Using Grad-CAM to explain Deep Learning Methods for Mapping Invasive Plants in a Biodiversity Hotspot

Neha Pattanaik, Dr. Abhijin Adiga

Network Systems Science and Advanced Computing Division (NSSAC)

The Project:

Mapping the spread of Parthenium Hysterophorus, Chromolaena Odorata, and Lantana Camara in the Chitwan-Annapurna Landscape (CHAL) in Nepal

BACKGROUND

Motivation:

- CHAL is a “biodiversity hotspot”
  - Contains at least 1500 endemic species
  - Has lost at least 70% of its primary native vegetation
- Threatened by invasive species
  - Parthenium hysterophorus, Lantana Camara, Chromolaena Odorata

Previous Work:

- Obtained high quality satellite images
- Developed and trained patch based convolutional neural network
  - square patches are extracted from the satellite images and the center pixel is used to determine presence or absence

Challenges

- Lack of usable data due to seasonality (plant appearance changes in different seasons), weather conditions, and mountainous terrain
  - Overcome using a multiple holdout method, which is an iterative process that uses a different model for each iteration

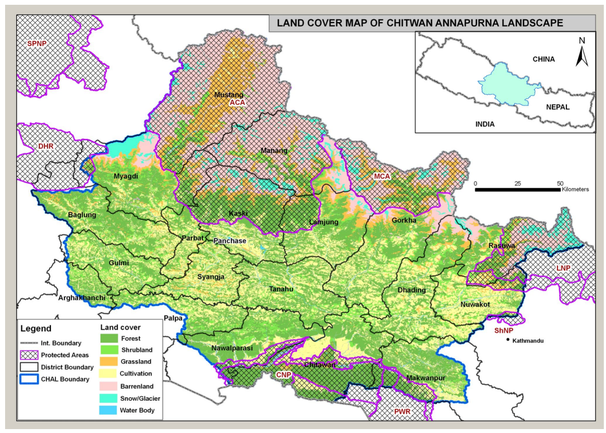

Figure #1 CHAL region in Nepal

Figure #2 Parthenium Hysterophorus

Figure #3 Chromolaena Odorata

Figure #4 Lantana Camara

Grad-CAM Process Visualization

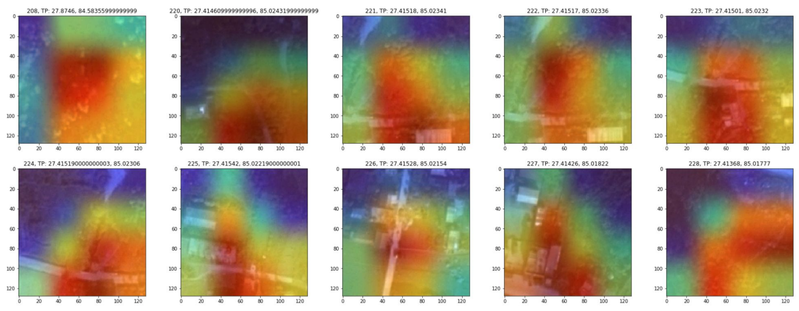

Figure #5: Grad-CAM results for Parthenium hysterophorus from a single model

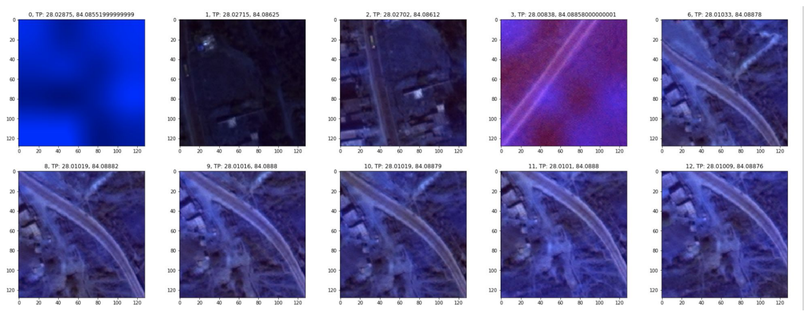

Figure #6: Grad-CAM attempt at model combination for Parthenium Hysterophorus, results were difficult to interpret

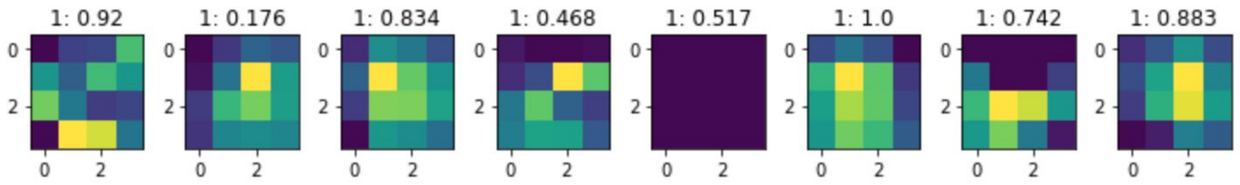

Figure #7: Analyzing heat maps produced by all models for Parthenium hysterophorus

CURRENT WORK

Gradient-weighted Class Activation Mapping (Grad-CAM):

- Explainability of deep learning models

- Uses gradients flowing into final convolutional layer to highlight regions in the image that were used for prediction
  - looking at what the model used to make the prediction helps to identify potential bias
- Produced Grad-CAM Analysis for one model
  - Analyzed images based on confusion matrix value (True positive, true negative)

Combining Models

- Multiple holdout method uses multiple models
  - 8 best models identified
- Determining which region of the image each model is using to make prediction
  - ideally all the models are looking at the same region
- Originally attempted to take average heatmap of all models and overlay (*figure 6)*
  - Difficult to interpret
  - No confusion matrix value since one model could predict “present” while another model could predict “not present”
- Interpreting results of heatmaps produced by different models
  - Conducting visual analysis based on heatmaps and softmax values
  - Analyzing mean square error (MSE) between models

FUTURE DIRECTIONS

- Continue to analyze combination of models
  - Similarities and differences in predictions, regions used to make predictions and softmax values

Big Picture:

- Evaluate how robust multiple holdout method is when using small datasets
  - Conduct experimentation using ImageNet Dataset

References

Global Invasive Species Database (2021) Species profile: *Parthenium hysterophorus*. Downloaded from http://www.iucngisd.org/gisd/species.php?sc=153 on 30-07-2021.

“Landcover Map of Chitwan Annapurna Landscape.” *WWF*, www.wwfnepal.org/about_wwf/where_we_work/chal_landcover/.

Selvaraju, Ramprasaath R., et al. “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization.” *ArXiv.org*, 3 Dec. 2019, arxiv.org/abs/1610.02391.