RÉPUBLIQUE FRANÇAISE
*Liberté*
*Égalité*
*Fraternité*

ONERA THE FRENCH AEROSPACE LAB

SONDRA Lab

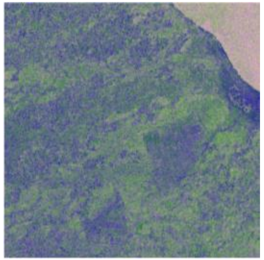
université PARIS-SACLAY

CentraleSupélec

# Multi-Branch Deep Learning model for detection of settlements without electricity

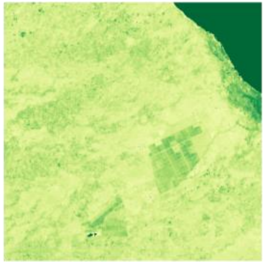
*Thomas Di Martino, Maxime Lenormand, Élise Colin Koeniguer*

# Data – Satellite Imagery



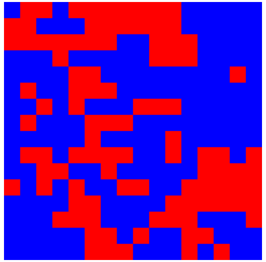
**S1**
**(VV, VH, VV-VH)**

**S2**
**(B04, B03, B02)**



**S2**
**(B08-B04)/(B08+B04)**
**“*RdYlGn*” Color Scale**

**VIIRS**
**(DNB)**

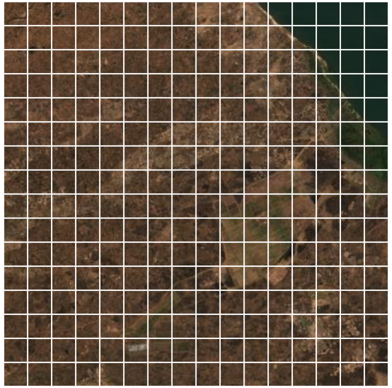


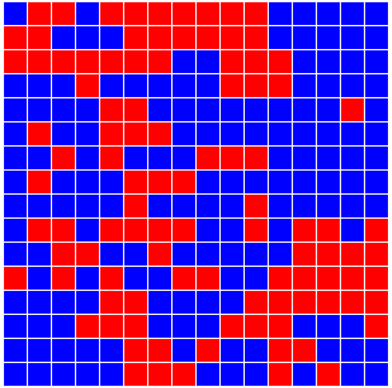
**Prediction Map**
**(Red: Human settlements without electricity, Blue: No human settlements without electricity)**

60 images of 800x800 pixels combining multimodality & multi-temporality:

- Sentinel 1**: 4 acquisitions** with **2 bands** each, resolution ~10m (GRD product)
- Sentinel 2: **4 acquisitions** with **12 bands** each, resolution of 10m, 20m & 60m
- Landsat 8: **3 acquisitions** with **11 bands** each, resolution of 15m, 30m, & 100m
- VIIRS: VNP46A1 product, for **9 acquisitions**, with the **Day-Night Band** only, resolution of 750m resampled to 500m

# Our perception of the dataset

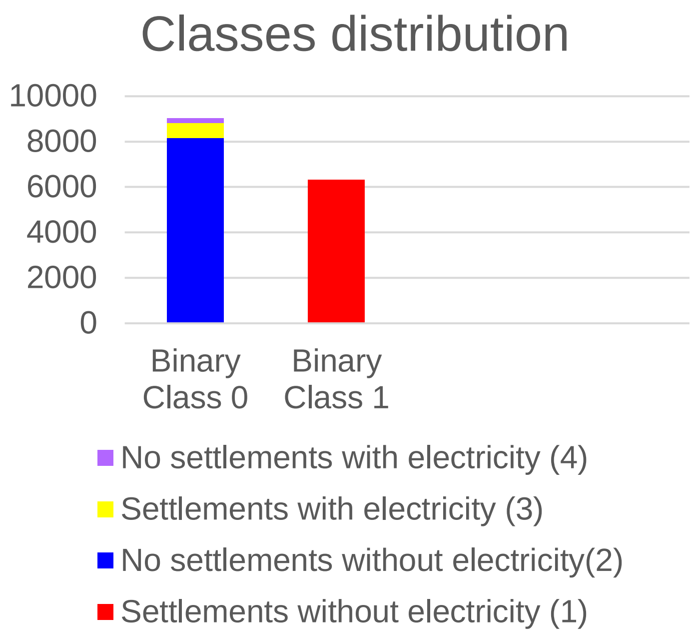




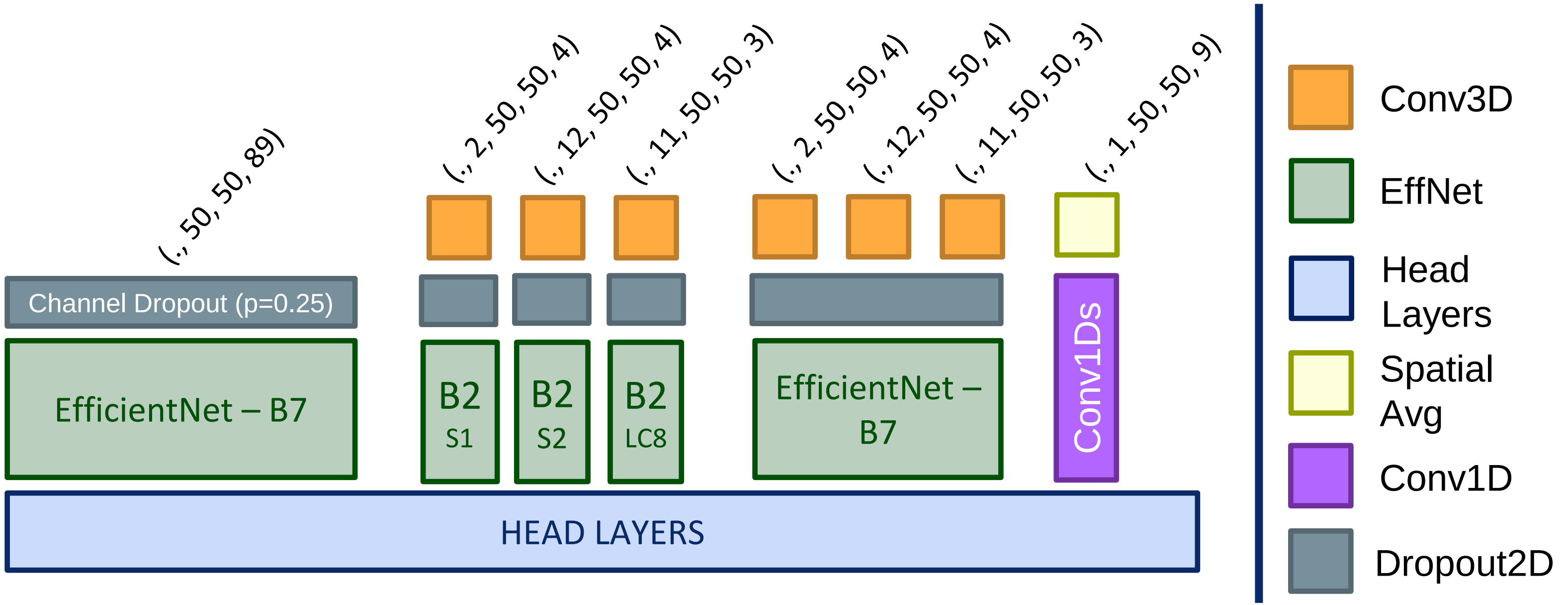
- Each label image consists of a **16x16 matrix of classes upsampled** to match the original dataset resolution of 800x800 pixels.
- We **split** the initial 800x800 images **into 256 50x50** patches, each having a single class value.
- We transform a segmentation task into a classification one.
- We split the new dataset of 15 360 images into 3 folds for cross-validation.

# Data – Ground Truth Labels

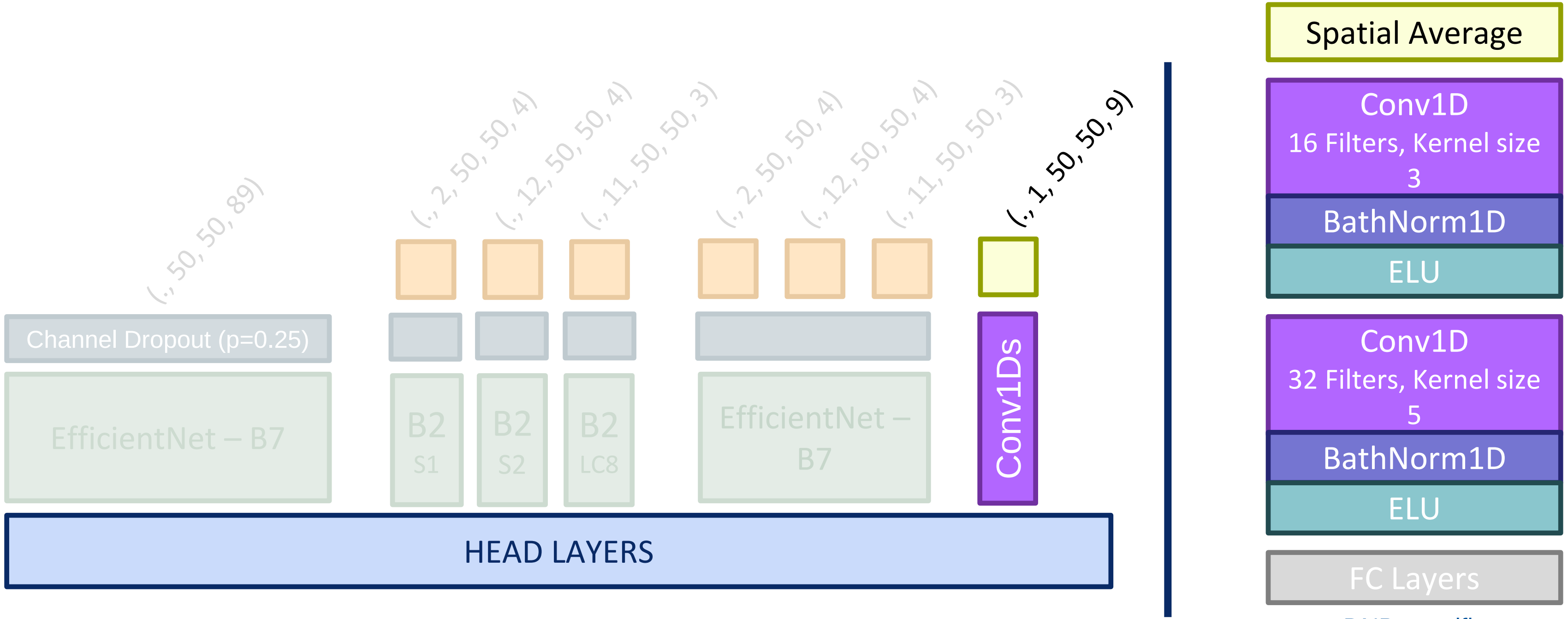
| | With **electricity** | Without **electricity** | |
|---|---|---|---|
| With **settlements** | 676 | 6318 **(ROI)** | 6994 |
| Without **settlements** | 211 | 8155 | 8366 |
| | 887 | 14473 | |



# Model Architecture

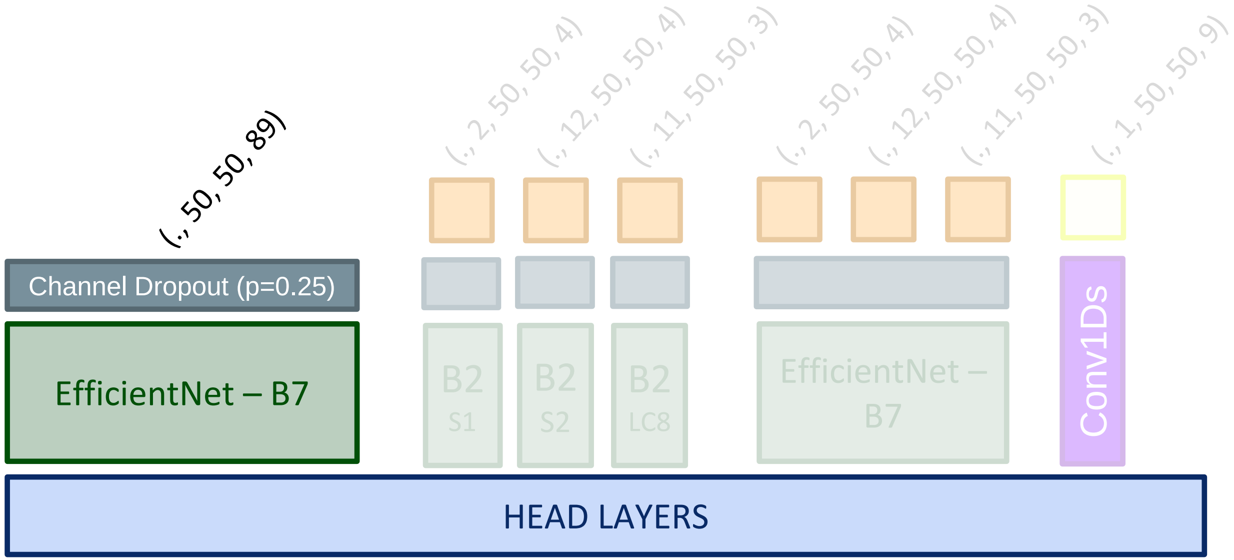


# Model Architecture: Day-Night Band processing branch



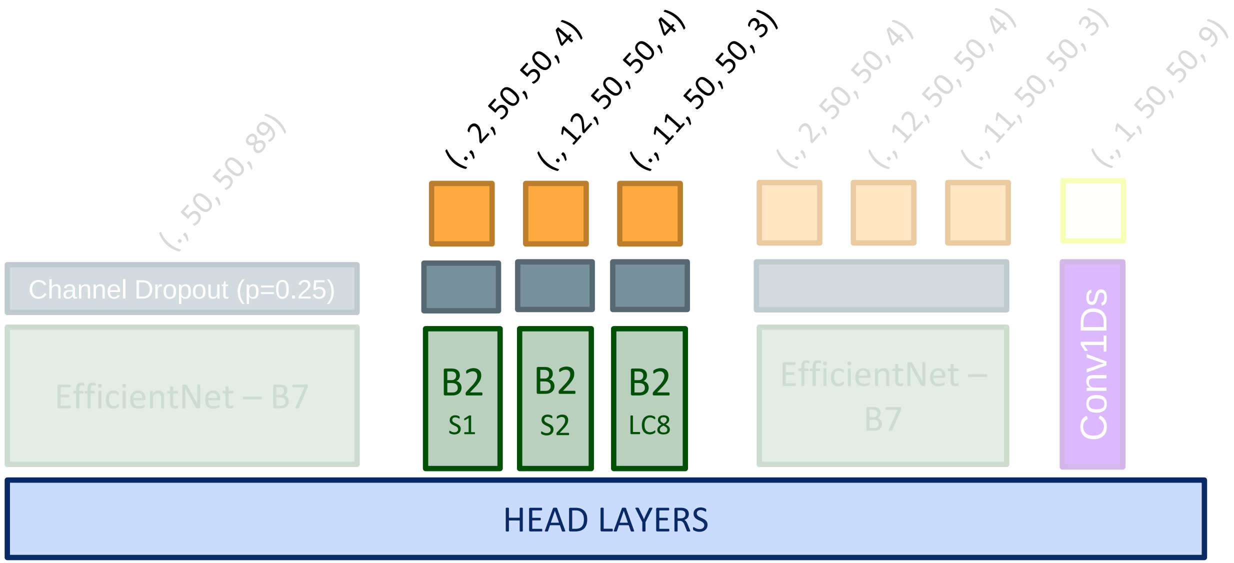

DNB-specific TempCNN model

# Model Architecture: Multimodal Branch

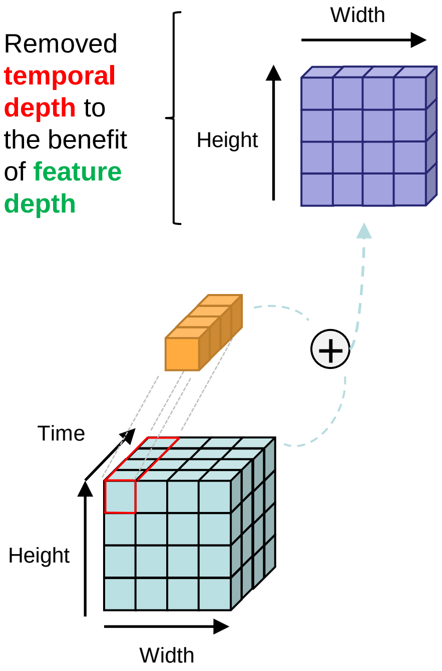



- 89 channels is a lot and can be detrimental to convergence
- Adding channel dropout with a rate of 25% helps for that matter

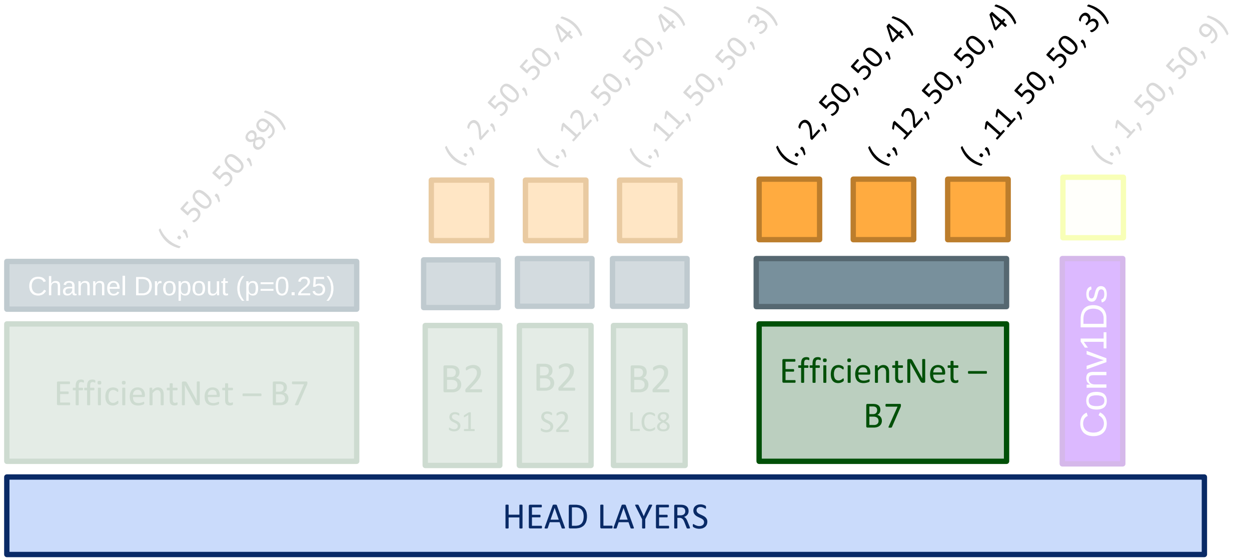
# Model Architecture: MultiUnimodal branch




Removed **temporal depth** to the benefit of **feature depth**

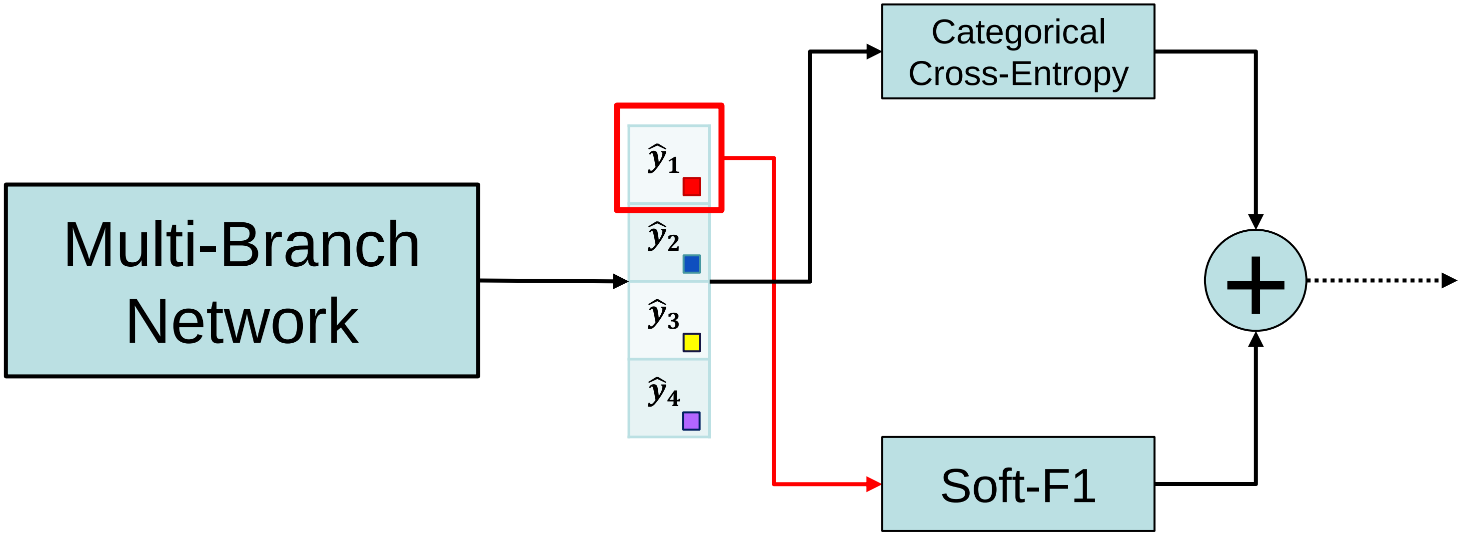


# Model Architecture: Temporal-Merged branch




- Separate temporal processing but merged texture feature extraction
- Mix between the Multimodal and the Multi-Unimodal branches

# Training setup: Loss calculation



# Training setup: Soft-F1 Loss

We define the equations for our Soft-F1 Loss as the following:

$$\left.\begin{aligned} \boldsymbol{Soft - precision(y, \hat{y})} &= \frac{\sum \boldsymbol{\hat{y}_1 * y_1}}{\sum \boldsymbol{\hat{y}_1 + \epsilon}} \\ \boldsymbol{Soft - recall(y, \hat{y})} &= \frac{\sum \boldsymbol{\hat{y}_1 * y_1}}{\sum \boldsymbol{y_1 + \epsilon}} \end{aligned}\right\} \boldsymbol{Soft - F1(y, \hat{y}) = 1 - \frac{2 * sp(y, \hat{y}) * sr(y, \hat{y}) + \lambda}{sp(y, \hat{y}) + sr(y, \hat{y}) + \lambda}}$$

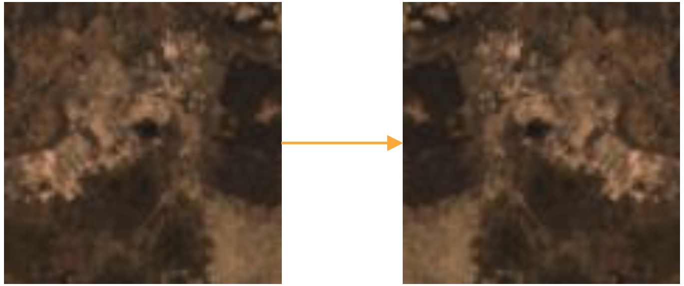
The two added hyperparameters are the following:

- $\epsilon$ is a term used to prevent any divison by zero, resulting in a Loss value equal to Nan. It is set to $1e-5$.
- $\lambda$ is a term used to smoothen the Soft-F1 loss value. It is set to $0.1$.

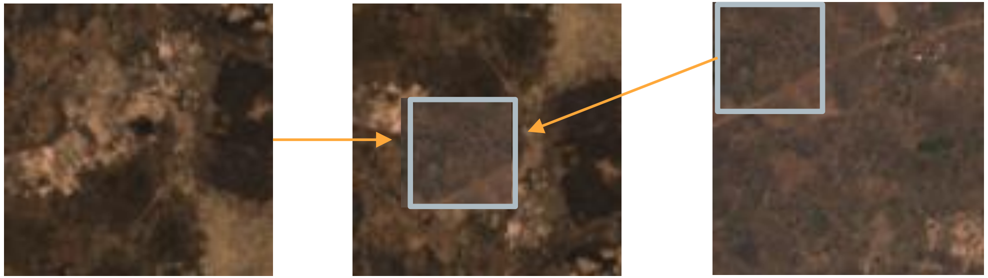
Inspired by: https://www.kaggle.com/c/planet-understanding-the-amazon-from-space/discussion/34484#191547

# Data Augmentation

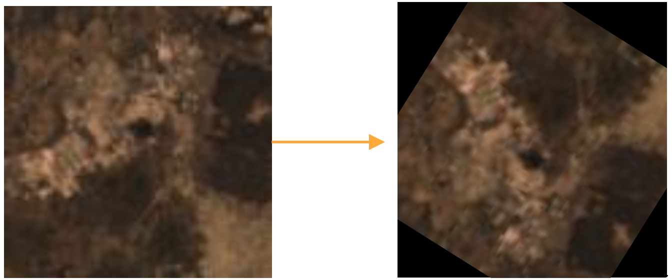
**Horizontal/Vertical Flip**



**CutMix** (restricted between same-class images)



**Random Rotation**



**Noisy Labels** (random offset in cropping)

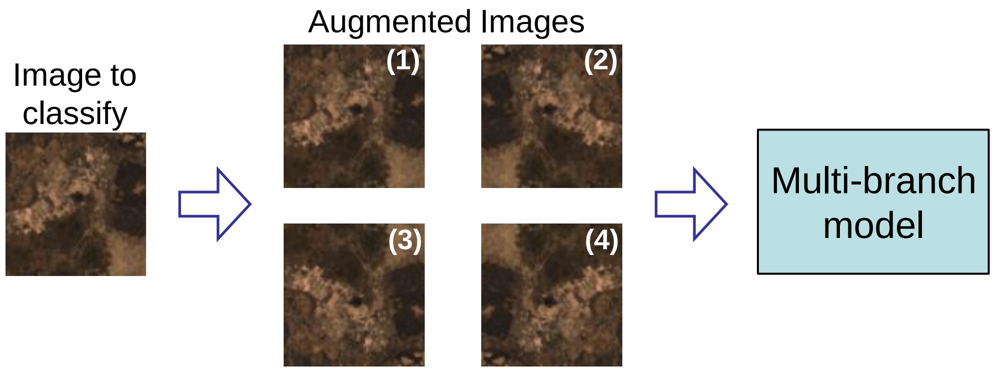
# Test-Time Augmentation & Ensembling

## INFERENCE

### Test-Time Augmentation

1. No augmentation
2. Horizontal Flip
3. Vertical Flip
4. Horizontal + Vertical Flip

### 3-Folds Prediction Ensembling




$$\boldsymbol{pred}(\boldsymbol{X}) = \begin{cases} \mathbf{1}, & \sum_{aug=1}^{4} \sum_{m=1}^{3} \boldsymbol{model}_m\left(\boldsymbol{f}_{aug}(\boldsymbol{X})\right)_1 > \boldsymbol{thresh} \\ \mathbf{0}, & \boldsymbol{else} \end{cases}$$

# Final results

| Fold 1 Val | Fold 2 Val | Fold 3 Val | Dev Phase | Test Phase |
|---|---|---|---|---|
| 0.8547 | 0.8533 | 0.8722 | 0.8877 (1st) | 0.8798 (3rd) |

Winning model submission F1 score

# Assessing VIIRS ability to detect electrification

- Using a binary version of the dataset with the classes: electrified, not electrified

- Isolation of the DNB-specific branch to train as a separate classifier of electrification

| With **electricity** | Without **electricity** |
|---|---|
| 887 | 14473 |

| Subset | Fold 1 | | Fold 2 | | Fold 3 | |
|---|---|---|---|---|---|---|
| | Train | Val | Train | Val | Train | Val |
| VIIRS | 0.712 | 0.718 | 0.765 | 0.462 | 0.674 | 0.75 |


VIIRS TempCNN model

# Assessing each sensor ability to detect the class of interest

- Using combinations of VIIRS and each sensor to assess and study their ability to detect *settlements without electricity*

| Sensor Subset | Fold 1 | | Fold 2 | | Fold 3 | |
|---|---|---|---|---|---|---|
| | Train | Val | Train | Val | Train | Val |
| S1, VIIRS | 0.681 | 0.653 | 0.657 | 0.665 | 0.677 | 0.672 |
| LC8, VIIRS | 0.766 | 0.758 | 0.750 | 0.745 | 0.768 | 0.776 |
| S2, VIIRS | **0.834** | **0.817** | **0.822** | **0.824** | **0.850** | **0.859** |
| S1, S2, LC8, VIIRS | 0.893 | 0.854 | 0.900 | 0.853 | 0.878 | 0.872 |

# General Conclusion

- Development of a multi-branch architecture, acknowledging the multimodal and multitemporal structure of the data

- Design of a custom training & testing environment (custom loss, data augmentation, TTA & ensembling)

- Experimentations displaying the contribution of each sensor to the final prediction (S2 > LC8 > S1)

- Potential axis of improvements: reflect regarding the type of data to be aggregated & how to combine them in a physically meaningful way (e.g., SAR Time Series & Interferometric products could be of interest)

# Acknowledgements


- Thank you to the IEEE GRSS IADF Technical Committee, Hewlett Packard Enterprise, SolarAid & Data Science Experts for organizing this Data Fusion Contest.

- Congrats to all participants for their results and thank you for the exciting *Development phase* and the thrilling and positively stressful *Test Phase*. ☺

- Many thanks to our colleagues at ONERA for their support during the challenge, especially Adrien Chan Hon Tong, Aurélien Plyer and Guy Le Besnerais.

- And many thanks to you for attending this presentation !

# Bibliography

[1] Naoto Yokoya, Pedram Ghamisi, Ronny Hansch, Colin Prieur, Hana Malha, Jocelyn Chanussot, Caleb Robinson, Kolya Malkin, and Nebojsa Jojic, “2021 data fusion contest: Geospatial artificial intelligence for social good [technical committees],” *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, no. 1, pp. 287–C3, 2021.

[2] Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk  Chun, Junsuk Choe, and Youngjoon Yoo,“CutMix: Regularization strategy to train strong classifiers with localizable features,” 2019.