



Deep Learning Image Classifiers for Automated Identification of Stored-Grain Beetles

Lerato Lindo^{1*}, Michal Kang², Isha Ndlovu³

¹Department of Agri, Danube Institute of Agri, Uganda

²Department of Agri, Ural School of Agri, Vietnam

³Department of Agri, Ebro Graduate Institute of Agri, Switzerland

Article Info

Volume 1, Issue 1, 2026

doi: [10.62715/ijes.2026.v1.i1.pp19-24](https://doi.org/10.62715/ijes.2026.v1.i1.pp19-24)

Received: 24 April 2026

Accepted: 14 June 2026

Published: 06 August 2026

Abstract

Beetles that infest stored grain are small, dark and alike, and telling them apart is a skilled task on which trading decisions, fumigation and quarantine action depend. Convolutional neural networks trained on images now match specialists in some insect identification tasks, and low-cost cameras and microscopes are within reach of most grain laboratories. What is lacking is an agreed way of building and testing a classifier so that its stated accuracy can be trusted at an intake point or in a warehouse. This article proposes such a procedure in eight steps: define the taxa and the use case, assemble specimens identified by a specialist, image them to a fixed protocol, split the data by specimen and source, fine-tune networks pretrained on general images, evaluate on sources held out from training, calibrate the outputs and set a rule for declining to answer, and monitor the classifier after deployment. For each step the article gives the reasoning, the decision to be made and the record to be kept. Particular attention is paid to species pairs that cannot be separated from external appearance and to regulated pests for which a missed detection is costly. The procedure is proposed from the published literature and has not been validated on a dataset here.

Keywords: *Stored-product insects, Coleoptera, Convolutional neural networks, Transfer learning, Image classification, Species identification*

© 2026 Lerato Lindo et al. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

1. Introduction

A sample of infested wheat or rice may contain half a dozen beetle species within a few millimetres of one another in size. The commonest are the weevils of the genus *Sitophilus*, the lesser grain borer *Rhyzopertha dominica*, the flour beetles *Tribolium castaneum* and *T. confusum*, the saw-toothed grain beetle *Oryzaephilus surinamensis* and the flat grain beetles of the genus *Cryptolestes* (Rees, 2004). They differ in whether they attack sound kernels, in their tolerance of phosphine and in their regulatory status, so that management depends on knowing which are present (Hagstrum and Subramanyam, 2006). Identification is done under a stereo microscope by trained staff, of whom there are few.

Image classification by deep learning offers a way of extending that expertise. Convolutional networks

* Corresponding author: Lerato Lindo, Department of Agri, Danube Institute of Agri, Uganda. E-mail: lerato.lindo@danube.edu

learn discriminating features directly from labelled images (LeCun et al., 2015), and in entomology they have reached the accuracy of specialists on some difficult tasks (Valan et al., 2019). Reviews have set out the general promise for insect monitoring (Høye et al., 2021), and detection networks have been applied to images of insects in grain (Shen et al., 2018).

Published accuracies are, however, hard to compare and often optimistic, because they depend on how the images were collected and divided. This article proposes a procedure for developing and testing a classifier of stored-grain beetles that guards against the commoner sources of over-statement. It is a design drawn from the literature. No classifier was trained for this article, and the procedure is offered for others to apply and test.

2. Background

Two bodies of knowledge inform the design: the taxonomy of the beetles, which fixes what an image can and cannot show, and the practice of transfer learning, which fixes how much data is needed.

2.1. What can be seen in an image

Some stored-grain beetles are distinctive at a glance. Others form pairs that are separated on fine characters. *Tribolium castaneum* and *T. confusum* differ in the form of the antennal club and in the distance between the eyes on the underside of the head. *Sitophilus oryzae* and *S. zeamais* overlap in external appearance and are reliably separated only by the genitalia, and the same is true of several *Cryptolestes* species (Rees, 2004). The khapra beetle, *Trogoderma granarium*, a quarantine pest in many countries, belongs to a genus whose species are difficult even for specialists. A classifier fed dorsal photographs cannot be expected to resolve what a taxonomist resolves by dissection, and the class list has to respect that limit.

2.2. Transfer learning

The networks that transformed image recognition were trained on ImageNet, a collection of more than a million labelled photographs (Deng et al., 2009). No grain laboratory will assemble a set of that size. It does not need to, because the early layers of a network trained on general images encode edges, textures and shapes that transfer to other tasks (Yosinski et al., 2014). Fine-tuning a pretrained network on a few hundred images per class is now the standard route for species identification (Wäldchen and Mäder, 2018). Residual networks (He et al., 2016) are the usual baseline. Compact architectures such as MobileNetV2 (Sandler et al., 2018) and the EfficientNet family (Tan and Le, 2019) are attractive where the model must run on a phone or an embedded board.

3. The Proposed Procedure

The procedure has eight steps, shown in sequence in Figure 1. Table 1 sets out for each step the decision it requires and the record it should leave, so that a reader of the eventual report can judge the result.

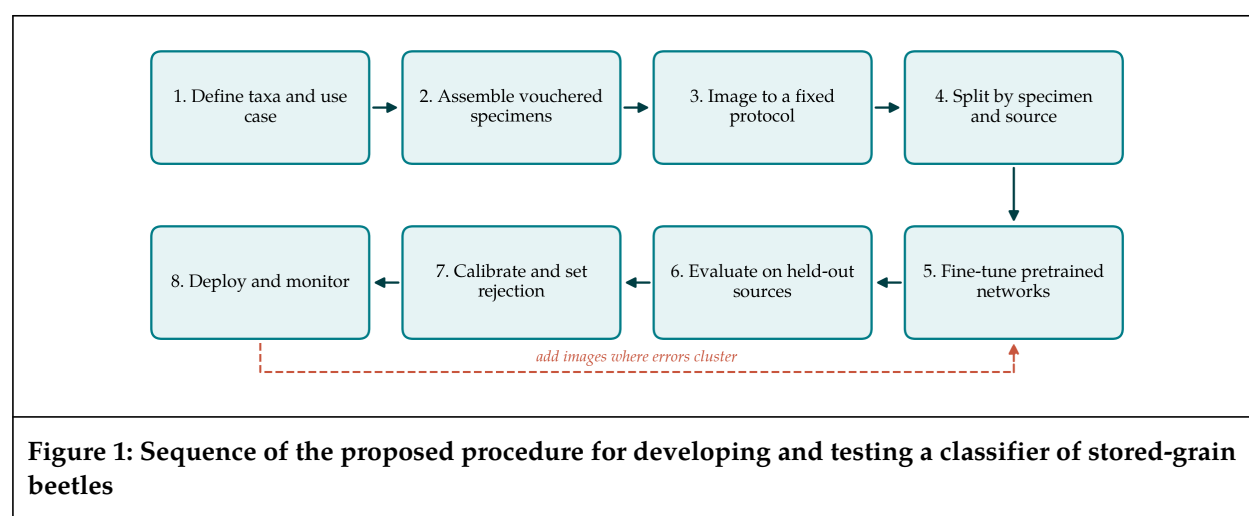
3.1. Step 1: Define the taxa and the use case

The first decision is what the classifier is for. Screening intake samples for any live insect, naming the species in a trap catch, and flagging a possible quarantine pest are different tasks with different costs of error. The class list follows from the task and from what images can show. Species that cannot be separated externally are merged into a single class named for the pair or the genus. A class for other insects and debris is included from the outset, since a deployed system will meet mites, psocids, moth larvae, fragments and weed seeds.

3.2. Step 2: Assemble vouchered specimens

Labels are only as good as the identifications behind them. Each specimen should be identified by a competent person using standard keys, with dissection where the species requires it, and should carry an identifier that links it to its images. Specimens should come from several laboratory cultures and field collections for each class. Beetles from one culture share age, diet and condition, and a network can learn

the culture in place of the species.



3.3. Step 3: Image to a fixed protocol

Images should be taken with the kind of device, magnification, background and lighting that will be used in service, and those conditions should be varied deliberately across all classes. If every *Rhyzopertha* is photographed on white card and every *Oryzaephilus* on grey, the background becomes the easiest feature to learn. Several views of each specimen, dorsal, lateral and ventral, add information, and the ventral view matters for the flour beetles. Dead, damaged and dusty individuals should be included because they are what real samples contain.

3.4. Step 4: Split by specimen and by source

Dividing images at random between training and test sets is the commonest cause of inflated accuracy. Several images of one beetle are nearly duplicates, and if they fall on both sides of the split the test measures memory. All images of a specimen must go to one partition. Better still, whole sources, a culture, a collection site or an imaging session, are held out, following the argument for blocked validation of structured ecological data made by Roberts et al. (2017). The test set is then set aside and consulted once.

3.5. Step 5: Fine-tune pretrained networks

Two or three pretrained architectures of different size are fine-tuned on the training partition, with a validation partition used to choose the learning rate and the stopping point. Augmentation by rotation, flipping, cropping and changes of brightness and contrast enlarges the effective sample and suits specimens that have no natural orientation (Shorten and Khoshgoftaar, 2019). Colour should be perturbed only mildly, because the reddish brown of *Tribolium* against the darker *Sitophilus* is itself diagnostic. Classes with few images are over-sampled or weighted in the loss.

3.6. Step 6: Evaluate on held-out sources

Overall accuracy conceals what matters when classes are unequal in size and in importance. The report should give the full confusion matrix, recall and precision for each class, and an average that weights classes equally (Sokolova and Lapalme, 2009). For a regulated species the figure of interest is the proportion of true specimens that the classifier misses. Saliency maps such as Grad-CAM (Selvaraju et al., 2017) should be inspected for a sample of images to confirm that the network attends to the insect and, ideally, to the parts a taxonomist would examine.

3.7. Step 7: Calibrate and set a rejection rule

Modern networks are often over-confident, and their output scores should not be read as probabilities without adjustment. Temperature scaling on the validation set is a simple correction (Guo et al., 2017).

With calibrated scores a threshold is set below which the system returns no identification and refers the specimen to a person. The same mechanism gives a first defence against species outside the training set, which tend to receive lower maximum scores ([Hendrycks and Gimpel, 2017](#)), although it is far from complete. The threshold is chosen from the cost of a wrong answer against the cost of a referral, and the share of images declined is reported with the accuracy on the remainder.

Table 1: Steps of the proposed procedure, the decision each requires and the record to keep		
Step	Decision to be made	Record to keep
1. Taxa and use case	Which classes; which species pairs to merge	Class list with reasons
2. Specimens	Sources per class; who identifies and how	Specimen register and vouchers
3. Imaging	Device, views, backgrounds and lighting	Written imaging protocol
4. Data split	Which sources are held out for testing	Partition list by specimen
5. Training	Architectures, augmentation, class weighting	Code, settings and random seeds
6. Evaluation	Metrics; which errors matter most	Confusion matrix and per-class recall
7. Calibration	Confidence below which no answer is given	Threshold and share of images declined
8. Deployment	How samples are re-checked by a person	Audit log and retraining dates

3.8. Step 8: Deploy and monitor

A classifier in service meets new cameras, new grain types and new populations. A fixed fraction of its identifications, and all referrals, should be checked by a trained person, and the agreement tracked over time. Images on which it failed are added to the training pool with verified labels, and the model is retrained at stated intervals against the same held-out test set, so that successive versions can be compared.

4. Fitting the Procedure to a Grain Laboratory

The demands of the procedure are modest in equipment and heavier in discipline. A stereo microscope with a camera port, or a phone with a clip-on macro lens, is adequate for adults of the species concerned. Training can be done on a single graphics card or a rented one. The scarce input is the taxonomist's time in step 2, and it is best spent on a few hundred well-documented specimens per class from varied sources, in preference to thousands of images of beetles from one jar. Where a laboratory handles a regulated pest, any positive from the classifier should be treated as a prompt for confirmation by the national plant protection service and never as a diagnosis.

5. Limitations and What Remains to Be Shown

The procedure has not been applied to a dataset in this article, and its benefits are argued, not demonstrated. It would need to be shown that classifiers built this way lose less accuracy when moved to a new site than classifiers built on randomly split data, and that the rejection rule catches a useful share of unfamiliar species. The treatment here is confined to adult beetles presented singly. Larvae, which do much of the damage inside kernels, mixed samples that require detection before classification, and insect fragments in flour are separate problems. Nor does image classification remove the need for dissection or molecular methods where species are externally alike; it only makes clear where that need lies.

6. Conclusion

Automated identification of stored-grain beetles from images is technically within reach of an ordinary laboratory. The risk lies less in building a classifier than in believing an accuracy figure that reflects how

the data were divided. The eight steps proposed here tie the class list to what the taxonomy allows, hold out whole sources for testing, report errors class by class, and give the system a way of declining to answer. Applied and reported in this form, classifiers from different laboratories could be compared, and their performance in service could be checked against what was claimed for them.

Conflict of interest

To be declared by the author(s) before publication.

Funding

To be declared by the author(s) before publication.

References

- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-Fei, L. (2009). [ImageNet: A large-scale hierarchical image database. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 248-255.](#) <https://doi.org/10.1109/CVPR.2009.5206848>
- Guo, C., Pleiss, G., Sun, Y. and Weinberger, K.Q. (2017). [On calibration of modern neural networks. *Proceedings of the 34th International Conference on Machine Learning*, PMLR 70: 1321-1330.](#)
- Hagstrum, D.W. and Subramanyam, B. (2006). [Fundamentals of Stored-Product Entomology](#). AACC International, St. Paul.
- 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*, 770-778.](#) <https://doi.org/10.1109/CVPR.2016.90>
- Hendrycks, D. and Gimpel, K. (2017). [A baseline for detecting misclassified and out-of-distribution examples in neural networks. *Proceedings of the International Conference on Learning Representations*.](#)
- Høye, T.T., Årje, J., Bjerger, K., Hansen, O.L.P., Iosifidis, A., Leese, F., Mann, H.M.R., Meissner, K., Melvad, C. and Raitoharju, J. (2021). [Deep learning and computer vision will transform entomology. *Proceedings of the National Academy of Sciences*, 118\(2\): e2002545117.](#) <https://doi.org/10.1073/pnas.2002545117>
- LeCun, Y., Bengio, Y. and Hinton, G. (2015). [Deep learning. *Nature*, 521\(7553\): 436-444.](#) <https://doi.org/10.1038/nature14539>
- Rees, D. (2004). [Insects of Stored Products](#). CSIRO Publishing, Collingwood.
- Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillerá-Arroita, G., et al. (2017). [Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40\(8\): 913-929.](#) <https://doi.org/10.1111/ecog.02881>
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L.-C. (2018). [MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4510-4520.](#) <https://doi.org/10.1109/CVPR.2018.00474>
- Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D. (2017). [Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision*, 618-626.](#) <https://doi.org/10.1109/ICCV.2017.74>
- Shen, Y., Zhou, H., Li, J., Jian, F. and Jayas, D.S. (2018). [Detection of stored-grain insects using deep learning. *Computers and Electronics in Agriculture*, 145: 319-325.](#) <https://doi.org/10.1016/j.compag.2017.11.039>
- Shorten, C. and Khoshgoftaar, T.M. (2019). [A survey on image data augmentation for deep learning. *Journal of Big Data*, 6: 60.](#) <https://doi.org/10.1186/s40537-019-0197-0>
- Sokolova, M. and Lapalme, G. (2009). [A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45\(4\): 427-437.](#) <https://doi.org/10.1016/j.ipm.2009.03.002>

- Tan, M. and Le, Q.V. (2019). [EfficientNet: Rethinking model scaling for convolutional neural networks](#). *Proceedings of the 36th International Conference on Machine Learning*, PMLR 97: 6105-6114.
- Valan, M., Makonyi, K., Maki, A., Vondráček, D. and Ronquist, F. (2019). [Automated taxonomic identification of insects with expert-level accuracy using effective feature transfer from convolutional networks](#). *Systematic Biology*, 68(6): 876-895. <https://doi.org/10.1093/sysbio/syz014>
- Wäldchen, J. and Mäder, P. (2018). [Machine learning for image based species identification](#). *Methods in Ecology and Evolution*, 9(11): 2216-2225. <https://doi.org/10.1111/2041-210X.13075>
- Yosinski, J., Clune, J., Bengio, Y. and Lipson, H. (2014). [How transferable are features in deep neural networks?](#) *Advances in Neural Information Processing Systems*, 27: 3320-3328.