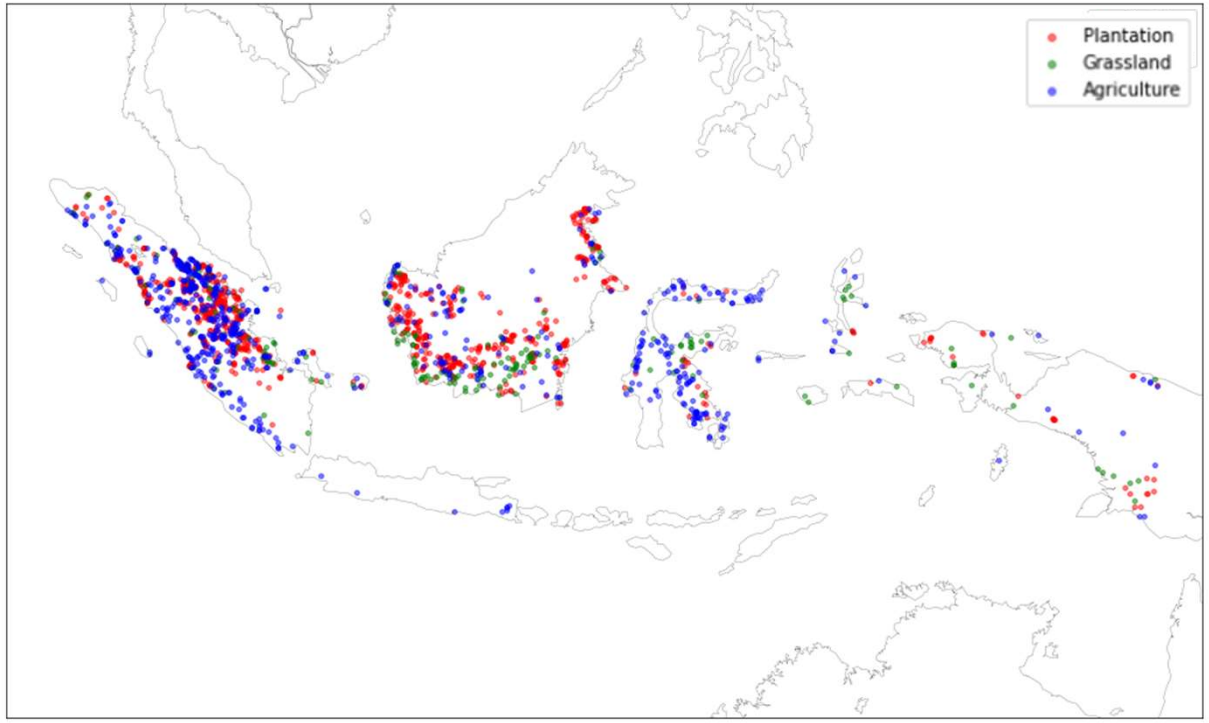
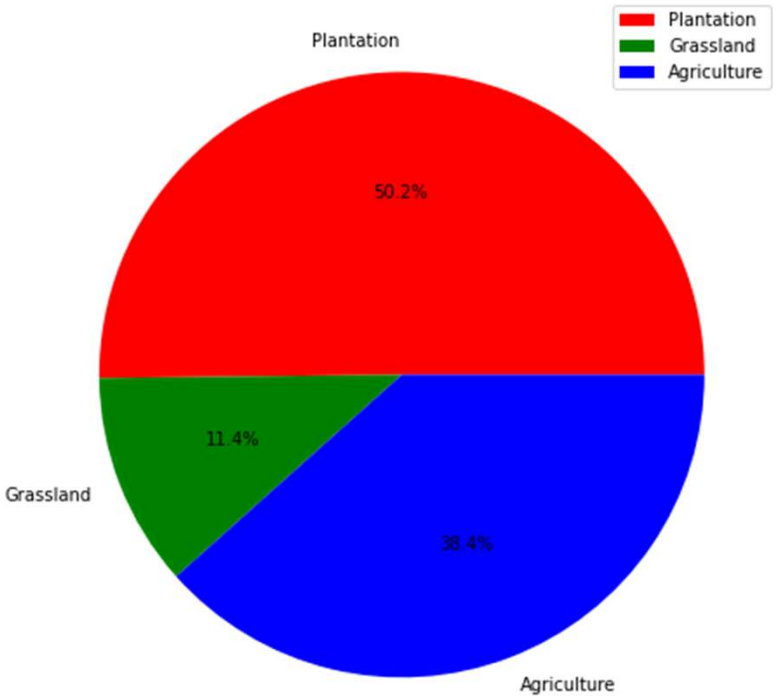
Exploratory Data Analysis

Explanation and reasoning of the Methodology used

Results and Conclusions

Exploratory analysis



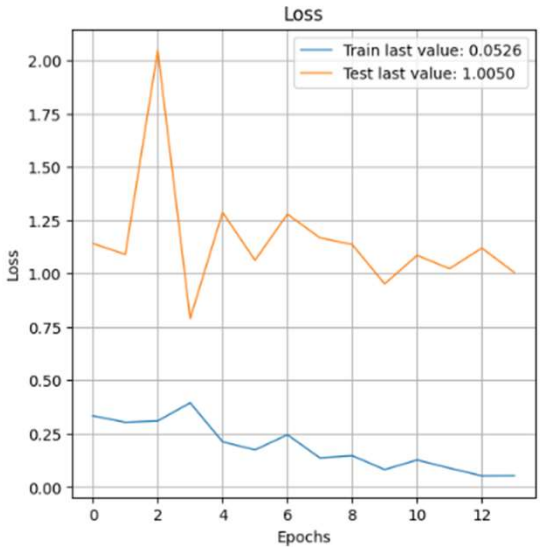


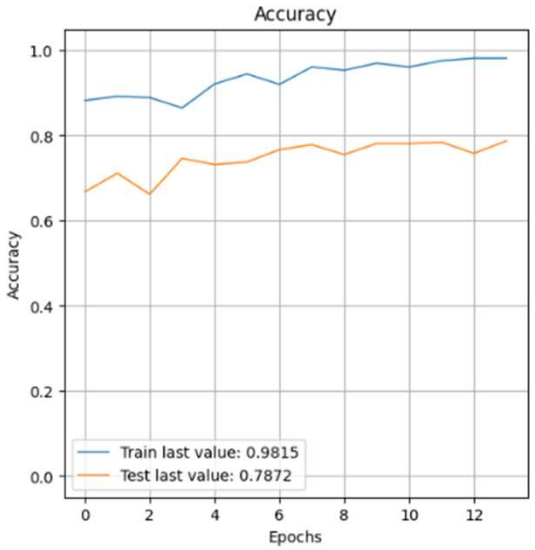
- **1714** images
- The distribution by categories is **unbalanced**, with a small quantity of Grassland pictures.
- Images between **2001 and 2016.** Most of the pictures are from the period 2007-2016, with approximately 120 pictures per year. **Grassland pictures** were taken **after 2012**.
- The pictures are mainly concentrated in the regions of **Sumatra, Kalimantan and Celebes at Indonesia.** All the categories have a **good grade of dispersion**.

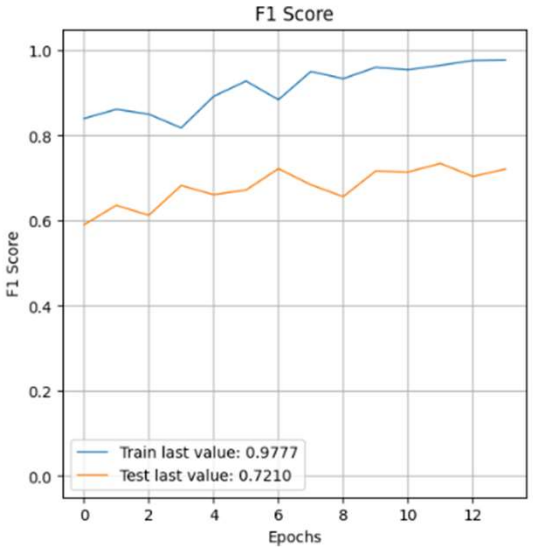
Explanation and reasoning of the Methodology used

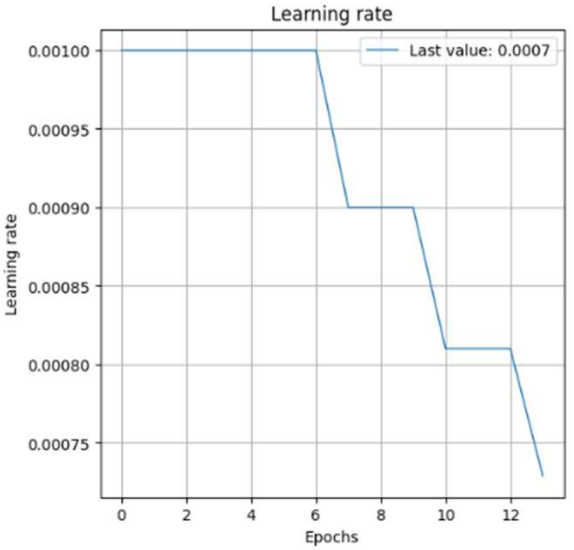
- Tensorflow was the framework to perform the study. Tensorboard was used to visualize training on real time.
- Images where used as data input, ignoring the location and year, since they apport most information to the model.
- Different Convolutional Neural Network pretrained models were tested: mobilenetv2, resnet_v1, ... We conclude that the best option was to use model **efficientnet_v2**.
- Proper callbacks were introduced to lower the learning rate when loss does stop improving between epochs.

Results and Conclusions









- **Maximum F1 scored 0.70** as best performance with the pretrained model **efficientnet_v2**.
- **Additional contextual data was no used** on the image classification.
- **Data augmentation was not applied** and could improve performance.