Deep Learning and Deforestation

Ian Avery Bick, Dennis Wang and Ben Mullet
{ianbick, dwang22, bmullet}@stanford.edu

Problem

Accurate classification of satellite imagery is a critical task for understanding the scope and manifestation of deforestation. Commonly used classification methods (available in ArcMap or Google Earth Engine) such as mean-shift, SVM, and random forest classifiers ultimately depend on a high degree of human correction. Our project further automates the classification process by using a fully convolutional neural networks for image segmentation. Specifically, we seek to ease classification of forest boundaries and forest density to support wildlife habitat and deforestation analyses.

Data

Two sets of satellite imagery were utilized to evaluate the CNN accuracy with both binary and multi-category forest classification. Both datasets were in .tiff format and normalized to Top of Atmosphere reflectance.

	RapidEye	LANDSAT[2]
Image Year	2015	2000
Resolution	5 m	30 m
Bands	5	4
Bit Depth	8-bit	16-bit
Forest Mask	Binary (0,1)	Continuous (0-100)

Rapideye data was accompanied by an unsupervised classification of forest pixels which was manually refined into a forest/no-forest binary mask through E-IPER PhD candidate Laura Bloomfield's thesis project. LANDSAT data and a corresponding continuous forest density mask was provided from a global forest density map[2]. The forest mask was separated into five categories (0-20, 20-40, 40-60, 60-80, and 80-100% tree canopy cover).

Models and Features

Our convolutional neural network was inspired by the U-Net originally proposed by Ronneberger et al[3]. The network is characterized by a "U-shaped" sequence of traditional CNN contracting layers followed by an equal number of expanding layers. Expanding layers combine deconvolutions with concatenations of the previously stored contracting output. Final classification was generated using a pixel-wise softmax classification layer with a cross entropy loss function. The input data was cut into 256x256 pixel tiles, and the training set was augmented by a factor of 16 with rotations, flips, and added Gaussian noise.

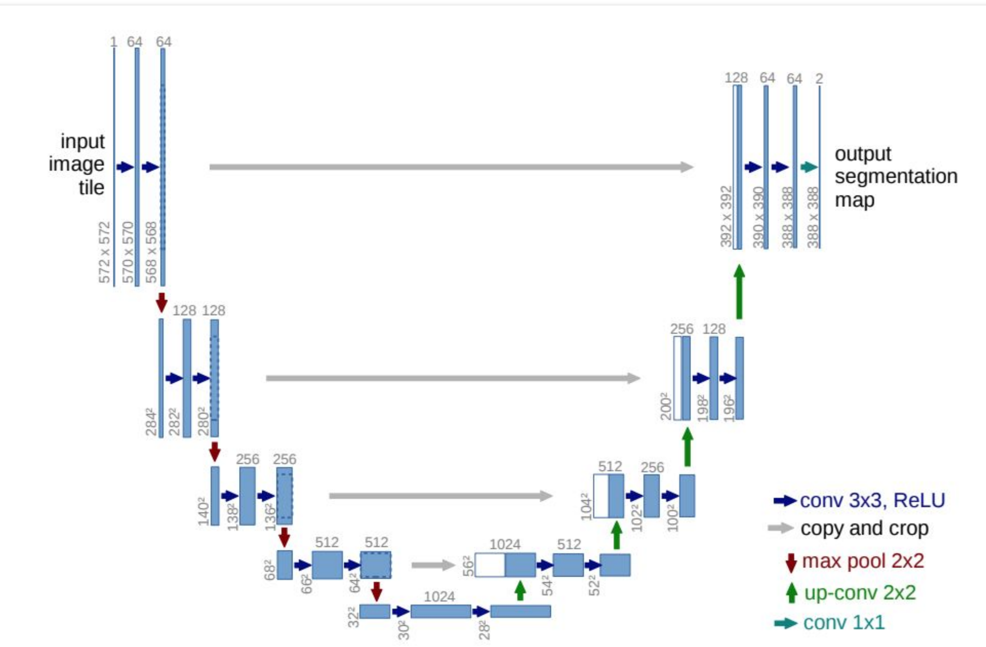
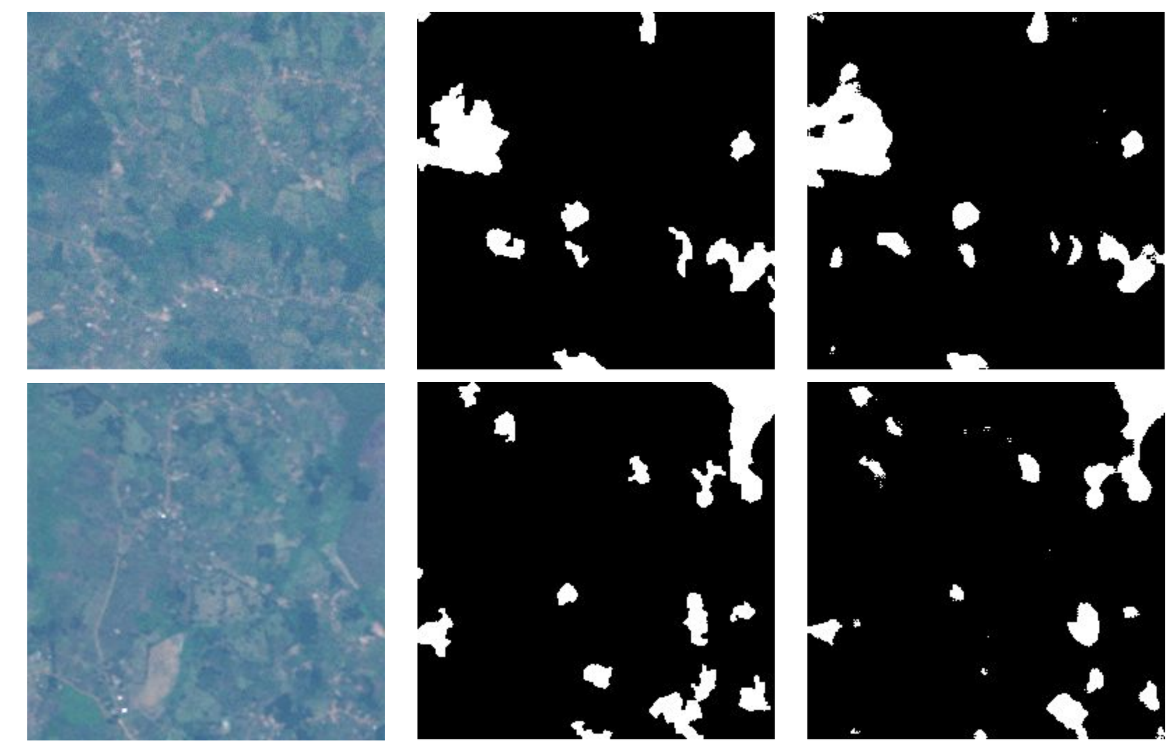


Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

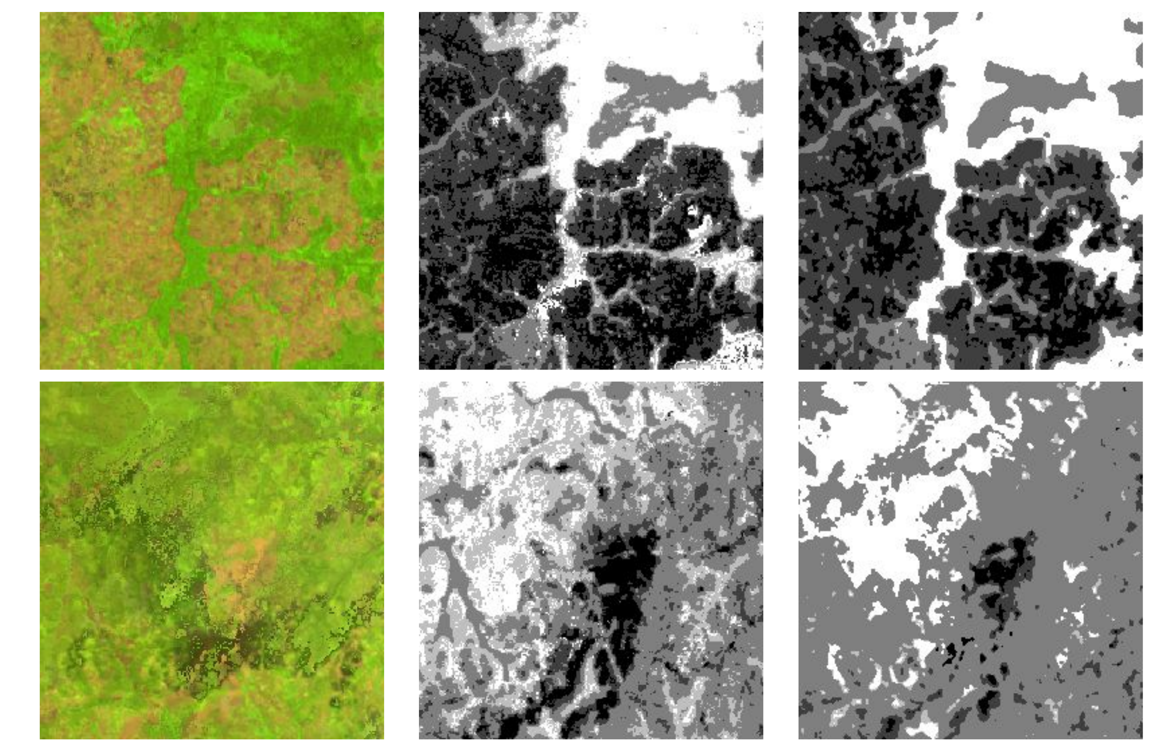
Fig. 1: Example U-Net, from [3]

Results

Model	Train Size	Dev Size	Test Size	Train Acc.	Dev Acc.	Test Acc.	Train IoU	Dev IoU	Test IoU
Binary	270	34	34	0.963	0.651	0.733	0.789	0.326	0.366
Binary Aug.	216 (x16)	122	122	0.946	0.901	0.878	0.828	0.809	0.768
Multi-Class	568 (x16)	71	71	0.724	0.692	0.685	0.475	0.462	0.462



Binary Classification Examples (left to right: image, label, pred)



Multi-Class Classification Examples (left to right: image, label, pred)

Discussion

The first conclusion we drew from our results was that given the small amount of labeled training data, data augmentation helped immensely to reduce overfitting. For the binary task, our achieved pixel accuracy is meets the generally accepted land cover classification accuracy criteria of 85%[4,5]. However, our metrics were much lower for the multi-class task. From the sample images, we can see that the model was able to learn general forest density structure but not precise details. We suspect a more complex model could improve this.

Future

The model as presented could be improved by obtaining additional and more diverse training data. Given time, we would like to try more complex semantic seg architectures (i.e. Tiramisu Densenets), especially for multi-class. We would also like to incorporate the methodology into a Land Use and Land Cover Change Model to make predictions.

References

1. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18-27.
2. Hansen, M. C. et al. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." Science 342 (15 November): 850–53. Data available on-line from:http://earthenginepartners.appspot.com/science-2013-global-forest.
3. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.
4. Thomlinson, John R, et al. "Coordinating Methodologies for Scaling Landcover Classifications from Site-Specific to Global." Remote Sensing of Environment, vol. 70, no. 1, 1999, pp. 16–28., doi:10.1016/s0034-4257(99)00055-3.
5. Foody, Giles M. "Status of Land Cover Classification Accuracy Assessment." Remote Sensing of Environment, vol. 80, no. 1, 2002, pp. 185–201., doi:10.1016/s0034-4257(01)00295-4.