DATA FUSION

MIAMI UNIVERSITY
OXFORD, OH • EST. 1809

College of Engineering and Computing | Benjamin Bowser, Matt Spinelli, Taylor Florence, and Dr. John Femiani

Data Fusion Background:

- **Computer vision** is a scientific field that focuses on how computers can be designed to gain a high-level of understanding from digital images or videos.
- **Machine learning** is a scientific field of study that uses algorithms and statistical models to perform a specific task without the use of explicit instructions, relying on observable patterns.
- **Data fusion** is the combination of these two studies, integrating multiple data sources to produce more consistent information

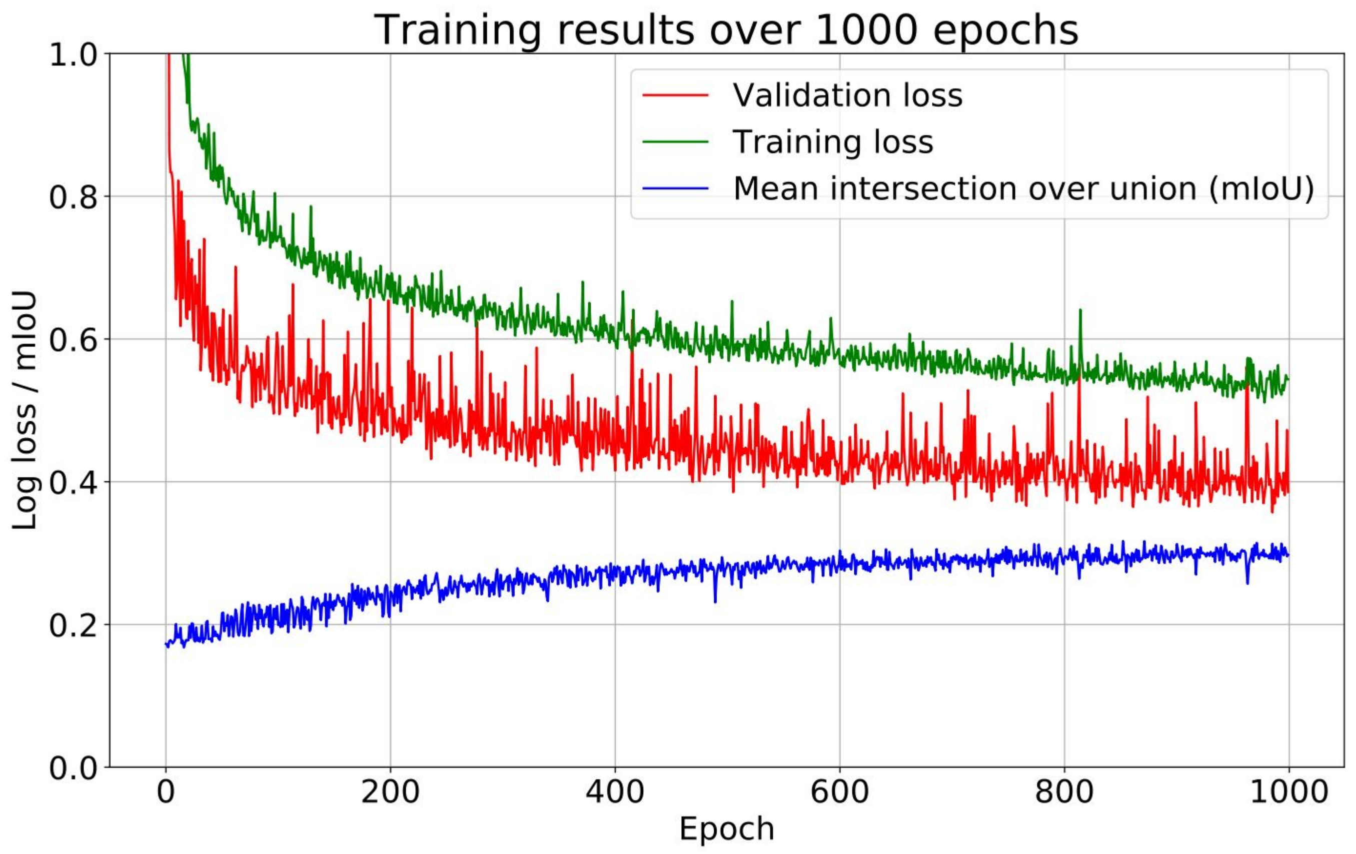


Figure 1. In green and red, the training and validation log loss is compared with mIoU in blue. Over time performance improves with a lower log loss and a higher mIoU.

Data Processing:

The code below is used to save the current best epoch. The method adds the specifications of the current model into a dictionary, then compares the current epoch with the best epoch. If the current epoch is better then the change is made.

```
def save_checkpoint():

    torch.save(dict(
        model = model.state_dict(),
        optimizer = optimizer.state_dict(),
        train_losses = train_losses,
        val_losses = val_losses,
        val_IoUs = val_IoUs,
        best_epoch = best_epoch,
        epoch  = epoch
    ), open(f'checkpoint-{slug}.pth.tar', 'wb'))

    if best_epoch == epoch:
        shutil.copyfile(f'checkpoint-{slug}.pth.tar', 'best-{slug}.pth.tar')
```

Competition Timeline:

- January 9th - Contest opens with release of training and validation data
- February 7th - Validation server opens and leaderboards open
- March 7th - Release of test data and opening of test server
- March 22nd - Deadline to submit results for contest
- March 29th - Winners announced

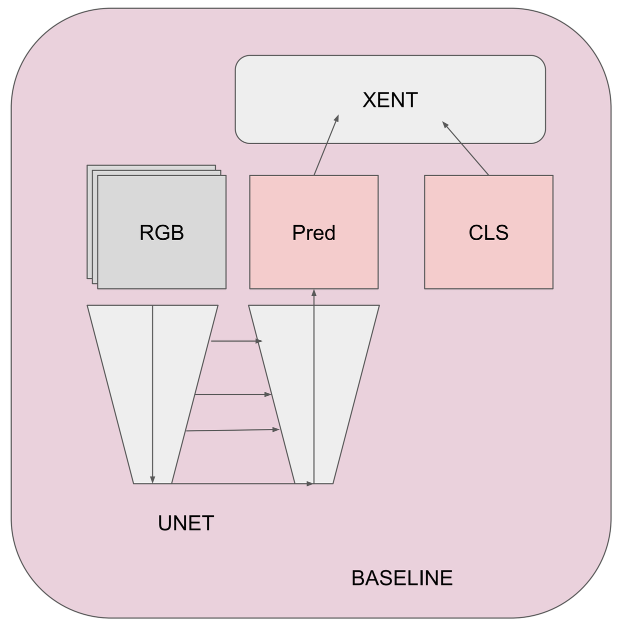
Competition Results:

IoU	Us	Winner
Overall (Mean)	0.29760	**0.7787**
Ground	**0.87620**	0.7974
High Vegetation	0.28345	**0.5654**
Building	0.29911	**0.7848**
Water	0.00000	**0.9417**
Elevated Road	0.02926	**0.8041**

Figure 2. The scores for the classifications of different objects, ranging exponentially from 0-1 (1 being the best) inclusive for the winner's results and ours.

U-Net:

U-net, a convolutional neural network, uses upsampling instead of traditional pooling methods. It creates layers of data through upsampling which can increase the output precision as more samples are run through.



Limitations:

- Time was a major constraint. The entire project was less than 3 months and the competition started halfway through winter break, putting us 3 weeks behind. We were competing against teams of PhD students.

Understanding the Data:

TYPES OF DATA:

- We had approximately 10GB of data to work with, including Lidar, RGB, and multispectral imagery taken from overhead. We had to take this data in image format and process it programatically to analyze segments of each image. We passed training data through the various sets of code we wrote to train our model to process the actual contest information.

USE OF SERVERS

- We had access to two main servers to process our data. Our lower powered machine, located on Miami's campus, did not have a large amount of graphical processing power for this job. We requested access to Ohio State's supercomputer system which would provide us with large amount of graphical processing power. These resources were accessed on demand in time increments as we worked on testing different aspects of our code.

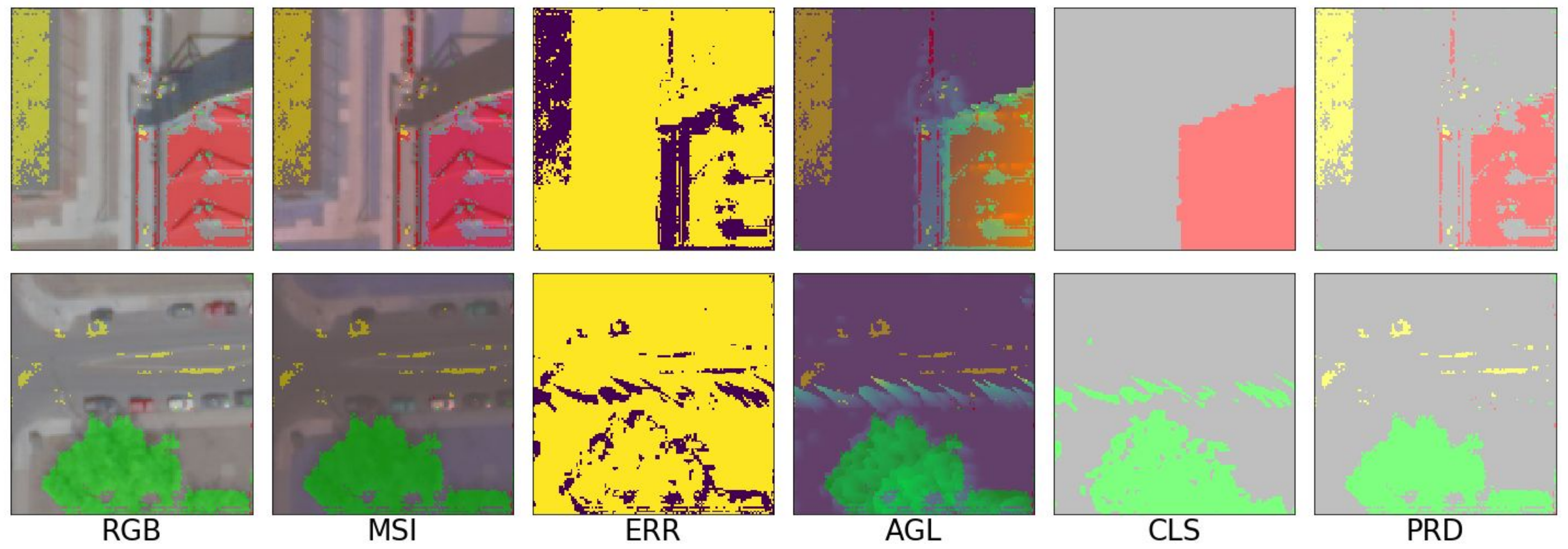


Figure 3. Predictions (PRD) are superimposed on RGB, MSI, and AGL images. Error (ERR) between PRD and ground truth classification (CLS) are shown as purple while yellow represents a correct prediction. For CLS and PRD, red=buildings, grey=ground, green=high vegetation, and yellow=elevated road.