

Exploring the Usefulness of Adding Auxiliary Preprocessed Image Layers with Convolutional Neural Networks



Jordan Goetze and Anne Denton

Department of Computer Science, North Dakota State University, Fargo, ND 58108-6050

ABSTRACT

For this study, we use the SegNet CNN architecture, along with imagery comprised of Red, Green, Blue, Near-Infrared and Normalized Difference Vegetation Index image bands. We also use image bands generated from the original 5 bands. The goal of this study is to discern whether the inclusion of several auxiliary image layers aids in the classification accuracy and quantitative quality of classifications in the realm of land use classification of orthoimagery.

PROBLEM

- Convolutional Neural Networks learn abstract representations of objects. It is difficult to determine what sorts of patterns a CNN is able to learn with reasonable time.
- Train models with and without the inclusion of auxiliary preprocessed image layers to see if the inclusion or exclusion of additional layers aids or hinders the model effectiveness.
- Model effectiveness must be determined quantitatively by model classification accuracy and qualitatively because of the low quality of available ground truth data.

APPROACH

- Using SegNet model which is an encoder-decoder CNN for per-pixel image classification
- Compare the results of 3 different models.
 - Control: Vanilla SegNet with RGB and NDVI as input layers.
 - Regression: Control Model with the addition of a regression based preprocessed image layer.
 - Gradient: Control Model with the addition of a gradient based preprocessed image layer.

APPROACH CONT.

- Regression Layer: Using a sliding window, take the slope of the line calculated by taking the linear-regression between two bands. Bands: Red and NIR. Selection based of insight from paper: *Multi-scalar Analysis of Geospatial Agricultural Data*.
- Gradient Layer: Using a sliding window, take the slope calculated with regard to spatial coordinates for a band.
- Image Data: Image data from National Aerial Imagery Program. 1m/pixel resolution. RGB and NIR image layers
- Ground Truth Data: National Agricultural Statistic Service Land Use Classifications. 50m/pixel resolution.
- Data set size: 2,000 256x256 pixel images.

RESULTS

Not Water	Water	Model Type	Accuracy
93%	7%	Control	92.3680%
Table 1: Class Label Percentage Breakdown		Regression	86.0488%
		Gradient	93.3380%

Table 2: Overall Model Accuracies

Model Type	Not-Water	Water
Control	96.8892%	37.6272%
Regression	89.7641%	40.8237%
Gradient	98.5329%	30.2602%

Table 3: Per-class Model Accuracies

- The overall accuracy of the gradient model is only slightly better than if we were to predict only Not-Water.
- By examining the per-class accuracies for each model we can see the trade offs made by each model.
- While the gradient model more reliably predicted Not-water, the regression model favored water accuracy.

RESULTS CONT.



Figure 1: Evaluation output of each model for the same image. First the input image, the ground truth, the model classification, the input image overlaid with the model classification, and the difference between the ground truth and model classification.

- As was noticed in *Multi-scalar Analysis of Geospatial Agricultural Data*, it appears that regions of foliage next to a water source tend to have a strong response when the regression image generation method is applied between the Red and NIR image bands. This leads the model to reliably detect coastlines of large water bodies that do not fit entirely within one image.

CONCLUSIONS

- While we did not see a significant difference in accuracy with the inclusion of the gradient or regression based image layers, we did observe that the inclusion of these additional layers affected the vanilla model's ability to classify certain types of features.
- This suggests that the vanilla CNN was unable to learn the sorts of features encapsulated by the preprocessing techniques within the allowed the training time. Further training time would have overfit the model.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation through grant IIA-1355466.