Early detection of crop type using LANDSAT & cover classification imagery

Ashish Masih, Quan Nguyen, Shubham Saraf, Salem Nyacyesa.

University of Nebraska-Lincoln

1 Introduction

Irrigation scheduling necessitates regular monitoring of water availability in the root zone. However, when dealing with vast expanses of land, relying solely on ground-based measurements becomes unfeasible due to the sheer scale of the coverage area.
In such cases, remote sensing-based models come into play, leveraging satellite imagery to estimate water consumption on a larger scale. These models rely on inputs like satellite imagery, data from the nearest weather station, and crop type to create accurate models of crop water usage.
This project captivated our interest by enabling precise agricultural planning without extensive paperwork, aiding local farmers in tracking crop health and growth for their well-being and national food security, and supporting researchers in understanding the environmental impact of local cultivation for informed land use planning and biodiversity conservation.
Satellite images and weather station data are relatively accessible at frequent intervals. On the other hand, information about the current crop in cultivation can be obtained by directly reaching out to the farmer. The USDA annually releases crop classification data for the entire United States between January and February of the previous year. Consequently, applying remote sensing water-use models to the current year's crops becomes challenging.
Therefore, our goal is to create machine learning models with the capability to classify land types throughout the United States, identify the crops that have been planted, and determine the current growth state of those crops.

2 Related Works

2.1 Mapping of Crop Cover in an Irrigated Semi-arid region:

There was a paper published in 2019 [1] about using the LANDSAT-8 & Sentinel-2A satellite data in supervised classification of crop cover in the semi arid irrigated ares. They used the satellite temporal data as time series input to the TIMESAT software to develop a crop phonological stage curve 1. The software calculates the moving average of the pixel values (temporally) and creates a smooth crop phenological stage curve for entire image series (one year) .

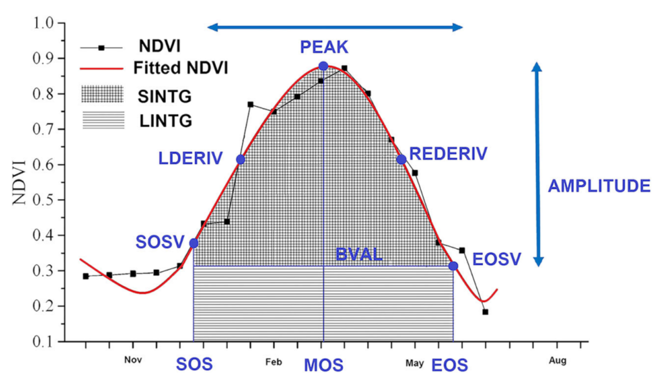


Figure 1: Crop Phenological stages [1].

Finally the satellite imagery (NDVI) was sub-divided based on the ground-truth data collected on each crop type. And several spectral signature curves each belonging to a specific crop were developed 2.

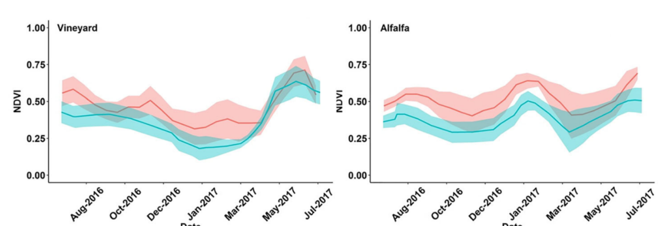


Figure 2: NDVI spectral crop signature [1].

The fact that each crop produces a different spectral signature during its growth stages was leveraged. And a random-forest machine learning model was created to implement supervised classification on the temporal NDVI data.

2.2 Object-based Image Classification for summer crop

The study by Kamble et al. [2] investigates the advantages of employing object-based approaches, as opposed to traditional pixel-based methods, for capturing spatial patterns and contextual information in the classification of summer crops. Various techniques, including Multilayer Perceptron (MLP), Convolutional Neural Network, Random Forest, feature extraction, and Support Vector Machines (SVM), were experimentally evaluated to comprehend their performance in crop classification.

Furthermore, the research delves into the impact of different input types on the classification process. The two primary input types explored were spectral features, derived from electromagnetic radiation characteristics of crops in images, and textural features, representing the distribution of pixel values in an image. The findings reveal that while textural features were beneficial for identifying fields of crops, their contribution to crop classification was not significant. In contrast, spectral features emerged as the primary contributors to accurate crop classification.

Although each strategy exhibited merits independently, the combination of SVM and MLP, along with spectral features, demonstrated a noteworthy improvement over any single strategy. The study also discerned that each crop exhibited a distinct sensitivity rate, largely influenced by the crop's representation in the dataset. Crops with higher representation demonstrated a more accurate identification by the model, emphasizing the importance of dataset composition in influencing sensitivity rates.

3 Image data:

We downloaded 10 years of LANDSAT imagery from 2013 - 2023 (available every 16 days) for the each year. Full dataset (9 bands per scene) from the Earth Explorer was downloaded and its size was approx. 89 GB (98 scenes in total). Due to limited processing and storage capabilities we had to filter out the least import bands. We kept only the NIR, Red & Thermal bands to create Normalized Difference Vegetation Index (NDVI) and thermal band to detect canopy temperature. Additionally, we also downloaded crop classification images for the same time-frame. Lastly, we used the current year's randomly selected images in the crop-growing phase to see how well our model performed.

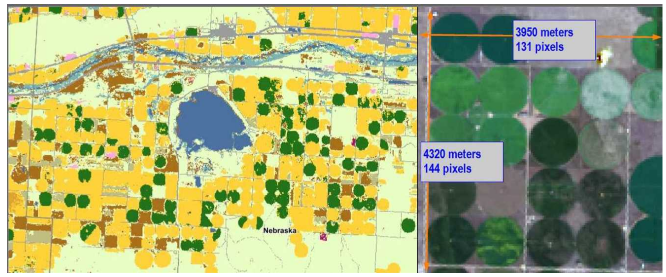


Figure 3: Crop classification & LANDSAT image (yellow is corn & green is soybeans).

4 Objectives

Objective 1*:

- Obtain the LANDSAT dataset and enhance those images using vegetation indices such as the NDVI for the identification of generic surfaces like water, bare soil, vegetation, etc. Additionally, perform data augmentation to improve the correlation when building the crop detection model.
- Create a segmentation map to get uniform crop classification boundaries i.e. avoid multiple crops in the same filed.

Objective 2*: Classify land type in broader categories e.g. agricultural land, wild grass, water bodies, bare soil, etc. in each image using a machine learning model or other techniques.

Objective 3*: Agricultural crops (a subset of the above objective) will be used to build a machine-learning prediction model that can identify the crop types e.g. corn, soybean, wheat, barley, etc.

Objective 4: Estimate the planting date of the crop and the current growth stage at any given date.

5 Problem definition

Given a dataset of satellite images, the goal is to identify and categorize crops in each image. The inputs are a set of LANDSAT images from single scene in Nebraska. Crop labels were obtained from USDA NASS Crop data [3] while crop growth stage was estimated from NDVI temporal images. Then a moving window filtering is applied to remove the crop signature it is used to implement the image segmentation

- A moving window filtering is applied to remove the mixed crop signature areas to improve correlation in the model generation.
- Generate a new image segmentation with cleaner boundaries to avoid the mixed crop signatures from the same field.
- Crop Vs No crop identification model by using crop's spectral signature from NDVI image.
- Create crop classification model using shortwave infrared image (brightness temperature) to detect the

crop as early as possible.

- Predict the crop growth stage using the NDVI Vs. Kc. (crop coefficient) [4]

6 Evaluation

For the first objective, the dataset will be evaluated based on the quality of each image, the number of images, and the distribution of labels in the dataset. Additionally, the image segmentation boundaries and mixed signature will also be qualitatively analyzed.

For the second, third, and fourth objective, the model will be evaluated based on accuracy (top-1, top-3, and/or top-5), F1-score, and other appropriate metrics. Other additional metrics may include how stable a model is, whether the model is over-fitting, and how long the inference takes for the model generation.

7 Implementation

7.1 Objective 1: Preprocessing & image segmentation

For dataset preparation, our team first identified LANDSAT scene (at path=31 & row=32) by selecting land surfaces with lots of agricultural land. Then, we downloaded each year's USDA crop classification image for the selected LANDSAT scene. LANDSAT images have a cloud percent property per scene. A cloud filter of 5 % was applied to remove images that are likely of creating problems later in the model building. After that, we obtained the crops in an $N \times N$ pixel window by searching the same $N \times N$ location in the CDL image of that year and saving it as an image only if all the pixels are of the same class (e.g. water, corn, soybean, grass, etc.). It was determined that the most common field size (center pivot) was approx. 27×27, therefore a moving window size of 9×9 was chosen to make sure a few $N \times N$ windows falls within the field. Additionally, the secondary objective of generating cleaner image boundary and avoiding mixed segmentation the meta AI's SAM model was implemented.

7.2 Objective 2: Crop vs no crop

The preprocessing of the data involves several steps to prepare the brightness temperature and landcover images for further analysis. The initial images are of size 857x853 pixels. However, to streamline the data and possibly reduce computational complexity, the images are cropped down to a smaller size of 460x460 pixels. This cropping operation focuses on retaining the relevant information while discarding unnecessary background or redundant pixels.

Following the cropping step, the preprocessed images are further divided into smaller patches of size 256x256 pixels. This patching process involves systematically segmenting the larger images into a grid of smaller subimages, each of which is 256 pixels in width and 256 pixels in height. This step is particularly useful in breaking down the images into more manageable and localized regions, facilitating subsequent analysis at a finer spatial scale. The non-overlapping nature of the patches ensures that each pixel in the original images is included in only one patch, avoiding redundancy in the data.

The entire preprocessing pipeline is crucial for optimizing the data for subsequent tasks of UNet Input. By cropping and patching the images, the preprocessing not only reduces computational demands but also enables the extraction of localized features, enhancing the granularity of the analysis on the brightness temperature and landcover data.

UNet is a convolutional neural network architecture designed for semantic segmentation tasks, which involve classifying and delineating different objects or regions within an image. It was introduced by Olaf Ronneberger, Philipp Fischer, and Thomas Brox in 2015 [5] and has since become a widely adopted architecture in the field of medical image analysis, as well as in various other computer vision applications.

The name "UNet" is derived from its U-shaped architecture, featuring a contracting path on one side and an expansive path on the other. The key innovation of UNet lies in its ability to capture context at multiple scales through a combination of downsampling and upsampling operations. This is achieved through a series of convolutional and pooling layers in the contracting path, which gradually reduces the spatial dimensions of the input, followed by upsampling layers in the expansive path that progressively restore the spatial resolution.

The architecture incorporates skip connections between corresponding layers in the contracting and expansive paths. These skip connections enable the network to retain high-resolution information during the down-sampling and up-sampling processes, facilitating the precise localization of objects. The skip connections help address the challenge of information loss during down-sampling, ensuring that fine-grained details are preserved for accurate segmentation.

Unlike the original UNet with a 512x512 input, we modified the original version to utilize a smaller input size of 256x256 pixels. This downsizing of the input dimensions offers computational efficiency, quicker training times, and a more streamlined model, especially in scenarios where computational resources are limited. Furthermore, the low-level resolution, initially set at 64x64 pixels in the original UNet, has been further reduced to 16x16 pixels in the modified architecture. This reduction aims to strike a balance between maintaining essential details and managing computational complexity.

7.3 Objective 3: Crop classification

7.3.1 Preprocessing and Data augmentation

To accomplish the third objective, I, Max, generated a collection of images that feature the Brightness Temperature of diverse regions across the USA. Brightness Temperature denotes the radiance level of short-wave infrared radiation detected by the satellite during the capture of LANDSAT images. Each image in this set has dimensions of 900 by 900 pixels. To create training and testing datasets, I randomly selected images of size 128 x 128 pixels from each larger image. Subsequently, I compared the locations of these sampled images with the USDA NASS Cropland Data to assign labels to each image [3]. This project involves working with 15 distinct labels.

Color	Description
#ffd300	Corn
#ff2626	Cotton
#00a8e2	Rice
#ff9e0a	Sorghum
#267000	Soybeans
#ffff00	Sunflower
#70a500	Peanuts
#00af49	Tobacco
#dda50a	Sweet Corn
#dda50a	Pop or Orn Corn
#7cd3ff	Mint
#e2007c	Barley
#896054	Durum Wheat
#d8b56b	Spring Wheat
#a57000	Winter Wheat

Figure 4: Crop labels

Each image sampled is assigned three labels corresponding to the three most cultivated crops. To promote an equitable distribution across 15 crops, I set a condition on the initial label, mandating a minimum representation of 6% for each crop type. This precautionary measure aims to prevent the dataset from being dominated by prevalent crops such as corn and soybeans.
Moreover, during the image generation process, I introduced diversity by incorporating random rotations. Additionally, certain images underwent cropping and resizing to produce blurred variants. Finally, the combination of two images was employed to generate novel composite images. These strategies collectively contribute to the diversification of the image dataset. In total, there are 90,000 images generated. Among these, 20% were set aside as test datasets.

7.3.2 Neural Network

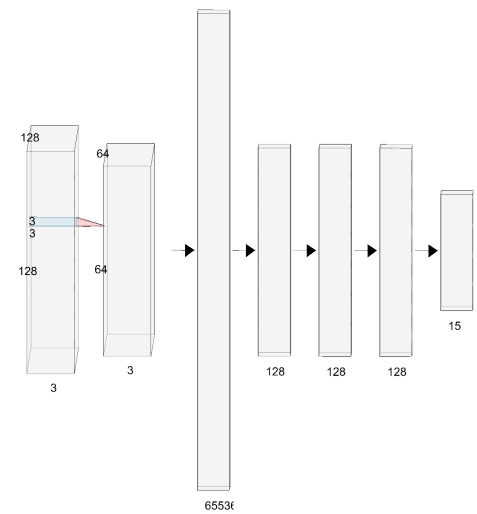


Figure 5: Neural network model

The neural network architecture comprises 2 convolutional layers followed by 4 fully connected layers. All layers, except the final one, employ a ReLU activation function, while the last layer utilizes softmax. The two convolutional layers possess 32 and 64 input channels, each succeeded by Batch Normalization and Max Pooling layers. The fully connected layers, except the last, consist of 128 hidden nodes. An extra Batch Normalization precedes the second fully connected layer, and a Dropout layer follows the third fully connected layer. The chosen loss function is Cross Entropy, and the optimizer function is Adam.
Batch normalization is incorporated to mitigate the impact of exploding gradients during training. Some irregularities in the training datasets, resulting from the combination of two different images, necessitate batch normalization to stabilize the input scale in specific layers. Furthermore, gradient clipping is employed to prevent spontaneous spikes in the learning rate, ensuring a stable training process.
Finally, due to the substantial size of the training dataset, it is practically unfeasible to load all images into memory using the available software. Consequently, the model initially trained on 80% of the training dataset, reserving one-fifth as a validation dataset. Every 5 epochs, a fresh

10% of the dataset is loaded into memory, and concurrently, 10% of the current training dataset is discarded. This strategy facilitates a systematic rotation of images throughout the training process.

7.4 Objective 4: Crop Growth

Using literature [6] & [4], we created a crop growth stage label for each crop type and further divided the images into four sub-folders, i.e., initial-stage, development-stage, mid-season-stage & end-stage. Depending on the crop coefficient (Kc) value and the Day of the Year. A particular location was assigned the appropriate class using the information provided in the figures below. Kc values were calculated using the following formula: Kc = 1.457 NDVI − 0.1725. The normalized difference vegetation index (NDVI) values were calculated using the following formula. NDVI = (NIR-Red)/(NIR+Red) where NIR is Near-Infrared.

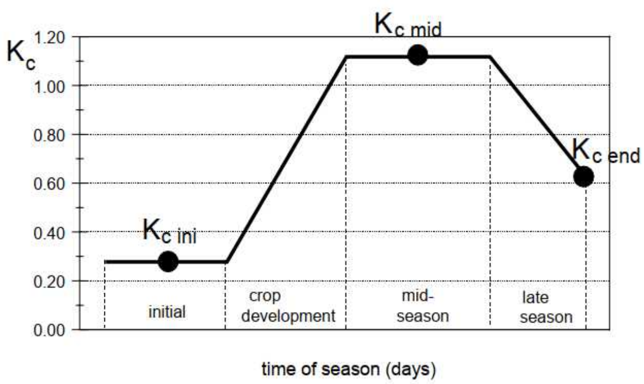


Figure 6: Crop coefficient Kc, it varies based on the crop growth stage.

For simplicity, our crop growth prediction model was trained, developed, and evaluated on corn only. Our first task was categorizing NDVI_images into their respective classes by matching the Day of Year Range and when the image had been captured. More information can be found in the table below. Secondly, we filtered out locations from the Cropland Data Layer where corn was the only crop being grown . That was accomplished by using the code for eliminating non-corn areas (corn_code = 1). The next step was applying the generated cdl_corn image mask onto the NDVI images so that we would obtain NDVI values for corn locations alone .

Kc range	DoY Range	Classes	St. Date	E. Date
0.3 ± 0.03	140-165	Initial	05/20	06/14
$0.4 - 0.9$	165-190	Devlop.	06/14	07/9
1 ± 0.2	190-220	Mid	07/09	07/29
0.6 ± 0.2	220-250	Late	07/29	08/27

Table 1: Breakdown of how NDVI images were selected to be placed in each class. DoY is Date of Year

In order to obtain locations where corn was in the correct growing stage, a technique similar to the NDVI filtration was used. At this point, only pixels with only corn and the matching Kc values were retained.

8 Result and analysis

8.1 Objective 1: Preprocessing & segmentation:

8.1.1 Post processing of LANDSAT images:

A qualitative analysis was done to see if the $N \times N$ filtering made any difference. A comparison of NDVI values scatter per crop class was done with original classes and $N \times N$ filtered classes 7.

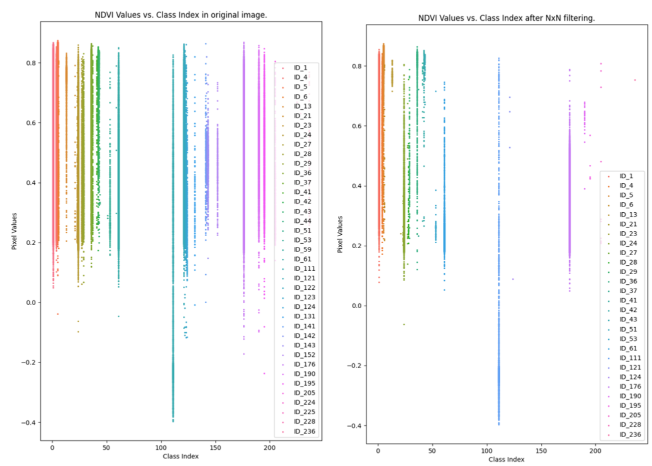


Figure 7: NDVI values scattering in original & filtered CDL image

It can be seen that original image had lot more classes but those classes were scattered all over the places and would make machine learning model building difficult. In the filtered image, there were lot less classes, but those classes have cleaner boundaries and no mixed class signature.

8.1.2 Crop area segmentation:

A segmentation algorithm was implemented [7] to see if it is possible to reduce the mixed signature from certain areas in the original CDL (Cropland Data Layer) imagery. SAM (Segment Anything Model, Meta AI) was implemented on a very small section of 7 Mega Pixels of the satellite image and it took approx. 9.5 min. to complete the processing on Google Colab. And the GPU memory utilization was 50%. The Colab kept crashing with larger images and it was going to take over 12.5 hours just to loop through all section of the single satellite scene. Therefore I (Ashish) decided to try segmentation on a small area to overcome the crash in GoogleColab due to limited resources. Nevertheless the output image is a good proof of concept.

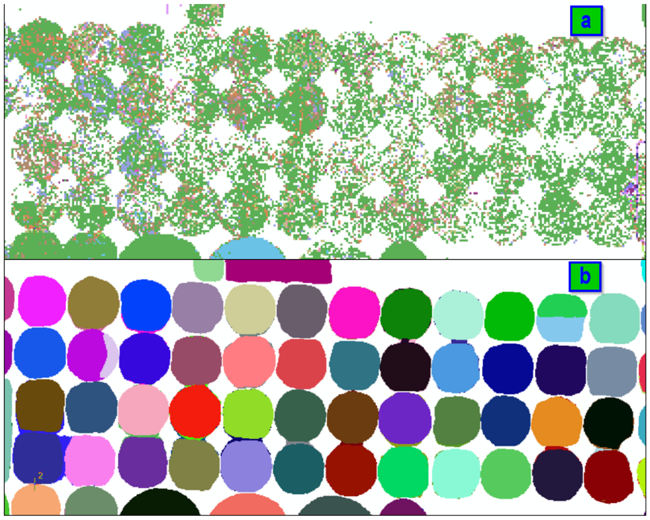


Figure 8: Image segmentation original Vs SAM by Meta AI. a) Original CDL classified image; b) SAM segmented image.

In the figure 8 the top part is the original CDL classification, whereas the bottom part (b) is newly segmented layer using SAM. It is evident that the SAM model works well, but it has to be implemented at a larger scale and that is not practical at free tier of Google Colab.
Note:

1. The image a & b doesn't match in color as the class ID values were different. To mach the class ID's, we might have to run some sort of logic where each feature will get the majority of pixel ID's from the original segment.

2. This segmentation was partially implemented at the time of presentation and it still needs improvement in order to be used as an input segment for our crop classification model.

8.2 Objective 2: Crop vs no crop

With the modified UNet architecture, featuring a reduced input size of 256x256 pixels and a lower-level resolution of 16x16 pixels, we conducted comprehensive training and evaluation on our dataset, ultimately achieving a commendable accuracy of 45%. This result signifies the effectiveness of the adapted architecture in successfully addressing the specific constraints and objectives of our segmentation task.

The decision to scale down the input size and resolution was driven by considerations of computational efficiency and training speed, which are particularly pertinent in scenarios with limited computational resources. Despite the reduction in spatial dimensions, the inclusion of skip connections within the UNet architecture played a pivotal role in preserving crucial details during the various stages of processing. This contributed to the model's ability to accurately segment and classify objects in the images, leading to the observed 45

It's essential to note that achieving optimal performance often involves a careful balance between model complexity, data characteristics, and computational constraints. The obtained accuracy serves as an encouraging validation of the adapted architecture's suitability for our specific segmentation task. Further fine-tuning and experimentation may lead to additional improvements, demonstrating the iterative nature of model development and refinement in the pursuit of enhanced performance.

8.3 Objective 3: Crop classification

The model is evaluated on 4 metrics: top1 accuracy, top3 accuracy, F1 score, and loss value. Regarding F1 score, even though restrictions are in place to prevent one label from dominating the dataset, such restrictions only affect the top label and not the second or third one. Therefore, F1 score is measured to ensure that there is not an uneven results across different labels.

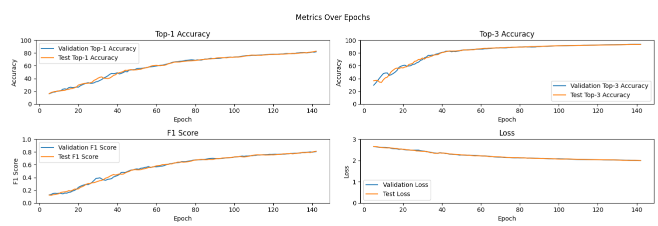


Figure 9: Metric over epochs

After undergoing training for more than 140 epochs, the model achieved a top-1 accuracy of 82% and a top-3 accuracy of 94%. For context, opting for corn as the label for all instances results in merely 7% top-1 accuracy, whereas selecting "corn, soybeans, barley" yields approximately 30% top-3 accuracy. It's important to note that these figures deviate from the actual crop distribution in the US due to a restriction on the number of labels each crop can have in the datasets. The F1 score reached 0.8, indicating a well-balanced distribution between precision and recall. Furthermore, there are no signs of overfitting, as the accuracy on the test dataset continues to improve with ongoing training.

However, a noteworthy observation emerges from the results. Despite the model's commendable performance, the accuracy graph suggests that the model has not yet converged. Both the validation and test datasets exhibit accuracy lines that have not plateaued, hinting that additional training may lead to further improvements. Additionally, some subtle fluctuations in accuracy during later epochs imply that the model parameters have not yet stabilized.

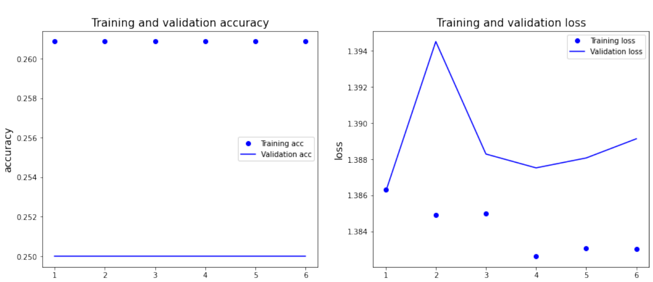
8.4 Objective 4: Crop Growth

A CNN-based model was trained, but it was extremely inaccurate, with an average accuracy of 25.0 percent. More data, new training, and testing techniques would be needed to improve the model. Ideally, the model should first be trained to recognize various NDVI values. The model can then be trained to recognize regions that are only exclusive to corn. Finally, the deeper layers of the model would be trained to recognize different crop coefficient (Kc) values and then make a prediction of which growth stage the crop is in.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 835, 843, 32)      320
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 417, 421, 32)      0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 415, 419, 32)      9248
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 207, 209, 32)      0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 205, 207, 64)      18496
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 102, 103, 64)      0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 100, 101, 64)      36928
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 50, 50, 64)        0
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 48, 48, 64)        36928
_________________________________________________________________
max_pooling2d_4 (MaxPooling2 (None, 24, 24, 64)        0
_________________________________________________________________
flatten (Flatten)            (None, 36864)             0
_________________________________________________________________
dense (Dense)                (None, 64)                2359360
_________________________________________________________________
dense_1 (Dense)              (None, 4)                 260
=================================================================
Total params: 2,461,540
Trainable params: 2,461,540
Non-trainable params: 0
```

Figure 10: Convolutional Neural Network Model

The model incorporates multiple Conv2D layers to perform hierarchical feature extraction, with each layer followed by MaxPooling2D to reduce spatial dimensions efficiently. Dense layers are then employed for final classification, providing a comprehensive understanding of the extracted features. The intermediate layers utilize ReLU activation to introduce non-linearity and capture complex patterns within the data. Finally, the output layer employs the softmax activation function, enabling the model to output probability distributions across multiple classes for effective classification.



As evident from figure , the CNN corn growth predictor had a very poor performance, as indicated by the confusion matrix and classification report. The model struggled particularly with the "0 Initial" class, failing to achieve any precision, recall, or F1-score for this category, suggesting a severe misclassification issue. Moreover, despite perfect recall for "1 Development," the low precision of 0.25 implies a tendency to incorrectly predict other instances as this class, leading to an unbalanced performance. Additionally, the model yielded zero metrics for classes "2 Mid" and "3 Late," further highlighting its inability to identify instances of these categories correctly. The overall accuracy was only 0.25, underscoring the model's limited ability to make reliable predictions across various corn growth stages.

```
Confusion Matrix:
[[0 1 0 0]
 [0 1 0 0]
 [0 1 0 0]
 [0 1 0 0]]

Classification Report:
               precision    recall  f1-score   support

    0 Initial       0.00      0.00      0.00         1
1 Development       0.25      1.00      0.40         1
        2 Mid       0.00      0.00      0.00         1
       3 Late       0.00      0.00      0.00         1

     accuracy                           0.25         4
    macro avg       0.06      0.25      0.10         4
 weighted avg       0.06      0.25      0.10         4
```

Figure 11: Confusion Matrixl

9 Summary and future work

For the first objective, a better strategy to curl all images and efficiently label each image may produce a higher-quality datasets. Additionally, for the secondary objective, the crop boundaries were clearly defined and there was very little mixed crop signatures for the crop parcels. Though, the barren land was left unclassified, therefore a further tweaking with the hyper parameters might add more classes to the barren land. Or some hybrid approach to keep the signatures from the original land cover image in case of not segmented areas.

For the second objective, The modified UNet architecture, with a reduced input size of 256x256 pixels and a low-level resolution of 16x16 pixels, This adaptation aimed to enhance computational efficiency while preserving essential details through skip connections. The success underscores the model's effective trade-off between reduced complexity and accurate segmentation. Ongoing refinement and experimentation may further optimize performance, emphasizing the iterative nature of adapting neural architectures to specific task constraints and objectives.

For the third objective, the model displayed remarkable performance in terms of accuracy with promising improvements if allowed to train further. Furthermore, future works, may encompass detecting where each crop is in an image. Additionally, larger iamge datasets, ideally including crops outside of US, can be the next stops. Lastly, better optimizations for faster training and inference such as quantization and transfer learning may allow the model to converge faster.
For the fourth objective, after examining the corn growth predictor model, it is evident that multiple things can be done to improve the model, such as expanding and diversifying the available dataset, incorporating robust data augmentation techniques, and refining the model architecture. Additionally, employing regularization techniques and fine-tuning the training process may mitigate overfitting issues. Future efforts should also explore advanced optimization algorithms to optimize convergence during training. Finally, the implementation of a more nuanced evaluation strategy with metrics tailored to the imbalanced dataset would provide a clearer understanding of the model's performance across diverse corn growth stages.

10 Note: About partial objective 4 details:

One of our team member was not able to produce the required details and consequently, in most places in this report, details related to objective-4 were missing or insufficient. He did tried to do everything in 24 hours before the deadline, but we as a team could not cross-check his works.
He (Salem) should be able to provide more details on that context.

References

[1] A. Htitiou, A. Boudhar, Y. Lebrini, R. Hadria, H. Lionboui, L. Elmansouri, B. Tychon, and T. Benabdelouahab, "The performance of random forest classification based on phenological metrics derived from sentinel-2 and landsat 8 to map crop cover in an irrigated semi-arid region," *Remote Sens. Earth Syst. Sci.*, vol. 2, pp. 208–224, Dec. 2019.

[2] J. M. Peña, P. A. Gutiérrez, C. Hervás-Martínez, J. Six, R. E. Plant, and F. López-Granados, "Object-based image classification of summer crops with machine learning methods," *Remote Sensing*, vol. 6, no. 6, pp. 5019–5041, 2014.

[3] USDA, "Usda nass cropland data," 1997.

[4] B. Kamble, A. Kilic, and K. Hubbard, "Estimating crop coefficients using remote sensing-based vegetation index," Mar 2013.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015.

[6] R. G. Allen, L. S. Pereira, M. Smith, D. Raes, and J. L. Wright, "Fao-56 dual crop coefficient method for estimating evaporation from soil and application extensions," *Journal of Irrigation and Drainage Engineering*, vol. 131, no. 1, p. 2–13, 2005.

[7] Q. Wu and L. P. Osco, "samgeo: A python package for segmenting geospatial data with the segment anything model (SAM)," *J. Open Source Softw.*, vol. 8, p. 5663, Sept. 2023.