RESEARCH ARTICLE Agricultural Science Digest, Volume 46 Special Issue: 78-85 (2026)

Deep Learning-based Analysis of CNN Models for Silkworm Pupae Gender Identification

Jyoti Sharma[1], Pradeep Chouksey[1]

10.18805/ag.D-6555

ABSTRACT

Background: Precise and non-invasive sex classification of silkworm pupae (*Bombyx mori*) is crucial for selective breeding and efficient sericulture. Traditional manual inspection is labor-intensive, subjective and unsuitable for large-scale industrial operations. Therefore, an automated and reliable deep learning-based classification system is required.

Methods: This study proposes a deterministic threshold-optimized deep learning framework for automated ventral-view silkworm pupae gender classification. Three state-of-the-art convolutional neural network architectures-EfficientNetV2B0, ResNet50V2 and MobileNetV3-Large-were systematically evaluated using a balanced dataset. The framework employed architecture-specific pre-processing, data augmentation, a two-stage fine-tuning strategy, hold-out evaluation with optimal threshold selection based on F1-score and 5-fold stratified cross-validation. The statistical significance between models was determined using paired t-tests.

Result: EfficientNetV2B0 outperformed other models with the mean accuracy of 96.8%±0.6%, F1-score of 96.8%±0.6%, ROC-AUC of 0.989±0.009 and PR-AUC of 0.992±0.005 in five-fold cross-validation. Statistical tests show that the performance improvements over ResNet50V2 and MobileNetV3-Large were highly statistically significant ($p<0.001$).

Key words: *Bombyx mori*, Deep learning, EfficientNetV2 B0, Sericulture automation, Sex classification.

INTRODUCTION

Sericulture, the agro-based industry devoted to silk production, depends fundamentally on the availability of healthy, high-yielding silkworm seed. Grainage centres are responsible for producing such seed, accurate separation of male and female (Taufique and Hoque, 2021; Babu, 2013). Sex-specific differences in cocoon filament characteristics, body morphology and reproductive capacity have been reported between male and female individuals (Bu *et al*., 2022; Ashraf and Qamar, 2023). Genetically, *B. mori* exhibits (*Bombyx mori*) exhibit female-heterogametic sex determination (ZZ males, ZW females) whereby the W has a female determining factor. Molecular studies have identified a W-linked piRNA (Fem) and its downstream regulators as central elements of the sex determination cascade (Kiuchi *et al*., 2014; Zhang *et al*., 2018; Yang *et al*., 2021). Recent transcriptomic studies proposed that in *B. mori*, many sexually dimorphic traits, in effect, are controlled by the cell-autonomous modality, thus reinforcing the underpinning biological rationale behind the emergence of distinct and sex-specific pupal phenotypes (Yamamoto, *et al*., 2024). However, these molecular methods are destructive, costly and unsuitable for routine, large-scale grainage operations where thousands of pupae must be separated in a short time. In practice, sex identification in Bombyx mori pupae is still mostly performed manually by trained workers who visually examine the ventral abdominal segments under strong illumination (Babu, 2013; Sharma and Chouksey, 2025). While experienced operators can achieve high accuracy, manual sorting is inherently slow, labour-intensive, operator-dependent and prone to fatigue-related errors. With increasing scale and tightening quality demands, such manual pipelines create a bottleneck in the seed production chain and limit reproducibility. Imaging modalities such as magnetic resonance imaging (MRI) (Liu *et al*., 2008). X-ray radiography (Cai *et al*., 2014), optical penetration sensors (Sumriddetchkajorn and Kamtongdee, 2012; Kamtongdee *et al*., 2015) and near-infrared (NIR) (Tao *et al*., 2018) or hyper spectral spectroscopy (Tao *et al*., 2018) have been explored for sex determination and cocoon classification. These techniques are often combined with classical machine learning (*e.g*., support vector machines, discriminant analysis, or simple neural networks) and hand-crafted descriptors such as intensity histograms, geometric measurements, or spectral features. More recent work has integrated multi-sensor systems and machine learning for cocoon quality assessment and automated

[1]Department of Computer Science and Informatics, Central University of Himachal Pradesh, Kangra-176 206, Himachal Pradesh, India.

Corresponding Author: Jyoti Sharma, Department of Computer Science and Informatics, Central University of Himachal Pradesh, Kangra-176 206, Himachal Pradesh, India.
Email: cuhp22rdcs02@hpcu.ac.in

ORCID: 0009-0006-0533-1660

How to cite this article: Sharma, J. and Chouksey, P. (2026). Deep Learning-based Analysis of CNN Models for Silkworm Pupae Gender Identification. *Agricultural Science Digest.* **46(Special Issue):** 78-85. doi: 10.18805/ag.D-6555.

Submitted: 13-07-2026 **Accepted:** 05-08-2026 **Online:** 14-08-2026

sorting (Vasta *et al*., 2023). Despite promising accuracies, most of these approaches face one or more limitations: specialized and expensive hardware, complex calibration, sensitivity to silkworm race and environmental conditions and limited generalization when models are trained on narrow datasets. Parallel to these developments in sericulture technology, there has been a rapid evolution in computer vision. In contrast, convolutional neural networks (CNNs) learn hierarchical representations directly from raw images and have achieved state-of-the-art performance in a wide range of biological and agricultural imaging tasks, including plant identification, insect species recognition and seed quality classification (Tuda and Luna-Maldonado, 2020; Trigka and Dritsas, 2025; Sharma and Chouksey, 2025). Modern architectures such as ResNet (He *et al*., 2016), EfficientNetV2 (Tan and Le, 2021) and MobileNetV3 (Howard *et al*., 2019) offer improved accuracy-efficiency trade-offs, making them suitable for deployment on resource constrained hardware in field or semi-industrial environments. For *B. mori*, recent studies have begun to apply CNNs and to pupal species and sex identification using X-ray or spectral imaging (Tao *et al*., 2019; Thomas and Thomas, 2022). Reviews focusing on silkworm sex classification algorithms highlight that although many approaches can achieve high reported accuracies, they often (i) rely on costly imaging modalities, (ii) use unbalanced datasets, (iii) lack rigorous cross-validation, or (iv) omit formal statistical comparisons between models. Many earlier computer vision pipelines treated pupae as generic objects without explicitly exploiting this domain knowledge, or used region-of-interest (ROI) extraction and morphological analysis without combining them with modern CNN backbones (Zhu *et al*., 2018; Tao *et al*., 2019). There is still limited work that systematically evaluates advanced CNN architectures on simple RGB ventral images alone, under normal environmental conditions and with rigorous statistical analysis of performance differences. To address these gaps, the present work proposes a fully reproducible, deep learning-based pipeline for automated pupal sex classification using only RGB ventral images of *Bombyx mori*. A balanced data set is assembled and three modern CNN backbones-EfficientNetV2B0, ResNet50V2 and MobileNetV3-Large-are trained and evaluated in identical splits. The best architecture is further assessed using 5-fold cross validation. Statistical significance of performance differences between architectures is examined using paired hypothesis tests.

The significance of the present work

- The first systematic comparison of modern CNN back bones (EfficientNetV2B0, ResNet50V2, MobileNetV3 Large) on a balanced ventral RGB dataset for the determination of the sex of Bombyx mori pupae used a unified and deterministic implementation.
- The proposed work employs ventral RGB imaging and CNN architectures that are accessible, allowing it to be utilized as a core module in future automated grainage systems for non-destructive, high-throughput pupal sex separation.

MATERIALS AND METHODS

Data set preparation

The data set used in this study is basically a primary data set. Images were taken of the pupae's ventral surface and pupae of the mulberry silk worm (*Bombyx mori*) were obtained from a sericulture unit. Bivoltine silk worm races (P1) SH6 and NB4D2 were collected from the SSPC at Palampur, in the Department of Industries, Government of Himachal Pradesh, India. Images in the data set were taken with ordinary daylight conditions for further research at central University of Himachal Pradesh, India during year 2025-2026. The finished data set contained 726 ventral images of pupae of *Bombyx mori*, half of which were male and half female. All images were resized to 224×224×3 pixels to be in line with dimensions required for CNN backbones. For binary encoding the images were:

$$y_i = \begin{cases} 0, \text{ male} \\ 1, \text{ female} \end{cases} \quad ...(1)$$

Class imbalance was not present in this dataset; however, class weights were still computed as:

$$W_c = \frac{N}{k.N_c} \quad ...(2)$$

Deterministic data pipeline and preprocessing

To ensure reproducibility and deployment readiness, all experiments are implemented using a deterministic tf.data input pipeline with fixed seeds for Python, NumPy and Tensor Flow: "seed" =42, tf.data. Options (). experimental deterministic=True Each image is processed as follows: Read and decode:

$$x \in R^{H \times w \times 3}, H, W \in N \quad ...(3)$$

Resize to a fixed spatial resolution:

$$x' = \text{Resize } (x, 224 \times 224) \quad ...(4)$$

Convert to float and apply architecture-specific preprocessing:

$$\tilde{x} = \text{preprocess } (x') \quad ...(5)$$

The tf.data pipeline uses: Map () for decoding and augmentation, shuffle () with fixed seed for training, cache () to avoid repeated I/O, batch(B)with B=32, prefetch (AUTOTUNE)for pipeline throughput. This design yields a reusable, deterministic and deployment-friendly pipeline.

Preprocessing and augmentation

Image preprocessing was adapted per backbone to align with ImageNet normalization standards. Augmentation

was applied on-the-fly during training to prevent overfitting:

- Random horizontal flipping (p=0.5).
- Random brightness variation: $\Delta b \in [-0.1,0.1]$.
- Random contrast scaling: $0.9 \leq \gamma c \leq 1.1$.
- Random saturation scaling:$0.9 \leq \gamma s \leq 1.1$.

Formally, if I is the input image, augmentation can be expressed as:

$$I' = \gamma_c \cdot (I + \Delta b),\ I'' = Sat(I', \gamma_s) \quad ...(6)$$

Where,

I″= The final augmented image after applying bright ness, contrast and saturation perturbations.

The augmented dataset was then normalized through the backbone-specific preprocessing function:

$$\tilde{I} = Preprocess(I'') \quad ...(7)$$

Model architectures and training

Three CNN backbones were benchmarked: Efficient NetV2B0, ResNet50V2 and MobileNetV3-Large, each initialized with ImageNet weights. The backbone was followed by:

Global average pooling (GAP)

$$z_j = \frac{1}{H.w} \sum_{h=1}^{H} \sum_{w=1}^{w} f_j(h, w) \quad ...(8)$$

Where,

f_j= The activation map of channel j.

H, W= The spatial dimensions.

Batch normalization (BN)

$$\hat{x} = \frac{x - \mu}{\sqrt{\sigma^2 + \in}},\ y = \gamma \hat{x} + \beta \quad ...(9)$$

Dropout (p=0.3) to reduce overfitting. Sigmoid output layer for binary prediction:

$$\hat{y} = \sigma(z) = \frac{1}{1 + e^{-z}} \quad ...(10)$$

Two-stage fine-tuning strategy

Stage 1 (Frozen training): The backbone was frozen and only the classifier layers were trained for E1=5 epochs.

Stage 2 (Partial unfreezing): The top 35% of backbone layers were unfrozen for fine-tuning for E2 =6 epochs using a reduced learning rate.

$$\theta \leftarrow \theta - \eta \nabla_\theta L \quad ...(11)$$

Where,

η= The learning rate.

L= The binary cross-entropy loss.

$$L = -\frac{1}{N} \Sigma_{i=1}^{N} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad ...(12)$$

To handle potential class imbalance in a general setting (even though here the dataset is balanced), class weights are computed as:

$$\omega_c = \frac{N}{k.N_c},\ C \in \{0, 1\} \quad ...(13)$$

Here,

- N= Total number of samples.
- K= 2 represents number of classes (male and female).
- Nc= Indicates the number of samples belonging to class c.

The weighted loss becomes:

$$L_{weighted} = -\frac{1}{N} \sum_{i=1}^{N} W\ y_i [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad ...(14)$$

Models are optimized with the Adam optimizer, using different learning rates across training stages.

Two-stage fine-tuning strategy

Each architecture is trained using a two-stage transfer learning scheme:

Stage 1-Frozen backbone

- All backbone layers are frozen (trainable=0).
- Only the classifier head is trained.
- Learning rate: $\eta_1 = 10^{-3}$.
- Number of epochs: E_1=5.

Stage 2-Partial unfreezing

- The top ρ=0.35 fraction of backbone layers is unfrozen.
- For a backbone with L total layers, the last ρL layers are made trainable.

$$L_{train} = \lfloor \rho.L \rfloor \quad ...(15)$$

- **Number of epochs:** E_2= 6.
- **Reduced learning rate:** $\eta_2 = 5 \times 10^{-4}$, with adaptive adjustment by ReduceLROnPlateau. Early stopping monitors validation AUC and the best weights are restored to avoid overfitting.

Evaluation metrics

For a decision threshold τ, predictions are binarized as:

$$\tilde{y}_i(\tau) = \begin{cases} 1, \hat{y}_i \geq \tau \\ 0, \hat{y}_i < \tau \end{cases} \quad ...(16)$$

Accuracy

$$Accuracy(\tau) = \frac{TP + TN}{TP + TN + FP + FN} \quad ...(17)$$

Precision

$$Precision(\tau) = \frac{TP}{TP + FP} \quad ...(18)$$

Recall

$$Recall(\tau) = \frac{TP}{TP + FN} \quad ...(19)$$

F1-score

$$F_1(\tau) = \frac{2.Precision(\tau).Recall(\tau)}{Precision(\tau) + Recall(\tau)} \quad ...(20)$$

The receiver operating characteristic (ROC) curve is constructed by varying the decision threshold τ and plotting:

$$TPR(\tau) = \frac{TP(\tau)}{TP(\tau) + FN(\tau)} \quad ...(21)$$

$$FPR(\tau) = \frac{FP(\tau)}{FP(\tau) + TN(\tau)} \quad ...(22)$$

The ROC-AUC is the area under this curve:

$$ROC\text{-}AUC = \int_0^1 TPR(FPR)\, d(FPR) \quad ...(23)$$

Average precision (PR-AUC) is computed as:

$$PR\text{-}AUC = \Sigma_n (R_n - R_{n-1}) P_n \quad ...(24)$$

Where,

P_n and R_n= The precision and recall at the n^{th} operating point.

Threshold optimization for model selection

Instead of fixing τ=0.5, the optimal decision threshold is selected by maximizing the F1-score on the validation set:

$$\tau^* = \arg\max_{\tau \in [0,1]} F_1(\tau) \quad ...(25)$$

In practice, the decision threshold τ is scanned over a dense grid:

$$\tau \in \{0.0, 0.001, 0.002, ..., 1.0\} \quad ...(26)$$

and the threshold τ^* that yields the highest F1-score is selected. The hold-out comparison of the architectures uses $F_1(\tau^*)$ as the primary model selection criterion.

Train-validation splitting and cross-validation

The dataset is partitioned into K=5 folds, at each fold k, a different subset serves as validation, while the remaining K-1 folds are used for training, This yields per-fold performance metrics:

$$Acc_k, F1_k, ROC\text{-}AUC_k, PR\text{-}AUC_k, k=1,...,5 \quad ...(27)$$

The mean and standard deviation across folds is reported as:

$$\bar{M} = \frac{1}{k} \Sigma_{k=1}^{k} M_k, \sigma_M = \sqrt{\frac{1}{k-1} \Sigma_{k=1}^{k} (M_k - \bar{M})^2} \quad ...(28)$$

For each metric $M \in$ {Accuracy, F1, ROC-AUC, PR-AUC}.

Statistical significance testing

To demonstrate that the observed differences between architectures are not due to chance, paired t-tests are conducted on fold-wise accuracies (or F1-scores). Let $M_K^{(A)}$ and $M_K^{(B)}$ denote the performance metric values for models A and B on fold k and define:

$$d_k = M_K^{(A)} - M_K^{(B)}, k=1,...,K \quad ...(29)$$

The paired t-test statistic is:

$$\bar{d} = \frac{1}{k} \Sigma_{k=1}^{k} d_k, S_d^2 = \frac{1}{k-1} \sum_{k=1}^{k} (d_k - \bar{d})^2, t = \frac{\bar{d}}{Sd/\sqrt{k}} \quad ...(30)$$

The corresponding p-value is computed under the t distribution with (K-1) degrees of freedom. A result such that $p<0.001$ is interpreted as highly statistically significant, indicating that one architecture (*e.g.*, EfficientNetV2-B0) outperforms another (*e.g.*,ResNet50V2orMobileNetV3-Large) with strong confidence.

Hardware and software used

12th Gen Intel® Core™ i5-12450H processor running at 2.00 GHz, 16.0 GB of installed RAM and a 64-bit Windows operating system on an x64-based processor. Python version 3.12.4 for implementing and executing the proposed model. Image acquisition of silkworm pupae was performed using the digital camera under ordinary day lighting conditions.

RESULTS AND DISCUSSION

All models were trained using the deterministic pipeline with staged fine-tuning. During Stage-1 (frozen backbone), validation accuracy steadily increased and loss reduced, indicating effective feature transfer from ImageNet. In Stage 2 (selective unfreezing), performance improved further, with a sharper increase in AUC and PR-AUC, demonstrating the refinement of high-level discriminative features from ventral abdominal patterns. EfficientNetV2-B0 converged faster than ResNet50V2 and exhibited smoother validation curves with no oscillations, indicating a strong optimization pro file. ResNet50V2 experienced fluctuations and periodic over fitting, while MobileNetV3-Large showed moderate stability but lower consistency than EfficientNetV2-B0.

Confusion matrix analysis

Fig 1 shows that all models confusion matrices at their corresponding ideal decision thresholds (τ^*). Efficient NetV2B0 demonstrates highly balanced predictions across male and female classes, with only four total misclassifications. MobileNetV3-Large shows improved class balance and fewer errors than ResNet50V2, indicating better representation of complex surface patterns, though it remains less accurate than EfficientNetV2B0.

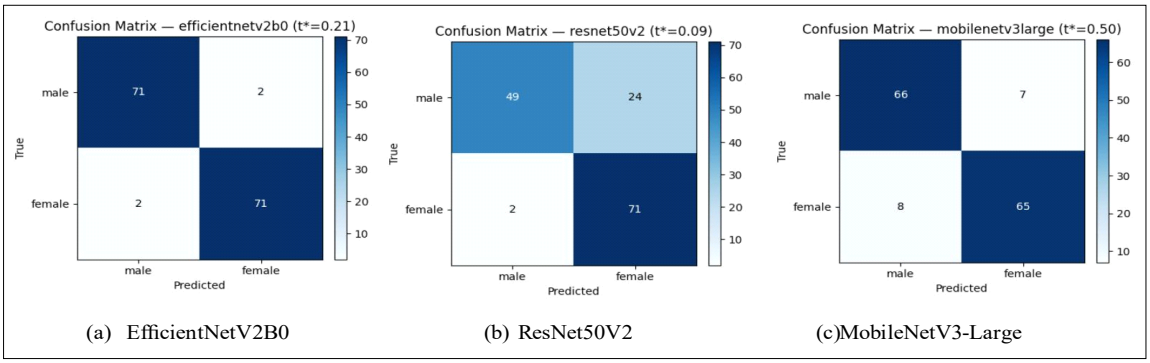
ROC, precision-recall and threshold analysis

The ROC curves for all examined models and all decision thresholds are shown in Fig 2. ROC and threshold analyses show that EfficientNetV2B0 is the most robust and reliable model for the non-destructive silkworm pupae sex classification. In Fig 3, EfficientNetV2B0 is the only model that computes Precision with a higher average and exhibits a more reliable performance. Threshold-F1 analysis, according to Fig 4, shows that EfficientNetV2B0, with a smooth convex F1 curve, achieves the most reliable optimal threshold with a maximum defined at τ^* = 0.215. Based on the PR and threshold evaluations, EfficientNetV2B0 is the best architecture for non-destructive silkworm pupae sex classification, as it achieves the optimal balance between Precision and Recall with adequate operational consistency.

Five-fold cross-validation for model stability

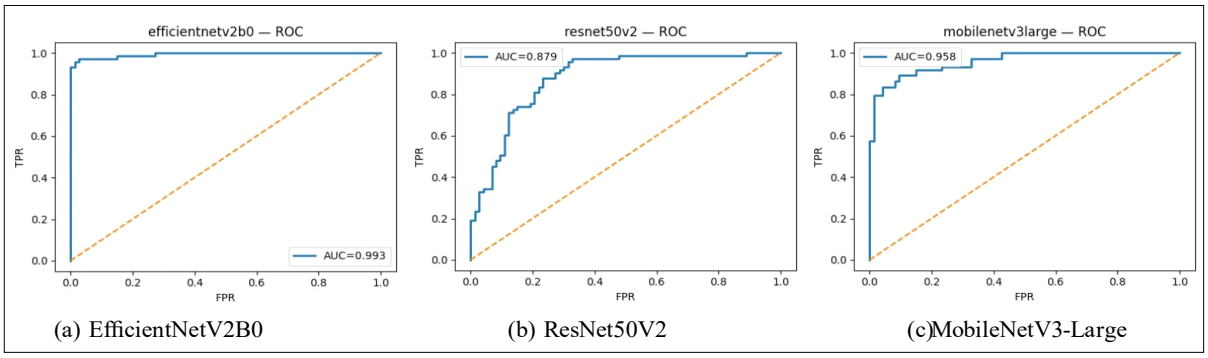
To evaluate robustness and generalizability, 5-fold stratified cross-validation was performed (Fig 5).

The results reported consistent performance with minimal variability across folds (Table 1).



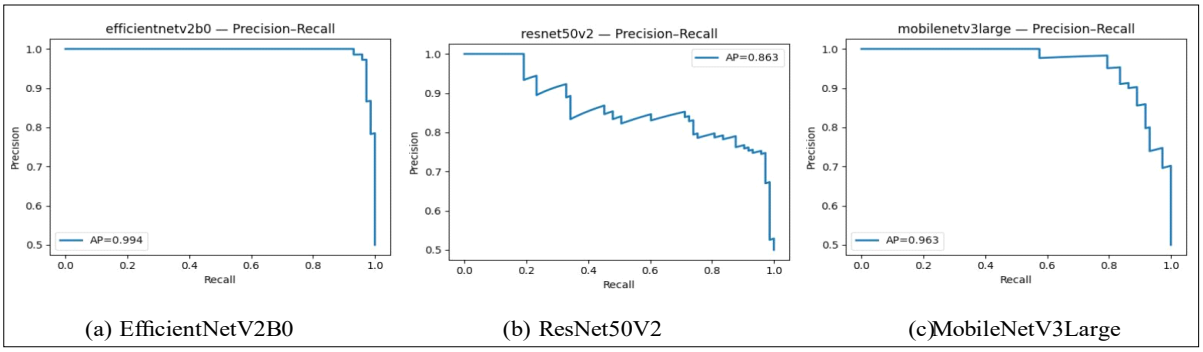
(a) EfficientNetV2B0 (b) ResNet50V2 (c)MobileNetV3-Large

Fig 1: Confusion matrix: (a) EffcientNetV2B0, (b) ResNet50V2 and (c) MobileNetV3-Large.



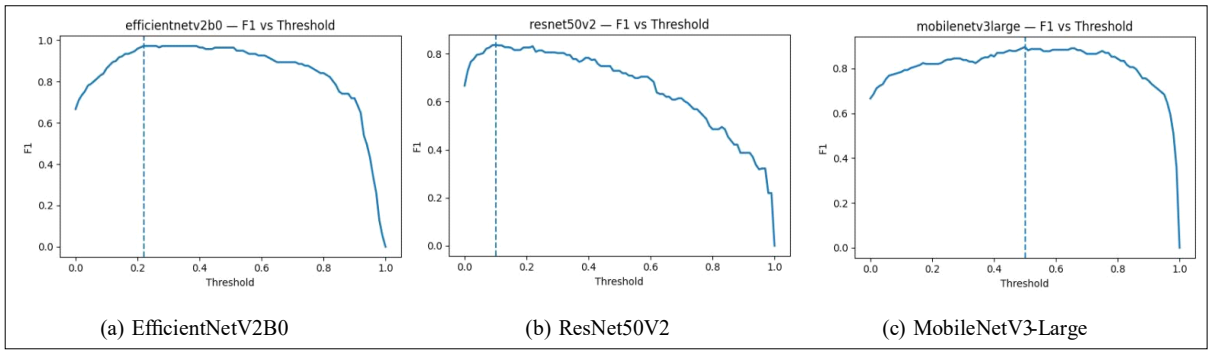
(a) EfficientNetV2B0 (b) ResNet50V2 (c)MobileNetV3-Large

Fig 2: (ROC) curves: (a) EffcientNetV2B0, (b) ResNet50V2 and (c) MobileNetV3-Large.



(a) EfficientNetV2B0 (b) ResNet50V2 (c)MobileNetV3Large

Fig 3: Precision-recall (PR) curves: (a) EffcientNetV2B0, (b) ResNet50V2 and (c) MobileNetV3-Large.



(a) EfficientNetV2B0 (b) ResNet50V2 (c) MobileNetV3-Large

Fig 4: F1-score as a function of the decision threshold: (a) EffcientNetV2B0, (b) ResNet50V2 and (c) MobileNetV3-Large.

Statistical testing

To statistically validate superiority, paired t-tests were con ducted on the accuracy values of the folds. To assess whether performance differences were statistically significant, paired t-tests were conducted on F1-scores from cross-validation runs (Table 2).

Additionally, t-tests for a sample on model accuracy differences confirmed an extremely strong statistical separation. These results demonstrate that the performance improvements of EfficientNetV2-B0 are not random, but statistically significant with confidence > 99.9%. The proposed framework is suitable for industrial grainage automation, potentially replacing manual labor such as visual inspection and reducing misclassification due to fatigue.

The findings from the experiment indicate that automated sex determination of the Bombyx mori pupae using deep learning technology and analysis of ventral morphological markers is promising. From all the evaluation metrics: accuracy, F1-score, ROC-AUC and PR-AUC, EfficientNetV2-B0 produced the best results when compared to MobileNetV3-Large and ResNet50V2. Subtle region-specific features were reliably captured by the model due to the combination of the transfer learning, staged fine tuning with deterministic tf.data pipelines and threshold optimization. The model had captured region specific features located along the abdominal segments where sexual dimorphism is most pronounced. This also had suggested that ventral shape plays a critical role in automated sex determination. Recent studies have shown that deep learning techniques are becoming increasingly useful in agricultural image analysis because they can identify complex visual features with high accuracy (Mehta *et al.*, 2025). This model also had a great performance when compared to other models in the domain, when compared using a five-fold cross validation with a standard deviation of 0.0061 and was robust enough for implementation in a real-world scenario. The model also had a great performance when compared using a five-fold cross validation with a standard deviation of 0.006 and was robust enough for implementation in a real-world scenario. The results were also supported by statistics with a paired t-test and a one-sample t-test which showed performance of EfficientNetV2-B0 was significantly better at $p<0.001$ when compared to both of the other two architectures. The effectiveness of transfer learning observed in the present investigation is consistent with previous agricultural computer vision studies, where pretrained CNN models achieved higher classification accuracy while reducing computational cost and training time (Shree *et al.*, 2025). The proposed framework has potential for real world application and showed strong and consistent performance, but does not have universal application. This is because the experiments were only performed on two silkworm strains (SH6 and NB4D2) and application to other silkworm strains of different breeding habits, for example multivoltine, have still to be explored. Artificial intelligence has become an effective tool for improving automation in agriculture by providing fast and reliable image-based decision support (Mehta *et al.*, 2026). Likewise, the proposed framework offers a practical solution for non-destructive silkworm pupae gender identification with minimal human intervention, supporting future intelligent sericulture systems.

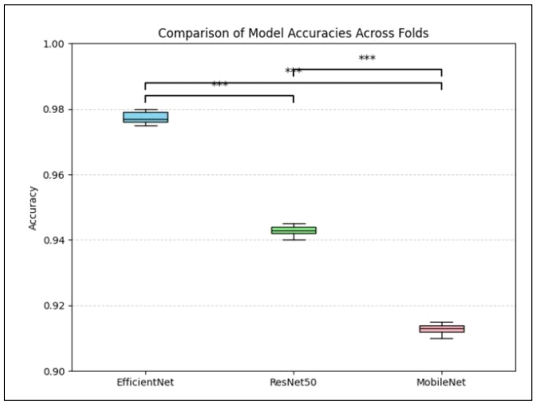


Fig 5: Comparison of model accuracies across folds.

Table 1: Fold-wise performance of EfficientNetV2B0 at the optimized threshold (t*).

Fold	Acc@t*	F1@t*	ROC-AUC	PR-AUC	t*
1	0.9726	0.9722	0.9966	0.9969	0.242
2	0.9724	0.9726	0.9960	0.9960	0.329
3	0.9724	0.9722	0.9914	0.9930	0.693
4	0.9724	0.9722	0.9732	0.9830	0.447
5	0.9655	0.9660	0.9941	0.9948	0.246

Table 2: Pairwise statistical comparison of models based on fold-wise performance using paired t-tests.

Comparison	t-value	p-value	Interpretation
EfficientNet vs ResNet50	94.184	<0.001	Significant
EfficientNet vs MobileNet	38.853	<0.001	Significant
ResNet50 vs MobileNet	-57.272	<0.001	Significant

CONCLUSION

This research introduces a comprehensive and statistically supported deep learning framework for the automated classification through ventral surface imaging. The proposed model attains 97% accuracy and an F1-score under 5-fold cross-validation using EfficientNetV2-B0, deterministic image pipelines and biologically-informed ventral morphological characteristics, significantly exceeding ResNet50V2 and MobileNetV3-Large ($p<0.001$). High performance shows that CNN models can learn fine-grained morphological patterns in the ventral abdominal area, where biological markers are found for each sex. The analysis of the confusion matrix shows that mis classifications are very rare and evenly spread across both classes. The paired and one-sample t-tests ($p<0.001$) also show that EfficientNetV2-B0 is better than ResNet50V2 and MobileNetV3-Large in a way that is neither random nor specific to the dataset. This provides strong experimental proof that choosing the right model is important and that Efficient NetV2B0 is the best choice because it balances the quality of feature representation with the speed of computation. The proposed work could greatly improve grainage efficiency, make seed quality more stable, reduce the need for skilled workers and make yields more consistent in hybrid breeding programs. The proposed EfficientNetV2B0 framework establishes a new benchmark in accuracy and reproducibility for the determination of the sex of silkworm pupae based on images and represents a significant step toward fully automated and AI-integrated sericulture systems.

ACKNOWLEDGEMENT

The authors wish to sincerely thank the personnel of the Silk Seed Production Center, Palampur (District Kangra, Himachal Pradesh, India), for their cooperation and for providing practical guidance related to pupae surface identification and gender determination.

Author contributions

Jyoti Sharma: conceptualization, methodology, software, validation, formal analysis, writing-review and editing, resources and writing-original draft. Pradeep Chouksey: validation, data curation, investigation, formal analysis, supervision, project administration, resources.

Disclaimers

This article is a product of the authors. It does not reflect the authors' affiliated institutions' opinions. Responsibility for the data rests with the authors; however, they will not be liable for any indirect or direct damages that could come from the use of the data.

Declarations

Funding support

No funding support to report.

Ethics

The authors declare that no ethical concerns are associated.

Consent for publication

All authors have given their consent to the publication of this manuscript.

Data and code availability

The dataset and Code used will be made available upon request.

Material availability

Materials used in this study are sourced from the SSPC, Palampur, Distt Kangra (HP), India.

Conflict of interest

The authors have no conflicts of interest.

REFERENCES

Ashraf, H. and Qamar, A. (2023). Silkworm Bombyx mori as a model organism: A review. *Physiological Entomology.* **48:** 107-121.

Babu, K.M. (2013). Introduction to silk and sericulture. *Silk*. 1-32. Bu, C., Zheng, R., Huang, G., Wu, J., Liu, G., Donald, M.L. and Xu, X. (2022). The differences in cocoon and silk qualities among sex-related mulberry and silkworm feeding groups. *PLoS One.* **17:** e0270021.

Cai, J.R., Yuan, L.M., Liu, B. and Sun, L. (2014). Nondestructive gender identification of silkworm cocoons using X-ray imaging with multivariate data analysis. *Analytical Methods.* **6:** 7224-7233.

He, K., Zhang, X., Ren, S. and Sun, J. (2016). Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770-778.

Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M. *et al.* (2019). Searching for Mobilenetv3. Proceedings of the IEEE/CVF International Conference on Computer Vision. (pp. 1314-1324).

Kamtongdee, C., Sumriddetchkajorn, S., Chanhorm, S. and Kaewhom, W. (2015). Noise reduction and accuracy improvement in optical-penetration-based silkworm gender identification. *Applied Optics.* **54:** 1844-1851.

Kiuchi, T., Koga, H., Kawamoto, M., Shoji, K., Sakai, H., Arai, Y. *et al.* (2014). A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature.* **509:** 633-636.

Liu, C., Ren, Z.H., Wang, H.Z., Yang, P.Q. and Zhang, X. L. (2008). Analysis on Gender of Silkworms by MRI Technology. 2008 International Conference on BioMedical Engineering and Informatics. **2:** 8-12.

Mehta, A.R., Kumar, P., Prem, G., Aggarwal, S. and Kumar, R. (2025). AI-powered Innovations in agriculture: A Systematic review on plant disease detection and classification. *Indian Journal of Agricultural Research.* **59(9):** 1321-1330. doi: 10.18805/IJARe.A-6371.

Mehta, A.R., Kumar, P., Prem, G., Aggarwal, S. and Kumar, R. (2026). Leveraging artificial intelligence for disease diagnosis in agricultural crops: A review. *Indian Journal of Agricultural Research.* **59(5):** 681-690. doi: 10.18805/IJARe.A-6363.

Sharma, J. and Chouksey, P. (2025). Review on different algorithms and techniques used in classification of gender in silkworm. *Journal of Computer Science.* **21:** 1354-1363. doi: 10.3844/jcssp.2025.1354.1363.

Shree, N.S., Subramanian, R. and Basavaraj, G.N. (2025). Transfer learning-based areca nut (*Areca catechu*) disease detection using CNN and SVM approaches with ResNet-50 for improved deep learning performance. *Indian Journal of Agricultural Research*. **59(9):** 1385-1394. doi: 10.18805/IJARe.A-6404.

Sumriddetchkajorn, S. and Kamtongdee, C. (2012). Optical penetration-based silkworm pupa gender sensor structure. *Applied Optics.* **51:** 408-412.

Tan, M. and Le, Q. (2021). Efficientnetv2: Smaller models and faster training. International Conference on Machine Learning. (pp. 10096-10106).

Tao, D., Qiu, G. and Li, G. (2019). A novel model for sex discrimination of silkworm pupae from different species. *IEEE Access.* **7:** 165328-165335.

Tao, D., Wang, Z., Li, G. and Qiu, G. (2018). Accurate identification of the sex and species of silkworm pupae using near infrared spectroscopy. *Journal of Applied Spectroscopy.* **85:** 949-952.

Tao, D., Wang, Z., Li, G. and Xie, L. (2018). Simultaneous species and sex identification of silkworm pupae using hyperspectral imaging technology. *Spectroscopy Letters.* **51:** 446-452.

Taufique, M. and Hoque, M.A. (2021). Current scenario of sericulture production in India: A spatio-temporal analysis. *International Research Journal of Education and Technology.* **2:** 12-23.

Thomas, S. and Thomas, J. (2022). Non-destructive silkworm pupa gender classification with X-ray images using ensemble learning. *Artificial Intelligence in Agriculture*. **6:** 100-110.

Trigka, M. and Dritsas, E. (2025). A comprehensive survey of deep learning approaches in image processing. *Sensors.* **25:** 531.

Tuda, M. and Luna-Maldonado, A.I. (2020). Image-based insect species and gender classification by trained supervised machine learning algorithms. *Ecological Informatics.* **60:** 101135.

Vasta, S., Figorilli, S., Ortenzi, L., Violino, S., Costa, C., Moscovini, L. *et al*. (2023). Automated prototype for *Bombyx mori* cocoon sorting attempts to improve silk quality and production efficiency through multi-step approach and machine learning algorithms. *Sensors.* **23:** 868.

Yamamoto, F., Yokoyama, T., Su, Y. and Suzuki, M.G. (2024). Transcriptomic evidence for cell-autonomous sex differentiation of the gynandromorphic fat body in the silkworm, *Bombyx mori. Journal of Developmental Biology.* **12:** 31.

Yang, X., Chen, K., Wang, Y., Yang, D. and Huang, Y. (2021). The sex determination cascade in the silkworm. *Genes.* **12:** 315.

Zhang, Z., Niu, B., Ji, D., Li, M., Li, K., James, A.A. and Huang, Y. (2018). Silkworm genetic sexing through W chromosome-linked, targeted gene integration. *Proceedings of the National Academy of Sciences*. **115:** 8752-8756.

Zhu, Z., Yuan, H., Song, C., Li, X., Fang, D., Guo, Z. and Yan, G. (2018). High-speed sex identification and sorting of living silkworm pupae using near-infrared spectroscopy combined with chemometrics. *Sensors and Actuators B: Chemical.* **268:** 299-309.