

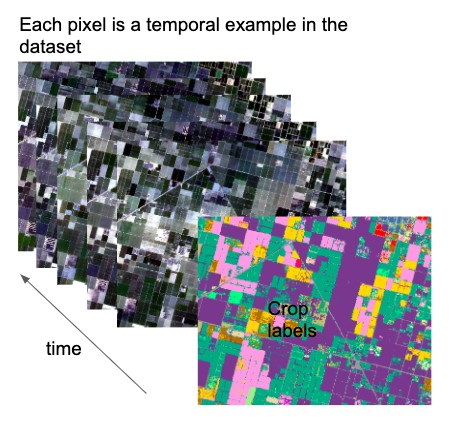
**Deep learning based maping of crops Using Satellite Imagery**

**18MI31008-Devanshu Pathak**

**18MI10012-Rahul Kumar Shaw**

# Introduction:-

Crop mapping is crucial for understanding the agricultural cover of our ever-increasingly crowded world. Remote sensing data, such as calibrated satellite photography, can be useful in tracking food growth throughout the world. Satellite imaging studies are frequently limited to publicly available data with low return rates and/or poor geographic resolution. A recent influx of satellite data from new-aerospace companies, on the other hand, offers daily imaging with reasonably high spatial resolution. The addition of temporal information into crop categorization algorithms is enabled by high revisit rates in satellite image acquisition. With high-frequency temporal data only recently becoming available, there is lots of space to investigate the data and categorization algorithms.



The image is a temporal signature of a pixel scene location throughout an agricultural growing season (time). The output are predictions for each agricultural crop type in the scene.

# Data Collections:--

The dataset consists of 10 RapidEye satellite images provided by the planet.com and 1 USDA Cropland data layer which provides the pixel level crop labels. With no current standard for temporal crop classification datasets, two separate datasets were constructed. Satellite data of crop fields was taken from the Planet Explorer API, and labels were constructed from the USDA Crop Data Layer database. Both sources were cropped to the same area of interest (AOI) defined by latitude and longitude coordinates, and the USDA data was sampled to match the Rapid Eye pixel resolution. Each Rapid Eye image has five spectral bands that represent blue, green, red, red edge, and near infrared spectral values. We get valid images at fifteen time stamps over two scenes, with at least one image per month. Scene 1 is a simpler dataset with six crop types, while Scene 2 is a more complicated, sporadic dataset with nine crop types. To define the pixels that will make up the datasets, we mask the top six and nine crops, respectively, from the USDA data, and only consider the pixels that fall under these masks. Scene 1 contains crop classes of cotton, safflower, tomato, winter wheat, durum wheat, and idle land in about 35 million pixels. Scene 2 contains crop classes of alfalfa, almonds, corn, cotton, idle land, pistachios, walnuts, winter wheat, and a corn/winter wheat mix in about 20 million pixels. To balance the datasets, we randomly select 100,000 pixels from each crop and add it to the corresponding dataset. We shuffle the selected examples and use a 90-05-05 split for the training, validation, and test sets, respectively.

For Scene 1, this leaves us with a training set of 540,000 examples, a validation set of 30,000 examples, and a test set of 30,000 examples, where each example (a pixel location in space) has 75 features (15 time stamps of 5 spectral bands). Scene 2 contains nine crop types, with a training set of 810,000 examples, a validation set of 45,000 examples, and a test set of 45,000 examples.

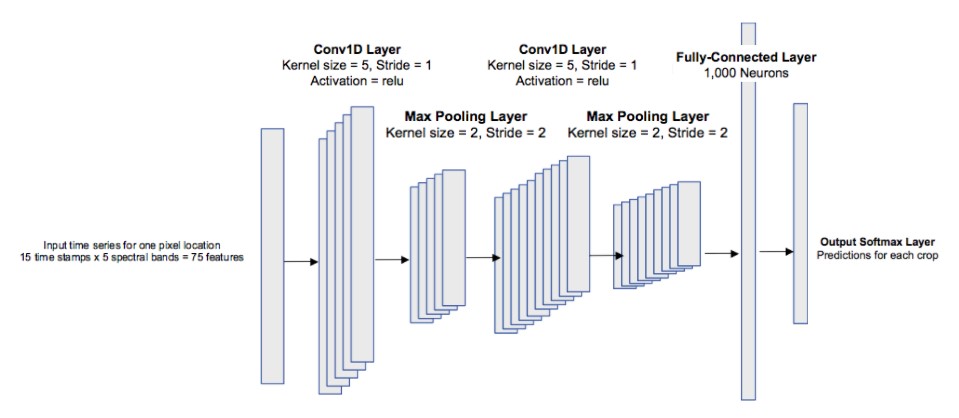
# Methods--

Multiple supervised machine learning algorithms, including softmax regression, support vector machines (SVMs) and deep learning algorithms like one-layer neural network (NN), and a convolutional neural network (CNN), have been used for the mapping of crop but the CNN stands out for its accuracy among them.

**Convolutional Neural Network (CNN)**:-

A CNN shares connection weights in each layer, rather than in the NN where each neuron has a unique connection to every other neuron in neighboring layers. Backpropagation is also used to find the network parameters, stored in matrices of convolutional kernels. CNNs are known for performing well in recognition tasks, and so we apply them here to identify features in the temporal crop signatures. The CNN architecture used is shown in the Figure. For the multi-temporal model, the convolutional layers in the network have 32 and 64 neurons, respectively, while the fully-connected layer has 1000 neurons. For the mono-temporal model, the convolutional layers have 8 and 16 neurons, respectively, while the fully-connected layer has 200 neurons. The model was simplified to give better performance for the lower-dimensional input feature space of the mono-temporal inputs. Weights were initialized with the Xavier initialization method, a cross entropy loss was used, and the network parameters were optimized with the Adam optimizer. In all cases, training ran for 10 epochs, with a batch size of 100 examples. Each example is a temporal signature, which is why the network inputs are vectors rather than matrices.

**The CNN architecture used for the project:-**



Code link:

Data Preprocessing:-

[https://colab.research.google.com/drive/1jr-eI-Wj6ZLVBTE2f6Li4BbszqZCXumr?us p=sharing](https://colab.research.google.com/drive/1jr-eI-Wj6ZLVBTE2f6Li4BbszqZCXumr?usp=sharing)

Deep learning implementation:[https://colab.research.google.com/drive/1Gr1xTE2HGFWbP9lcvurzRe0KdAw0Plpt? usp=sharing](https://colab.research.google.com/drive/1Gr1xTE2HGFWbP9lcvurzRe0KdAw0Plpt?usp=sharing)

# Results and Conclusion:----

Tables show the results of the multi-temporal models for Scene 1 and Scene 2, respectively. This provides an important insight that temporal information is helpful for crop classification, showing that crops have distinguishable temporal signatures throughout the growing season.

**Scene 1**:--- Multi-Temporal Results

Softmax Reg. SVMs SimpleNN CNN

Train Accuracy 92.38 96.46 92.88 95.66

Test Accuracy 92.06 92.21 92.06 92.64

**Scene 2**:---Multi-Temporal Results

Softmax Reg. SVMs Simple NN CNN

Train Accuracy 76.21 85.78 81.74 87.14

Test Accuracy 76.05 81.07 81.43 85.50

A multi-temporal result shows that temporal information can be helpful for successful crop classification. Surprisingly, all methods did well on Scene 1, with test accuracies around 92% for the multi-temporal. It is evident that temporal information is crucial for successful classification results of this more complicated scene.