# Enhancing Satellite Image Classification with Data Augmentation Techniques to Address Class Imbalance

Yutai Ke Hangze Wu

## Introduction

Classifying satellite images accurately is vital for applications in precision agriculture and industry, yet class imbalance in these datasets can lead to biased models that misidentify minority classes.(Mulla, 2012; Lobell et al., 2003)

To tackle this, we use a subset of the DeepSat-4 dataset, undersampling the 'barren land' class to create imbalance while keeping other classes unchanged. We enhance our convolutional neural network (CNN) model with data augmentation techniques such as rotations and flips, along with SMOTE and RandomOverSampler to generate synthetic samples for the minority class. This approach aims to improve classification accuracy and robustness, ultimately aiding in better decision-making and resource management.

## Literature Review

1. Industrial and Agricultural Impact
Satellite imagery is crucial for precision agriculture and industrial applications, improving spatial resolution, spectral diversity, and temporal frequency to enhance crop and soil monitoring (Mulla, 2012; Lobell et al., 2003). These advancements are vital for optimizing agricultural productivity and sustainable resource management (Bégué et al., 2018; Jiang et al., 2008). Therefore, accurate classification of satellite images supports better decision-making in response to environmental challenges and resource allocation.
2. Suitability of CNN for Image Classification
Convolutional Neural Networks (CNNs) are highly effective for image classification due to their ability to learn spatial hierarchies of features. Krizhevsky et al. (2012) demonstrated this in their groundbreaking results on the ImageNet dataset. Also, CNNs have been shown to perform well on satellite images by extracting relevant spatial features and improving classification accuracy (Rawat & Wang, 2017).

3. Imbalanced class prediction problems are also vital to agricultural tasks. In imbalanced datasets, oversampling methods, such as SMOTE, can generate synthetic samples to balance the dataset. He & Garcia (2009) and Batista et al. (2004) discussed these techniques' pros and cons. DeepSMOTE, as proposed by Dablain et al. (2021), effectively combines deep learning with SMOTE to generate high-quality synthetic images, enhancing the classification of minority classes.
4. Image Augmentation Techniques
Data augmentation methods like rotation and flipping are essential for enhancing the diversity of training datasets. Shorten & Khoshgoftaar (2019) provide a detailed review of these techniques, while Wong & Yeh (2018) specifically explore their application in satellite image classification, demonstrating significant performance improvements. Additionally, methods like MixChannel and other advanced augmentations have shown to improve the robustness of models in multispectral satellite images (Illarionova et al., 2021).

## Data Description

The DeepSat-4 dataset consists of 28x28 pixel satellite image patches extracted from the NAIP dataset, labeled into four classes: barren land, trees, grassland, and other land covers. The dataset includes 400,000 training samples and 100,000 test samples, with each image comprising four spectral bands (red, green, blue, and near-infrared). Accessible through Kaggle link: https://www.kaggle.com/datasets/crawford/deepsat-sat4

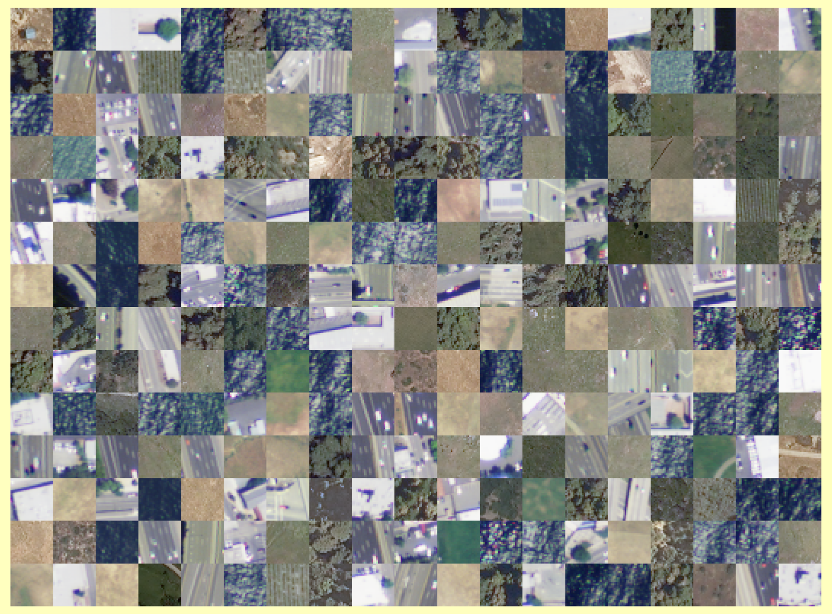

Figure 1: DeepSta-4 satellite images

We used a subset of DeepSat-4 for our train and test datasets. For the aim of address class imbalance dataset, we undersampled the 'barren land' class (class 0) to a minority class.

## Methodology

### Basic CNN setup

To start with, we adapted a CNN model setting from a kaggle project(Rahul, 2000) with some custom modifications.

The CNN we designed for classifying images into four categories consists of the following layers: two convolutional layers with 128 filters for feature detection, each followed by Leaky ReLU activation and max-pooling layers to introduce non-linearity and reduce spatial dimensions. The network then flattens the feature maps into a one-dimensional vector, which is fed into a fully connected layer with 128 neurons for complex pattern learning. The final layer is a softmax output layer for classification. The model uses the Adam optimizer and categorical cross-entropy loss function, achieving an accuracy of 0.950 and a categorical cross-entropy of 0.236 on the test set. Unsatisfied with these metrics, we performed a grid search over a set of hyperparameters for further improvement.

### Grid Search optimized CNN

After the grid search, the basic setup of the model remains unchanged, but some interesting phenomena happened: in the optimal combination of hyper parameters we have, the two convolutional layers now each have only 32 filters, and the second-last layer now also only has 32 neurons. This result suggests that in the initial model we are using a too complicated setup which causes overfitting, resulting in a bad out-of-sample performance. The improved accuracy and Categorical Cross-Entropy are 0.963 and 0.109, so we do achieved a strict improvement from the new model.

### Undersampling-Upsampling analysis

Here starts the key part of our analysis: the effect of class imbalance and the efficiency of the various upsampling methods. To carry our analysis systematically, we designed this workflow, as shown in the figure bellow:

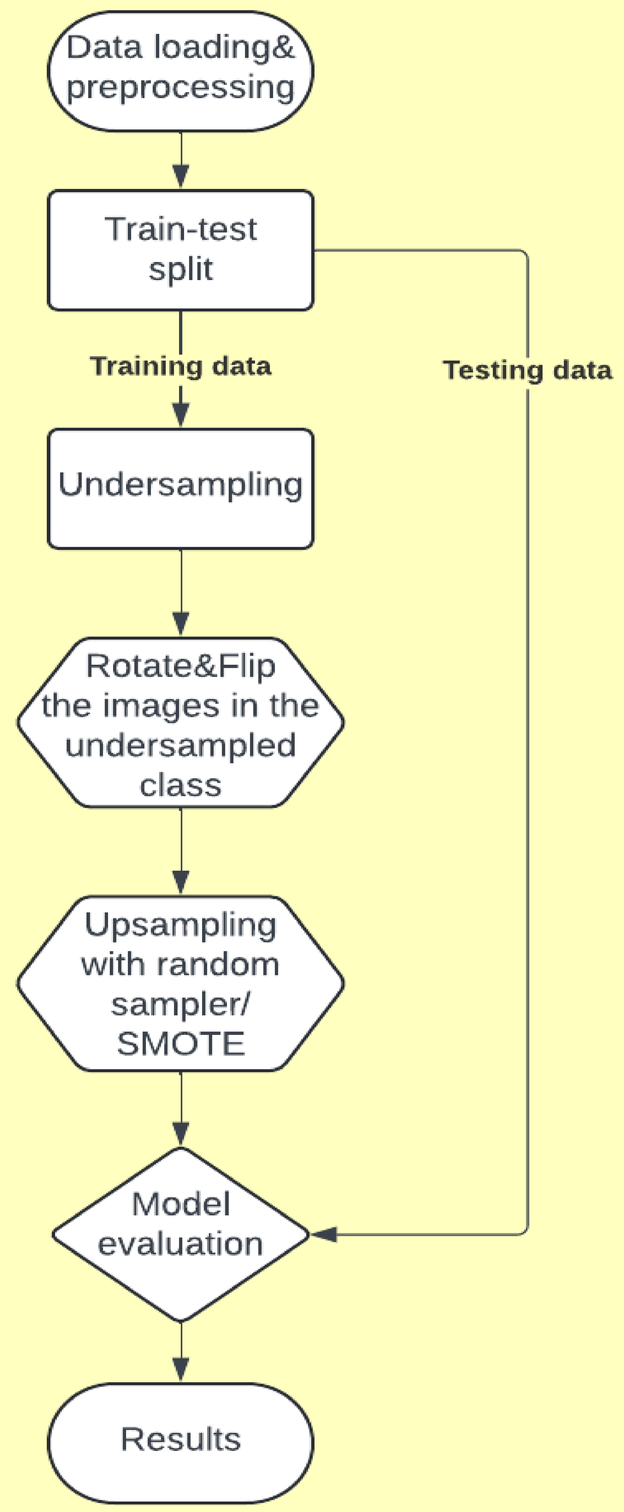


Figure: Workflow of out analysis

## Methodology

Where:

- We have two types of process boxes: rectangular ones are obligatory and the hexagons are optional. We will focus on different combinations of the methods in the hexagons and evaluate their (combined) performance.
- When doing the train-test split, we will make the training set perfectly balanced in the sense that the 4 classes in the dataset all have exactly 10,000 samples.
- In the undersampling feature box, we will pick one of the four classes we have (in our code we just picked the first one, corresponding to the Barren land (drought) class in the dataset) and delete data from it until it only represents $0.1\%$ of the data. Now this class is extremely imbalanced.
- In the first optional hexagon, for the underrepresented class label, we will seek to exploit the image nature of the data to populate the dataset. First, we will rotate the images 90, 180, and 270 degrees, and include these pictures into the dataset. Then, for each picture with the minority class in the updated dataset, we will flip the image vertically, horizontally, and across both diagonals. The result will be that each image in the minority class will become:

$$1\ \text{(original)} + 3\ \text{(rotations)} + 4 \times 4\ \text{(flipping each of the 4 images)} = 1 + 3 + 16 = 20$$

- In the second optional hexagon, we will upsample the training data (after reshaping it to a 2D array) with a random sampler or SMOTE technique, then reshape it back to a image vector format.

## Results

| Method | Accuracy (overall) | Recall for minority class |
|---|---|---|
| Class imbalanced | 0.931 | 0.76 |
| random sampler only (second hexagon) * | 0.9467 | 0.81 |
| SMOTE only (second hexagon) * | 0.962 | 0.86 |
| image rotating & flipping + Random sampler ** | 0.964 | 0.87 |
| image rotating & flipping + SMOTE ** | 0.973 | 0.91 |
| image rotating & flipping only (first hexagon) * | 0.951 | 0.83 |

Table 1: Methods comparison: One star (*) indicates the use of one optional hexagon, two stars (**) indicate the use of both. No stars represent the baseline model with imbalanced data fitted directly without preprocessing.

From the table above we can observe that:

- In the class-imbalanced dataset, the accuracy is negatively affected by the class imbalance. A poor recall for the minority class indicates that the model fails to detect all class 1 samples in the test set due to an insufficient learning process.
- When using data upsampling methods only (like random sampler and SMOTE in our case) without considering the image nature of the dataset, we can improve the performance of the model, but only to a limited extent.
- When using both image rotating and flipping strategies, we achieve the best performance, with accuracy even better than the original model trained with a balanced dataset. We deduce that this is because adding rotated and flipped images to the dataset effectively teaches the model to be invariant to images taken from different directions.
- Lastly, if we only use the rotating and flipping strategies, the performance is still suboptimal because the class is still relatively imbalanced.

## Conclusions and future steps

In image datasets with extreme class imbalance, combining the strategies that leverage the image nature of the data and standard multivariate upsampling methods will yield the best outcome.

In the future, we would like to extend our work to bigger image samples with higher resolution and more classes, and also with deep neural networks. Apart from that, we can consider SMOTE not only before feeding the data to the model(where the dimension is usually high), but also after convolutional layers where features has been extracted and dimension has been reduced. According to Dablain et.al, (2021), this yields better results.

## References

Referrence is included in next page.

References


- Batista, G. E., Prati, R. C., & Monard, M. C. (2004). A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data. *SIGKDD Explorations*. https://doi.org/10.1007/s10115-004-0170-3
- Bégué, A., Arvor, D., Bellon, B., Betbeder, J., De Abelleyra, D., Ferraz, R. P. D., ... & Simoes, M. (2018). Remote sensing and cropping practices: A review. *Remote Sensing*, 10(1), 99. https://doi.org/10.3390/rs10010099
- Basu, S., DiBiano, R., Karki, M., Mukhopadhyay, S., Ganguly, S., & Nemani, R. R. (n.d.). DeepSat (SAT-4) Airborne Dataset. Retrieved from http://csc.lsu.edu/ saikat/deepsat/http://csc.lsu.edu/s̃aikat/deepsat/
- Dablain, D., Krawczyk, B., & Chawla, N. V. (2021). DeepSMOTE: Fusing Deep Learning and SMOTE for Imbalanced Data. *arXiv preprint arXiv:2105.02340*. https://doi.org/10.48550/arXiv.2105.02340
- He, H., & Garcia, E. A. (2009). A Study of Techniques for Handling Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*. https://doi.org/10.1109/TKDE.2008.239
- Illarionova, S., Nesteruk, S., Shadrin, D., Ignatiev, V., Pukalchik, M., & Oseledets, I. (2021). MixChannel: Advanced Augmentation for Multispectral Satellite Images. *Remote Sensing*. https://doi.org/10.3390/rs13112181
- Jiang, Z., Huete, A. R., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, 112(10), 3833-3845. https://doi.org/10.1016/j.rse.2008.06.006
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *NIPS*. https://doi.org/10.1145/3065386
- Lobell, D. B., Asner, G. P., Ortiz-Monasterio, J. I., & Benning, T. L. (2003). Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. *Agricultural and Forest Meteorology*. https://doi.org/10.1016/S0167-8809(02)00021-X
- Mulla, D. J. (2012). Twenty-five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*. https://doi.org/10.1016/j.biosystemseng.2012.08.009
- Rahul, D. S. (n.d.). Agricultural drought prediction, Kaggle Notebook. Retrieved June 25, 2024, from https://www.kaggle.com/code/dsrhul/agricultural-drought-predictionhttps://www.kaggle.com/code/dsrhul/agricultural-drought-prediction
- Rawat, W., & Wang, Z. (2017). Deep Learning for Image Classification: A Review. *IEEE Access*. https://doi.org/10.1109/ACCESS.2017.2696520
- Rumelhart, D., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. 1, pp. 318-362). MIT Press. https://doi.org/10.1016/B978-1-4832-1446-7.50035-2
- Shorten, C., & Khoshgoftaar, T. M. (2019). A Survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*. https://doi.org/10.1186/s40537-019-0197-0
- Wong, T. T., & Yeh, P. Y. (2018). Data Augmentation for Satellite Image Classification Using Rotation and Flipping. *Remote Sensing*. https://doi.org/10.3390/rs10060925