

DEVELOPMENT AND TESTING OF AN AI-BASED PESTICIDE RECOMMENDER SYSTEM

Juvelekian, Alexander G., El Hajj, Ahmad A., Raffoul, Nour G., Khodor, Hashem M., Charif, Yehya M., and Khater, Christopher K.¹

Abstract— Pesticide overuse poses a long-standing threat to Lebanese public health, and is a primary source of water and soil poisoning. A major reason behind pesticide overuse is the overhead it reduces when it comes to pest detection. The main strategy in Lebanese rural areas seems to be widespread and unorganized spraying of a wide variety of pesticides, some of which have been banned or are now deemed obsolete, rather than strategic and limited application. Automating pest detection through AI-based solutions could empower farmer to reduce pesticide use, through the promotion of targeted pesticide application. By limiting pesticides to affected areas, and the type of pesticide to the detected threat, the amount of pesticide residues found in both soil and water could finally be lowered, improving both public health and local ecosystems. This paper details the implementation of a pest detection algorithm, as well as the development of a pesticide recommender system, which could further assist farmers in making responsible and adequate pesticide choices. The limitations imposed by the computational power available appear to severely impact the performance of the pest detection algorithm, although existing literature indicates that the idea is viable. The pesticide recommender appears to be a promising way to assist farmers in choosing pesticide, and could be further developed into a fully-fledged tool.

I. PROBLEM STATEMENT

PESTICIDE overuse in Lebanon is an old and complicated issue. In 1978, a paper by A.S. Talhouk mentioned in its abstract the “messy an unorderly use of pesticides to combat pests in apple farms in Lebanon” [2]. Little seems to have changed since, as a June 2022 report by the FAO (Food and Agriculture Organization) notes that Lebanon has one of the highest uses, per hectare, of fertilizers (331 kg/ha) and pesticides (7 kg/ha) in the world [3].

Today, agriculture is one of the most productive sectors of the Lebanese economy. In Akkar for example, agriculture takes up nearly half of the land [4]. According to the FAO, agriculture-related activities in the rural areas of Lebanon account for up to 80% of the local GDP [5]. It is therefore especially devastating to note that this region has recently witnessed several pest infestations, such as several new invasive species targeting oak trees [6]. This problem has been exacerbated by the temperature changes brought about by climate change [7].

Given the persistent threat that pests pose to crop production, it makes sense that farmers would then resort to pesticides. Usually, to prevent profit-seeking from undermining safety, concerned organizations would set strict regulations regarding pesticide use. In the United States for example, it is the Environmental Protection Agency that overviews pesticides, in accordance with federal laws like the Federal Insecticide, Fungicide, and Rodenticide Act [8]. While regulations do exist in Lebanon, there is no clear unified law. Instead, pesticide management jurisdictions are divided among various ministries [7]. On top of the scattered nature of the available legislation, the plans and strategies developed by the Ministry of Agriculture to reduce pesticide use are not being followed, and evaluations based on assessments do not show further achievements or accomplishments [7]. The attitude of laissez-faire when it comes to public safety and environmental concerns is so drastic that illegal or obsolete pesticides are routinely smuggled into the country [9] and used on several crops, such as Loquat [10].

Monitoring crops is a labor-intensive process, and infestations can spread quickly and have devastating economic and environmental consequences in no time. The difficulty of adequate monitoring might be why many farmers prefer the simpler option of indiscriminately spraying pesticide, as economic factors are one of the driving forces behind pesticide dependency [11]. Nonetheless, pesticides represent a long-term threat to the environment and to public health. They can lead to increased amounts of heavy metals in soil as well as an increased nitrification rate, both of which are detrimental to soil health and constitute an ecological hazard [12]. Pesticides can also damage natural resources that many civilians rely on. Indeed, the groundwater in Akkar is significantly contaminated by nitrates and nitrites due to agricultural fertilizers, and by a consequential amount of OCPs (Organochlorine Pesticides), and OPPs (Organophosphate Pesticides) [4]. Those compounds are associated with neurological disorders, liver tumors, fertility risks, and many other health concerns [13]. They also affect birds, farm animals, and reptiles, which gives them the potential to disrupt entire ecosystems [13]. Clearly, excessive reliance on pesticides will not diminish unless farmers are provided with suitable solutions which encourage responsible and safe pesticide use, while still protecting crop yields.

This paper serves as proof of concept for a proposed alternative: AI-pest detection through drone monitoring, supplemented with a pesticide recommender system.

¹ Authors are listed in increasing order of height. Studies have shown that authors with alphabetically “early” last names are cited more often [1], so this choice aims to add some genetic randomness to the citation process to hopefully correct this bias. The caveat of a more widespread use of this citation order is the obvious fact that men cannot be trusted to state their height honestly.

AI technology, specifically image processing and recognition, can save on human labor and to improve detection [14], as AI can identify and filter through data faster and with greater accuracy than people [15]. Models can be developed and trained on databases specific to the local agricultural landscape, using satellite images or data collected by UAVs (Unmanned Aerial Vehicles). Those same tools used to generate training data can then be used as monitoring devices, and the now-trained AI model can identify visual signs of a pest infestation and allow the farmers to spray pesticides only when needed.

The choice of which pesticide to spray is further informed by an item-based recommender system, which takes into account a variety of metrics gathered on each kind of pesticide, and uses the individualized pesticide history of each farmer to select the most suitable one.

II. METHODOLOGY

The proposed framework combines two machine learning models are combined: a Convolutional Neural Network (CNN) and a variation of a content-based recommender system.

The system works as follows: the user inputs an image of a plant that contains a pest. This image is passed through the CNN which gives us up to 3 predictions. These predictions are then passed through a database that maps pests to appropriate pesticides, from which all relevant pesticides are filtered. Those pesticides are then passed to the recommender system, which, based on user purchase history, recommends the top K pesticides from the input.

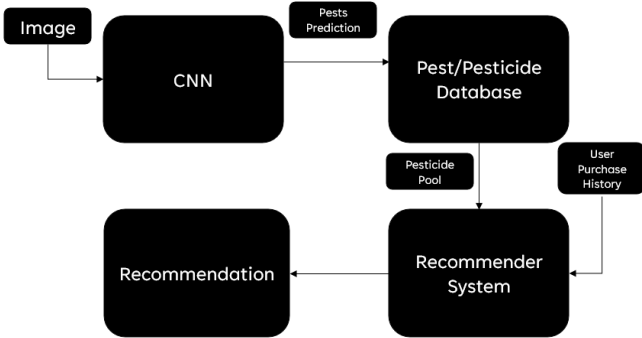


Fig. 1. Main architecture of the project

A CNN is chosen for pest detection, as CNNs are famously well-suited to image detection applications.

CNNs are a kind of Neural Networks, comprised of three types of layers: convolutional layers, pooling layers, and fully connected layers. It is the specific architecture of the CNN which allows it to perform image detection.

Convolutional layers perform feature extraction. In a convolutional layer, a filter of a determined size slides over the image and performs a convolution (hence the name of the

layer). Different types of filters perform different actions (some filters might, for example, detect vertical edges). Early convolutional layers (those closer to the input of the CNN) extract low level features such as edges, while later convolutional layers (those closer to the output) detect high level features such as faces. A typical activation function for a convolution layer is the RELU function, which replaces negative values by 0, and passes positive values as they are.

Pooling layers simply reduce the size of the input they are given, which both improves the robustness of the model and makes computations faster. They “summarize” the features extracted by convolutional layers.

Fully-connected layers are typically placed at the output of the CNN. These layers perform classification, so they attribute a label to the image based on the extracted features. The activation function of the last layer is typically the softmax function, which can be interpreted as choosing for the image the label with the highest associated probability. [16]

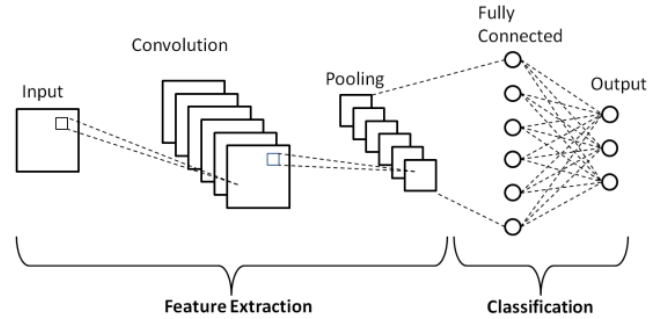


Fig. 2. Schematic diagram of a basic CNN [17].

For the second part, a recommender system is implemented. It takes as input a list of pesticides, which is obtained according to the output of the CNN (the detected pest). The recommender uses cosine similarity to equate the average similarity of each pesticide with the purchase history of the user, and returns the top K most-likely-to-be-purchased pesticides. The idea seems novel, as a preliminary review of the literature showed no implementation of a similar concept. Most papers claiming to implement AI-based pesticide recommendation simply use AI to detect pest, and the pesticide recommendation is done in a deterministic manner, rather than being dynamically update according to purchase history.

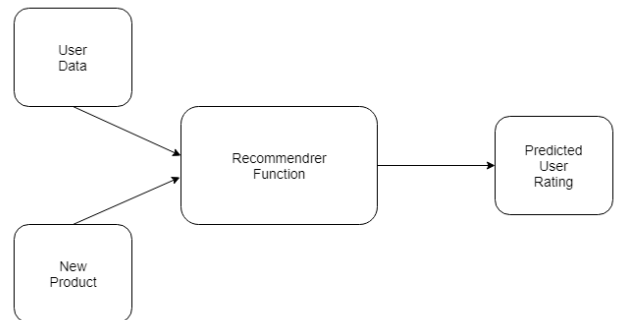


Fig. 3. Schematic diagram of a simplified Recommender System [18]

III. DATA COLLECTION AND PREPROCESSING

Three datasets are used to train the full model. Dataset 1 is used to train the CNN algorithm, while datasets 2 and 3 are used to train the recommender system.

Dataset 1

Dataset 1 is taken directly from “Pestopia: Indian Pesticides and Pesticides Datasets”, uploaded to Kaggle by user Feature_Finder. This dataset is made up of two directories, Pest_Dataset and Pesticide_Dataset. The former, which consists of a total of 56,685 labelled images of 132 different types of pests, is used to train the CNN pest detection algorithm.

Due to the sheer amount of images and lack of computational power available, the size of the dataset had to be reduced in order for the CNN algorithm to run smoothly.

Different numbers of pests and training images per pest are selected, in an attempt at striking a balance between speed and accuracy. This endeavor is further detailed in section IV.



Fig. 4. Example of a Gall Fly image in Dataset 1

Dataset 2

Dataset 2, an excerpt of which is shown in Table II for the Jute aphid and Rice Stemfly only, is adapted from the second directory of the Pestopia dataset (an excerpt of which is shown in Table I). While this dataset does provide a list of the most commonly used pesticides for each pest in Dataset 1, several processing steps are needed in order to make the dataset internally consistent.

The case of the Rice Stemfly can be taken as an example of the inconsistencies within the dataset. As shown in Table I, Clothianidin is listed as a possible pesticide for the Jute Aphid, but not for the Rice Stemfly. This might lead to the assumption that Clothianidin must not be used for the Rice Stemfly. “Neonicotinoids”, though, are listed, and Clothianidin belongs to the class of neonicotinoid pesticides, meaning it is actually adequate to use Clothianidin in this case. Imidacloprid is also mentioned for the Rice Stemfly, even though it is a kind of neonicotinoid, making its presence redundant. This is something an uninformed user of the Dataset might easily miss, if simply naively iterating through it. Most of the pesticides listed are individual active ingredients, except for Neonicotinoids, Pyrethroids, and Organophosphates, which are classes of pesticides. The choice to write down both individual active ingredients and whole classes of pesticides makes the table unusable as is. Instead, those three classes of pesticides

are replaced by the individual active ingredients that belong to it. Neonicotinoids, for example, are replaced when they are mentioned by Acetamiprid, Dinotefuran, Thiamethoxam, Clothianidin and Imidacloprid, which are the individual neonicotinoids that appear in the table. The list format is also replaced by a one-hot encoding table, as shown in Table II.

A choice is also made to remove the 16 active ingredients which work only for a total of one or two pests. This choice is made under the assumption that a versatile pesticide is preferable to a highly specific one. This is also done to reduce the total number of active ingredients under consideration as a way to shorten the data collection stage, as a significant amount of data must be collected by hand for each active ingredient in order to build Dataset 3. Fenthion, which for example only works against one pest (*Dacus dorsalis*), is removed from the list of pesticides under study. It was of course verified that removing those 16 pesticides did not leave any pest with no corresponding pesticide.

TABLE I

Excerpt from the second directory of the Kaggle Pestopia dataset

Pest name	Most commonly used pesticide
...	
Jute aphid	Thiamethoxam, Imidacloprid, Clothianidin
Jute hairy	Imidacloprid, Chlorpyrifos, Lambda-cyhalothrin
...	
rice shell pest	Cartap, Methomyl, Carbaryl
Rice Stemfly	Neonicotinoids, Fipronil, Imidacloprid
...	

TABLE II

Excerpt from Dataset 2

Pesticide / Pest	Jute aphid	Rice Stemfly
Acetamiprid		1
Clothianidin	1	1
Dinotefuran		1
Fipronil		1
Imidacloprid	1	1
Thiamethoxam	1	1
...		

Dataset 3

Dataset 3, finally, is fully built by hand. It consists of information gathered on each active ingredient seen in Dataset 2 (having removed those that target a total of one or two pests only). The metrics under consideration are versatility, cost, environmental impact, impact on non-invasive species, efficacy and dangers to human health.

Versatility is directly obtained from Dataset 2, as a measure of how many pests in total the pesticide is effective against. It is normalized through division by the maximum.

Cost is found by looking up each pesticide on DoMyOwn.com, and calculating the cost it would take to treat five acres of land with this compound. The calculation is done in several steps. First the cost per ounce is calculated. Then the dilution ratio (number of ounces to be diluted in 100 gallons of water) is taken from the manufacturer's usage manual. The total amount of water to be sprayed on five acres of land is either found from the manual as well, or, when not mentioned by the manufacturer at all, taken to be 100 gallons. For dry pesticides, which are not diluted, the amount necessary to treat five acres is taken directly from the manual. These amounts obviously vary with the severity of the infestation and the nature of the pest threat. An informed average is taken here, as a mild infestation is assumed, and the type of pest is informed by the list of pests in Dataset 2 for which each compound is listed. The cost is then normalized through division by the maximum.

Six compounds, either because they were banned in the US, withdrawn, or are deemed obsolete, were not listed active ingredients in any pesticide. The cost for these is noted as -1, and is later replaced by the average cost.

The impact on noninvasive species (such as fish or honey bees), the danger to human health, as well as the environmental impact of each pesticide, are taken from three databases: The University of California Integrated Pest Management program (UC IPM), the University of Hertfordshire Pesticide Properties DataBase (PPDB) and the University of Hertfordshire BioPesticides DataBase (BPDB). These databases give each pesticide an impact rating, ranging from low to high. A score of 0 is given for low impact, a score of 1 for moderate, and a score of 2 for high, with scores of 0.5 and 1.5 being allowed as well to indicate an impact in between low and moderate or moderate and high respectively. Each active ingredient is found on at least one of these three databases. When more than one database provides a score, a "pessimistic average" is taken and rounded to the nearest half integer, with the worst (so highest) score being given twice the weight of the others. This is done in the name of caution, as carefulness is always preferred when it comes to environmental and health precautions. The databases used to calculate the score of each compound are kept track of in the dataset.

The efficacy score is calculated in a more complicated fashion. The three databases (UP ICM, PPDB and BPDB) do not consistently give an efficacy score to all available pesticides, so a significant amount of compounds are left

without one. Several papers, referenced separately in both Dataset 3 and Appendix A, are analyzed in order to determine an accuracy score for these pesticides. These papers were also used to verify and potentially update the efficacy score when the databases did provide one. The final score given is 0 for low efficacy, 0.5 for low to moderate efficacy, 1 for moderate, 1.5 for moderate to high and 2 for high efficacy. The sources used to determine each score, as well as the changes made to the database score after conducting research, are kept track of in the dataset.

IV. TRAINING, SIMULATION RESULTS AND OPTIMIZATION

Pest Detection

Several issues are encountered while training the CNN. The size of Dataset 1, as mentioned in section II, exceeds available computational power. The dataset has to be limited to five pest categories, with about 200 images per pest, in order to prevent crashing of the program.

Attempts to include more types of pests were made at first, but the testing accuracy obtained was abysmal. For 20 kinds of pests, with 100 images each, the accuracy obtained on the validation set was 4.5%. Five pests seems like the most realistic number to consider.

A CNN algorithm, whose architecture is detailed in fig. 5, is trained on a total of 897 images: 186 images of *Adristyrannus*, 200 images of *Aleurocanthus spiniferus*, 200 images of the alfalfa plant bug, 111 images of the alfalfa seed chalcid, and 200 images of the alfalfa weevil. The evolution of the training and testing accuracies with each epoch (iteration through the entire training set) is shown in fig. 6. After 19 epochs, the obtained training accuracy is 88.3%, and the testing accuracy is 52.5%. The training accuracy consistently goes down with each epoch, while the testing accuracy stagnates and even starts increasing, indicating that the model is likely overfitting the data. The confusion matrix for this run, which details the correct and wrong classification predictions, is shown in fig. 7. It can be observed that class 0 is often misclassified as 4, and class 4 is often misclassified as 2. Class 3 seems to be classified almost at random by the algorithm.

Convolutional Neural Network Architecture
Model: "sequential_3"

Layer (type)	Output Shape	Param #
conv2d_9 (Conv2D)	(None, 110, 110, 32)	896
max_pooling2d_9 (MaxPooling2D)	(None, 55, 55, 32)	0
conv2d_10 (Conv2D)	(None, 53, 53, 64)	18496
max_pooling2d_10 (MaxPooling2D)	(None, 26, 26, 64)	0
conv2d_11 (Conv2D)	(None, 24, 24, 128)	73856
max_pooling2d_11 (MaxPooling2D)	(None, 12, 12, 128)	0
flatten_3 (Flatten)	(None, 18432)	0
dense_6 (Dense)	(None, 512)	9437696
dropout_5 (Dropout)	(None, 512)	0
dense_7 (Dense)	(None, 5)	2565

=====
Total params: 9533509 (36.37 MB)
Trainable params: 9533509 (36.37 MB)
Non-trainable params: 0 (0.00 Byte)

Fig. 5. Architecture of the CNN model for run 1

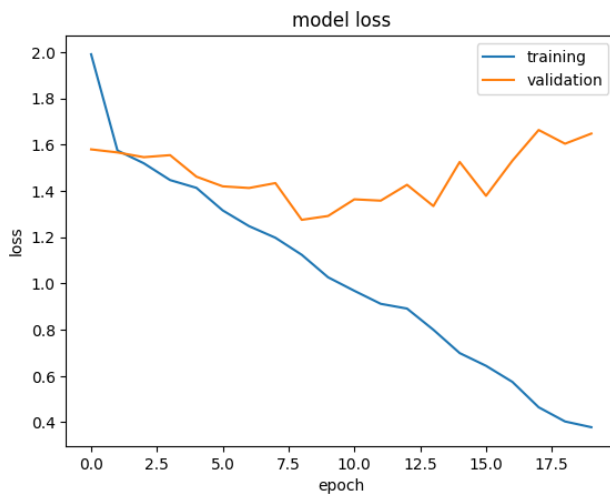


Fig. 6. Training and validation errors for run 1

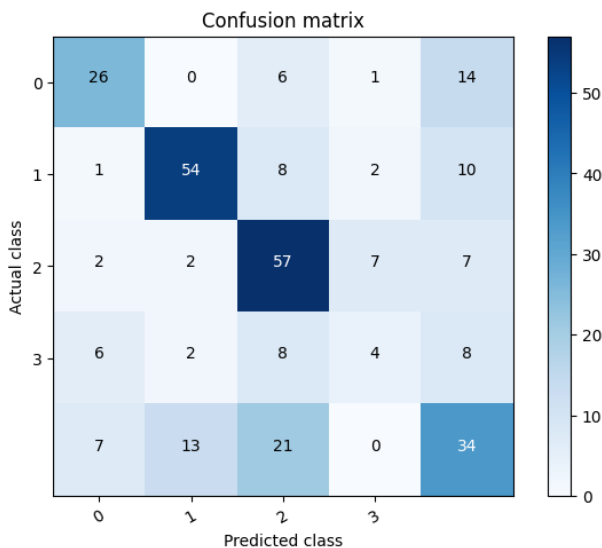


Fig. 7. Confusion matrix for run 1.

The CNN architecture is subsequently updated in an attempt to improve prediction accuracy. The architecture is as shown in fig. 8, with vgg16 being a deep convolutional layer imported from Keras, and containing thirteen convolutional layers and three fully connected ones. The model ends up performing worse, with a final training accuracy of 63.2%, and testing accuracy of 40.5%. The confusion matrix (fig. 10) shows that the misclassification of class 4 has worsened, and the model most often predicts classes 1 and 2 instead.

Model: "sequential_5"

Layer (type)	Output Shape	Param #
vgg16 (Functional)	(None, 3, 3, 512)	14714688
flatten_5 (Flatten)	(None, 4608)	0
dense_11 (Dense)	(None, 1024)	4719616
dropout_8 (Dropout)	(None, 1024)	0
dense_12 (Dense)	(None, 512)	524800
dropout_9 (Dropout)	(None, 512)	0
dense_13 (Dense)	(None, 132)	67716

=====
Total params: 20026820 (76.40 MB)
Trainable params: 5312132 (20.26 MB)
Non-trainable params: 14714688 (56.13 MB)

Fig. 8. Architecture of the CNN model for run 2

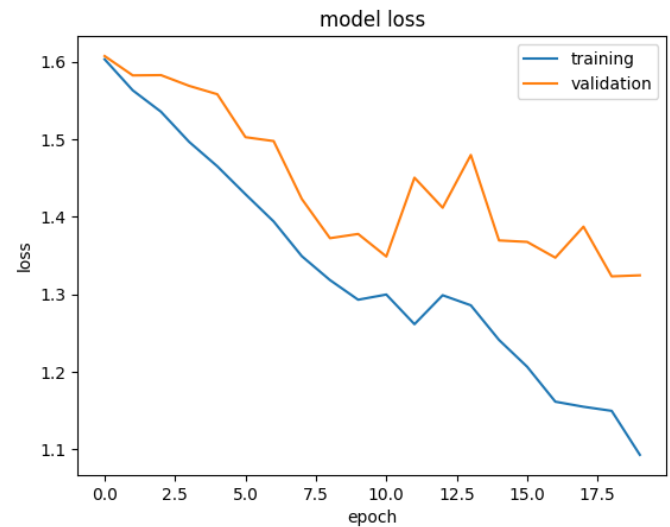


Fig. 9. Training and validation errors for run 2

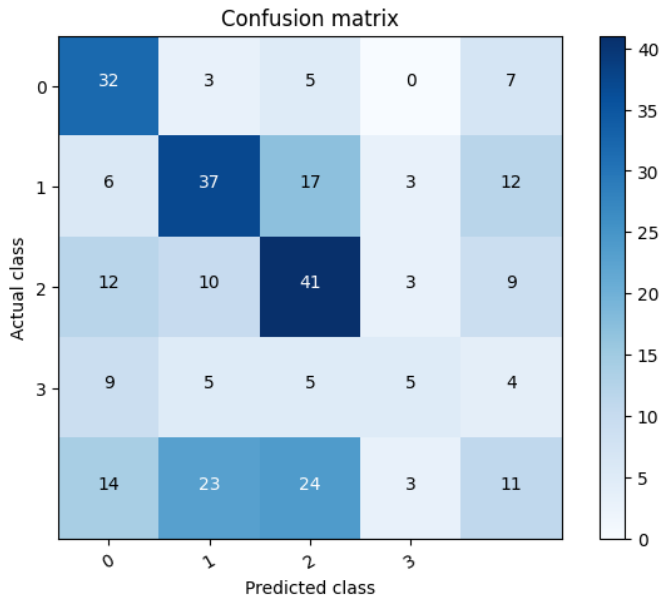


Fig. 10. Confusion matrix for run 2

There are several reasons to consider for the bad performance of the CNN. First and foremost, the quality of the dataset was scrutinized, with some of the images including text, others being drawings rather than pictures, or some even including several images in one, as seen in fig. 11, fig. 12 and fig. 13 respectively. This inconsistency in the image format, especially when a smaller number of images per pest is randomly selected in order to reduce dataset size, might be behind the unsatisfactory accuracy.

The images also are at varying closeness to the insect in question. Some images are close-ups, while others are taken from much farther away. Different life stages are also grouped together for the same insect, like for the Jute hairy, which is at times shown in its caterpillar stage and at times in its moth stage, as in fig. 14 and fig. 15 respectively. It's unclear how much this variation affects the CNN's performance. Nonetheless, an attempt at manually selecting more uniform training images is made.

All in all, by taking a higher quality subset of the dataset, consisting of more consistent and similar images of each pest, a better accuracy can be obtained.



Fig. 11. Example of an Aleurocanthus spiniferus image

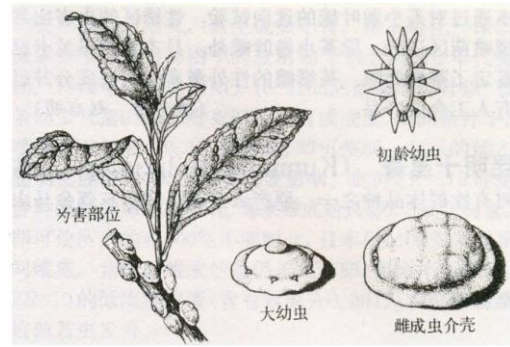


Fig. 12. Example of a Ceroplastes rubens image



Fig. 13. Example of a fruit piercing moth image



Fig. 14. Example of a Jute hairy image, with the insect in its caterpillar life stage

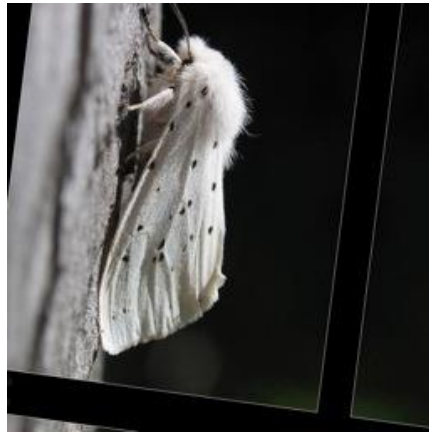


Fig. 15. Example of a Jute hairy image, with the insect in its moth life stage

On the third run, a total of 862 images are taken: 200 for the Amelophaga, 79 for the beet spot flies, 189 for the field cricket, 200 for the jute hairy, and 200 for the jute stem girdler. The data is, this time, curated by hand. The obtained training and testing accuracies are 23% and 27% respectively for 19 epochs. The model seems to be underfitting the data this time, as both accuracies are barely better than random

guessing. This model is ran for more iterations than the previous 2, with up to 200 epochs. The testing accuracy does not improve above 27%, and rather gets significantly worse as the model overfits the data more. The training accuracy, on the other hand, reaches over 99%.

Convolutional Neural Network Architecture
Model: "sequential_11"

Layer (type)	Output Shape	Param #
conv2d_33 (Conv2D)	(None, 109, 109, 512)	25888
max_pooling2d_33 (MaxPooling2D)	(None, 54, 54, 512)	0
conv2d_34 (Conv2D)	(None, 51, 51, 256)	2097408
max_pooling2d_34 (MaxPooling2D)	(None, 25, 25, 256)	0
conv2d_35 (Conv2D)	(None, 22, 22, 128)	524416
max_pooling2d_35 (MaxPooling2D)	(None, 11, 11, 128)	0
flatten_11 (Flatten)	(None, 15488)	0
dense_58 (Dense)	(None, 1024)	15860736
dropout_46 (Dropout)	(None, 1024)	0
dense_59 (Dense)	(None, 512)	524800
dropout_47 (Dropout)	(None, 512)	0
dense_60 (Dense)	(None, 512)	262656
dropout_48 (Dropout)	(None, 512)	0
dense_61 (Dense)	(None, 512)	262656
dropout_49 (Dropout)	(None, 512)	0
dense_62 (Dense)	(None, 512)	262656
dropout_50 (Dropout)	(None, 512)	0
dense_63 (Dense)	(None, 5)	2565

Total params: 19822981 (75.62 MB)
Trainable params: 19822981 (75.62 MB)
Non-trainable params: 0 (0.00 Byte)

Fig. 16. Architecture of the CNN model for run 3

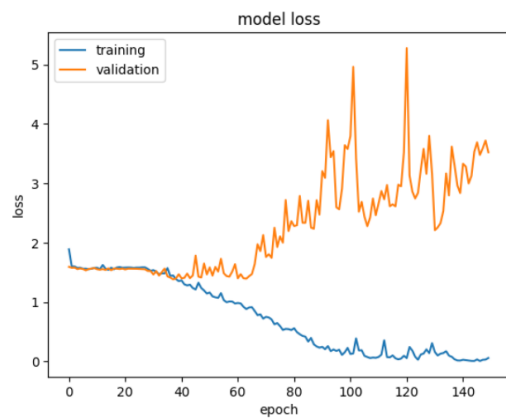


Fig. 17. Training and validation errors for run 3

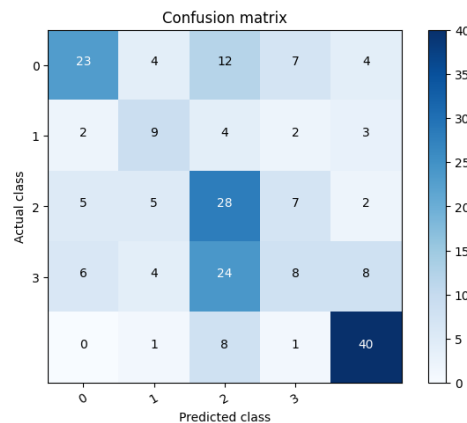


Fig. 18. Confusion matrix for run 3

Pesticide Recommender

The pesticide recommender is developed according to the detailed methodology. A sample run, with a made up user history, is shown in fig. 19.

```
Pesticide pool:
{'Adiristyrannus': ['Acephate', 'Malathion', 'Permethrin'],
 'alfalfa plant bug': ['Chlorpyrifos', 'Imidacloprid', 'Lambda-cyhalothrin']}

For user with history:
{'Bacillus thuringiensis': 3,
 'Carbaryl': 7,
 'Chlorpyrifos': 2,
 'Permethrin': 10}

The recommendations sorted in non-increasing order for each pest are:
pest : recommendation
Adiristyrannus : Permethrin, Malathion, Acephate
alfalfa plant bug : Lambda-cyhalothrin, Imidacloprid, Chlorpyrifos

Raw Similarities Scores
[('Permethrin', 2.05260857672653), ('Malathion', 2.0496886224813773), ('Acephate', 1.9622628451587834)]
[('Lambda-cyhalothrin', 2.0550246472851215), ('Imidacloprid', 2.0519714971481946), ('Chlorpyrifos', 2.021229623188594)]
```

Fig. 19. Sample run of the recommender

V. DISCUSSION AND CONCLUSION

The CNN model provides overall disappointing results. The testing accuracy barely exceeds 50%, and the model seems to consistently overfit the data. Choosing the training images by hand does not help improve accuracy. Of course, this could simply mean that not enough image were curated by hand, and computational limitations might after all be the main problem. A different study also attempting pest detection through CNNs has established that an accuracy of over 90% can be reached [19], although the dataset considered contained 12600 images, which goes way beyond the data used in this paper. Another study, computing the performance of the RetinaDet, EfficientDet, YOLOv5, Yolov8, FasterRCNN and MaskRCNN models when tested on the IDADP dataset, obtained accuracy scores ranging from 81% to 91% [20], but this dataset once again contains 3622 images, divided amongst 7 categories only.

The recommender system seems to perform well, although, unlike the CNN, there is no hard-set way to test it without real-life implementation. A recommendation is more subjective task, unlike classification which can be determined to have failed or succeeded with certainty. Nonetheless, the results appear promising, and further research might prove the idea useful.

As a preliminary example of how the idea might be further developed into a user-friendly tool, a website is developed and attached separately, as part of the wider course project.

REFERENCES

- [1] Abramo, G., & D'Angelo, C. A. (2017). Does your surname affect the citability of your publications? *Journal of Informetrics*, 11(1), 121-127. <https://doi.org/10.1016/j.joi.2016.12.003>
- [2] Talhouk, A. S. (1978). Contributions to the knowledge of almond pests in East Mediterranean countries: VIII. The establishment of an integrated program of almond pest management in Lebanon 1. *Zeitschrift für Angewandte Entomologie*, 87(1-4), 1-14.
- [3] FAO. (2022, June). Pesticides contamination and exposure reduction. Retrieved from <https://www.fao.org/documents/card/fr/c/cb9953en/>

- [4] Al Haj Ishak Al Ali, Roukaya, Mondamert, L., Halwani, J., Jandry, J., Nassif, N., Shaban, A., Berjeaud, J., & Labanowski, J. (2023). Temporal evolution of organochlorine and organophosphate pesticide residues in wells in the Akkar region (Lebanon). *Environmental Monitoring and Assessment*, 195(1), 121-121. <https://doi.org/10.1007/s10661-022-10671-y>
- [5] FAO (2023). Lebanon at a Glance. Retrieved from <https://www.fao.org/lebanon/fao-in-lebanon/lebanon-at-a-glance/fr/>
- [6] Moussa, Z., Choueiri, E., & Hanna, A. (2021). Research Paper (Survey: Insects) New Invasive Insects Associated with Oak Forests in Lebanon. *Arab. J. Plant Prot*, 39, 164-172.
- [7] Abou Zeid, M. I., Jammoul, A. M., Melki, K. C., Abou Jawdah, Y., & Awad, M. K. (2020). Suggested policy and legislation reforms to reduce deleterious effect of pesticides in Lebanon. *Heliyon*, 6(12).
- [8] EPA (2023). About Pesticide Registration. <https://www.epa.gov/pesticide-registration/about-pesticide-registration>
- [9] Houssari, N. (2023, August 23). "Banned pesticides 'smuggled into Lebanon as cleaning products,' MP claims". *Arab News*. <https://arab.news/8u5z7>
- [10] Abou Zeid, M. I., Awad, M. K., Melki, K. C., Abou Jawdah, Y., & Jammoul, A. M. (2021). Pesticides residues on Loquat: A minor crop in Lebanon. *Food Control*, 130, 108297.
- [11] Hu, Z. (2020). What socio-economic and political factors lead to global pesticide dependence? a critical review from a social science perspective. *International Journal of Environmental Research and Public Health*, 17(21), 8119.
- [12] Mukhles, M. B., Rahman, M. M., Rana, M. R., Huda, N., Ferdous, J., Rahman, F. & Biswas, S. K. (2022). Effect of pesticides on nitrification activity and its interaction with chemical fertilizer and manure in long-term paddy soils. *Chemosphere*, 304, 135379.
- [13] Jayaraj, R., Megha, P., & Sreedev, P. (2016). Organochlorine pesticides, their toxic effects on living organisms and their fate in the environment. *Interdisciplinary toxicology*, 9(3-4), 90-100. <https://doi.org/10.1515/intox-2016-0012>
- [14] He, Y., Zeng, H., Fan, Y., Ji, S., & Wu, J. (2019). Application of deep learning in integrated pest management: A real-time system for detection and diagnosis of oilseed rape pests. *Mobile Information Systems*, 2019.
- [15] Ai, Y., Sun, C., Tie, J., & Cai, X. (2020). Research on recognition model of crop diseases and insect pests based on deep learning in harsh environments. *IEEE Access*, 8, 171686-171693. <https://doi.org/10.1109/ACCESS.2020.3025325>
- [16] Mishra, M. (2020, August 26). Convolutional Neural Networks, Explained. Towards Data Science. <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939>
- [17] Phung, V. H. (2019). A high-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets. *Applied Sciences*, 9(21), 4500. <https://doi.org/10.3390/app9214500>
- [18] An In-Depth Guide to How Recommender Systems Work. BuiltIn. <https://builtin.com/data-science/recommender-systems>
- [19] Selvaraj, M. G., Vergara, A., Ruiz, H., Safari, N., Elayabalan, S., Ocimati, W., & Blomme, G. (2019). AI-powered banana diseases and pest detection. *Plant Methods*, 15(1), 1-11. <https://doi.org/10.1186/s13007-019-0475-z>
- [20] Kang, H., Ai, L., Zhen, Z., Lu, B., Man, Z., Yi, P., Li, M., & Lin, L. (2023). A novel deep learning model for accurate pest detection and edge computing deployment. *Insects* (Basel, Switzerland), 14(7), 660. <https://doi.org/10.3390/insects14070660>

APPENDIX A

The papers used to estimate the efficacy score are referenced here in alphabetical order.

- [i] Agumba, S., Gimnig, J. E., Ogonda, L., Ombok, M., Kosgei, J., Munga, S., Guyah, B., Omondi, S., & Ochomo, E. (2019). Diagnostic dose determination and efficacy of chlorfenapyr and clothianidin insecticides against anopheles malaria vector populations of western kenya. *Malaria Journal*, 18(1), 243-243. <https://doi.org/10.1186/s12936-019-2858-z>
- [ii] Bi, J. L., & Toscano, N. C. (2007). Efficacy of spiromesifen against greenhouse whitefly (homoptera: Aleyrodidae) on strawberry. *Hortscience*, 42(2), 285-288. <https://doi.org/10.21273/hortsci.42.2.285>
- [iii] Kumar, S., & Pandey, A. K. (2022). Bio-efficacy of various insecticides against white grubs (coleoptera: Scarabaeidae) infesting sugarcane. *International Journal of Tropical Insect Science*, 42(5), 3319-3325. <https://doi.org/10.1007/s42690-022-00820-8>
- [iv] Leskey, T. C., Lee, D., Short, B. D., & Wright, S. E. (2012). Impact of insecticides on the invasive halyomorpha halys (hemiptera: Pentatomidae): Analysis of insecticide lethality. *Journal of Economic Entomology*, 105(5), 1726-1735. <https://doi.org/10.1603/EC12096>
- [v] Liu, T. (2004). Toxicity and efficacy of spiromesifen, a tetrone acid insecticide, against sweetpotato whitefly (homoptera: Aleyrodidae) on melons and collards. *Crop Protection*, 23(6), 505-513. <https://doi.org/10.1016/j.cropro.2003.10.006>
- [vi] Mihretie, A., Yimer, D., Wudu, E., & Kassaw, A. (2020). Efficacy of insecticides against african bollworm (helicoverpa armigera hubner) on chickpea (cicer arietinum) in the lowlands of wollo, northeastern ethiopia. *Cogent Food & Agriculture*, 6(1), 1833818. <https://doi.org/10.1080/23311932.2020.1833818>
- [vii] Ochiai, N., Mizuno, M., Mimori, N., Miyake, T., Dekeyser, M., Canlas, L. J., & Takeda, M. (2007). Toxicity of bifenazate and its principal active metabolite, diazene, to Tetranychus urticae and Panonychus citri and their relative toxicity to the predaceous mites, Phytoseiulus persimilis and Neoseiulus californicus. *Experimental & applied acarology*, 43(3), 181-197. <https://doi.org/10.1007/s10493-007-9115-9>
- [viii] Rossitto De Marchi, B., & Smith, H. (2021). Efficacy of buprofezin, pyriproxyfen and spirotetramat against bemisia tabaci MEAM1 nymphal field populations in florida. *Crop Protection*, 149, 105756. <https://doi.org/10.1016/j.cropro.2021.105756>