

IDENTIFYING IMAGES OF INVASIVE HYDRANGEA USING PRE-TRAINED DEEP CONVOLUTIONAL NEURAL NETWORKS

Belal A. M. Ashqar and Samy S. Abu-Naser*

*Department of Information Technology,
Faculty of Engineering & Information Technology,
Al-Azhar University-Gaza, Palestine
sabunaser@gmail.com*

Abstract— Invasive species are threatening habitats of native species in many countries around the world. The current methods of monitoring them depend on expert knowledge. Trained scientists visit designated areas and take note of the species inhabiting them. Using such a highly qualified workforce is expensive, time inefficient and insufficient since humans cannot cover large areas when sampling. In this paper, machine learning based approach is presented for identifying images of invasive hydrangea (a beautiful invasive species original of Asia) with a dataset that contains approximately 3,800 images taken in a Brazilian national forest and in some of the pictures there is Hydrangea. A deep learning technique that extensively applied to image recognition was used. Our trained model achieved an accuracy of 99.71% on a held-out test set, demonstrating the feasibility of this approach.

Keywords— Invasive Species, Classification, Deep Learning

1.INTRODUCTION

In North America, brown marmorated stink bugs are quickly becoming a nuisance to people in their homes and to the agriculture industry while cogongrass is now spreading through the Southeast, displacing native plants. It provides no food value for native wildlife, and increases the threat of wildfire as it burns hotter and faster than native grasses. These are just two invasive species of many which can have damaging consequences on the environment, the economy, and even human health. Despite widespread impact, efforts to track the location and spread of invasive species are so costly that they're difficult to undertake at scale. Because scientists cannot sample a large quantity of areas, some machine learning techniques are used in order to predict the presence or absence of invasive species in areas that have not been sampled. The accuracy of this approach is far from optimal, but still contributes to approaches to solving ecological problems.

Ecologists have hosted a Kaggle competition in order to check if computers can accurately identify invasive species in photos. In this work, we showed that transfer learning of Deep Convolutional Neural Network (CNNs Trained on imagenet dataset) does well in identifying images of invasive hydrangea. In computer vision, CNNs have been known to be powerful visual models that yield hierarchies of features enabling accurate segmentation. They are also known to perform predictions relatively faster than other algorithms while maintaining competitive performance at the same time [8].

Received: January 7, 2019

Reviewed: March 25, 2019

Accepted: April 2, 2019

* Corresponding Author



2.BACKGROUND

2.1 INVASIVE SPECIES

An invasive species is a species that is not native to a specific location (an introduced species), and that has a tendency to spread to a degree believed to cause damage to the environment, human economy or human health [1]. The criteria for invasive species have been controversial, as widely divergent perceptions exist among researchers as well as concerns with the subjectivity of the term "invasive" [2].

While all species compete to survive, invasive species appear to have specific traits or specific combinations of traits that allow them to out compete native species. In some cases, the competition is about rates of growth and reproduction. In other cases, species interact with each other more directly.

Researchers disagree about the usefulness of traits as invasiveness markers. One study found that of a list of invasive and noninvasive species, 86% of the invasive species could be identified from the traits alone.[3] Another study found invasive species tended to have only a small subset of the presumed traits and that many similar traits were found in noninvasive species, requiring other explanations[3]-[5]. Common invasive species traits include the following:

- Fast growth.
- Rapid reproduction.
- High dispersal ability.
- Phenotype plasticity (the ability to alter growth form to suit current conditions).
- Tolerance of a wide range of environmental conditions (Ecological competence).
- Ability to live off of a wide range of food types (generalist).
- Association with humans [6].
- Prior successful invasions [7].

2.2 DEEP LEARNING

Deep learning (also known as deep structured learning or hierarchical learning) is part of a broader family of machine learning methods based on learning data representations, as opposed to task-specific algorithms. Learning can be supervised, semi-supervised or unsupervised [9]-[11].

In deep learning, each level learns to transform its input data into a slightly more abstract and composite representation. In an image recognition application, the raw input may be a matrix of pixels; the first representational layer may abstract the pixels and encode edges; the second layer may compose and encode arrangements of edges; the third layer may encode a nose and eyes; and the fourth layer may recognize that the image contains a face. Importantly, a deep learning process can learn which features to optimally place in which level on its own. (Of course, this does not completely obviate the need for hand-tuning; for example, varying numbers of layers and layer sizes can provide different degrees of abstraction)[9],[12].

2.3 CNNs

In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of deep neural networks, most commonly applied to analyzing visual imagery. CNNs use a variation of multilayer perceptrons designed to require minimal preprocessing [13]. They are also known as shift invariant or space invariant artificial neural networks (SIANN), based on their shared-weights architecture and translation invariance characteristics [14],[15].

2.4 TRANSFER LEARNING

Transfer learning is a research problem in machine learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem. [16] For example, knowledge gained while learning to recognize cars could apply when trying to recognize trucks. This area of research bears some relation to the long history of psychological literature on transfer of learning, although formal ties between the two fields are limited.

2.5 IMAGENET

ImageNet is an image dataset organized according to the WordNet hierarchy. Each meaningful concept in WordNet, possibly described by multiple words or word phrases, is called a "synonym set" or "synset". There are more than 100,000 synsets in WordNet, majority of them are nouns (80,000+). In ImageNet, we aim to provide on average 1000 images to illustrate each synset. Images of each concept are quality-controlled and human-annotated. In its completion, we hope ImageNet will offer tens of millions of cleanly sorted images for most of the concepts in the WordNet hierarchy [17],[18].

3. RELATED WORK

The Authors in [19] used deep learning to detect five tomato leaves diseases. They achieved a high accuracy in detecting the tomato disease.

The authors in [20] used Convolutional Neural Network (CNN) to classify plants that belong to 12 species at a few developing phases. The authors in [21] used VGG16 CNN architecture, pretrained on the ImageNet dataset, to identify hydrangea plants in images of forest.

The authors in [22] used transfer learning for Image Classification.

4. METHODOLOGY

In this section we describe the proposed solution as selected convolutional network (ConvNet) architecture and discuss associated design choices, evaluation methods and implementation aspects.

4.1 DATASET

The dataset provided by Kaggle has 2295 training images and 1531 test images. All images are 866×1154 pixels in size. A typical positive sample can be seen in Fig. 1. About two-thirds of the training set is positives (the samples distribution can be seen in Fig. 2). There are a significant number of images that have only small portions of the image occupied by hydrangeas. The negative images are mostly forest scenes without hydrangea, as can be seen in Fig. 3. Sometimes they contain other features such as lakes, horses, sheds, or people.



Fig. 1 A typical positive sample

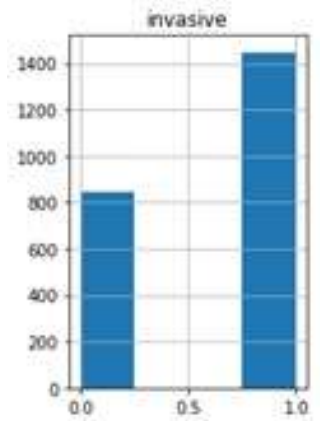


Fig. 2 Samples distribution



Fig. 3 Negative image

4.2 EVALUATION

The competition was a binary classification problem with area under the ROC curve between the predicted probability and the observed target as an evaluation matrix, since we want our submissions to be as close as possible to the actual probabilities, we used the binary crossentropy loss function with sigmoid as the last layer's activation function.

4.3 VALIDATION METHOD

In order to evaluate our model, we have to split the available dataset into training and validation sets, so we conducted an experiment to find the proper splitting method between simple hold-out validation and k-fold cross validation. In our experiment we trained and evaluated our model 4 times using simple hold-out validation method, and in each time we chose a different validation set.

We observed a large diversity between the scores in the 4 times as shown in Fig. 4, so we concluded that the simple hold-out validation method is not suitable for this dataset, and we decided to use k-fold cross validation method instead.

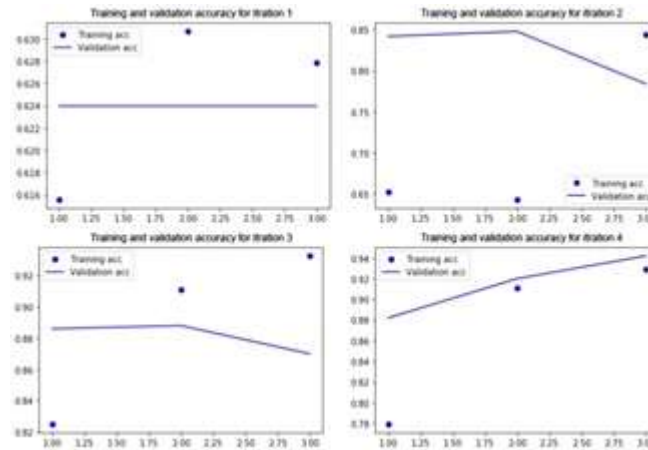


Fig. 4 simple hold-out validation results

4.4 BASE LINE

To have a baseline for which we could compare how well we should expect our later models to reasonably perform, we use pre-trained VGG16 with the following structure:

Table I. Base model structure

| Layer (type) | Output shape | Param # |
|-----------------------------|------------------|----------|
| InputLayer | (5, 128, 128, 3) | 0 |
| BatchNormalizationLayer | (5, 128, 128, 3) | 12 |
| vgg16 (Model) | (5, 4, 4, 512) | 14714688 |
| GlobalAveragePooling2DLayer | (5, 512) | 0 |
| DropoutLayer | (5, 512) | 0 |
| DenseLayer | (5, 1) | 513 |

In all of the experiments we have done, the only changes we did is replacing the Model layer and increasing the input size.

4.5 TRANSFER LEARNING

Since the dataset is relatively small, we had to use transfer learning with a set of pre-trained models, focusing on VGGNets[23], DenseNets[24] and Inception[25] of various sizes that were pre-trained on ImageNet[17],[18].

In order to apply these pre-trained models to anew dataset, there is some steps we had to do.

The first step is replacing the last fully-connected layer; this is necessary to output the classes of the new inputs. In our case, we replaced the last fully-connected layer with the following layers in order:

- GlobalAveragePooling2D layer.
- Dropout layer.
- And the last layer was fully-connected dense layer with output size of 1 (for binary classification) and sigmoid activation function.

After that we had to choose between three approaches of training the model:

- Freeze all the pre-trained layers.
- Freeze some of the pre-trained layers, and re-train the others.
- Re-train all the pre-trained layers.

And we decided to re-train all of them.

4.6 DATA AUGMENTATION

In order to make the most of our few training examples and increase the accuracy of the model, we augmented the data via a number of random transformations. The selected data augmentation techniques were: rotations of 180, horizontal and vertical flipping. Furthermore, it is expected that data augmentation should also help prevent overfitting (a common problem with small datasets, when the model, exposed to too few examples, learns patterns that do not generalize to new data) and, for this reason, improving the models ability to generalize.

5. EXPERIMENTS

We have experimented with a variety of pre-trained models from Keras application library, namely VGG16, DenseNet169, and InceptionV3 that were trained on Imagenet dataset with results reported in Table II.

In all of the experiments we have done, we have used 5-folds cross-validation, and we trained the models at maximum of 50 epochs.

In order to achieve maximum accuracy and avoid overfitting, we've used the following hyperparameters values:

- Adam optimizer with learning rate of 1e-4.
- Early stopping with the following setup:
 - Monitor = validation loss.
 - Patience = 3 epochs.
 - Verbose = 1.
 - min_delta = 1e-4.
- ReduceLROnPlateau with the following setup:
 - Monitor = validation loss.
 - Factor = 0.1.
 - Patience = 1.
 - Cooldown = 1.
 - Verbose = 1.
 - min_lr = 1e-7.

VGG16

For VGG16 base model we resized the images to 128×128 pixels, the mode achieved an accuracy of 97.6% on Kaggle test set.

Progress for both validation and training loss and accuracy for each fold can be seen through Figures 5 to 9.

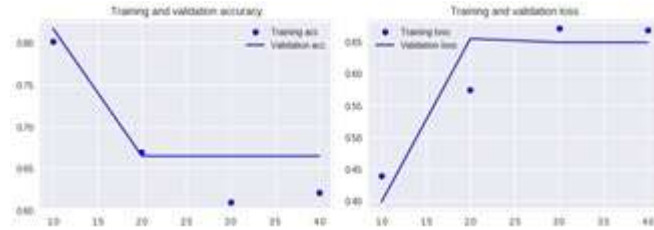


Fig. 5 VGG16 fold-1

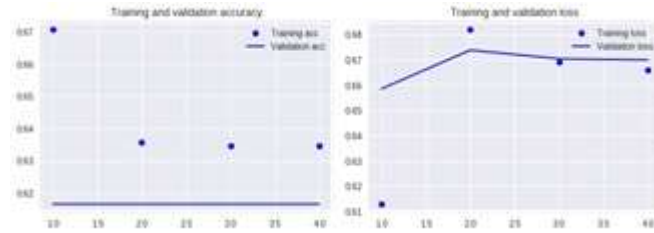


Fig. 6 VGG16 fold-2

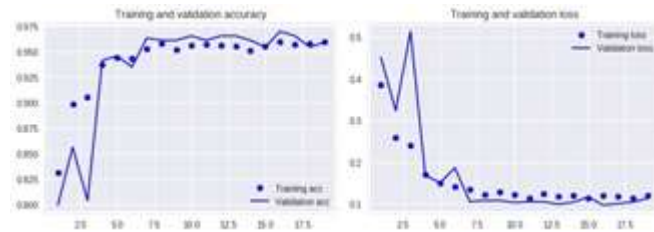


Fig. 7 VGG16 fold-3

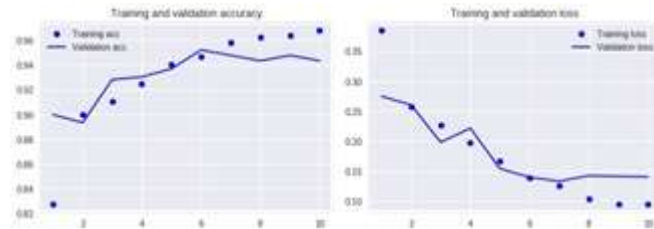


Fig. 8 VGG16 fold-4

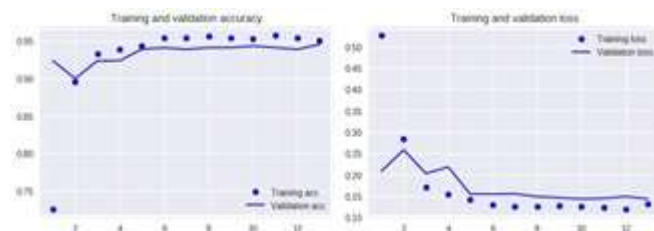


Fig. 9 VGG16 fold-5

5.1 DENSENET169

It was our second best model, in this experiment we resize the images to 600×600 pixels, the model scored an accuracy of 99.44% on Kaggle test set.

On this experiment, we've used the ROC curve on both training and validation sets for evaluating our model in order to get out scores as close as possible to the actual score for Kaggle test set.

Progress for validation and training loss, accuracy and ROC curve for each fold can be seen through Figures 10 to 14.

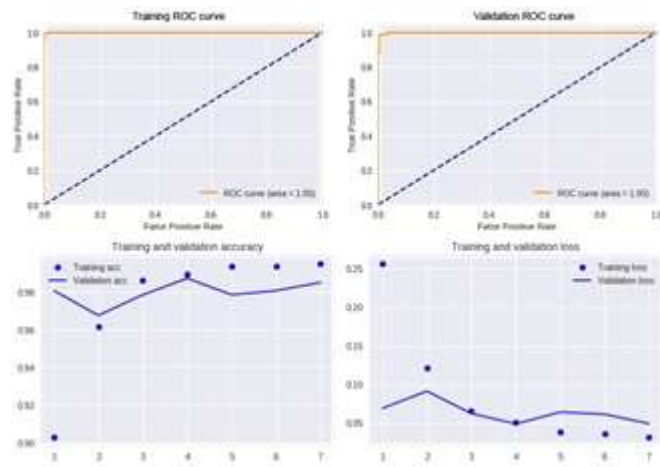


Fig. 10 DenseNet169 fold-1

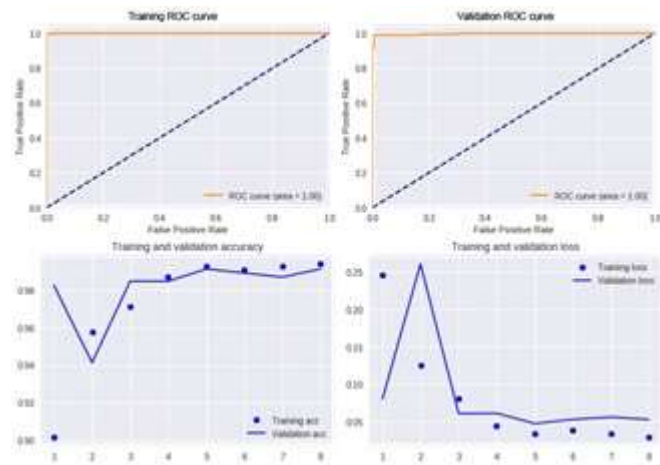


Fig. 11 DenseNet169 fold-2

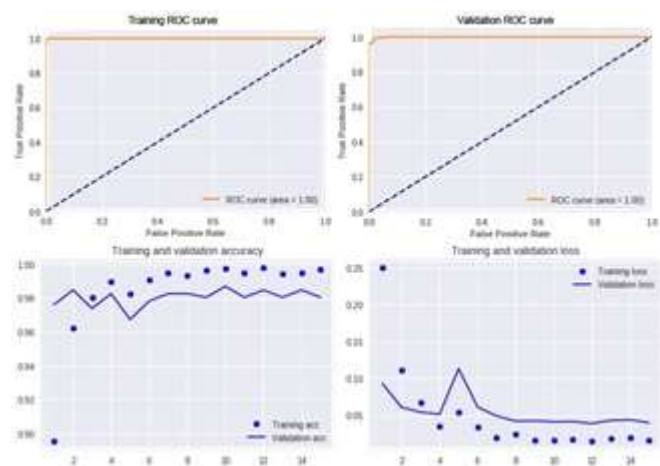


Fig. 12 DenseNet169 fold-3

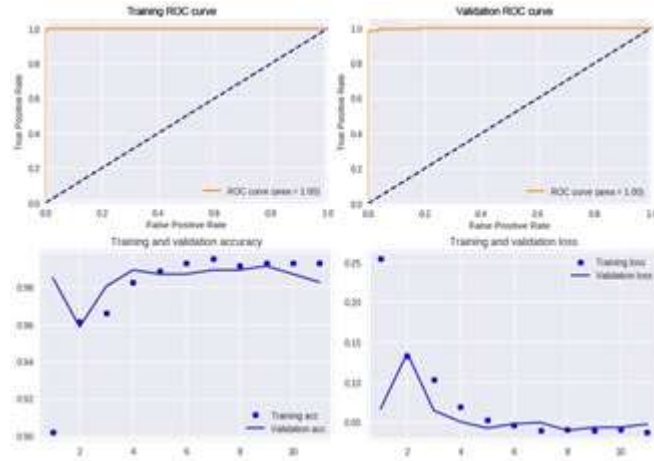


Fig. 13 DenseNet169 fold-4

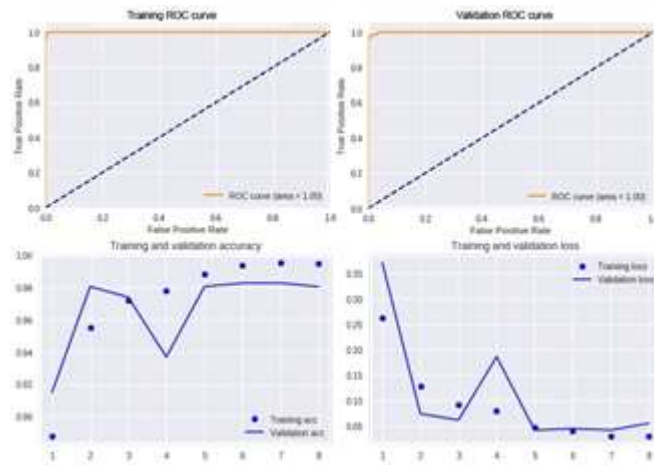


Fig. 14 DenseNet169 fold-5

5.2 INCEPTION V3

It was our best model, in this experiment we resize the images to 800×800 pixels, the model scored an accuracy of 99.71% on Kaggle test set.

And also we've used the ROC curve on both training and validation sets for evaluating our model in order to get out scores as close as possible to the actual score for Kaggle test set.

Progress for validation and training loss, accuracy and ROC curve for each fold can be seen through Figures 15 to 19.

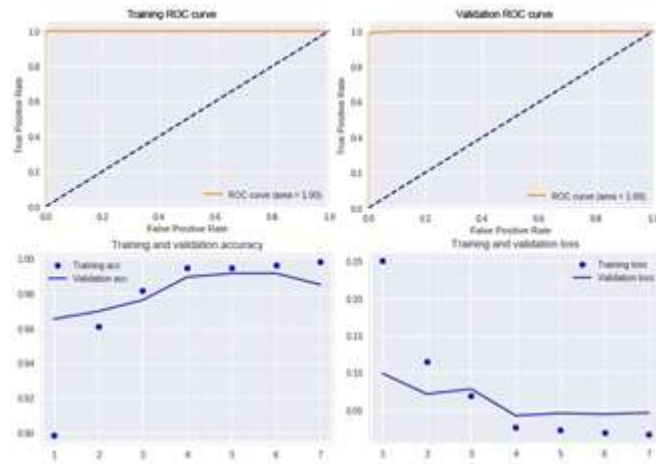


Fig. 15 InceptionV3 fold-1

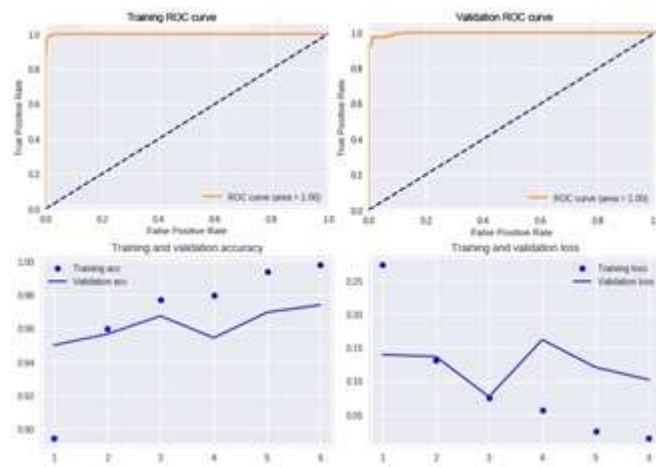


Fig. 16 InceptionV3 fold-2

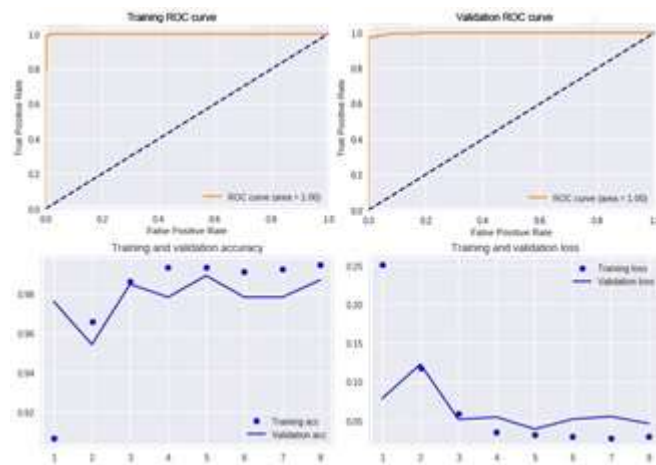


Fig. 17 InceptionV3 fold-3

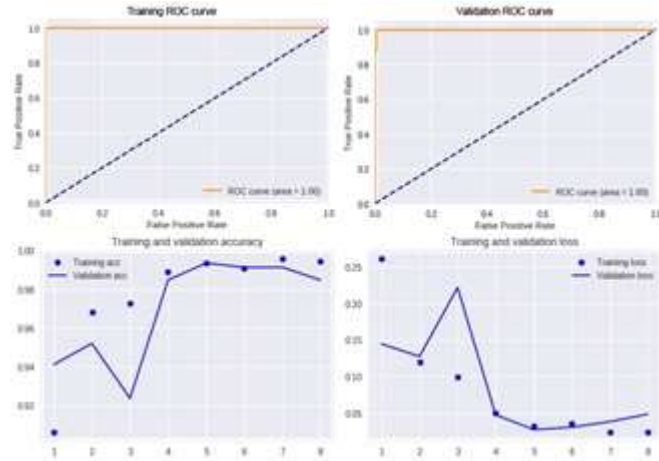


Fig. 18 InceptionV3 fold-4

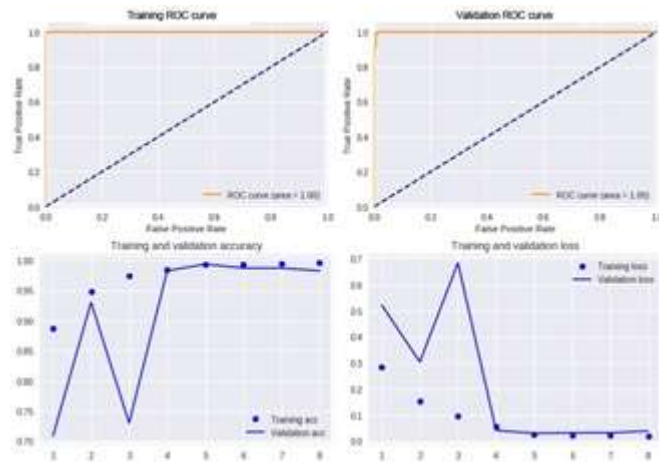


Fig. 19 InceptionV3 fold-5

6.DATA VISUALISATION

To see how the model works and what exactly learns we choose to visualize intermediate activation that consists of displaying the feature maps that are output by various convolution and pooling layers in a network, given a certain input (the output of a layer is often called its activation, the output of the activation function). This gives a view into how an input is decomposed into the different filters learned by the network.

As shown in Figures 20-24, the model learned how to identify the invasive hydrangea, especially its flowers.



Fig. 20 Input Image

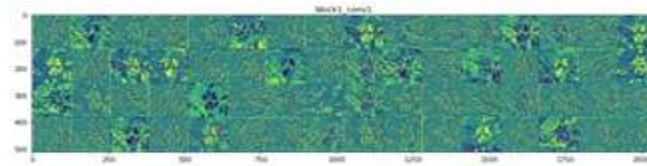


Fig. 21 Feature Map

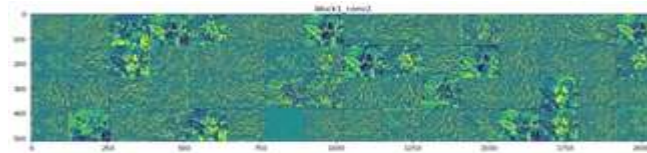


Fig. 22 Feature Map

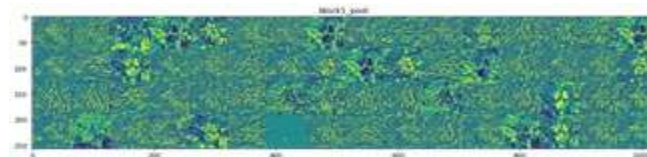


Fig. 23 Feature Map

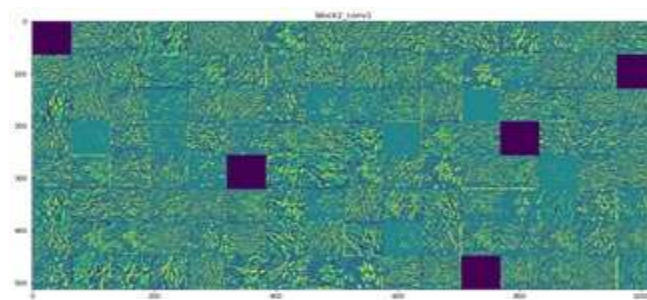


Fig. 24 Feature Map

7. CONCLUSION

We proposed a solution for assisting scientists to more accurately identifying and monitoring invasive hydrangea, with an accuracy of 99.71% for our best model.

Our solution could help making invasive species monitoring cheaper, faster, and more reliable.

Table II. Results of the experiments

| Model setup | Test |
|---|---------------|
| VGG16 | 97.6% |
| DenseNet169 | 99.44% |
| InceptionV3 | 99.71% |
| VGG16 + DenseNet169, ave. | 98.98% |
| VGG16 + InceptionV3, ave. | 99.42% |
| DenseNet169 + InceptionV3, ave. | 99.64% |
| VGG16 + DenseNet169 + InceptionV3, ave. | 99.5% |

REFERENCES

- [1] Joan G. Ehrenfeld. (2010). "Ecosystem Consequences of Biological Invasions", *Annual Review of Ecology, Evolution, and Systematics*, 41: 59–80.
- [2] "A neutral terminology to define 'invasive species'", Robert I. Colautti et al., *Diversity and Distributions*, (Diversity Distrib.) (2004) 10, 135–141.
- [3] Kolar, C.S.; D.M. Lodge (2001). "Progress in invasion biology: predicting invaders". *Trends in Ecology & Evolution*. 16 (4): 199–204.
- [4] The baud, C.; A.C. Finzi; L. Affre; M. Debussche; J. Escarre (1996). "Assessing why two introduced *Conyza* differ in their ability to invade Mediterranean old fields". *Ecology*. 77 (3): 791–804.
- [5] Reichard, S.H.; C. W. Hamilton (1997). "Predicting invasions of woody plants introduced into North America". *Conservation Biology*. 11 (1): 193–203.
- [6] Williams, J.D.; G. K. Meffe (1998). "Non indigenous Species". *Status and Trends of the Nation's Biological Resources*. Reston, Virginia: United States Department of the Interior, Geological Survey. 1.
- [7] Ewell, J.J.; D.J. O'Dowd; J. Bergelson; C.C. Daehler; C.M. D'Antonio; L.D. Gomez; D.R. Gordon; R.J. Hobbs; A. Holt; K.R. Hopper; C.E. Hughes; M. LaHart; R.R.B. Leakey; W.G. Wong; L.L. Loope; D.H. Lorence; S.M. Louda; A.E. Lugo; P.B. McEvoy; D.M. Richardson; P.M. Vitousek (1999). "Deliberate introductions of species: Research needs - Benefits can be reaped, but risks are high". *BioScience*. 49 (8): 619–630.
- [8] Werbos, P.J. (1975). *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*.
- [9] Bengio, Y.; Courville, A.; Vincent, P. (2013). "Representation Learning: A Review and New Perspectives". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 35 (8): 1798–1828. arXiv:1206.5538.
- [10] Schmidhuber, J. (2015). "Deep Learning in Neural Networks: An Overview". *Neural Networks*. 61: 85–117. arXiv:1404.7828.
- [11] Bengio, Yoshua; LeCun, Yann; Hinton, Geoffrey (2015). "Deep Learning". *Nature*. 521 (7553): 436–444.
- [12] El_Jerjawi, N. S., & Abu-Naser, S. S. (2018). Diabetes Prediction Using Artificial Neural Network. *International Journal of Advanced Science and Technology*, 124, 1-10.
- [13] LeCun, Yann. "LeNet-5, convolutional neural networks"..
- [14] Zhang, Wei (1988). "Shift-invariant pattern recognition neural network and its optical architecture". *Proceedings of Annual Conference of the Japan Society of Applied Physics*.
- [15] Zhang, Wei (1990). "Parallel distributed processing model with local space-invariant interconnections and its optical architecture". *Applied Optics*. 29 (32): 4790–7.
- [16] West, Jeremy; Ventura, Dan; Warnick, Sean (2007). "Spring Research Presentation: A Theoretical Foundation for Inductive Transfer". Brigham Young University, College of Physical and Mathematical Sciences. Archived from the original on 2007-08-01. Retrieved 2007-08-05.
- [17] L. Fei-Fei and O. Russakovsky, Analysis of Large-Scale Visual Recognition, Bay Area Vision Meeting, October, 2013.
- [18] L. Fei-Fei, ImageNet: crowdsourcing, benchmarking & other cool things, CMU VASC Seminar, March, 2010.
- [19] Ashqar, B. A. M., & Abu-Naser, S. S. (2019). Image-Based Tomato Leaves Diseases Detection Using Deep Learning. *International Journal of Academic Engineering Research (IJAER)*, 2(12), 10-16.
- [20] Ashqar, B. A. M., Abu-Nasser, B. S., & Abu-Naser, S. S. (2019). Plant Seedlings Classification Using Deep Learning. *International Journal of Academic Information Systems Research (IIAISR)*, 3(1), 7-14.
- [21] Jobson, Elliott, and Andres Hernandez. "Visual Classifier for Invasive Plant Species."
- [22] Quattoni, Ariadna, Michael Collins, and Trevor Darrell. Transfer learning algorithms for image classification. Diss. Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 2009.
- [23] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.
- [24] Huang, Gao, et al. "Densely connected convolutional networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- [25] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

