

Using Predictive Uncertainty for Cleaning Noisy Annotations

Jonas Gütter^a (jonas.guetter@dlr.de), Hannah Ulman, Julia Niebling^a

Introduction

Predictive Uncertainty during model training can be used to assess whether a sample is correctly annotated or not. To see if this is also possible on remote sensing data, we applied the method on a building segmentation task.

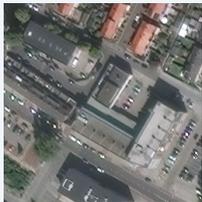
Data

Clean dataset

- Input Images: Optical imagery from the SpaceNet6 dataset
- Labels: Cadastral Data
- Size of the dataset: 1048 images
- Resolution: 256 * 256 pixels
- Location: Port area of the city of Rotterdam

Introducing Noise

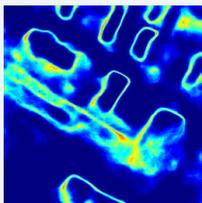
- Buildings were randomly removed from the label masks until a prespecified threshold of removed pixels was reached
- 11 datasets with noise levels between 0 % and 100 % were created
- Experiments were conducted on each of those 11 datasets



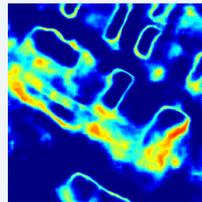
Input Image



Clean Label



Mean Softmax of Predicted class



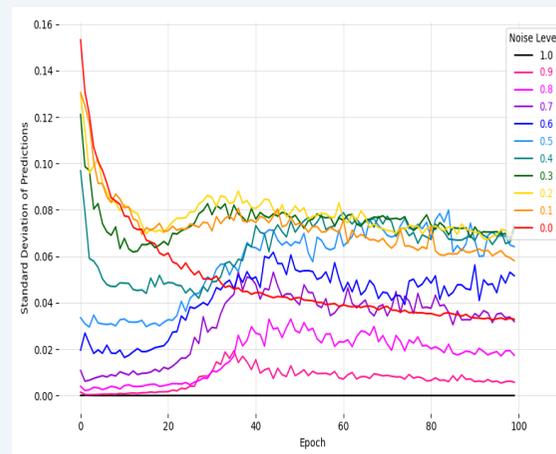
Standard Deviation of Softmax of Predicted Class

Model Training

- Model architecture: DeepLabV3+
- Number of training epochs: 100
- Metrics are evaluated on a clean test set
- We calculate predictive uncertainties in each epoch

Finding a suitable epoch

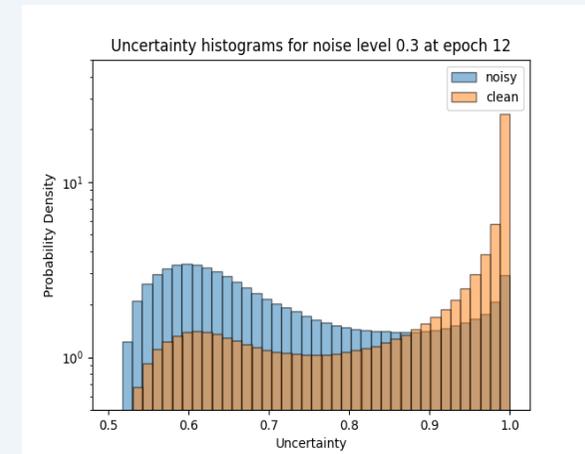
- Predictive Uncertainty during training shows characteristic Behaviour
- It was empirically shown that epoch of first local minimum is well suited for identifying noisy labels



Uncertainty Development during training for different noise levels

Identifying noisy labels

- In the chosen epoch, uncertainty distributions ideally should look different for clean and noisy samples
- Those distributions could potentially be used to assess trustworthiness of a sample
- Below, uncertainty distributions are shown for one of the datasets



Uncertainty distributions of clean and noisy samples

Conclusion

Our results indicate that predictive uncertainty is not able to fully distinguish between clean and noisy samples in our dataset. However a difference in distributions between clean and noisy samples can be observed for some of the noise levels, which could potentially be combined with existing methods for label cleaning to increase their effectiveness.