Land Cover Change Detection via Semantic Segmentation

Renee Su[1], Hank Magnuski[2]
[1]Henry M. Gunn High School, [2]Ncast

INTRODUCTION

- Land cover change detection is detecting when the land cover at a given location has been converted from one type to another; uses include quantifying:
  - Agriculture land loss to urban development
    - Urban land has increased from 3.1% to 3.6% from 2000 to 2010 (USDA, 2014)
  - Deforestation
    - More than one million kilometers of tropical forests were cleared from 2000 to 2012 (Norris, 2016)
- The goal was to design a change detection system that takes in aerial images and identifies changes in land cover, via a machine learning approach
- Used for managing natural resources and monitoring environmental changes
- Detect land cover change via semantic segmentation—predicting the class of each pixel in an image.
- Deeplab v3+ is a state-of-the-art semantic segmentation algorithm open sourced by Google (Figure 1)

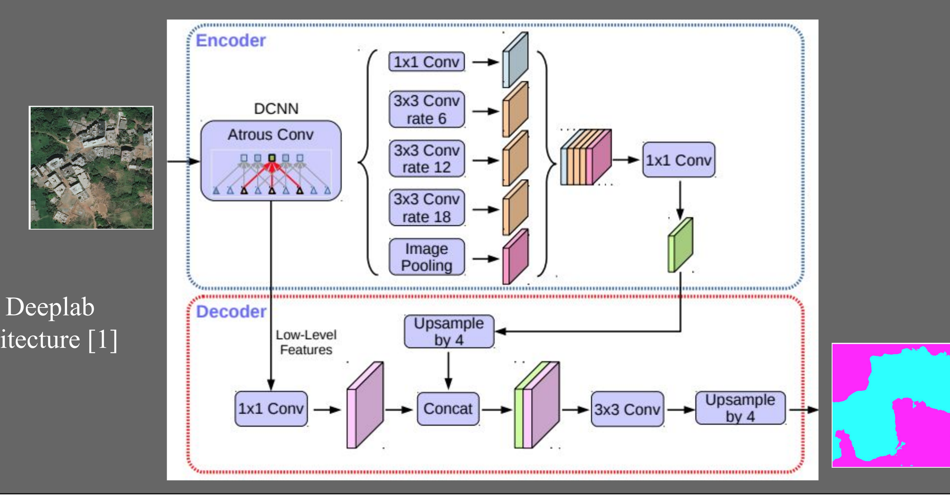


Figure 1: Deeplab v3+ Architecture [1]

RESEARCH METHODOLOGIES

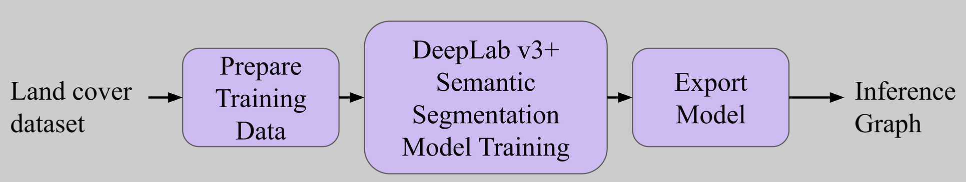


Figure 2: Land Cover Semantic Segmentation Training Pipeline

The training pipeline (Figure 2) illustrates each module used. We started with our land cover dataset and followed by preparing the training data (Figure 4). Then, we used the DeepLab model to train our land cover semantic segmentation model. Finally, we exported the trained model to an inference graph, which was used for prediction afterward. In order to measure the accuracy of our model, we used a metric called the Intersection over Union (Figure 3).

- Training dataset from DeepGlobe Land Cover Classification Challenge
  - Ground resolution of image pixels is 50 cm/pixel
  - 803 satellite images of size 2448 by 2448 pixels
  - Each image paired with a mask image for land cover annotation
  - 7 land cover types:
    - Urban land
    - Agricultural land
    - Rangeland
    - Forest land
    - Water
    - Barren Land
    - Unknown

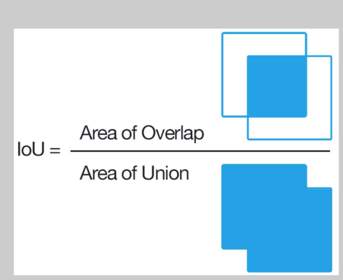
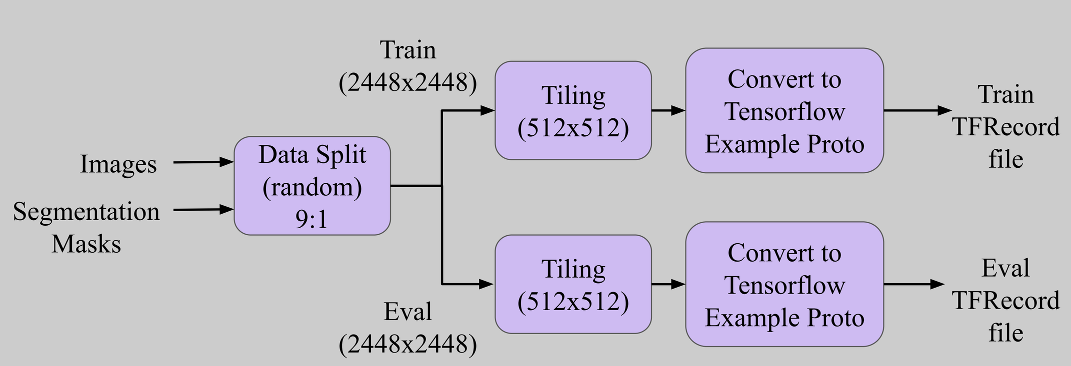


Figure 3: Visual of IoU metric

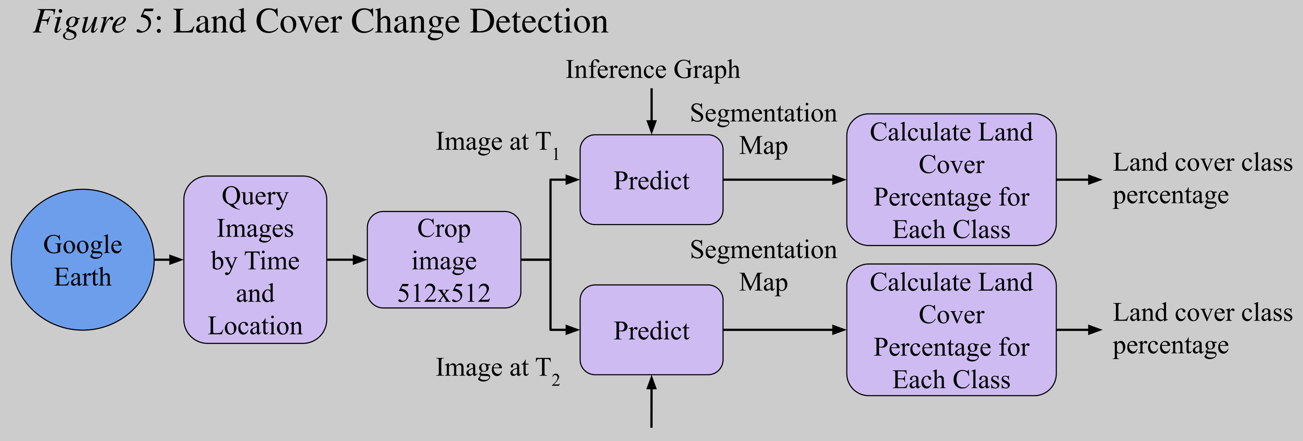
The Intersection over Union (IoU) metric measured the area of overlap between a class label in the ground truth segmentation map and the area of the same class label in the predicted segmentation mask, divided by the area of the union between the two. For this model, we took the average of the IoUs from each of the seven land cover classes (called the mIoU).

Figure 4: Training Data Preparation



For data preparation, we paired the images with their corresponding segmentation masks, then proceeded to split the images randomly into two groups: training and evaluation. 90 percent of the data were trained, and the remaining 10 percent were used for evaluation. Since the images in the dataset were originally 2448 by 2448 pixels, we divided each image (and its corresponding segmentation mask) into tiles of 512 by 512 pixels, resulting in 25 images of a much smaller size to be used for training. Afterward, we converted each pair into the Tensorflow example proto, which were subsequently written to a TFRecord file.

Figure 5: Land Cover Change Detection



After the model had been trained, we were able to utilize it for change detection. Using the Historical Image feature on Google Earth, we retrieved images from two different years (T_1 and T_2) at a location in the United States. Then, we cropped the images to 512 by 512 pixels, and applied the inference graph to predict the segmentation map. Using the segmentation map, we calculated the land cover percentage for each class type present in each image.

DATA AND FINDINGS

	mIoU
DeepGlobe land cover classification challenge	43.3% [2]
Ours	**75.6%**

Figure 6: mIoU comparison

From training the model, we achieved a mean IoU of 75.6% on our evaluation set. This was significantly higher than the mean IoU of 43.3% from the DeepGlobe challenge (Figure 6). This difference was due to the fact that the model from the challenge used Resnet18 as its network backbone, while our model used the improved Xception backbone. Furthermore, the newer version of Deeplab uses atrous convolutions, which effectively enlarge the field of view and allow for dense feature extraction. Finally, a decoder module is included, which refines the segmentation results, especially along object boundaries.

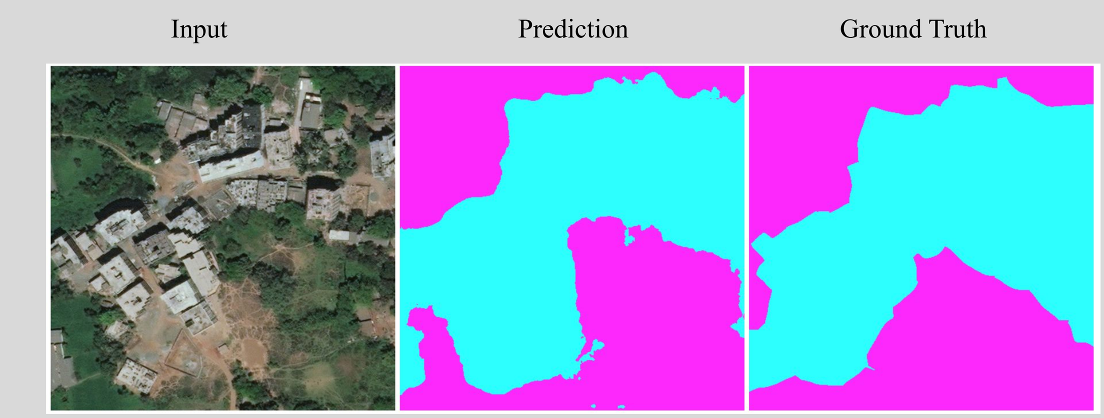


Figure 7: Evaluation example

Figure 7 illustrates a side-by-side comparison of the prediction and ground truth masks of an example input image from the evaluation set. In the lower left corner of the input image, there is a stretch of green land that was incorrectly labeled as urban land in the ground truth and was predicted correctly as rangeland in the inference. In this case, the model is more accurate than the person who labeled the ground truth, because it was able to recognize the vegetation in the area.

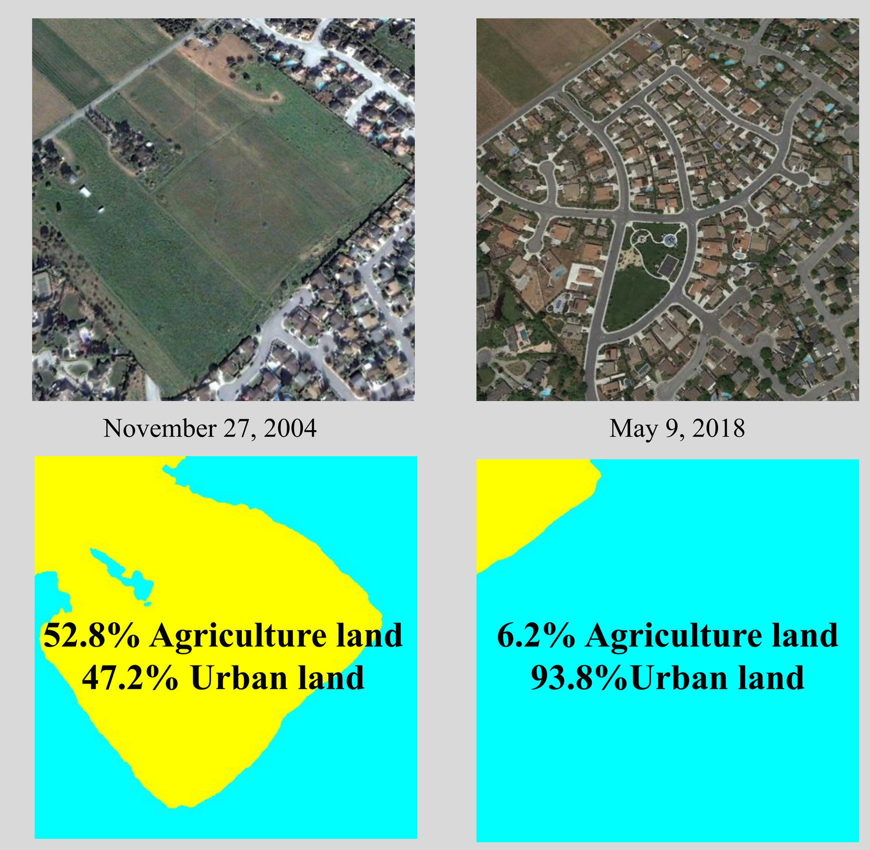


Figure 8: Morgan Hill, California area over a fourteen year period

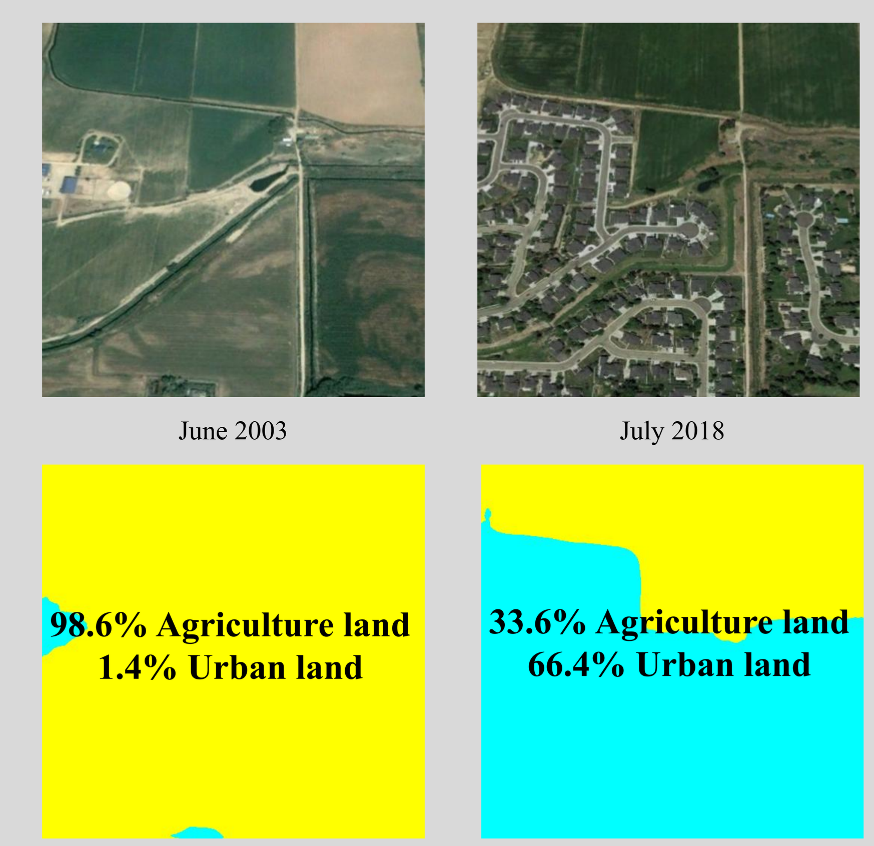


Figure 9: Star, Idaho area over a fifteen year period

Figures 8 and 9 are two locations retrieved from Google Earth to which we applied change detection. In the Morgan Hill area, one can see that in 2004, agriculture and urban land each occupied a roughly equal percentage of the land. However, in 2018, our model predicted that urban land percentage had almost doubled, leaving only 6.2 percent of the land in the image as agriculture land. Similarly, in Star, 2003, most of the land was for agriculture. In 2018, about two-thirds of that land had been built into a residential area. The model prediction in the second location roughly outlines the boundaries of the neighborhood shown.

DISCUSSION

- Achieved significant improvement in land cover semantic segmentation — mIOU 0.756 compared to 0.433 (reported in DeepGlobe land cover classification challenge)
- Proposed the first land cover change detection system that leverages the state-of-the-art semantic segmentation method and can be used for land management, in regard to natural resources, deforestation analysis, and changes in the environment
- For model generalization, add data augmentation to simulate images from Google Earth without proper color correction or haze removal
- For change detection in a large satellite image, we would need to divide the image into overlapped "tiles," apply inference on each tile, then study how to merge the segmentation results in the overlapped area.

ACKNOWLEDGEMENTS / REFERENCES

***Special thanks to Hank Magnuski and Angela Merchant for helping to make this project possible.

*****Works Cited:**

1. Chen, L., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. ECCV 2018.
2. Demir, I., Koperski, K., Lindenbaum, D., Pang, G. and Huang, J., Basu, S., Hughes, F., Tuia, D., Raskar, R. (2018). DeepGlobe 2018: A Challenge to Parse the Earth through Satellite Images. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2018.
3. Norris, K. (2016). Ecology: The Tropical Deforestation Debt. *Current Biology*, *26*(16), R770-R772.
4. Oswalt, S. N., & Smith, W. B. (Eds.). (2014). *U.S. Forest Resource Facts and Historical Trends*. Retrieved from United States Department of Agriculture website: https://www.fia.fs.fed.us/library/brochures/docs/2012/ForestFacts_1952-2012_English.pdf