Accurate species identification of food-contaminating beetles with quality-improved elytral images and deep learning

Halil Bisgin[1†], Tanmay Bera[2†], Leihong Wu[2], Hongjian Ding[3], Neslihan Bisgin[1], Zhichao Liu[2], Monica Pava-Ripoll[4], Amy Barnes[3], James F. Campbell[5], Himansi Vyas[3], Cesare Furlanello[6], Weida Tong[2] and Joshua Xu[2]*

[1]Department of Mathematics and Applied Sciences, University of Michigan-Flint, Flint, MI, United States, [2]Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR, United States, [3]Food Chemistry Lab 1, Arkansas Regional Laboratory, Office of Regulatory Affairs, US Food and Drug Administration, Jefferson, AR, United States, [4]Office for Food Safety, Center for Food Safety and Applied Nutrition, US Food and Drug Administration, College Park, MD, United States, [5]Stored Product Insect and Engineering Research Unit, US Department of Agriculture, Manhattan, KS, United States, [6]HK3 Lab, Milan, Italy

Food samples are routinely screened for food-contaminating beetles (i.e., pantry beetles) due to their adverse impact on the economy, environment, public health and safety. If found, their remains are subsequently analyzed to identify the species responsible for the contamination; each species poses different levels of risk, requiring different regulatory and management steps. At present, this identification is done through manual microscopic examination since each species of beetle has a unique pattern on its elytra (hardened forewing). Our study sought to automate the pattern recognition process through machine learning. Such automation will enable more efficient identification of pantry beetle species and could potentially be scaled up and implemented across various analysis centers in a consistent manner. In our earlier studies, we demonstrated that automated species identification of pantry beetles is feasible through elytral pattern recognition. Due to poor image quality, however, we failed to achieve prediction accuracies of more than 80%. Subsequently, we modified the traditional imaging technique, allowing us to acquire high-quality elytral images. In this study, we explored whether high-quality elytral images can truly achieve near-perfect prediction accuracies for 27 different species of pantry beetles. To test this hypothesis, we developed a convolutional neural network (CNN) model and compared performance between two different image sets for various pantry beetles. Our study indicates improved image quality indeed leads to better prediction accuracy; however, it was not the only requirement for achieving good

accuracy. Also required are many high-quality images, especially for species with a high number of variations in their elytral patterns. The current study provided a direction toward achieving our ultimate goal of automated species identification through elytral pattern recognition.

Introduction

A large group of nuisance insects that contaminate grains and other food items are commonly termed pantry beetles (Bell, 2013). They are notorious for spoiling stored grain and processed food products, leading to significant economic damage (Belluco et al., 2013). Some of these pantry beetles are aggressively invasive and can cause damage to local agriculture and ecological insects if they spread through the transportation of contaminated food products (Heeps, 2016). Some of the pests also pose a serious threat to public health, as they are active carriers of pathogens (Olsen et al., 2001).

To counter such adversities, food grains and products are monitored and routinely screened for pantry beetles or their remains (Bell, 2013; Belluco et al., 2013). The most common and widely-used method involves highly-trained analysts manually screening food samples for insect remains using optical microscopes. Any insect or insect remains found are then scrutinized using a comparison optical microscope to match the patterns from the insect fragments with reference images to identify the exact insect species, genus, or family. This identification step is crucial, as each species poses different threat levels and their contamination may require different methods of management and regulatory procedures. Currently, no reliable alternatives to the manual screening method are available, as spectroscopic or PCR-based detection techniques have remained challenging for this application. Moreover, due to the manual nature of the microanalysis, the current method is highly dependent on the experience and expertise of the individual analyst, making it more susceptible to human error and higher variation across institutions. Also, manual methods are difficult to scale up, hindering the screening of a larger number of samples in a shorter time frame, especially in the absence of experienced and dexterous analysts.

Species identification through image analysis has been explored for efficient taxonomical and environmental applications for several years (Norouzzadeh et al., 2018; Terry et al., 2019; Høye et al., 2020). These computer-aided applications have tried to address a wide range of problems from food safety to identification of insect pests (Daly et al., 1982; Weeks et al., 1997; O'Neill et al., 2000; Larios et al., 2008; Yalcin, 2015). With the advent of machine learning

methods, image-based species identification has gained further momentum and well-known discriminative models such as support vector machines (SVM) (Cortes and Vapnik, 1995) and generative models have been widely adopted for insect classification (Martineau et al., 2017). Examples of these models include, but are not limited to: insect or pest identification using SVM (Qing et al., 2012; Wang et al., 2012; Yang et al., 2015), honeybee and moth identification with decision trees (Mayo and Watson, 2007; da Silva et al., 2015), and red palm weevil and insect recognition systems through neural networks (Al-Saqer and Hassan, 2011; Wang et al., 2012). With increasing computational power, more complex neural network architectures, i.e., deep learning (DL) approaches have recently helped in tackling more challenging tasks in the field of food and agricultural science (Lee et al., 2015; DeChant et al., 2017; Lu et al., 2017; Zhang et al., 2018). Although there have been relatively fewer DL studies to identify filth elements for food contamination (Reinholds et al., 2015; Bansal et al., 2017), variations of DL designs such as Region-based Fully Convolutional Network (R-FCN), convolutional block attention module (CBAM), convolutional neural network (CNN) and pre-trained models have shown promising performances for pest, stored-grain insect, and fly classification (Chen et al., 2020; Kuzuhara et al., 2020; Shi et al., 2020). The DL models have not only achieved high classification accuracies, but also offered a new way of feature extraction embedded in the process as an alternative to conventional features such as domain-dependent, global, local, and mid-level features (Martineau et al., 2017).

We have also investigated similar approaches, i.e., machine learning techniques, with the aim of automating the identification process of pantry beetles whose elytra (hardened forewing) have unique patterns that can be considered as fingerprints or features. In a previous study, we demonstrated that a specific pantry beetle species could indeed be identified through elytral pattern recognition using machine learning (Martin et al., 2016). In our subsequent study, we observed that classical machine learning techniques such as artificial neural network (ANN) and SVM could be used for this application (Bisgin et al., 2018). However, optimized ANN and SVM models yielded about 80 and 85% of average accuracies, respectively. We observed that some species consistently performed less than others; which could be attributed to their misidentification with

another species from the same genus or family with similar or near-identical elytral patterns. We further studied more advanced machine learning techniques such as CNN, which also performed similarly (Wu et al., 2019).

Our findings in our earlier studies led us to scrutinize the image set and observe that images lacking visual clarity due to the reflective glare of the elytra surface were more prone to misidentification. To remedy this, we amended the optical and imaging settings and optimized the imaging conditions to obtain a high-quality image set unaffected by artifacts and showing the finer details of an elytron (Bera et al., 2021). We hypothesized that using such a high-quality image set would help us achieve a near-perfect prediction accuracy in identifying each pantry beetle species. In the current study, therefore, we tested this hypothesis by using a CNN model on an extended dataset which consisted of high-quality images of 27 species. We further shed light on the impact of enhanced images of 12 species in the same dataset that were previously studied. Our experiments showed both the utility of the prediction framework and the improvement in species identification due to image quality which could potentially guide any future efforts for auto-detection tools.

The rest of the paper is organized as follows: Section Material and methods details the dataset for 27 species and introduces our approach, Section Results and discussion presents our results, Section Discussion discusses our findings, and Section Conclusion concludes our work.

Materials and methods

Beetle sample collection and image acquisition

We elaborated on the details of sample collection, preparation, and imaging technique in our previous publications on imaging optimization (Bisgin et al., 2018; Wu et al., 2019; Bera et al., 2021). Briefly, we used 12 different pantry beetle species harvested from our in-house collection. We chose these species due to their prevalence and significance in food contamination, especially in North American food samples. Another 15 different species were collected from the U.S. Department of Agriculture's (USDA) Animal and Plant Health Inspection Service (APHIS) laboratory. Elytra from each beetle specimen were harvested, thoroughly cleaned through sonication in an ethanol solution, and subsequently preserved in 70% ethanol prior to imaging. Table 1 shows the full list of 27 species which include both our in-house collection (12 species) and additional 15 species.

The harvested elytra were then air-dried and imaged using stereo microscopes (Leica M205, Allendale, New Jersey). Unlike the older image set, which was subjected to varied magnification (in the 75–100× range) and two-point reflected light, we used a

fixed magnification of 100× and transmitted light for this study. These amendments significantly reduced glare spots and other imaging artifacts, and drastically improved the clarity of elytral patterns (Bera et al., 2021). We used a Leica MC170HD camera to acquire the images with an image resolution set to 2,592 × 1,944 dpi (dots per inch, the highest resolution available). In this study, only images from the ventral side (underside) of the elytra were used. The concave shape of the elytra naturally preserves the ventral side elytral patterns. This selection allowed us to focus our attention on only the pattern recognition aspect without having to worry about such artifacts as variation or loss of setae (surface hair) or other sample damages that often occur on the frontal side of the elytra during food or sample preparation steps.

We used 20 elytral images per species. Each image subsequently was divided into smaller subimages (tiles) to simulate physical fragmentation of the elytra that are often observed in contaminated samples. This simulated fragmentation step was critical to our application, as it allowed us to increase the sample size and to validate our algorithms in close to real-life scenarios, in which elytral fragments are the only viable remains found in contaminated food samples.

Image preprocessing

Each image frame (captured at 100× magnification) had the elytra at the center of the white background. Thus, in the first step of preprocessing, we removed the white background by determining the elytral border (line of maximum change in contrast). Next, we randomly split images belonging to the same species to construct training and test sets by observing a 4:1 ratio, as shown schematically in Supplementary Figure 1, which was the same practice we used in our previous studies. Since an early study showed the utility of images with a size of 448 × 448 (Wu et al., 2019), we randomly cropped 100 regions so that each image was the same size. These sub-images guaranteed they would be inside the borders detected in the previous step and allowed to have overlap. This resulted in 46,300 training and 10,800 test images. By following such an exercise, we ensured that all sub-images of a particular image were put either in the training or test set in order to prevent information leak. This "blind" cross-validation strategy reduced bias and minimized the possibility of overfitting.

Convolutional neural network and the model structure

For the classification task here, we adopted CNNs, which have been widely used in the research community for image classification and segmentation in recent years (Lawrence et al., 1997; Krizhevsky et al., 2012; LeCun, 2021). The ability of

TABLE 1 The complete list of pantry beetles used in this study, listed alphabetically by their family, genus, species and common names, with abbreviations.

	Family	Genus	Species	Common Name	SP Id new
1	Anthribidae	*Araecerus*	*fasciculatus*	Coffee Bean Weevil	AAF
2	Anobiidae	*Lasioderma*	*serricorne*	Cigarette Beetle	ALS
3	Anobiidae	*Stegobium*	*paniceum*	Drugstore Beetle	ASP
4	Bostrichidae	*Rhyzopthera*	*dominica*	Lesser Grain Borer	BRD
5	Chrysomelidae	*Callosobruchus*	*maculatus*	Cowpea Weevil	CCM
6	Curculionidae	*Sitophilus*	*granarius*	Granary Weevil	CSG
7	Curculionidae	*Sitophilus*	*oryzae*	Rice Weevil	CSO
8	Curculionidae	*Sitophilus*	*zeamaise*	Maize Weevil	CSZ
9	Dermestidae	*Attagenus*	*Unicolor*	Black Carpet Beetle	DAU
10	Dermestidae	*Trogoderma*	*inclusum*	Cabinet Beetle	DTI
11	Laemophloeidae	*Cryptolestes*	*ferrugineus*	Rusty Grain Beetle	LCF
12	Laemophloeidae	*Cryptolestes*	*pusillus*	Flat Grain Beetle	LCP
13	Laemophloeidae	*Cryptolestes*	*turcicus*	Flour Mill Beetle	LCT
14	Silvanidae	*Ahasverus*	*advena*	Foreign Grain Beetle	SAA
15	Silvanidae	*Ahasverus*	*species*	Fungus Beetle	SAS
16	Silvanidae	*Cathartus*	*quadricollis*	Squarenecked Grain Beetle	SCQ
17	Silvanidae	*Oryzaephilus*	*mercator*	Merchant Grain Beetle	SOM
18	Silvanidae	*Oryzaephilus*	*surinamensis*	Saw-toothed Grain Beetle	SOS
19	Tenebrionidae	*Cynaeus*	*angustus*	Larger Black Flour Beetle	TCA
20	Tenebrionidae	*Gnatocerus*	*cornutus*	Broad-horned Flour Beetle	TGC
21	Tenebrionidae	*Latheticus*	*oryzae*	Longheaded Flour Beetle	TLO
22	Tenebrionidae	*Lophocateres*	*pusillus*	Siamese Grain Beetle	TLP
23	Tenebrionidae	*Palorus*	*ratzeburgii*	Smalleyed Flour Beetle	TPR
24	Tenebrionidae	*Tribolium*	*castaneum*	Red Flour Beetle	TTCa
25	Tenebrionidae	*Tribolium*	*confusum*	Confused Flour Beetle	TTCo
26	Tenebrionidae	*Tribolium*	*Destructor*	Dark Flour Beetle	TTD
27	Tenebrionidae	*Tribolium*	*madens*	Black Flour Beetle	TTM

CNN to learn features while applying convolutional filters during the training stage makes it appealing and different from conventional image classification methods (Zheng et al., 2006). These types of deep neural network structures comprise cascaded convolutional and pooling layers in which filters are utilized to attain the most informative features that eventually provide significantly reduced image sizes. The CNN final output is then passed to a dense layer in a flattened representation, allowing passage to subsequent dense layers that finally terminate in another fully-connected layer with a number of neurons equal to the number of classes (i.e., species, in our case).

We constructed a CNN by using Keras (Chollet, 2015), which is an application programming interface (API) that runs the Tensorflow machine learning platform (Abadi et al., 2016) in the backend and offers further image preprocessing utilities for more generalizable models. Specifically, our network architecture consists of four convolutional layers along with corresponding pooling layers. These perform downsampling,

usually by either choosing the maximum or average value in a given region, and two additional dense layers. We employed 3×3 filters in the convolutional layers that were followed by max pooling layers using 2×2 windows to choose the maximum value. In order to avoid overfitting, we further adapted the dropout approach that randomly ignores some units at a desired level to prevent coadapting (Srivastava et al., 2014). In Figure 1, we illustrate the details of our network structure, listing all six layers and the number of nodes for each layer. We used Rectified Linear Unit (ReLU) activation function in the first five layers. In the final layer, we used a softmax function due to the multi-class nature of our predictions. For the optimizer, we used the Adam algorithm because of its efficient management of larger datasets and parameters (Kingma and Ba, 2014).

Keras's data augmentation features enabled us to artificially increase the sample size (i.e., number of subimages). Additionally, it helped generalize the model by applying image processing functions to the existing training samples. These functions perform image manipulations, such as rotations, that
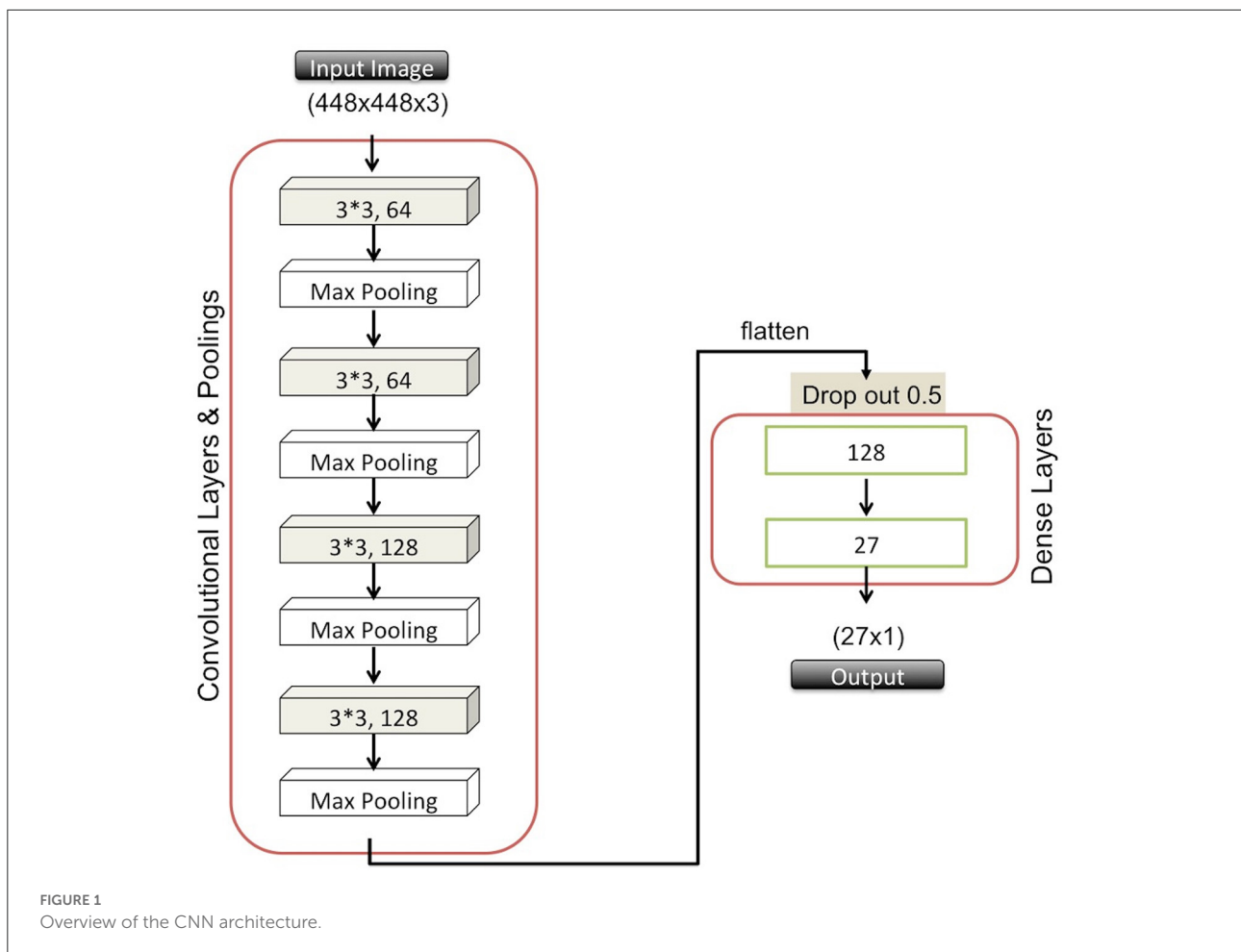
FIGURE 1
Overview of the CNN architecture.

TABLE 2 List of augmentation options and parameter values used in our study.

Option	Explanation	Value
rotation_range	Creates images with random rotations up to N degrees.	40
width_shift_range	Handles off-center objects by artificially creating shifted versions of the training data	0.2
height_shift_range		0.2
shear_range	Shear angle in counterclockwise direction in degrees	0.2
zoom_range	Random zoom range	0.2
horizontal_flip	Creates random flips of the image (supposes you feed a mirror image)	True
fill_mode	Helps in filling values outside the boundaries of an image	nearest

lead to a more diverse and larger set of images derived from the original set. We list details about the augmentation options and parameter values used in our study in Table 2. From the details shown in Table 2, we derived an augmented training set which could include additional images that might be shifted 20%, rotated 30 degrees, magnified 15%, sheared 10%, and horizontally flipped. If any pixels were lost due to the operations and needed to be filled to keep the image integrity, the nearest pixels could be used.

Model training and validation

Keras offers a user-friendly interface for data augmentation and experimental design, including the arrangement of training and test sets consisting of image folders maintained by the *ImageDataGenerator* module of the *keras_preprocessing* library. In our case, for 27 species we created a training directory that included 27 folders, from which class labels were inherited.

Similarly, we created a validation directory using the *flow_from_directory* function.

We passed these settings to the *fit_generator* function, along with the compiled neural network detailed above, with the *categorical_crossentropy* loss function, *adam* optimizer, and the default batch size (Bisgin et al., 2018). We trained our model for 100 epochs and tested its performance on the validation images after each epoch.

Model evaluation

As in our previous studies, we first computed the accuracy values for each species by computing the mean and standard deviation for each round of validation (Bisgin et al., 2018; Wu et al., 2019). This yielded a confusion matrix after the cross-validation from which true positive (TP), false positive (FP), true negative (TN), and false negative (FN) were computed. These were subsequently used to calculate the prediction parameters, namely Precision, Sensitivity (Recall), Specificity, Matthews Correlation Coefficient (MCC) using the standard formula, which can also be found in our previous report (Bisgin et al., 2018). Average prediction accuracy was also calculated by averaging species-wise accuracies.

Code and experimental environment

Given the significantly increased image size (average 14 mb per full elytra image and 600 kb per sub-images), we used the NCTR/FDA High-Performance Computing Cluster containing approximately 1100 CPU cores. The script used in this study can be found in github[1].

Results

Beetle species and the classifier

We initiated the study with 15 species of food-contaminating beetles most prevalent to North America. In the later part of the study, this number was expanded to 27 species. Table 1 contains a list of test species alphabetized by their family names with details on their nomenclature; namely, family, genus, species, and common names, along with their abbreviations. Those abbreviations were used to refer to each tested species. Supplementary Figure 2 shows some of the representative elytra images. For comparison, we provided images obtained though both the traditional and optimized methods. It was quite evident that imaging optimization significantly improved image quality and clarity of the elytral patterns. Compared to the traditionally

---

1 https://github.com/hbisgin/beetleCNN

acquired image set, the optimized image set was devoid of such artifacts as glare spots and other surface anomalies. Details on the imaging improvements, described elsewhere, are beyond the scope of this discussion (Bera et al., 2021). This image set subsequently was processed to obtain the set of sub-images used for our model.
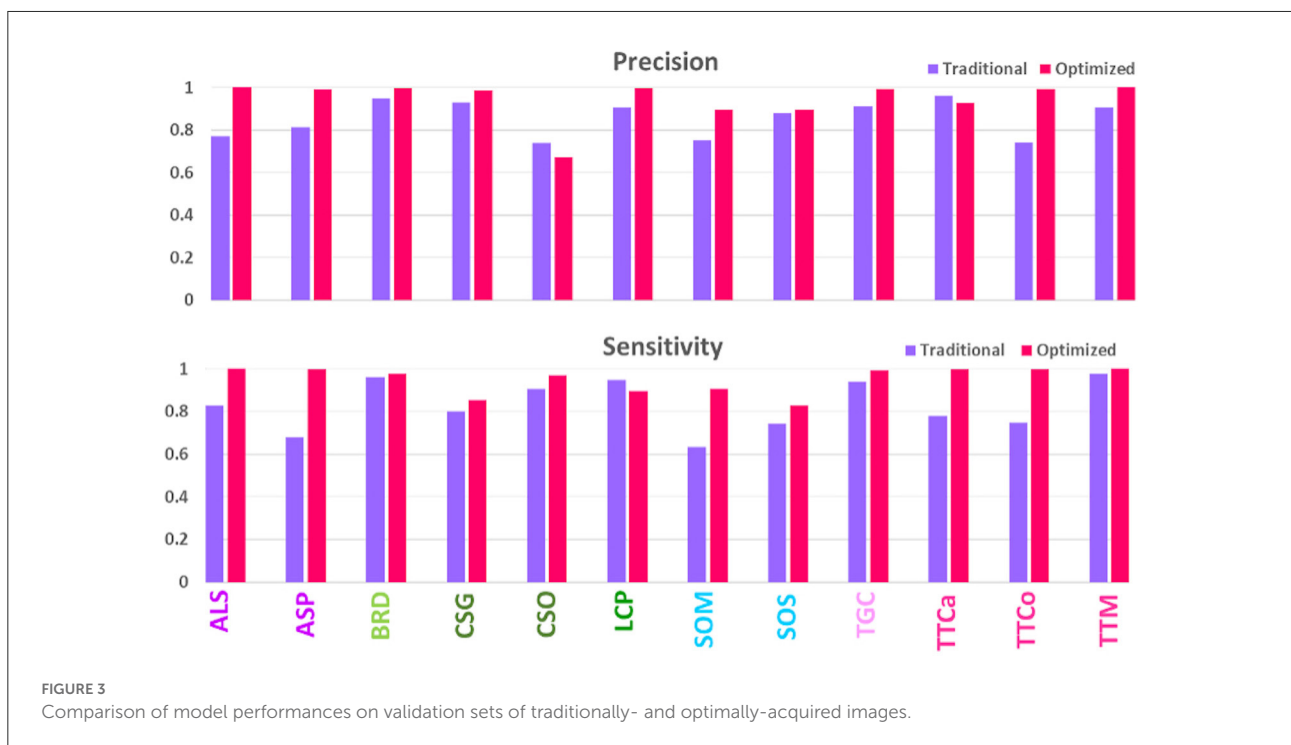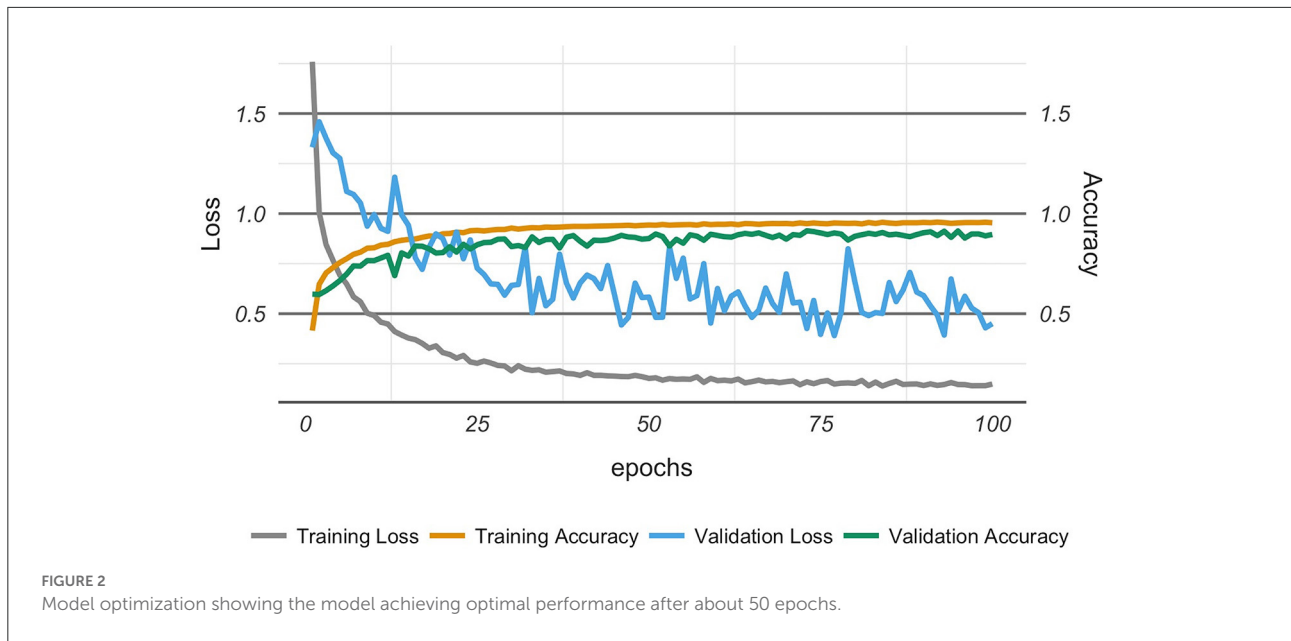
Model summary

The analysis of training and validation progress of the 27 classes along epochs is reported in Figure 2. We observed that the training loss (i.e., categorical cross-entropy) began to stabilize after ~50 epochs, beyond which the decrease was much more gradual. Also, we observed that testing accuracy approached saturation after ~50 epochs. Both observations might indicate that the model had reached nearly optimal accuracy, and that 50 epochs would have been enough, which was close to our earlier observations. However, the loss function for the testing (validation loss) fluctuated, but tended to stay in a limited bandwidth around the value at 50 epochs.

Species-wise performance and comparison

To test the hypothesis that a high-quality image set may increase prediction accuracy, we made a head-to-head comparison of the prediction results (Recall and Precision) for the same 12 species for earlier and current image sets, as shown in Figure 3. Evidently, the newer high-quality image set improves the prediction performance for most species, with an average prediction accuracy increasing from 80% to above 90%. The improvements were particularly notable for such species as ALS and ASP, SOM and SOS, and TTCa and TTCo; these had previously been difficult to accurately identify, however, can now be identified with >90% accuracy. These 12 species, especially, SOM, SOS, TTCa and TTCo, are some of most commonly encountered pantry beetles in North America. Therefore, improving the accuracy of their prediction identification will have regulatory significance. The traditionally-obtained images with higher artifacts and lower quality lacked the pattern clarity to distinguish one species from another. This was particularly true for species with near-identical elytral patterns (due to their genetic similarity) and belonging to the same genus and/or family [referred to as "difficult pairs" in our previous works (Bisgin et al., 2018; Wu et al., 2019)]. The high-quality images significantly improved the pattern clarity, allowing for distinct identification of each species, even within the difficult pairs. To our surprise, we observed exceptions to this general trend, especially for the species CSO. Of all 12 species, this one performed the poorest and showed a significant decrease in prediction accuracy

FIGURE 2
Model optimization showing the model achieving optimal performance after about 50 epochs.

FIGURE 3
Comparison of model performances on validation sets of traditionally- and optimally-acquired images.

compared to the traditionally-acquired images. The image set for this particular species possibly contained an anomaly, resulting in this decrease.

Expansion to more species—performance parameters

Expanding the number of species to 27 enabled us to verify the observations made with the initial 12 species with our

newly built model in this study. Four prediction parameters, namely Precision, Recall (or Sensitivity), Specificity and MCC, for these species are presented in Figure 4. The general trend of improved prediction is evident from this figure. Specificity values for all the species validate our hypothesis that high-quality images can improve prediction accuracy. However, there were exceptions to the general trend; as some species, such as CSO and LCP, performed quite poorly. Several other species, namely AAF, CSG, LCF, SAA, SOM, SOS, and TCA, performed below average, i.e., 90%. This suggests poor performance is not a
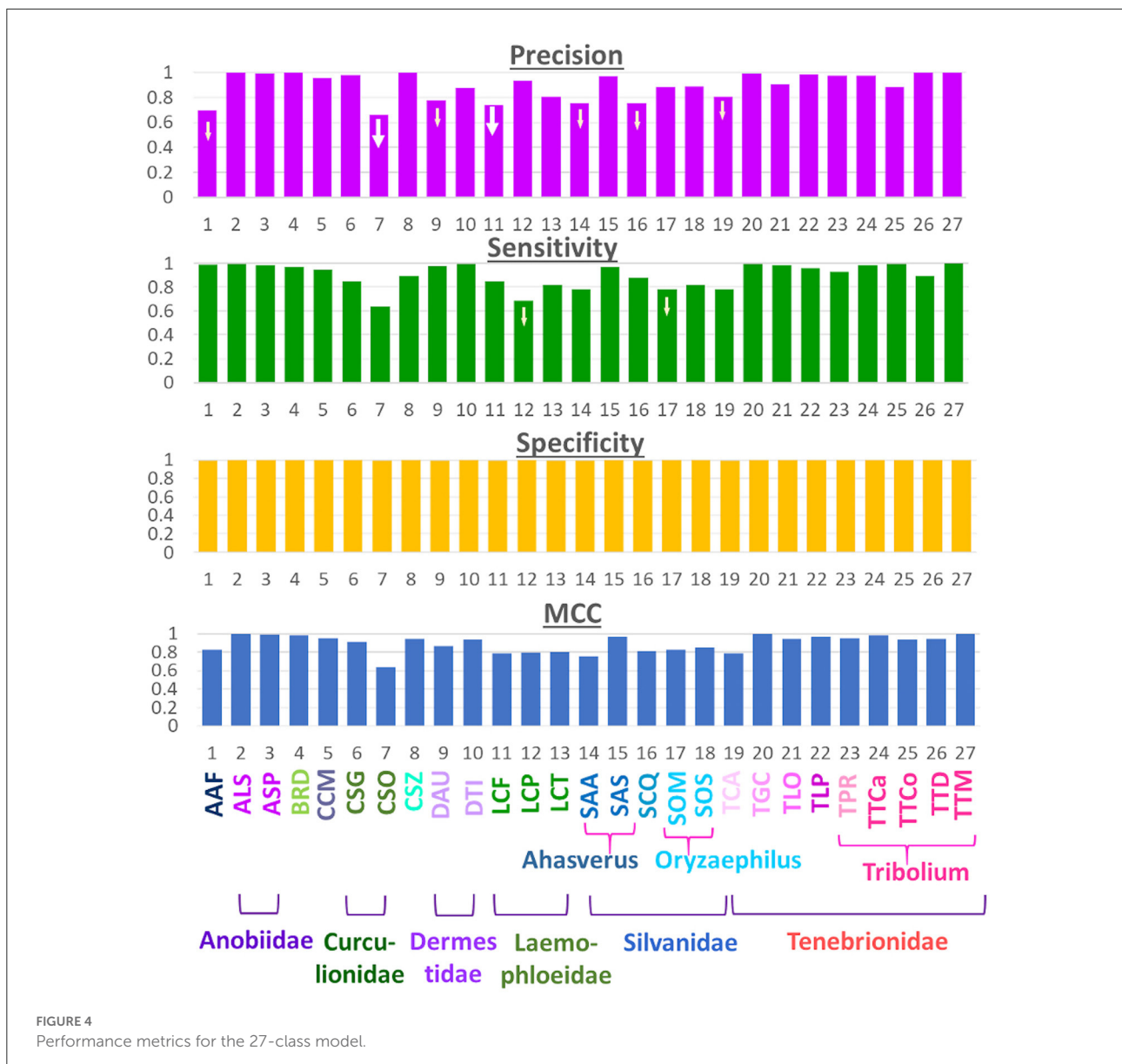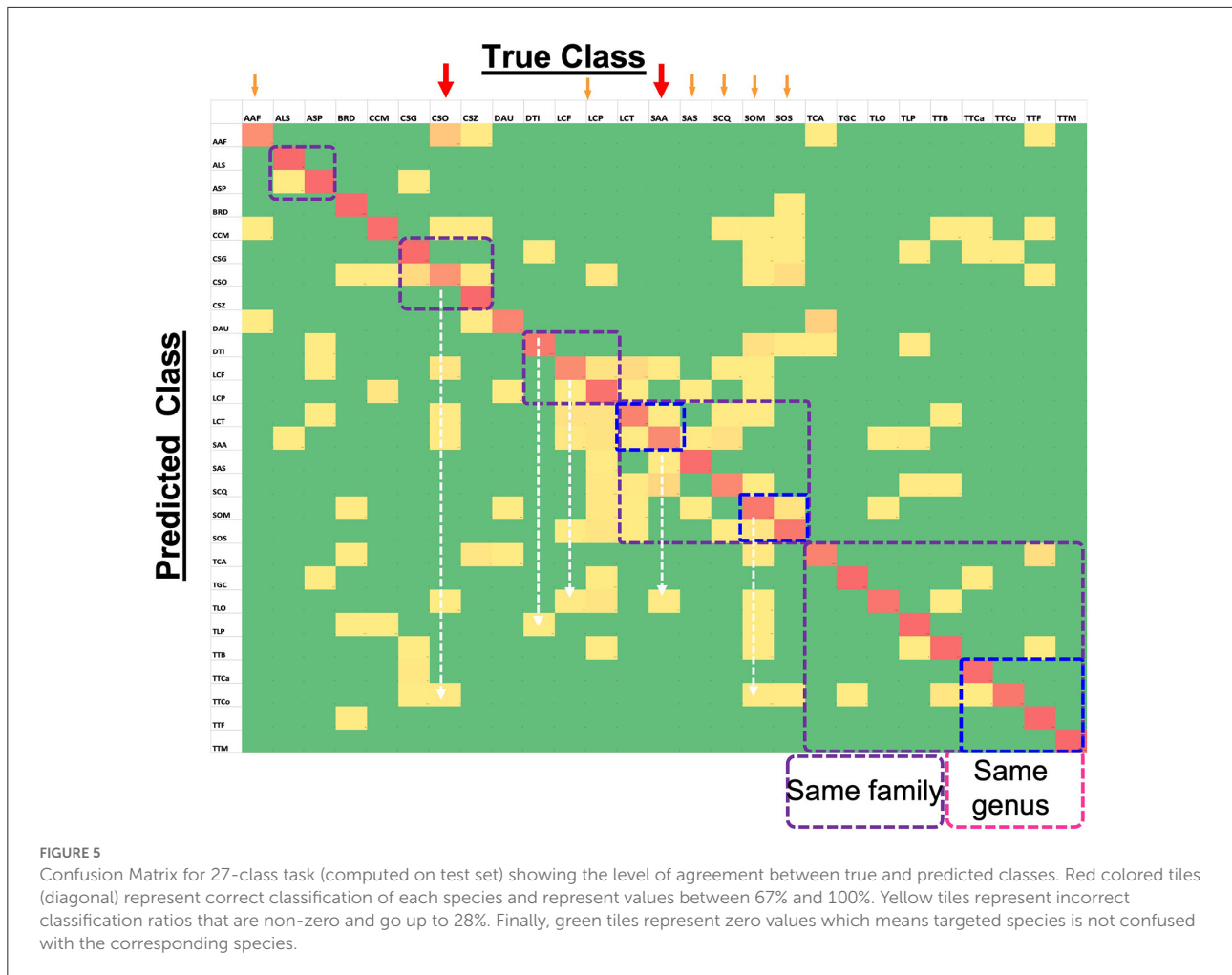
FIGURE 4
Performance metrics for the 27-class model.

singular anomaly in the image set of one species. Instead, there may be underlying factors that play a crucial role in a species' prediction performance and these need further research. One possibility, as we observed previously, is that species with similar elytral patterns (belonging to the same genus and/or family) were confused with one another during the prediction.

Confusion matrix

Figure 5 shows the confusion matrix for all the species, with horizontal rows showing the True class and vertical rows the Predicted class. It is evident that overall performance of the model is quite accurate, as the red diagonal entities are clearly

prominent. Although the model is far from perfect, as one can observe several non-diagonal entities in yellow, it is a good working model since the deviations were fairly low (mostly yellow and not orange non-diagonal entities) as indicated by the color scale. A closer look at the matrix, especially for the poorly-performing species (marked with red arrows) such as CSO, indicated that its low prediction performance was not due to the similarity of elytral patterns with a species from the same genus or family (marked by dotted squares). Rather, it was being predicted for several different species across various families. For instance TTCo was predicted as SOM and SOS for 9 and 6 times, respectively, compared to TGC, which is in the same family. This suggests that the image quality that showed distinction (or resulted in confusion) between

FIGURE 5
Confusion Matrix for 27-class task (computed on test set) showing the level of agreement between true and predicted classes. Red colored tiles (diagonal) represent correct classification of each species and represent values between 67% and 100%. Yellow tiles represent incorrect classification ratios that are non-zero and go up to 28%. Finally, green tiles represent zero values which means targeted species is not confused with the corresponding species.
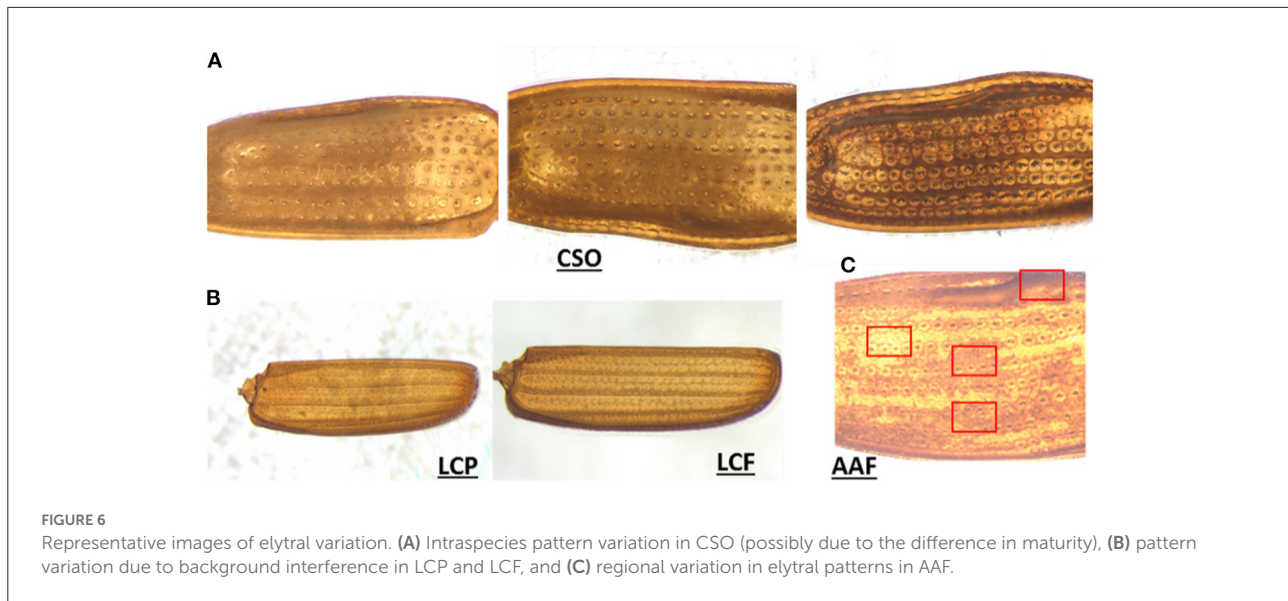
similar elytral patterns is not the major factor at play on our data. We made a similar observation for the second-lowest performer, LCF, which was also predicted beyond its own genus and/or family. Other low-performing species, such as AAF, SAA, SCQ, SOM, SOS, and TCA, showed comparable trends. The two-dimensional UMAP representation of all classes based on their extracted 128 features from the last layer of the network (Supplementary Figure 3) also illustrates misclassified species. This observation further bolstered our speculation that something other than pattern clarity may be affecting the prediction performance, and deserved detailed discussion.

Discussion

In most academic and research settings, the architecture of the model often receives more attention than does the quality of the data, possibly because cleaning the dataset often is beyond the scope of many researchers. This has been found to be true, particularly in image classification for species identification applications. Users of prediction models, even models with the best-known architecture, have found achieving good accuracy for noisy datasets challenging as quality of the data has impact on the classifier performance (Sáez et al., 2016). Our study also highlights this fact in the context of species identification and food safety, as the prediction performance showed improvement when a better-quality dataset was used to build the model.

Furthermore, our results indicated the importance and relevance of other factors beyond data quality. As discussed previously, we observed that species performing below average were not being inaccurately predicted or confused with another species from their own genus and/or family due to elytral pattern similarity, but were being misclassified into various different and unrelated species. To better understand this problem, we delved deeper and looked through the images of those species. Figure 6A shows three different elytral images of the same species, CSO. The difference in elytral patterns are obvious, and believed to be mostly due to age of the beetle. However, differences

FIGURE 6
Representative images of elytral variation. **(A)** Intraspecies pattern variation in CSO (possibly due to the difference in maturity), **(B)** pattern variation due to background interference in LCP and LCF, and **(C)** regional variation in elytral patterns in AAF.

could also be due to sex and/or individual variation, as the older beetles tend to develop a darker elytral color and prominent pattern, possibly to attract a mate. These variations are not uncommon and were found in such other species as SAA and SOS (Supplementary Figure 4), which also performed poorly.

Surprisingly, these species did not perform so poorly in our previous models using ANN and SVM. Supplementary Figure 5 shows the Recall and Precision comparisons for ANN, SVM, and CNN models using a conventionally-obtained (lower quality) image set than that of the present model (CNN, using a higher-quality image set). In comparison to ANN and SVM models, the performances of CSO dropped in both CNN models (using conventional and high-quality image sets), even though the CNN model is known to, in most cases, outperform ANN and SVM (Shin and Balasingham, 2017; Senyurek et al., 2019). We argue that this anomaly is due to the difference between explicitly defining features or trusting the CNN to develop its own feature extraction internally. In both ANN and SVM, the image features (such as size, shape, distribution, and color of the elytral pattern), were preprocessed before being used for training and testing the model. It is during this feature selection process that the intraspecies variations in elytral patterns probably did not get selected in the top-ranked features, as they appeared in only a handful of species. Subsequently, they remained unused in the ANN and SVM models and showed no influence in performance. On the other hand, output of convolutional layers served as the feature set in the current model, which could not take advantage of earlier select features, possibly causing a decrease in performance.

Unlike CSO, the species LCP (the second-lowest performer) did not show significant intraspecies variation. On minute observation, we found they contained imaging artifacts. LCP belongs to the family Laemophloeidae, which is one of the smallest species of pantry beetles. They also have extremely thin elytra and faint patterns, which when imaged on filter papers (a common practice in food filth analysis), in some cases resulted in a fibrous paper background getting embedded in the elytral images (see Figure 6B). This imaging artifact was prominent in some parts of a few of the elytral images, which appeared quite different from the actual elytral pattern and could very well be the reason for their poorer performance. AAF was another species performing below average. In this case, each elytron had regions that appeared different from one another (variegated pattern). In some areas, the elytra appeared much brighter, while in other regions they appeared much darker. Some regions had more prominent patterns compared to others (see Figure 6C). When the images of the whole elytra were divided into subimages, the subimage set had much more pattern diversity. Some of the randomly-selected subimages used in testing probably appeared quite different from the training subimages, yielding a lower prediction value. It can also be noted that the AAF had a high Recall value but low Precision values. This indicated that our model was impressive in choosing relevant species, but in this case was slightly less exact due to highly diverse subimages.

While our collective results showed that model performance improved significantly when using better-quality images, thus validating our initial hypothesis, they indicated that species with higher intraspecies elytral diversity or with enhanced variegated

elytral patterns do not perform as well. These observations seemed reasonable and have room for improvement without needing significant change in the model architecture. They are also aligned with a known general limitation of CNN models, which require training sets with both high-quality and large-quantity of images to yield better prediction accuracies (Valan et al., 2019; Høye et al., 2020).

While the cropped subimages were a way of imitating the actual beetle fragments and artificially increasing the size of the dataset, the limited number of elytral images remained one of the challenges in this study. Adding Keras image augmentation became a possible solution, as it has been used to solve imaging issues in domains such as medical image analysis (Shorten and Khoshgoftaar, 2019). In this step, several other scenarios, such as rotation, shearing, and zooming to some extent, were incorporated. During the training stage, the model was exposed to data augmentation to prepare it for possible variations, including likely presence of fragmented patterns. Even though this approach worked very well both for training accuracy and training loss, slightly lower accuracy and fluctuating loss observed in the validation stage also indicates that high variability of novel patterns is much harder to control and beyond the reach of data augmentation.

The broader objective of our work is to automate the process of elytral pattern recognition to better alleviate insect food contamination. We foresee this can only be achieved by concatenating the following three steps: (1) establishing a mechanism of acquiring high-quality images, (2) accumulating beetle images with proper labels in a repository with a growing number of samples for species with high variability, and (3) making them accessible for model development/improvement. Before developing a full-force effort to implement the whole process, it was critical to validate with a proof of concept the hypothesis that high-quality images can significantly improve predictive accuracy. The present study served this purpose and indicated that a high number of high-quality images is indeed a promising way forward in achieving precise identification over a large number of species. In our recent report on imaging optimization techniques, we elaborated on the method for acquiring high-quality images of pantry pests. Through this study, we developed a step-by-step procedure and a detailed instruction manual for high-quality image acquisition, which we will make publicly available. We currently are in the process of developing a high-quality image database containing 40 images per species for about 40 different pantry beetles, which will also be made public. Efforts currently are underway to construct a graphical user interface (GUI), from which any user can upload elytral images (preferably obtained by following the SOP and imaging manual) of pantry beetles in order to identify species using a CNN model similar to the one reported here. This use of the GUI will further enhance the high-quality image database and will provide a large number of high-quality, well-labeled image sets which can be used to further improve this CNN model in the future. At this point, the present work explores advantages and limitations of using a CNN model for classifying various species of pantry pests through elytral pattern recognition. We are optimistic that the current study has put us a step closer to achieving automated species identification of pantry pests, and thus toward a more efficient regulatory system to better manage food contamination scenarios.

Conclusions

In this study, we aimed at scouring the landscape and moving closer to achieving near-perfect species-level identification. We set out to explore whether high-quality elytral images were sufficient for improving the prediction accuracy of pantry beetle species identification. To test this hypothesis, we first compared two CNN models; one developed with traditionally-obtained, low-resolution images, and another with optimized imaging conditions, yielding high-quality images. Overall, we observed an improvement in average prediction accuracy due to the improved image quality. When we extended the analysis to 27 different pantry beetles, we achieved an average accuracy of ∼90%; however, several species fell below that average accuracy. A data review elucidated that below-average performance was not due to poor image quality, but rather to significant intraspecies variation of elytral pattern, and in some cases, to enhanced regional variation of patterns within one elytron. Detailed analysis indicated that greater numbers of high-quality images are necessary to account for these variations and achieve higher accuracy of the model. In future studies, we aim to achieve this objective using a publicly-available GUI for pantry beetle identification, allowing us to accumulate larger quantities of high-quality images through user participation. We hope this exploratory study will help achieve our ultimate goal of automated species identification of food-contaminating beetles.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://github.com/hbisgin/BeetleCNN.

Author contributions

TB acquired the images. HB performed the calculations. TB and HB analyzed the results with help from LW, NB, ZL, CF, and JX. HD and AB provided the in-house entomological support including samples and imaging facility. MP-R and JC provided the external entomological support and consultation. JX and HD led the projects. HV and WT managed and supported the study. All authors reviewed and approved the manuscript.

Funding

Acknowledgments

Conflict of interest

Author CF was employed by company HK3 Lab.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frai.2022.952424/full#supplementary-material

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. (2016). "Tensorflow: A system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation*, 265–283.

Al-Saqer, S. M., and Hassan, G. M. (2011). Artificial neural networks based red palm weevil (*Rynchophorus ferrugineous*, Olivier) recognition system. *Am. J. Agric. Biol. Sci.* 6, 356–364. doi: 10.3844/ajabssp.2011.356.364

Bansal, S., Singh, A., Mangal, M., Mangal, A. K., and Kumar, S. (2017). Food adulteration: sources, health risks, and detection methods. *Critic. Rev. Food Sci. Nutr.* 57, 1174–1189. doi: 10.1080/10408398.2014.967834

Bell, C. H. (2013). *Food Safety Management: Chapter 29. Pest Management.* Amsterdam: Elsevier Science. doi: 10.1016/B978-0-12-381504-0.00029-9

Belluco, S., Losasso, C., Maggioletti, M., Alonzi, C. C., Paoletti, M. G., Ricci, A., et al. (2013). Edible insects in a food safety and nutritional perspective: a critical review. *Comprehens. Rev. Food Sci. Food Saf.* 12, 296–313. doi: 10.1111/1541-4337.12014

Bera, T., Wu, L., Ding, H., Semey, H., Barnes, A., Liu, Z., et al. (2021). Optimized imaging methods for species-level identification of food-contaminating beetles. *Sci. Rep.* 11, 1–13. doi: 10.1038/s41598-021-86643-y

Bisgin, H., Bera, T., Ding, H., Semey, H. G., Wu, L., Liu, Z., et al. (2018). Comparing SVM and ANN based machine learning methods for species identification of food contaminating beetles. *Sci. Rep.* 8, 6532. doi: 10.1038/s41598-018-24926-7

Chen, Y., Zhang, X., Chen, W., Li, Y., and Wang, J. (2020). Research on recognition of fly species based on improved RetinaNet and CBAM. *IEEE Access.* 8, 102907–102919. doi: 10.1109/ACCESS.2020.2997466

Chollet, F. (2015). *Keras.* Available online at: https://keras.io (accessed May 11, 2022).

Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018

da Silva, F. L., Sella, M. L. G., Francoy, T. M., and Costa, A. H. R. (2015). Evaluating classification and feature selection techniques for honeybee subspecies identification using wing images. *Comput. Electron. Agric.* 114, 68–77. doi: 10.1016/j.compag.2015.03.012

Daly, H. V., Hoelmer, K., Norman, P., and Allen, T. (1982). Computer-assisted measurement and identification of honey bees (Hymenoptera: Apidae). *Ann. Entomol. Soc. Am.* 75, 591–594. doi: 10.1093/aesa/75.6.591

DeChant, C., Wiesner-Hanks, T., Chen, S., Stewart, E. L., Yosinski, J., Gore, M. A., et al. (2017). Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning. *Phytopathology.* 107, 1426–1432. doi: 10.1094/PHYTO-11-16-0417-R

Heeps, J. (2016). *Insect Management for Food Storage and Processing.* Amsterdam: Elsevier Science.

Høye, T. T., Ärje, J., Bjerge, K., Hansen, O. L., Iosifidis, A., Leese, F., et al. (2020). Deep learning and computer vision will transform entomology. *bioRxiv.* 2020.07.03.187252.doi: 10.1101/2020.07.03.187252

Kingma, D. P., and Ba, J. (2014). Adam: a method for stochastic optimization. *arXiv preprint arXiv:*14126980.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "ImageNet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'12)* (Red Hook, NY: Curran Associates Inc), 1097–1105.

Kuzuhara, H., Takimoto, H., Sato, Y., and Kanagawa, A. (2020). Insect pest detection and identification method based on deep learning for realizing a pest control system. *IEEE 2020,* 709–714. doi: 10.23919/SICE48898.2020.9240458

Larios, N., Deng, H., Zhang, W., Sarpola, M., Yuen, J., Paasch, R., et al. (2008). Automated insect identification through concatenated histograms of local appearance features: feature vector generation and region detection for deformable objects. *Mach. Vis. Appl.* 19, 105–123. doi: 10.1007/s00138-007-0086-y

Lawrence, S., Giles, C. L., Tsoi, A. C., and Back, A. D. (1997). Face recognition: a convolutional neural-network approach. *IEEE Trans. Neural Netw.* 8, 98–113. doi: 10.1109/72.554195

LeCun, Y. (2021). *LeNet-5, Convolutional Neural Networks* (2021). Available online at: http://yann.lecun.com/exdb/lenet (accessed August 01, 2021).

Lee, S. H., Chan, C. S., Wilkin, P., and Remagnino, P. (2015). Deep-plant: plant identification with convolutional neural networks. *IEEE 2015,* 452–456. doi: 10.1109/ICIP.2015.7350839

Lu, Y., Yi, S., Zeng, N., Liu, Y., and Zhang, Y. (2017). Identification of rice diseases using deep convolutional neural networks. *Neurocomputing*. 267, 378–384. doi: 10.1016/j.neucom.2017.06.023

Martin, D., Ding, H., Wu, L., Semey, H., Barnes, A., Langley, D., et al. (2016). "An image analysis environment for species indentification for food contaminating beetles," in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, vol. 16, 4375–4376. doi: 10.1609/aaai.v30i1.9846

Martineau, M., Conte, D., Raveaux, R., Arnault, I., Munier, D., Venturini, G. A., et al. (2017). survey on image-based insect classification. *Pattern Recogn.* 65, 273–284. doi: 10.1016/j.patcog.2016.12.020

Mayo, M., and Watson, A. T. (2007). Automatic species identification of live moths. *Knowledge-Based Syst.* 20, 195–202. doi: 10.1016/j.knosys.2006.11.012

Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., et al. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proc. Natl. Acad. Sci. USA* 115, E5716. doi: 10.1073/pnas.1719367115

Olsen, A. R., Gecan, J. S., Ziobro, G. C., and Bryce, J. R. (2001). Regulatory action criteria for filth and other extraneous materials v. strategy for evaluating hazardous and nonhazardous filth. *Regul Toxicol Pharmacol.* 33, 363–392. doi: 10.1006/rtph.2001.1472

O'Neill, M. A., Gauld, I. D., Gaston, K. J., and Weeks, P. J. D. (2000). "Daisy: an automated invertebrate identification system using holistic vision techniques," in *Proceedings of the Inaugural Meeting BioNET-INTERNATIONAL Group for Computer-Aided Taxonomy (BIGCAT)*, 13–22.

Qing, Y. A., Jun, L. V., Liu, Q. J., Diao, G. Q., Yang, B. J., Chen, H. M., et al. (2012). An insect imaging system to automate rice light-trap pest identification. *J. Integr. Agric.* 11, 978–985. doi: 10.1016/S2095-3119(12)60089-6

Reinholds, I., Bartkevics, V., Silvis, I. C., van Ruth, S. M., and Esslinger, S. (2015). Analytical techniques combined with chemometrics for authentication and determination of contaminants in condiments: a review. *J. Food Compos. Anal.* 44, 56–72. doi: 10.1016/j.jfca.2015.05.004

Sáez, J. A., Luengo, J., and Herrera, F. (2016). Evaluating the classifier behavior with noisy data considering performance and robustness: the equalized loss of accuracy measure. *Neurocomputing.* 176, 26–35. doi: 10.1016/j.neucom.2014.11.086

Senyurek, V. Y., Imtiaz, M. H., Belsare, P., Tiffany, S., and Sazonov, E. A. (2019). comparison of SVM and CNN-LSTM based approach for detecting smoke inhalations from respiratory signal. *IEEE* 2019, 3262–3265. doi: 10.1109/EMBC.2019.8856395

Shi, Z., Dang, H., Liu, Z., and Zhou, X. (2020). Detection and identification of stored-grain insects using deep learning: a more effective neural network. *IEEE Access.* 8, 163703–163714. doi: 10.1109/ACCESS.2020.3021830

Shin, Y., and Balasingham, I. (2017). Comparison of hand-craft feature based SVM and CNN based deep learning framework for automatic polyp classification. *IEEE* 2017, 3277–3280. doi: 10.1109/EMBC.2017.80 37556

Shorten, C., and Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *J. Big Data.* 6, 60. doi: 10.1186/s40537-019-0197-0

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1929–1958.

Terry, J. C. D., Roy, H. E., and August, T. A. (2019). Thinking like a naturalist: enhancing computer vision of citizen science images by harnessing contextual data. *bioRxiv.* 730887. doi: 10.1101/730887

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. *Syst. Biol.* 68, 876–895. doi: 10.1093/sysbio/syz014

Wang, J., Lin, C., Ji, L., and Liang, A. A. (2012). new automatic identification system of insect images at the order level. *Knowledge-Based Syst.* 33, 102–110. doi: 10.1016/j.knosys.2012.03.014

Weeks, P. J., Gauld, I. D., Gaston, K. J., and O'Neill, M. A. (1997). Automating the identification of insects: a new solution to an old problem. *Bull. Entomol. Res.* 87, 203–211. doi: 10.1017/S0007485300 02736X

Wu, L., Liu, Z., Bera, T., Ding, H., Langley, D. A., Jenkins-Barnes, A., et al. (2019). A deep learning model to recognize food contaminating beetle species based on elytra fragments. *Comput. Electron. Agric.* 166, 105002. doi: 10.1016/j.compag.2019.105002

Yalcin, H. (2015). Vision based automatic inspection of insects in pheromone traps. *IEEE;* 2015, 333–338. doi: 10.1109/Agro-Geoinformatics.2015. 7248113

Yang, H.-P., Ma, C.-S., Wen, H., Zhan, Q.-B., and Wang, X.-L. (2015). A tool for developing an automatic insect identification system based on wing outlines. *Sci. Rep.* 5, 1–11. doi: 10.1038/srep12786

Zhang, X., Qiao, Y., Meng, F., Fan, C., and Zhang, M. (2018). Identification of maize leaf diseases using improved deep convolutional neural networks. *IEEE Access.* 6, 30370–30377. doi: 10.1109/ACCESS.2018.28 44405

Zheng, C., Sun, D.-., W., and Zheng, L. (2006). Recent developments and applications of image features for food quality evaluation and inspection-a review. *Trends Food Sci. Technol.* 17, 642–655. doi: 10.1016/j.tifs.2006. 06.005