

Deep Learning Based Species Classification of Crabs Using Efficient Net and Mobilenetv2

Viveka¹, Meena Kowshalya²

¹Department of Computer Science and Engineering, Government College of Technology, Coimbatore

²Department of Computer Science and Engineering, Government College of Technology, Coimbatore

Abstract-Crab species identification plays a crucial role in fisheries management, aquaculture, and biodiversity conservation. Traditional identification methods rely on manual observation of morphological features, which are time-consuming, subjective, and require expert knowledge. This project proposes an automated deep learning-based system for accurate crab species classification from images. The proposed model utilizes a combination of MobileNetV2 and EfficientNet architectures to extract both lightweight and fine-grained visual features from crab images. To enhance feature discrimination, a hybrid loss function combining ArcFace Loss and Center Loss is employed. This approach ensures that features belonging to the same species are closely clustered, while those of different species are clearly separated in the embedding space. Each image is represented as a 128-dimensional feature vector, and classification is performed using cosine similarity, enabling efficient and scalable prediction. The model is trained on a diverse dataset of 358 crab images with variations in pose, lighting, and background, along with data augmentation techniques to improve generalization. Experimental results demonstrate that the hybrid loss model achieves a high classification accuracy of 96% with a corresponding confidence score of approximately 96%, outperforming individual loss functions. The developed system provides a fast, reliable, and automated solution for crab species identification, making it suitable for real-world applications in fisheries, aquaculture, and ecological monitoring.

Keywords: Crab Species Classification, Deep Learning, MobileNetV2, EfficientNet, ArcFace Loss, Center Loss, Cosine Similarity, Aquaculture.

1. INTRODUCTION

Crabs are an important part of aquatic ecosystems and play a major role in fisheries, aquaculture, and coastal biodiversity. In countries such as India, crab farming and fisheries contribute significantly to the economy and food resources. Accurate identification of crab species is essential for sustainable fisheries management, biodiversity conservation, ecological

monitoring, and scientific research. Traditional species identification methods mainly depend on manual observation of morphological characteristics such as carapace shape, texture, claw structure, and color patterns. However, these methods are time-consuming, labor-intensive, and highly dependent on expert taxonomists. In addition, visually similar crab species and environmental variations such as lighting, background, and image quality make manual identification difficult and inconsistent. Recent advancements in deep learning and computer vision have provided effective solutions for automated image-based species classification. Convolutional Neural Networks (CNNs) are widely used for extracting discriminative visual features and achieving high accuracy in image recognition tasks. In this research, an automated crab species classification system is proposed using hybrid deep learning techniques. The proposed framework combines MobileNetV2 and EfficientNet architectures to capture both lightweight and fine-grained image features efficiently. To further improve feature discrimination, a hybrid loss function integrating ArcFace Loss and Center Loss is utilized. ArcFace Loss enhances inter-class separability, while Center Loss minimizes intra-class variations, enabling the model to learn more robust feature embeddings. The extracted features are represented as compact 128-dimensional embeddings, and classification is performed using cosine similarity-based matching for efficient and scalable prediction. The proposed model is trained and evaluated on a dataset containing crab images with variations in pose, illumination, and background conditions. Data augmentation and preprocessing techniques are applied to improve model generalization and robustness. Experimental results demonstrate that the hybrid loss-based model achieves high classification accuracy and confidence, outperforming traditional individual loss approaches. The developed system provides a fast, reliable, and automated solution for crab species identification, which can support real-world applications in fisheries management, aquaculture monitoring, and biodiversity

conservation. The major contributions of this research work are summarized as follows:

- Developed an automated deep learning system for crab species identification using image data.
- Combined MobileNetV2 and EfficientNet for effective feature extraction.
- Implemented a hybrid loss function (ArcFace Loss + Center Loss) to improve feature discrimination and classification accuracy.
- Generated compact 128-dimensional feature embeddings with cosine similarity-based classification.
- Achieved approximately 96% classification accuracy and confidence score.
- Designed a scalable and reliable system for fisheries, aquaculture, and biodiversity monitoring applications.

2. MATERIALS AND METHODS

2.1 Feature Extraction and Recognition of Chinese Mitten Crab Carapace Based on Improved MobileNetV2

This study proposes a lightweight deep learning model based on MobileNetV2 for recognizing Chinese mitten crab carapace features. The model integrates a coordinate attention mechanism to focus on important morphological features such as grooves and textures of the shell. Additionally, ArcFace loss is used to improve feature discrimination by increasing inter-class separation and reducing intra-class variation. The system generates 128-dimensional feature embeddings and performs similarity based verification. The model achieves a high accuracy of 98.56%, while maintaining low computational complexity and fast processing speed.

2.2 Identification of “Bathed” Chinese Mitten Crabs Using Carapace Morphology

This study focuses on identifying “bathed” Chinese mitten crabs—crabs that are artificially treated to mimic high-quality species—using geometric morphological analysis of the carapace. The method analyzes structural features such as shell shape, curvature, and spatial proportions using mathematical and statistical models. By extracting geometric descriptors, the system attempts to differentiate natural crabs from modified ones. The study demonstrates that carapace morphology

plays a crucial role in crab identification and quality assessment. However, the approach relies heavily on manual or semi-automated feature extraction techniques, making it sensitive to noise, lighting variations, and image quality. It also lacks the ability to learn complex patterns automatically, limiting its performance in large-scale and real-world scenarios.

2.3 Summary and Research Gap

Deep learning techniques such as MobileNetV2 and EfficientNet have shown promising performance in image classification tasks even when image conditions such as lighting, background, and orientation are not perfect. Feature extraction and hybrid loss functions help in improving classification accuracy by learning discriminative crab features more effectively. Although many studies focus on crab species classification, there are still several limitations in existing research, such as:

1. Existing methods struggle to accurately classify visually similar crab species.
2. Traditional image processing approaches are affected by variations in lighting, background, and image quality.
3. Earlier deep learning models do not effectively capture important crab features such as carapace texture and spatial details.
4. Some advanced models require high computational resources, making them less suitable for real-time and lightweight applications.

This creates the need for an automated, accurate, and lightweight crab species classification system. In this work, the proposed study addresses these gaps by integrating improved MobileNetV2 and EfficientNet architectures with a Hybrid Loss Function (ArcFace + Center Loss) using crab image datasets collected from Kaggle for efficient and reliable crab species identification.

3. PROPOSED METHODOLOGY

The proposed system is a deep learning-based approach for automatic crab species classification from images. First, the input crab images are preprocessed by resizing, normalization, and data augmentation to improve quality and make the model robust to variations such as lighting, background, and orientation. Then, feature extraction is performed using two powerful CNN architectures, MobileNetV2 and EfficientNet, where MobileNetV2 extracts basic and structural features like edges and shapes with low computational cost, while EfficientNet captures more detailed and fine-grained texture patterns on the crab carapace. The extracted features are then converted into a compact 128-dimensional embedding vector that represents the unique characteristics of each crab species, ensuring that similar images have closer representations in feature space. To improve classification performance, a hybrid loss function combining ArcFace Loss and Center Loss is used, where ArcFace enhances inter-class separation by increasing angular distance between different species, and Center Loss reduces intra-class variation by pulling same-class samples closer together. Finally, classification is performed using cosine similarity, where the feature vector of a test image is compared with stored embeddings, and the species with the highest similarity is selected as the output. This combined approach ensures efficient feature learning, high accuracy, and strong generalization, making the system suitable for real-time applications in fisheries, aquaculture, and biodiversity conservation. This is depicted in Fig 1.

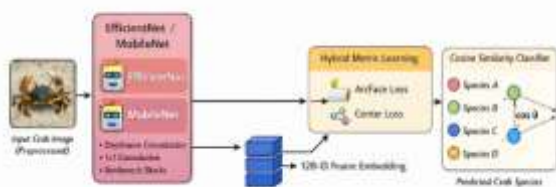


Fig-1: Proposed Methodology

3.1 Architecture of EfficientNet and MobileNetV2

The proposed model takes an RGB image of a crab carapace as input with dimensions $H \times W \times 3$, containing raw pixel information such as color, brightness, and texture, without any semantic understanding at the initial stage. The image first passes through an initial convolution layer (3×3 kernel with

stride 2), which extracts low-level visual features such as edges, corners, color gradients, and simple textures. These features are then processed through deep feature extraction blocks using both MobileNetV2 and EfficientNet architectures. In MobileNetV2, inverted residual bottleneck blocks are used, consisting of expansion, depthwise convolution, and projection layers; the expansion layer increases channel depth to capture richer feature representations, the depthwise convolution extracts spatial patterns like shell ridges, grooves, and fine surface textures independently for each channel, and the projection layer reduces channels while retaining only discriminative shape and texture information. Stride-1 bottleneck blocks preserve spatial resolution to capture fine texture continuity and small surface variations, while stride-2 blocks downsample feature maps to learn broader structural patterns such as overall shell shape, carapace contour, and relative part arrangement. In EfficientNet, MBConv blocks further enhance feature learning through depthwise convolutions with larger kernels (3×3 or 5×5) that capture both local and global spatial context, while the squeeze-and-excitation (SE) module learns channel-wise importance by emphasizing biologically relevant features like texture density and structural irregularities and suppressing background noise. Multiple MBConv blocks gradually build high-level semantic representations such as species-specific carapace morphology and unique texture signatures. The extracted features are then passed through a 1×1 convolution layer in both networks to combine spatial and channel information and convert them into compact high-dimensional descriptors with improved discriminability. Next, global average pooling (GAP) is applied to convert feature maps into a single global feature vector while removing spatial location bias, representing overall texture distribution, shape consistency, and shell structure. This is followed by a fully connected layer that maps the features into species-specific embeddings by integrating all learned visual cues. Finally, a softmax output layer generates class

probability scores, producing the final crab species classification result.

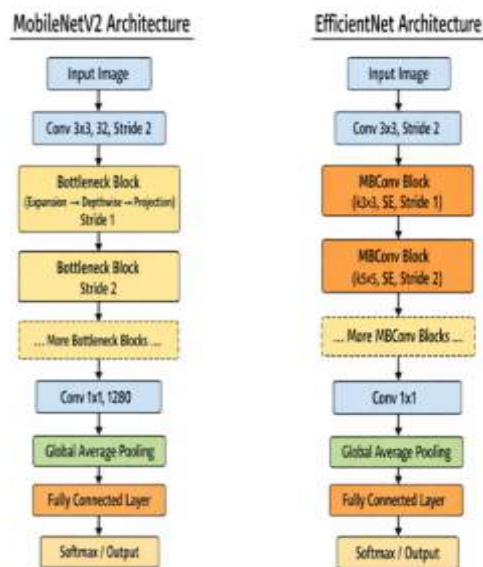


Fig-2: Architecture of EfficientNet and MobileNetV2

3.2 Data Augmentation

Data augmentation is applied to improve the generalization ability of the model and reduce overfitting. Several transformation techniques are used on the training images to simulate real-world variations in crab appearance. Random flip is used to handle different orientations of crab images, while random rotation (up to 20%) helps the model learn from different viewing angles. Random zoom (up to 20%) is applied to manage variations in object scale, and random contrast adjustment (up to 20%) is used to handle different lighting conditions. These augmentation techniques help the model become more robust and reliable when dealing with real-time and diverse image conditions, improving overall classification performance. This process is illustrated in Fig 3.

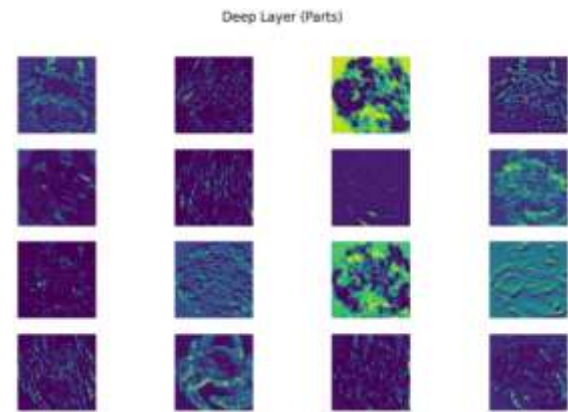


Fig-3: Data Augmentation

3.3 Feature Extraction

MobileNetV2 and EfficientNetB0 are used as pretrained models for feature extraction in the proposed system. The parameters of both models are frozen so that the learned weights are not updated during training, ensuring efficient and stable feature extraction. The features obtained from both models are then passed through an attention block to highlight the most important regions of the crab images. After this, the features are concatenated to form a combined representation that captures both lightweight structural details and fine-grained texture information. The combined features are then passed through a dense layer followed by batch normalization and dropout, which help improve learning stability, reduce overfitting, and enhance model performance. Finally, a 128-dimensional normalized embedding is generated, which serves as a compact representation of each crab image and is used for accurate classification based on similarity measures. Fig 4,5 and 6 illustrates the features extracted at each stage of the network, showing how the representation becomes more detailed from early to deeper layers.

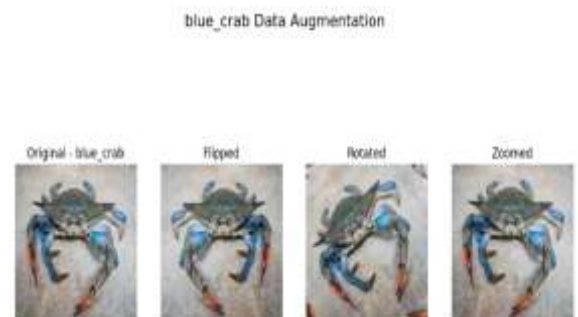


Fig-4: Feature Extraction in Early Layer

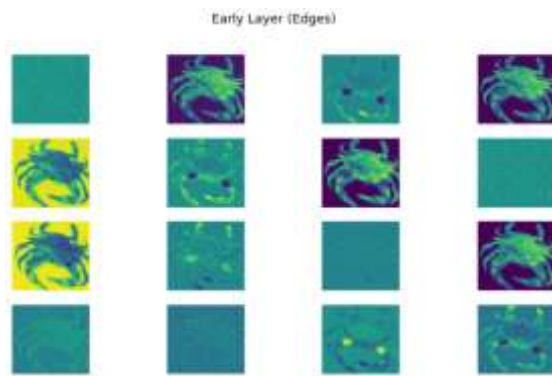


Fig-5: Feature Extraction in Middle layer

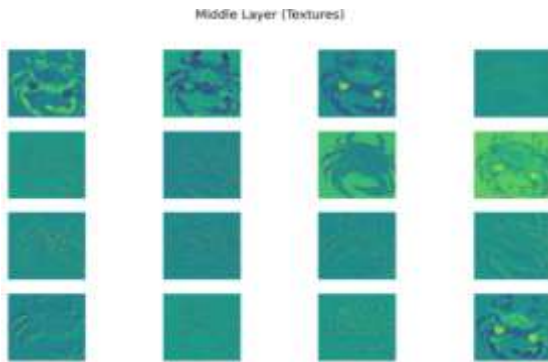


Fig-6: Feature Extraction in Deep layer

4. EVALUATION METRICS

ACCURACY

To evaluate the performance of the proposed model, accuracy is used as a quantitative performance metric. Accuracy measures how correctly the model classifies crab species based on its predictions. It is defined using the following equation:

Where:

$$\text{ACCURACY} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

- True Positive (TP): Number of crab images correctly identified as the same species
- True Negative (TN): Number of crab images correctly identified as different species
- False Positive (FP): Number of crab images incorrectly identified as the same species

- False Negative (FN): Number of crab images incorrectly identified as different species

Accuracy provides an overall measure of the model's performance by considering both correct and incorrect predictions. A higher accuracy indicates better classification performance.

4.1 Quantitative Results

The performance of different loss functions used in the proposed system was evaluated using classification accuracy as the primary metric.

Table -1: Comparison of Loss Models

Model/ Loss Type	Accuracy (%)
ArcFace Loss	93.00%
Center Loss	91.00%
HybridLoss (Combined)	96.00%

The results show that each loss function contributes differently to the model's ability to learn discriminative features. ArcFace Loss achieved an accuracy of 93.00%, indicating strong inter-class separation by improving angular margin between different crab species. Center Loss achieved an accuracy of 91.00%, which helps in reducing intra-class variation by pulling similar class samples closer together in the feature space. However, when both loss functions were combined as a Hybrid Loss (ArcFace + Center Loss), the model achieved the highest accuracy of 96.00%, demonstrating improved feature discrimination, better clustering of embeddings, and enhanced overall classification performance. These results clearly indicate that the hybrid loss function outperforms individual loss functions by effectively balancing inter-class separation and intra-class compactness, leading to more accurate and reliable crab species classification.

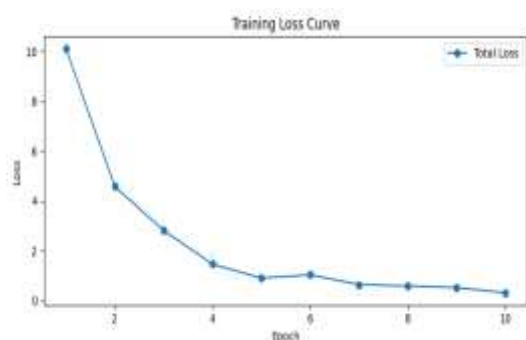


Fig-7: Training Loss Curve

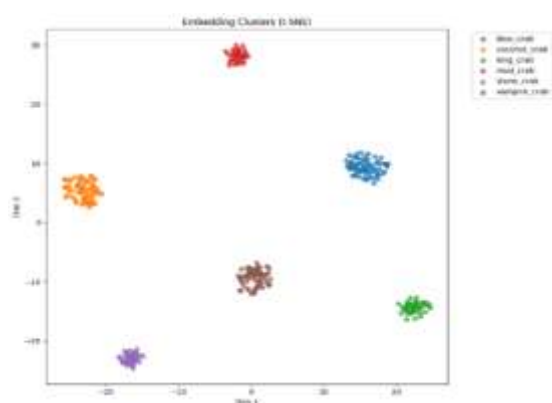


Fig-8: t-SNE Visualization

The training loss curve demonstrates that the proposed model learns effectively throughout the training process. Initially, the loss value is high at approximately 10 during the first epoch, indicating large prediction errors at the beginning of training. As the epochs progress, the loss decreases rapidly, showing that the model is successfully extracting meaningful features and improving its classification capability. A slight fluctuation is observed around the sixth epoch, but the overall trend remains consistently decreasing. Towards the final epochs, the loss stabilizes at a very low value close to 0.3, indicating proper convergence and reduced prediction error. This continuous reduction in loss confirms the stability, effectiveness, and good learning performance of the proposed deep learning model. This is illustrated in Fig 7.

This t-SNE plot visualizes high-dimensional embeddings of five crab species reduced to two dimensions, and the results show strong, well-separated clustering- meaning the model's embeddings have learned meaningfully distinct representations for each species. Each cluster is tight and compact with minimal internal scatter, indicating high intra-class consistency, while the clusters are spatially distant from one another, suggesting low inter-class confusion. Notably, no two

clusters overlap, which implies that the embedding space cleanly separates all five species (blue crab, coconut crab, king crab, mud crab, stone crab, and vampire crab) without ambiguity — a strong signal that whatever model or feature extractor produced these embeddings is highly effective at distinguishing between these crab types. The Visualization is depicted in Fig 8.

5.CONCLUSION

The proposed system successfully demonstrates an automated approach for crab species classification using deep learning techniques. By combining MobileNetV2 and EfficientNet, the model effectively extracts both lightweight and fine grained features, improving overall classification performance. The integration of a hybrid loss function (ArcFace + Center Loss) enhances feature discrimination by ensuring better separation between different species and tighter clustering within the same species. The use of 128-dimensional feature embeddings along with cosine similarity enables efficient and scalable classification, even for visually similar crab species. The model achieves high accuracy of around 96%, indicating strong performance and reliability. Additionally, the system shows good generalization by handling variations in lighting, background, and image conditions. Overall, the proposed model provides a fast, accurate, and scalable solution for crab species identification. It can be effectively applied in real-world scenarios such as fisheries management, aquaculture, and biodiversity conservation, reducing dependency on manual identification and expert knowledge.

ACKNOWLEDGEMENT

Authors thank Department of Computer Science and Engineering, Government College of Technology, Coimbatore for the support.

REFERENCES

- [1] Feature Extraction and Recognition of Chinese Mitten Crab Carapace Based on Improved MobileNetV2-2024
- [2] Song, X.; Huang, F.; Liang, F.; Hu, J.; Lu, B.; Cai, C. Identification of Geographical Origin and Culture Water Body of Chinese Mitten Handed Crab *Eriocheir sinensis* Based on Mineral Element Fingerprints. Chin. J. Fish-2024

- [3]Zhou, L.; Gao, J.; Yang, Y.; Nie, Z.; Liu, K.; Xu, G. Genetic Diversity and Population Structure Analysis of Chinese Mitten Crab (*Eriocheir sinensis*) in the Yangtze and Liaohe Rivers. *Fishes* -2023
- [4]Xu, Y.; Xue, J.; Liu, H.; Jiang, T.; Chen, X.; Yang, J. Identification of “Bathed” Chinese Mitten Crabs (*Eriocheir sinensis*) Using Geometric Morphological Analysis of the Carapace. *Fishes* - 2023
- [5]Minaee, S.; Abdolrashidi, A.; Su, H.; Bennamoun, M.; Zhang, D. Biometrics recognition using deep learning: A survey. *Artif. Intell. Rev*- 2023
- [6]Valero-Carreras, D.; Alcaraz, J.; Landete, M. Comparing two SVM models through different metrics based on the confusion matrix. *Comput. Oper. Res.* 2023.
- [7]Ahmad,M.;Abbas, S.; Fatima, A.; Issa, G.F.; Ghazal, T.M.; Khan, M.A. Deep transfer learning-based animal face identification model empowered with vision-based hybrid approach. *Appl. Sci.* **2023**.
- [8]Feng, Y.; Yang, X.; Xu, D.; Luo, N.; Chen, F.; Sun, C. Recognition method of individual images of river crabs based on transfer learning and pyramid convolutional network. *Fish. Mod.* **2022**.
- [9]Kim,J.H.; Poullose, A.; Colaco, S.J.; Neethirajan, S.; Han, D.S. Enhancing animal welfare with interaction recognition: A deep dive into pig interaction using xception architecture and SSPD-PIR method. *Agriculture* 2023, 13, 1522. [CrossRef]
- [10]Wan, Z.; Tian, F.; Zhang, C. Sheep face recognition model based on deep learning and bilinear feature fusion. *Animals* 2023, 13, 1957. [CrossRef] [PubMed]
- [11]Shoaib, M.; Shah, B.; Ei-Sappagh, S.; Ali, A.; Ullah, A.; Alenezi, F.; Gechev, T.; Hussain, T.; Ali, F. An advanced deep learning models-based plant disease detection: A review of recent research. *Front. Plant Sci.* 2023, 14, 1158933. [CrossRef] [PubMed]
- [12]Liu, J.; Wang, X. Early recognition of tomato gray leaf spot disease based on MobileNetv2-YOLOv3 model. *Plant Methods* 2020, 16, 1–16. [CrossRef] [PubMed]
- [13]Feng, Y.; Yang, X.; Xu, D.; Luo, N.; Chen, F.; Sun, C. Recognition method of individual images of river crabs based on transfer learning and pyramid convolutional network. *Fish. Mod.* 2022, 49, 52–60.
- [14]Shi, G.-Z.; Chen, M.; Zhang, C.-Y. Accurate traceability system of crab based on improved deep residual network. *Chin. J. Liq. Cryst. Disp.* 2019, 34, 1202–1209.
- [15]Huang, G.B.; Mattar, M.; Berg, T.; Learned-Miller, E. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. In *Proceedings of the Workshop on Faces in ‘Real-Life’ Images: Detection, Alignment, and Recognition*, Marseille, France, 16 September 2008.
- [16]Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520. [CrossRef]
- [17]Hu,L.; Ge, Q. Automatic facial expression recognition based on MobileNetV2 in Real-time. *J. Phys. Conf. Ser.* 2020, 1549, 022136. [CrossRef]
- [18]Gulzar, Y. Fruit image classification model based on MobileNetV2 with deep transfer learning technique. *Sustainability* 2023, 15, 1906. [CrossRef]