5. The YOLOv8-L model recorded the highest values for mAP@0.5 at 0.9244 and mAP@0.5:0.95 at 0.6628. When compared to the Large variant of the previous generation (YOLOv5-L), there is a significant increase in mAP of 3.48%, indicating sharper and more accurate localization. This enhancement suggests that the transition from anchor-based to anchor-free paradigms allows the model to better adapt to the non-rigid geometries of biological palm features. From the perspective of palmistry knowledge representation, improved localization accuracy enables more faithful computational encoding of subtle morphological characteristics that traditionally depend on manual observation.

To further understand the model's effectiveness on each specific feature, an analysis of the Average Precision (AP) per class was conducted. The detection results for the twelve palmar feature classes are presented in Figure 10.

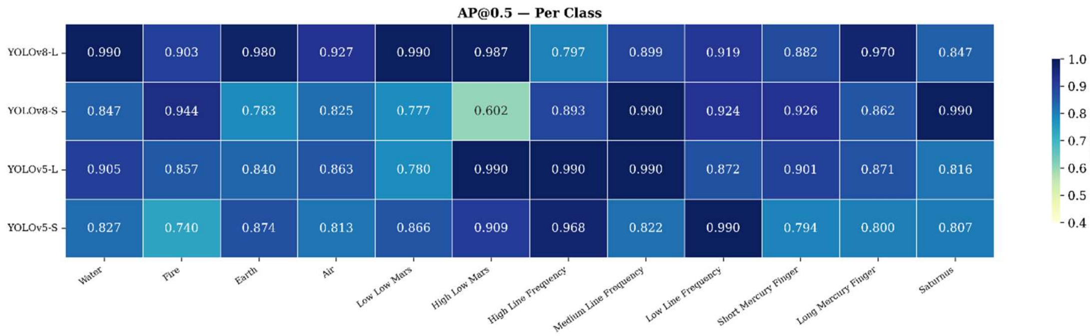


Figure 10: Class-wise AP@0.5 Comparison

The data in Figure 10 demonstrates that YOLOv8- L is highly superior in detecting element features and Lower Mars areas. The most significant improvement is observed in the Low Lower Mars class, with a notable increase of 21.01% compared to the baseline. However, for the High Line Frequency class, the YOLOv5-L model actually recorded a higher score. This phenomenon indicates that certain feature characteristics exist where the anchor-based architecture remains competitive, specifically in patterns with high-frequency edges and consistent aspect ratios. Despite this, YOLOv8-L remains the most stable model in terms of overall mean Average Precision.

This indicates that the proposed model has a very low rate of false positives, which is critical for biometric and morphological analysis. Although the Recall for YOLOv8-L at 0.8293 is slightly below that of the Small variant, the integration of these metrics resulted in an F1-Score of 0.8670. This balance reinforces YOLOv8-L's position as the most optimal model for identifying palmar morphological features in this study, successfully mitigating the visual challenges inherent in fine-grained object classes.

Comparative Analysis of YOLOv8: Anchor-Free Transformation in Complex Objects

The transition from YOLOv5 to YOLOv8 marks a fundamental paradigm shift in palmar feature detection, moving from a reliance on static reference boxes to an anchor-free approach centered on object points. Based on the experimental results, YOLOv8-L achieved peak performance with a mean Average Precision at 0.5 (mAP@0.5) of 0.9244, representing a significant 3.48% improvement over the highest YOLOv5 variant. This analysis reveals how architectural evolution directly correlates with increased accuracy in biometric objects characterized by high spatial complexity.

One of the most critical findings is the ability of YOLOv8 to overcome localization barriers in features with extreme aspect ratios through coordinate regression flexibility. In the previous generation, features such as the Mercury Finger often encountered detection failures due to the mismatch between rigid anchor boxes and the long, slender morphology of the fingers. In contrast, YOLOv8-L successfully achieved an Average Precision (AP) of 0.9702 in this class. By directly predicting the object's center point, YOLOv8 eliminates the geometric bias inherent in fixed-anchors, enabling coordinate regression that is more adaptive to variations in respondents' finger poses and lengths.

Furthermore, YOLOv8 also demonstrates superior robustness in handling biometric objects with abstract boundaries or soft edges, such as palmar mounts, thereby resolving spatial ambiguity. For instance, the Low Lower Mars class was detected with a perfect AP score of 0.9900. This success is driven not only by the integration of C2f modules that enrich gradient flow but also by the implementation of Distribution Focal Loss (DFL). This loss function allows the model to represent coordinate boundaries as a continuous probability distribution, which is highly effective in distinguishing the topographical elevations of the palm from the surrounding dermal texture. This architectural advancement successfully addresses the challenge that frequently caused false negatives in the single-value regression mechanism of the baseline model.

Finally, a notable performance surge was observed in the classification of global morphologies, particularly the Fire and Water hand types, which achieved AP scores of 0.9029 and 0.9900, respectively. This achievement proves that the YOLOv8 architecture is superior in performing feature disentanglement between finger and palm components through classification-localization synergy. The implementation of the Task-Aligned Assigner plays a pivotal role by aligning classification scores with localization accuracy during the training process. In the context of fine-grained detection, this mechanism ensures that class identity remains consistent and also minimizes predictive ambiguity across wide and heterogeneous palmar regions. The improved localization capability strengthens the representation of heterogeneous palmistry features by reducing ambiguity in the digital interpretation of morphological structures.

Model Scalability and Operational Efficiency

Beyond accuracy, computational efficiency and parameter scalability are crucial factors in standardizing real-time palmar analysis. This subsection dissects the relationship between model capacity or scaling and its ability to perform feature disentanglement on low-contrast objects. Cross-scale testing indicates that increasing parameters in the Large (L) variant yields a non-linear impact on the accuracy of specific classes. While features with low visual complexity, such as Low Line Frequency, reach accuracy saturation in the Small (S) variant, significant performance gaps emerge in abstract biometric objects that require a deeper hierarchy of features.

A prominent case analysis can be observed in the High Lower Mars class, where a performance degradation occurred in YOLOv8-S with a score of 0.6018, which subsequently surged to 0.9874 in the YOLOv8-L variant. This phenomenon suggests that the Small variant suffers from limitations in the network depth required to construct a feature hierarchy capable of distinguishing palmar topographical elevations from skin texture noise. In contrast, the Large variant, with its increased number of bottleneck blocks, possesses the representational capacity necessary to precisely extract palmar mount areas that lack distinct edges. This confirms that scalability in network depth is essential for mapping functional features with high organic variability.

Inference speed data obtained from testing demonstrates the efficiency of the YOLOv8 architecture. YOLOv8-S recorded a peak speed of 106.7 frames per second (FPS), while the YOLOv8-L variant maintained a speed of 97.2 FPS. This performance is impressively faster than the smallest variant of the previous generation, YOLOv5-S, which was limited to 69.1 FPS. These achievements confirm that optimizations in the C2f blocks and the Decoupled Head successfully reduced computational redundancy while maintaining high throughput.

This speed advantage is fundamentally driven by the implementation of the anchor-free paradigm. By eliminating thousands of predefined reference boxes or anchors, the model reduces the computational overhead during the post-processing stage, specifically minimizing latency in the Non-Maximum Suppression (NMS) process. This provides the flexibility for implementation in high-resolution video-based biometric systems without the risk of data throughput bottlenecks. The reduction in complexity allows the Large variant to achieve superior accuracy while still performing well above the 30 FPS threshold required for real-time applications.

Evaluation of the mAP@0.5:0.95 metric, often referred to as the COCO mAP, further reveals that the Large variant provides significantly superior localization stability, with YOLOv8-L scoring 0.6628 compared to 0.6222 for the Small variant. In the domain of palmar biometrics, the precision of bounding boxes at strict Intersection over Union (IoU) thresholds is a determinant of reliability for advanced morphological analysis. A shift of just a few pixels can alter the interpretation of micro-features, making high localization stability a critical requirement for automated identification systems.

The empirical superiority of YOLOv8-L in this metric validates that the synergy between Distribution Focal Loss (DFL) and the anchor-free paradigm produces more precise object boundary estimations compared to the deterministic coordinate regression in YOLOv5. Although a speed margin of approximately 9 FPS exists between the S and L variants, this increase in spatial precision is a justifiable trade-off. The minimal reduction in operational speed is outweighed by comprehensive detection reliability, solidifying the Large variant as the optimal configuration for standardizing automated hand feature identification that demands high accuracy without sacrificing real-time performance. The combination of localization precision and computational efficiency demonstrates that the proposed architecture is suitable for scalable palmistry knowledge representation in future digital documentation systems.

Limitations and Future Directions

Despite achieving high localization precision, this research acknowledges constraints that provide a trajectory for future study. The current dataset, while optimized through transfer learning as established in the introduction, remains limited in demographic diversity and environmental variance. Future work should implement Generative Adversarial Networks (GANs) to synthesize high-fidelity samples, addressing data scarcity and enhancing robustness against rare topographical features.

Furthermore, while this study successfully automates a traditional feature taxonomy into a digital framework, the transition from detection to functional application, such as behavioral biometrics or medical genetic marker screening, requires deeper interdisciplinary validation. Integrating Human-in-the-Loop (HITL) protocols with domain specialists and employing Explainable AI (XAI) will be crucial to verify that the model's feature prioritization aligns with anatomical truth. Such interdisciplinary validation would further improve the semantic reliability of palmistry knowledge representation by ensuring consistency between computational outputs and expert morphological interpretation. Lastly, migrating these high-parameter architectures to edge-computing platforms remains a vital step for the real-time, field-based identification of complex palmar morphologies.

IV. Conclusion

The comparative analysis conducted in this study demonstrates that YOLOv8 significantly outperforms YOLOv5 in the detection of 12 distinct palmar morphological features. Beyond improving detection performance, the proposed framework contributes to palmistry knowledge representation by providing a structured computational mechanism for encoding and standardizing heterogeneous palmar morphological information. Achieving a mean Average Precision (mAP@0.5) of 0.9244, the YOLOv8 architecture proves that the transition from anchor-based to anchor-free paradigms provides superior flexibility in localizing organic hand features characterized by extreme variability in scale and aspect ratios. This performance leap is technically supported by the integration of the Task-Aligned Assigner and Decoupled Head, which dynamically synchronize classification scores with localization precision through a dedicated alignment metric. Furthermore, the implementation of the C2f module successfully preserves high-frequency spatial information, ensuring that fine-grained dermal textures are maintained throughout the feature extraction process. While these findings validate the model's efficacy, this study acknowledges limitations concerning the controlled environment and constrained dataset scale, which may not fully reflect the stochastic noise of 'in-the-wild' biometric scenarios. To mitigate data scarcity, future research should implement Generative Adversarial Networks (GANs) to synthesize high-fidelity samples, thereby improving robustness against rare topographical variations. Furthermore, as this work serves the digitalization of intangible cultural heritage, future validation should involve a Human-in-the-Loop (HITL) approach with cultural and behavioral specialists. Integrating Explainable AI (XAI) will be essential to ensure that feature prioritization aligns with anatomical landmarks, while optimizing the architecture for edge-computing remains a critical step for real-time field deployment. Overall, this study establishes that fine-grained object detection can function not only as an accurate recognition tool but also as a computational foundation for palmistry knowledge representation, supporting objective, reproducible, and standardized digital interpretation of complex palmar morphology.

V. References

Afifi, M. (2019). 11K Hands: Gender Recognition and Biometric Identification Using a Large Dataset of Hand Images. *Multimedia Tools and Applications*, *78*(15). https://doi.org/10.1007/s11042-019-7424-8

Ariyanto, A., Djamal, E. C., & Ilyas, R. (2018). Personality Identification of Palmprint Using Convolutional Neural Networks. *2018 International Symposium on Advanced Intelligent Informatics (SAIN)*, 90–95. https://doi.org/10.1109/SAIN.2018.8673353

Autonomous Institute of Government of Maharashtra. (2024). *COEP Palm Print Database*. Retrieved February 2, 2024, from http://www.coep.org.in/resources/coeppalmprintdatabase

Azevedo, P., & Santos, V. (2024). Comparative Analysis of Multiple YOLO-based Target Detectors and Trackers for ADAS in Edge Devices. *Robotics and Autonomous Systems*, *171*, 104558. https://doi.org/10.1016/j.robot.2023.104558

Casas, E., Ramos, L., Bendek, E., & Rivas-Echeverria, F. (2024). YOLOv5 vs. YOLOv8: Performance Benchmarking in Wildfire and Smoke Detection Scenarios. *Journal of Image and Graphics*, *12*(2), 127–136. https://doi.org/10.18178/joig.12.2.127-136

Cummins, H., & Midlo, C. (1943). *Finger Prints, Palms and Soles: An Introduction to Dermatoglyphics.* The Blakiston Company.

Gao, P., & Li, Z. (2025). YOLO-S3DT: A Small Target Detection Model for UAV Images Based on YOLOv8. *Computers, Materials & Continua*, *82*(3), 4555–4572. https://doi.org/10.32604/cmc.2025.060873

Goldberg, E., & Bergen, D. (2016). *The Art and Science of Hand Reading Classical Methods for Self-Discovery through Palmistry*. Destiny Books.

Guo, T., Peng, Y., Han, L., Jia, T., Zhang, C., Liu, W., Yang, Q., Huang, H., & Hu, D. (2025). MAF-YOLOv8: A Lightweight, High-Precision Deep Learning Model Applied to Real-Time Detection and Counting of Betula Luminifera Seedling Leaves. *Industrial Crops and Products*, *235*, 121716. https://doi.org/10.1016/j.indcrop.2025.121716

Hanger, J., & Higgins, J. (2025). Anatomy of the hand. *Surgery (Oxford)*, *43*(2), 115–119. https://doi.org/10.1016/j.mpsur.2024.12.006

Huang, Y., Zhang, Y., Yan, J., Zhang, C., Gao, P., & Lv, X. (2025). Automated Detection of Cotton Aphids (Aphis Gossypii Glover) Using a Bi-Enhanced Attention Mechanism and Domain-Adaptive Transfer Learning Strategy. *Artificial Intelligence Review*, *59*(2), 76. https://doi.org/10.1007/s10462-025-11457-7

Khanam, R., & Hussain, M. (2024). *What is YOLOv5: A Deep Look Into the Internal Features of the Popular Object Detector*. http://arxiv.org/abs/2407.20892

Li, J., Glover, J. D., Zhang, H., Peng, M., Tan, J., Mallick, C. B., Hou, D., Yang, Y., Wu, S., Liu, Y., Peng, Q., Zheng, S. C., Crosse, E. I., Medvinsky, A., Anderson, R. A., Brown, H., Yuan, Z., Zhou, S., Xu, Y., … Wang, S. (2022). Limb Development Genes Underlie Variation in Human Fingerprint Patterns. *Cell*, *185*(1), 95-112.e18. https://doi.org/10.1016/j.cell.2021.12.008

Li, X., Zhang, Z., & Zhao, P. (2024). *Improvement of YOLOv8 Detection Algorithm for Worker-Related Objectives in Construction Scenarios*. https://doi.org/10.3233/ATDE240141

Mahdie Izadpanah. (2019). *Birjand University Mobile Palmprint Database(BMPD)*. https://www.kaggle.com/datasets/mahdieizadpanah/birjand-university-mobile-palmprint-databasebmpd

Patil, S. (2025). *Palmistry-Informed Feature Extraction and Analysis Using Machine Learning*. Cornell University.

Reddy, K. G., & Basha, S. S. (2025). Real Time Object Identification: A Study on COCO Dataset. *Proceedings of the International Conference on Advanced Materials, Manufacturing and Sustainable Development (ICAMMSD 2024)*, 860–870. https://doi.org/10.2991/978-94-6463-662-8_70

Reis, D., Kupec, J., Hong, J., & Daoudi, A. (2023). *Real-Time Flying Object Detection with*

YOLOv8. http://arxiv.org/abs/2305.09972

Rusjayanthi, N. K. D., Putra, I. K. G. D., Sudarma, M., & Sudana, A. A. K. O. (2025). Hand Objects Classification for Personality Recognition Using YOLOV5. *2025 International Conference on Smart-Green Technology in Electrical and Information Systems (ICSGTEIS)*, 289–294. https://doi.org/10.1109/ICSGTEIS68532.2025.11284389

Shukhratov, I., Pimenov, A., Stepanov, A., Mikhailova, N., Baldycheva, A., & Somov, A. (2024). Optical Detection of Plastic Waste through Computer Vision. *Intelligent Systems with Applications*, *22*, 200341. https://doi.org/10.1016/j.iswa.2024.200341

Van, T. P., Nguyen, S. T., Doan, L. B., Tran, N. N., & Thanh, T. M. (2020). Efficient Palm-Line Segmentation with U-Net Context Fusion Module. *2020 International Conference on Advanced Computing and Applications (ACOMP)*, (October), 23–28. https://doi.org/10.1109/ACOMP50827.2020.00011

Wang, G., Chen, Y., An, P., Hong, H., Hu, J., & Huang, T. (2023). UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios. *Sensors*, *23*(16). https://doi.org/10.3390/s23167190

Wang, J., Yang, P., Liu, Y., Shang, D., Hui, X., Song, J., & Chen, X. (2023). Research on Improved YOLOv5 for Low-Light Environment Object Detection. *Electronics*, *12*(14), 3089. https://doi.org/10.3390/electronics12143089

Wu, Z., Wang, X., Jia, M., Liu, M., Sun, C., Wu, C., & Wang, J. (2024). Dense Object Detection Methods in RAW UAV Imagery Based on YOLOv8. *Scientific Reports*, *14*(1), 18019. https://doi.org/10.1038/s41598-024-69106-y

Yaseen, M. (2024). *What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector*. Cornell University

Zhou, H., Zou, H., & Hu, G. (2025). An Efficient and Lightweight Algorithm for Detecting Surface Defects of Steel Based on SCCI-YOLO. *Scientific Reports*, *15*(1), 36276. https://doi.org/10.1038/s41598-025-20154-y