Invasive Species Detection

Authors: Christy Dennison and Charles Hale
Convolutional Neural Networks for Visual Recognition, Stanford University

Motivation

Conservationists around the world are greatly concerned by invasive species, since they can destroy delicate ecological balances in the habitats they invade. The ability to quickly identify invasive species is critical to their swift removal. The Kaggle Competition “Invasive Species Monitoring” provides a dataset of images that either do or do not contain invasive Hydrangea. This competition boils down to a simple binary image classification problem. As such, we can comfortably apply CNNs for excellent results.

Evaluation

The “Invasive Species Monitoring” Kaggle Competition uses the area under the ROC curve to evaluate submissions.

Methods

Transfer learning was critical to our approach. We explored a range of techniques for training modified pretrained ImageNet architectures and ensuring size compatibility.

Learning Methods:

- Train last layer only.
- Train first and last layers only.
- Train entire network.


Average training set image

Resizing Methods:

- Crop input randomly to square size (866x866), then rescaling using Lanczos interpolation to 224x224.
- Insert convolutional layer whose output is 224x224.
- Insert Spatial Pyramid Max Pool layer at beginning as a form of preprocessing to scale the image to 224x224.
- Resize the image to 224x224 using Lanczos interpolation.

Preprocessing methods:

- Normalize images with means and stds from ImageNet
- Normalize images with means and stds from dataset itself

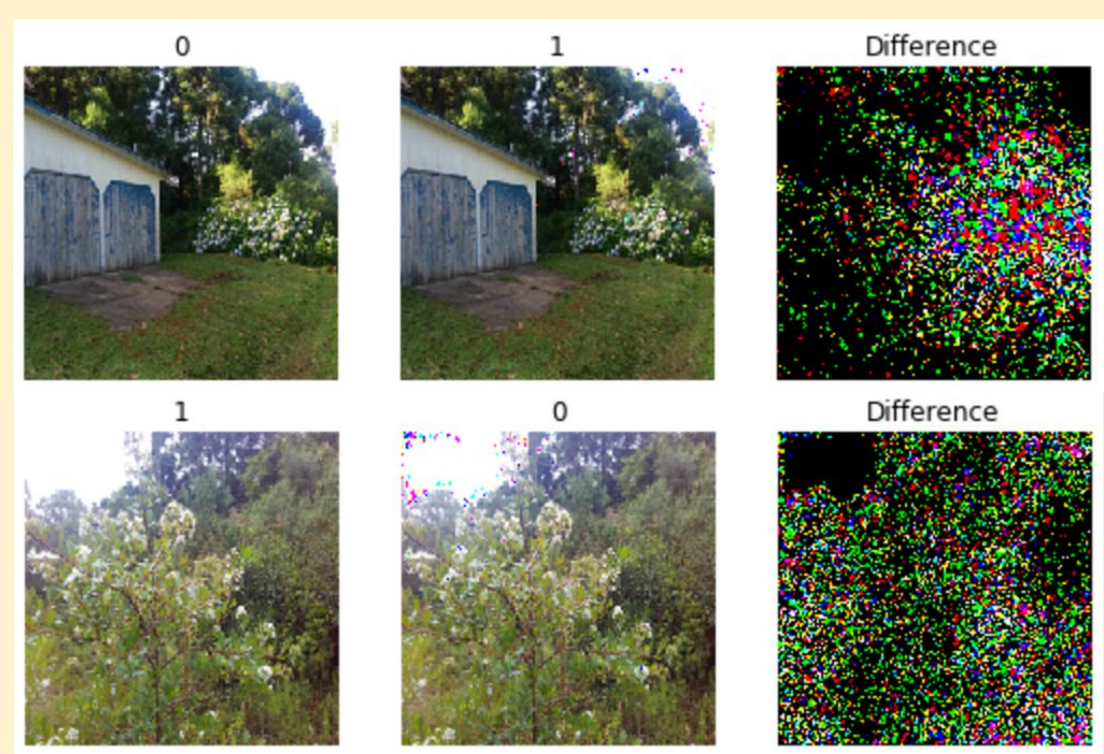
Dataset

- Training set of 2295 images (we split 80/20 for train/val)
- Images are outdoor scenes labeled 1 or 0 depending on whether they contain invasive Hydrangea
- Classes are evenly represented in the training set
- Test set of 1531 images (labels withheld by Kaggle)
- Images are size 866x1154
- Most images are full-frame photos of Hydrangea flowers

Learning Results

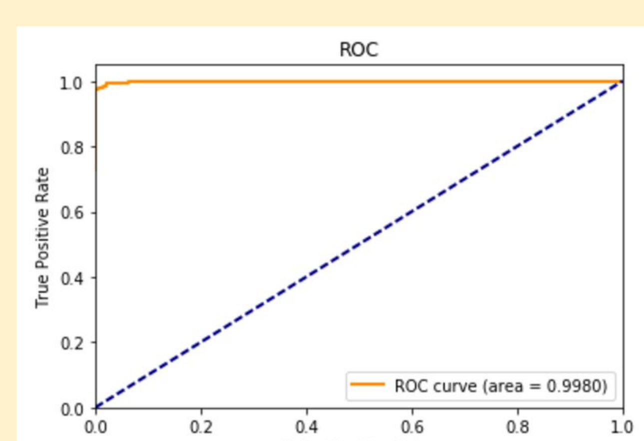
Saliency Maps

False negatives categorized images from the best ensemble show that the net is correctly focusing on the flowers but isn’t confident enough to classify it, given that most flowers are not so far in the background. False negatives also showed focus on the incorrect flowers.



Fooling Images

The false negative fooling shows that the net emphasizes where the flowers are, and the false positive fooling shows the net deemphasizes where the incorrect flowers are.



ROC Curve

The ROC curve for the best ensemble on the validation set achieves an AUC of 0.9979.

Results

Model (all using transfer learning)	Val Acc.	Test Acc.
ResNet18 Insert conv layer for scaling, only change first and last	92.94	
ResNet18 Insert conv layer for scaling	90.21	
ResNet18 Random crop w/o normalization	96.81	
ResNet18 Random crop w/ normalization (ImageNet)	95.44	
ResNet18 Random crop w/ normalization (dataset)	97.72	
ResNet18 Random crop w/o norm., change only last layer	92.48	
ResNet18 Spatial Pyramid Pooling w/o norm	98.18	
ResNet18 Spatial Pyramid Pooling w/ norm (ImageNet)	97.27	
ResNet18 Spatial Pyramid Pooling w/ norm (dataset)	97.72	
ResNet18 Lanczos Scaling w/o normalization	98.18	95.84*
ResNet18 Lanczos Scaling w/ normalization (ImageNet)	97.49	
ResNet18 Lanczos Scaling w/ normalization (dataset)	97.49	
ResNet18 Lanczos Scaling w/o norm., 90/10 train/val	98.23	98.79
ResNet18 Lanczos Scaling w/o norm., plus train on train errors	97.95	
ResNet18 Lanczos Scaling w/o norm., plus train on val errors	99.32	98.81
ResNet34 Lanczos Scaling w/o norm	99.32	
ResNet52 Lanczos Scaling w/o norm	98.86	
ResNet101 Lanczos Scaling w/o norm	99.31	
ResNet152 Lanczos Scaling w/o norm	98.86	
VGG13 Lanczos Scaling w/o norm	98.17	

*Binary (not prob.) submission. Note that Kaggle restricts the number of submissions per day, so attempts were not made for most models.

Ensemble	Val Acc.	Test Acc.
ResNet (18, 34, 50, 101, 152), max probability	99.54	98.33
ResNet (18, 34, 50, 101, 152), average probability	99.32	99.16
VGG13 + ResNet (18, 34, 50, 101, 152), max probability	99.77	98.39
VGG13 + ResNet (18, 34, 50, 101, 152), average probability	99.32	99.32

5th place (of 131) in Kaggle Competition!

Conclusions

By analyzing the false negatives and positives of our network, we conclude that Hydrangea detection is very reliably and accurately solved using transfer learning. However, this method is less reliable for photos where Hydrangea are far in the background or are at a small scale. We believe that an attention-based model with a larger dataset would be able to overcome these limitations.