Efficient Detection of Invasive Stink Bugs Using Convolutional Neural Networks

Alexandru Dinca
Faculty of Automatic Control and Computer Science
National University of Science and Technology Politehnica Bucharest
Bucharest, Romania
marius.dinca1411@stud.acs.upb.ro

Nicoleta Angelescu
Faculty of Electrical Engineering, Electronics, and Information Technology
Valahia University of Targoviste
Targoviste, Romania
nicoleta.angelescu@valahia.ro

Loretta Ichim
Faculty of Automatic Control and Computer Science s
National University of Science and Technology Politehnica Bucharest
Bucharest, Romania
loretta.ichim@upb.ro

Trandafir-Liviu Serghei
Faculty of Automatic Control and Computer Science
National University of Science and Technology Politehnica Bucharest
Bucharest, Romania
ORCID: 0000-0001-9181-0743

Dan Popescu
Faculty of Automatic Control and Computer Science
National University of Science and Technology Politehnica Bucharest
Bucharest, Romania
dan.popescu@upb.ro

Abstract—**Timely pest detection and identification is critical as part of modern agriculture. Halyomorpha Halys is a prevalent pest with proven harmful impacts on numerous crops and agricultural regions. The paper proposes an efficient model to improve the detection of two invasive stink bugs: Halyomorpha halys and Nezara Viridula. Automatic detection of these two bugs is essential in various fields, such as precision agriculture and integrated pest management. The high performances obtained in the present study open new perspectives for the further development of insect pest detection systems and can serve as a basis for future modifications and improvements of these models.**

Keywords—convolutional neural networks, insect pest detection, agriculture

I. Introduction

Two invasive stink bugs: Halyomorpha halys (HH) (brown marmorated stink bug) and Nezara Viridula (NV)(southern green stink bug), were introduced relatively recently in Europe. These species are very polyphagous and cause important damage to crops, orchards, and ornamental plants [1], [2]. To reduce economic damages produced by insects it is very important to detect and monitor them in the early phase, and then, use professional expertise to ecologically destroy these insects. Since the use of insecticides can have a negative impact on the environment and human health, an integrated management of pests through biological control is preferable.

Today, machine learning technology and, especially, neural networks are recommended to identify these insect species and their abundance to prevent the diseases they can cause as well as to monitor their spread in crops or orchards [3]. The creation and organization of the dataset followed the introduction of digital images illustrating the pest insects in various poses which were taken from public data sets, although in real applications particular databases adapted to the specifics of the application must be created and used

Although few, there are several recent research articles [4], [5], and international projects [6] that focus on HH detection using convolutional neural networks and UAVs.

Paper [3] investigates the potential of deep learning models to classify pest species with high interspecies similarity and intraspecies variability. It introduces a modified SSD model with an IoU value of 70.2% as a performance indicator for detecting four species of Phyllocephalidae insects. This approach has the potential to reduce costs, enhance performance, and increase scalability in pest monitoring and analysis. Finally, it was noted that monitoring the Pentatomidae family of pests is crucial in modern agriculture for identifying variations in infection levels and enhancing integrated pest management strategies. The study [4] emphasizes the necessity of early pest detection and identification in precision agriculture, with a particular focus on the common pest Halyomorpha Halys and its harmful impact on crops. To recognize the insect pests, four advanced neural networks were used in the study. Despite the inherent difficulty of automated insect identification in natural contexts, without the use of traps, the results are encouraging. The study of time and accuracy metrics provided validation accuracy values ranging from 85% to 88%, demonstrating that applying deep learning architectures for pest detection and classification in agriculture has tremendous potential and development chances. The study [5] focuses on detecting the brown marmorated stink bug (BMSB), a pest in fruit production. It compares three sophisticated image recognition and classification models: MobileNetV2, Xception, and EfficientNet. These models were slightly customized for the task and achieved good accuracy results, ranging from 96% to 97.5% in BMSB detection.

The highlighted information from the previous studies is that early insect pest detection, particularly focusing on Halyomorpha halys (the brown marmorated stink bug), is a significant and compelling research topic. The invasive nature of this pest and its ability to cause substantial damage to various crops underscores the importance of developing advanced detection methods. Deep learning approaches, as demonstrated in the discussed research, hold great promise in addressing this agricultural challenge by improving accuracy and efficiency in pest monitoring and management. Early

979-8-3503-0313-1/23/$31.00 ©2023 IEEE

insect pest detection using deep learning techniques, with a specific focus on Halyomorpha halys, is a highly relevant and valuable area of research in modern agriculture. The paper presents some performant neural networks adapted for HH and NV detection in precision agriculture. Considering that these insects are relatively small compared to the images obtained in the field, the obtained performances are good.

II. Materials and Methods

A. Datasets Used

The Xie. Dataset [7] and the Maryland Biodiversity Project database [8] were chosen for the dataset of the present work. Representative images from the datasets were taken for the reference insects Halyomorpha Halys and Nezara Viridula. Related to the topic of insect pest identification, the mentioned dataset was divided into 3 directories for training, validation, and testing. The ratio for organizing the data set aimed at percentages of 70% for testing, accumulating 764 images, 20% for validation - 109 images, and 10% for testing - 54 images. The total dataset for this paper comprised 927 images, with 488 images categorized under the Halyomorpha Halys class and 439 images representing Nezara Viridula. As input part for CNNs images chosen were resized to 224x224 pixels, using RGB images.

This dataset organization pattern is a typical practice for training and evaluating deep-learning models based on working with digital images. Finally, the representative classes that are attached to the data set were created by the acronyms of the two insects - we have two classes HH and NV. Examples of representative images for the mentioned dataset can be viewed in Fig. 1 for both datasets queried, Maryland and Xie. Dataset. Since the number of representative images for the organized pest classes in the dataset of the present work is relatively small, it was aimed to introduce augmentation operations for the size and robustness of the final dataset. For this paper, image augmentation techniques involved making controlled modifications to the input images gathered, and they included: flipping horizontally and vertically, crop, rotation between -15° and +15°, shear ±10° horizontal, ±10° vertical, brightness: between -25% and +25%, blur: up to 1.5px, random noise (salt and pepper): up to 2% of pixels. In this sense, image augmentation was chosen as having a positive impact on image classification tasks creating improved generalization, robustness, data efficiency, and translation invariance.

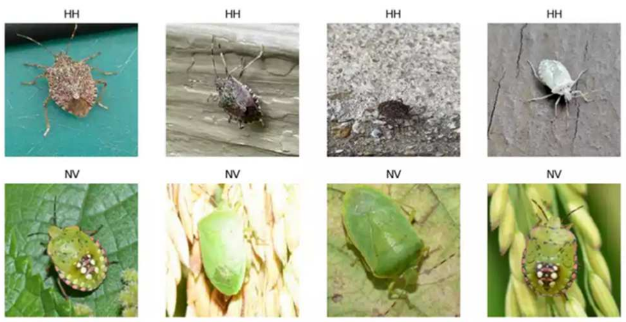

Fig. 1. Example images from the dataset created.

As a result, the CNN is exposed to a wider range of variation in the data, helping each CNN model to learn based on robust and generalized features. Translation, rotation, and cropping are augmentation approaches that assist CNNs in learning translation-invariant features. This is necessary for identifying insects in various locations within a picture. Finally, CNNs grow more resilient to numerous transformations, noise, and variances found in real-world settings through training on augmented data. This improves the model's ability to handle a variety of lighting situations, views, and object orientations. Example images from the augmented dataset for both classes are shown in the figure below. Fig. 2 shows representative samples that have undergone various data augmentation techniques to enhance the diversity and robustness of the dataset.

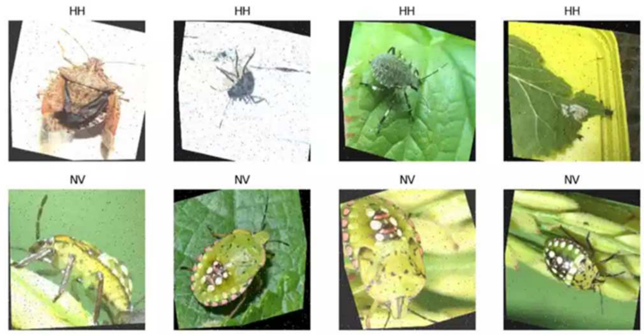

Fig. 2. Example images from the augmented dataset.

It is important to note that the images chosen to create the data set illustrate real poses of the reference insects, and no artificially generated images have been introduced as is the case in various approaches to working with digital images. The motivation of this fact denotes the orientation of the present work toward the automatic detection of harmful insects in real frames, usually represented by orchards and agricultural areas.

B. Neural Networks Used

The performance comparison of multiple ImageNet [9] pre-trained models was pursued. For the present study, the models used followed the implementation of several architectures as follows: VGG19 with Batch Normalization (BN) [10], ResNet152 [11], InceptionResNetV2 [11], DenseNet201 [12], Xception [13], MNasNet [14], NasNetLarge [15], and ConvNextLarge [16]. These models differ in depth, complexity, and architectural innovations, making them appropriate for a variety of applications and resource restrictions, especially as new trends for the modern digital agriculture [17].

Starting with the first one, VGG19 is a 19-layer deep CNN architecture. To stabilize and speed training, batch normalization is used. It features a straightforward and consistent design with 3x3 convolutional layers. Although useful for feature extraction, it may be computationally costly [10]. Next, ResNet152 is part of the ResNet (Residual Networks) family. This kind of architecture introduces skip connections named residual blocks, which aid in the resolution of the vanishing gradient problem. In general, ResNets are beneficial for highly deep networks, allowing hundreds of layers to be trained [11]. InceptionResNetV2 combines Google's Inception and ResNet concepts. It improves feature extraction by utilizing residual connections and inception modules. Suitable for applications requiring both depth and feature variety [11].

DenseNet201 is a deep neural network with a special type of dense layer connection. This model promotes feature reuse and mitigates the vanishing gradient issue. In terms of parameter utilization, it is quite efficient, and it is suitable for jobs requiring a small amount of data and resources [12]. Another model, Xception is an abbreviation for "Extreme Inception." It employs computationally efficient depthwise separable convolutions. With fewer settings, this model achieves competitive performance, and it is especially handy when computing resources are constrained [13].

In the same topic, MNasNet is a neural network design that is mobile-friendly. Designed for low-resource mobile and embedded devices, MNasNet is efficient and lightweight while retaining high accuracy for a wide range of operations [14]. NasNetLarge is an architecture for convolutional neural networks that stands for "Neural Architecture Search Network Large." It is a member of the NasNet model family. NasNetLarge was created utilizing the NAS approach, which includes searching for the optimum neural network architecture for a given job. This automated architectural search approach aids in the development of efficient and effective models. It strikes a good mix between model size and performance, making it an excellent choice for applications with limited resources [15]. Finally, part of the novel architectures, ConvNeXt combines principles from the transformer architecture into classic convolutional neural networks such as ResNet. The objective is to improve CNN performance for computer vision tasks. To accomplish this improvement, many major adjustments were made, including changes to the stem, the inclusion of inverted bottlenecks, changes to activation functions and normalizing layers, and the addition of distinct downsampling layers [16]. ConvNeXt combines the advantages of standard CNNs and transformers while requiring fewer parameters and computing resources, giving it a high-performance choice for a wide range of computer vision applications. This novel technique enables for straightforward deployment while outperforming conventional models [16].

C. Hardware and Software Used

To ensure the reliability and replication of our results, we used a custom system setup for our tests and research on deep learning for insect pest detection. Our tests were carried out on a computer running the Ubuntu 20.04 LTS operating system. The Python programming language, version 3.10.12, was used as the primary scripting language for our study. An Intel Core i9-11900K CPU, noted for its speed and multi-core capabilities, served as the system's computational backbone. PyTorch, version 2.0.1, was used for our deep learning research. PyTorch was chosen for its adaptability, broad community support, and full collection of deep-learning research tools. Next, we used the computing capability of an NVIDIA GeForce RTX 2080 Ti GPU with 11020MiB of video memory and CUDA to speed up the training and inference procedures of our deep learning models.

III. Experimental Results and Discussion

To note the results and discussion for the models used, one method is to use these models as feature extractors first, and then train a classifier on top of the retrieved features using a subset of the dataset. This method can aid in determining which model has the best feature extraction capabilities. Following that, each model may be fine-tuned using the complete dataset and its performance compared. To avoid overfitting during training, techniques like data augmentation and dropout were used. Each network was evaluated across a number of epochs, and the accuracy value attained on the validation dataset was noted for each.

For the evaluation of the models and part of the comparative study, several settings were followed in studying the impact of the chosen architectures. Starting from the created data set, convolutional neural network models were trained and evaluated for the automatic identification of the two classes HH and NV. Implementing numerous CNN models with transfer learning and fine-tuning is one technique for enhancing the accuracy and efficiency of insect categorization operations. Transfer learning is the process of fine-tuning a pre-trained CNN model for a specific task. This is performed by adjusting the weights of the pre-trained model to reflect the new dataset. As part of the training and evaluation, the model is moved to the GPU device, and the training is performed for a total of 30 epochs. As the optimization criteria, the Cross-Entropy Loss is utilized, which is often used in classification tasks to assess the difference between predicted and actual class probabilities. The neural network's parameters are optimized using Stochastic Gradient Descent with a learning rate of 0.001 and a momentum of 0.9. During training, the learning rate is modified using a step-wise decay technique. This aids model convergence by gradually reducing the step size for parameter updates. These options together configure the neural network training procedure.

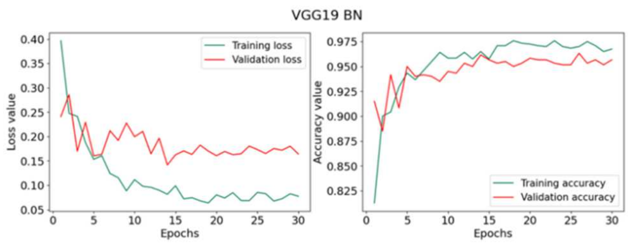

Fig. 3. VGG19 BN Results

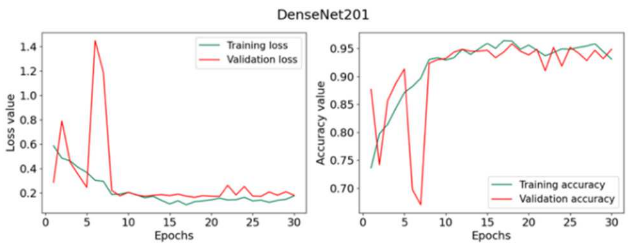

Fig. 4. DenseNet201 Results

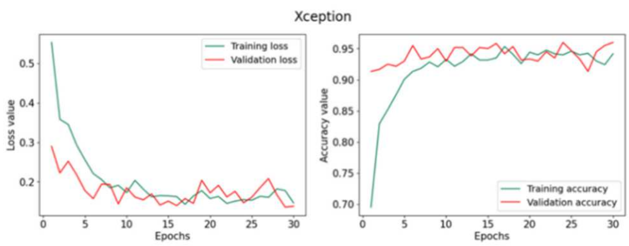

Fig. 5. Xception Results

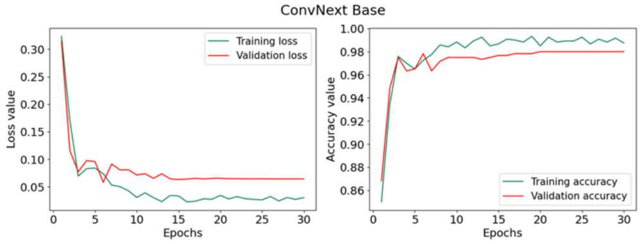

Fig. 6. ConvNeXt Base Results

The validation accuracy values obtained varied from 94% to 98%, indicating that deep learning architectures have a considerable influence and provide chances for advancement in pest detection and categorization. Table 1 displays the metrics for the highest validation accuracy achieved. The successful performance of such models in the task of recognizing the harmful insects from digital photos is observed using the initial set of data and the results obtained. To improve performance, the models may be tuned on a large data set in the future. Furthermore, new models may be combined, and approaches like combining neural networks utilizing different ensemble methods can be used to generate fusion models by integrating decisions of independently trained and evaluated models.

TABLE I. VALIDATION ACCURACY METRICS FOR THE BEST MODELS

Model used	Best accuracy value
VGG19 BN	96.83%
DenseNet201	95.17%
Xception	94.83%
ConvNeXt Base	98.00%

IV. CONCLUSIONS

This study optimizes and compares deep neural networks for the identification and categorization of the pests Halyomorpha Halys and Nezara viridula with a focus on the HH pest. The models and architectures utilized have the potential to advance insect identification and categorization, notably in agriculture. The good loss and accuracy results indicate that these models can be successfully used to automate pest monitoring and identification activities. More instances can be added to the dataset to improve the work. Furthermore, in the future, approaches such as model fusion might be investigated and used to optimize architectures and enhance performance, especially for increasingly complicated structures. Finally, the most effective model for insect pest detection, particularly HH, would be determined by the unique demands and limits of the real-world application, as well as the availability of data and computing resources. Based on the present results, the models can be developed along the way to improve the identification and automatic detection part, considering real scenarios and personal contributions on the applicability part.

ACKNOWLEDGMENT

This work was supported by HALY.ID project. HALY.ID is part of ERA-NET Co-fund ICT-AGRI-FOOD, with funding provided by national sources [Funding agency UEFISCDI, project number 202/2020, within PNCDI III] and co-funding by the European Union's Horizon 2020 research and innovation program, Grant Agreement number 862665 ERA-NET ICT-AGRI-FOOD (HALY-ID 862671).

REFERENCES

[1] M. Kavran, A. Konjević, D. Petrić, and A.I. Ćupina, "The introduction and establishment of four invasive insect species in Serbia", Insects 2023, 14, 728.

[2] R Ciceoi, M Dumbrava, IO Jerca, CM Pomohaci, I Dobrin, "Assessment of the damages on maize crop by the invasive stink bugs Halyomorpha halys (Stål, 1855) and Nezara viridula (Linnaeus, 1758)(Hemiptera: Pentatomidae)", Acta Zool Bulg. 2017, Suppl 9, pp. 211-217

[3] R. Trufelea, M. Dimoiu, L. Ichim, and D. Popescu, "Detection of Harmful Insects for Orchard Using Convolutional Neural Networks", U.P.B. Sci. Bull., Series C, Vol. 83, Iss. 4, 2021, pp. 85-96.

[4] A. Dinca, N. Angelescu, L. Ichim, and D. Popescu, "Deep Neural Networks for Halyomorpha Halys Detection," 31st Mediterranean Conference on Control and Automation (MED), Limassol, Cyprus, 2023, pp. 317-322.

[5] A. Vicol, L. Ichim, and D. Popescu, "Brown Marmorated Stink Bug Detection Using Multiple Neural Networks," 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Tenerife, Canary Islands, Spain, 2023, pp. 1-6

[6] HALYomorpha halys IDentification: Innovative ICT tools for targeted monitoring and sustainable management of the brown marmorated stink bug and other pests, https://www.haly-id.eu/, accessed on 6 September 2023.

[7] C. Xie, et al., "Multi-level learning features for automatic classification of field crop pests", Comput. Electron. Agric. 152, 2018/09/01, pp. 233–241.

[8] Maryland Biodiversity Database, Online, https://www.marylandbiodiversity.com, Accessed: September 2023.

[9] J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, and F.F. Li, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248-255, 2009.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," International Conference on Learning Representations (ICLR), San Diego, CA, 2015.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770–778.

[12] G. Huang, et al., "Densely connected convolutional networks" In Proceedings of the IEEE Conference on computer vision and pattern recognition, pp. 4700-4708, 2017.

[13] F. Chollet, "Xception: Deep learning with depthwise separable convolutions", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1251-1258).

[14] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q.V. Le, "Mnasnet: Platform-aware neural architecture search for mobile", In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 2820-2828.

[15] B. Zoph, V. Vasudevan, J. Shlens, and Q.V. Le, Q. V., "Learning transferable architectures for scalable image recognition", In Proceedings of the IEEE Conference on computer vision and pattern recognition, 2018, (pp. 8697-8710.

[16] Z. Liu, H. Mao, C.Y. Wu, C. Feichtenhofer, T. Darrell, T., and S. Xie, " A convnet for the 2020s", In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 11976-11986.

[17] S. Coulibaly, B. Kamsu-Foguem, D. Kamissoko, and D. Traore, "Deep learning for precision agriculture: A bibliometric analysis", Intelligent Systems with Applications, Volume 16, 2022, 200102.