A Recurrent Neural Network (RNN) based approach for reliably classifying land usage from satellite imagery

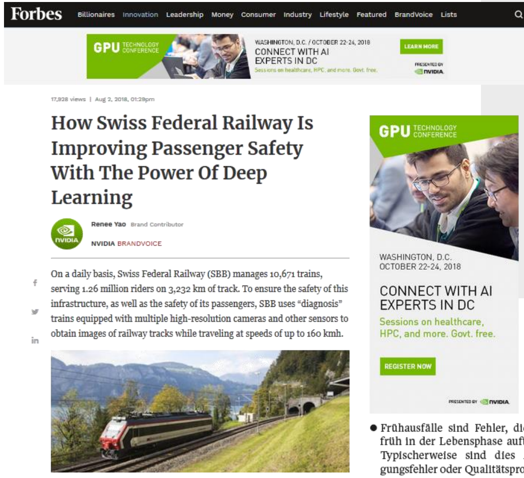
I. Kastanis, P. Purwar, Ph. Schmid, A. Steinecker

Scientific Symposium FAIR Data Sciences for Green Life Sciences, Wageningen, 12.12.2018

Data is never perfect

- Precise method
- Reliability
- Structured approach
- Cost-efficient

- **Integrate human knowledge**

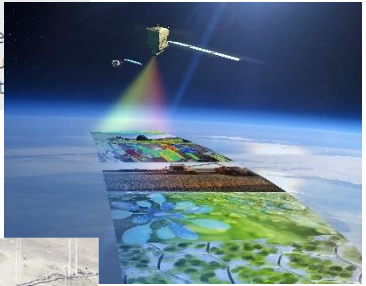


Forbes Billionaires Innovation Leadership Money Consumer Industry Lifestyle Featured BrandVoice Lists

How Swiss Federal Railway Is Improving Passenger Safety With The Power Of Deep Learning

Renee Yao Brand Contributor
NVIDIA BRANDVOICE

On a daily basis, Swiss Federal Railway (SBB) manages 10,671 trains, serving 1.26 million riders on 3,232 km of track. To ensure the safety of this infrastructure, as well as the safety of its passengers, SBB uses "diagnosis" trains equipped with multiple high-resolution cameras and other sensors to obtain images of railway tracks while traveling at speeds of up to 160 kmh.



- Frühausfälle sind Fehler, die sehr früh in der Lebensphase auftreten. Typischerweise sind dies Auslegungsfehler oder Qualitätsprobleme in der Produktion, die sogenannten Kinderkrankheiten.
- Zufallsausfälle sind unsystematische Fehler, die unerwartet über die komplette Lebensdauer einer Maschine auftreten können. Auch Bedienungsfehler, äussere Einwirkungen und Wartungsfehler zählen zu dieser Kategorie. Diese Ausfälle sind am stärksten gefürchtet – sie sind sozusagen die Erdbeben einer Maschine.
- Spätausfälle sind Fehler, die gegen Ende des Lebenszyklus auftreten. Keine Maschine kann unendlich lange funktionieren. Ursache sind oft Alterung (Korrosion), Abnutzung (Verschleiss) und Ermüdung (Bruch).

Bei Predictive Maintenance liegt der Fokus primär darauf, Anzeichen für Zufallsausfälle frühzeitig zu erkennen

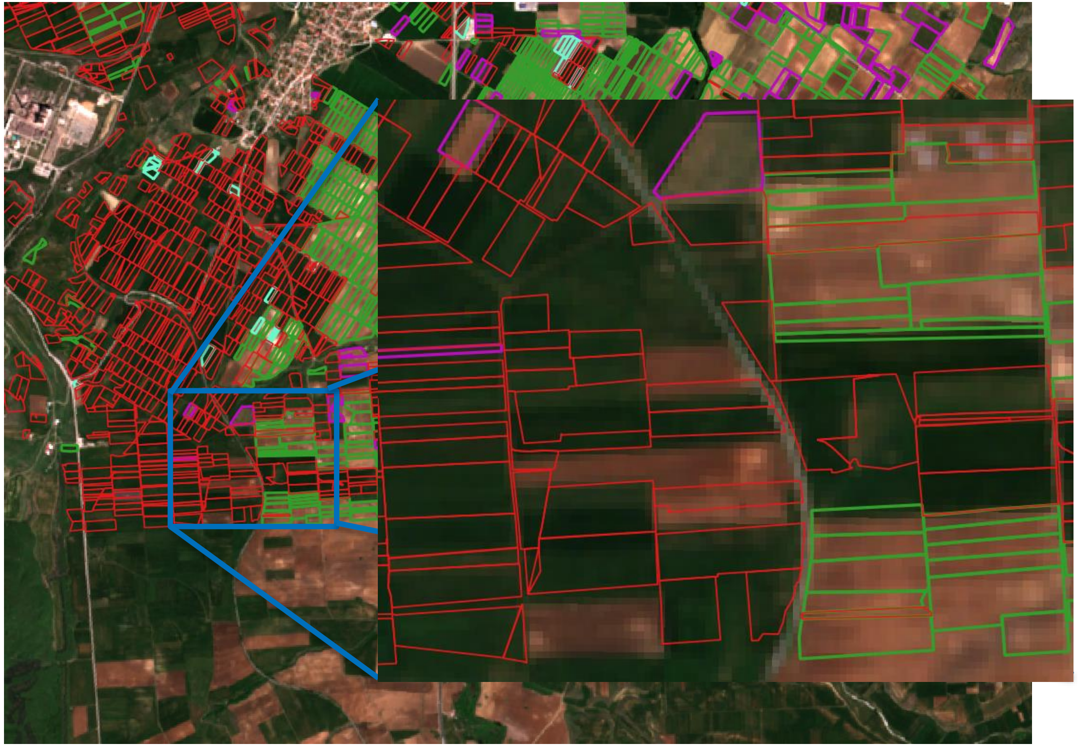
Eine getriebelose Erzmühle von ABB: Anhand einer solchen Anlage testet das CSEM im Rahmen einer Forschungszusammenarbeit die Predictive-Maintenance-Lösung und entwickelt sie weiter. Die Algorithmen lassen sich auch auf kleine Systeme übertragen.

DataBio project

https://www.databio.eu

- Budget: 16.2 m Euro
- Lighthouse project
- 48 partners
- 100 organization involved in demonstrations
  - Agriculture
  - Forestry
  - Fishery

Big data



- *Data from Neuropublic SA*
- *Images available for **16 days** for year 2016*
- *Labelled at parcel level*
- *27k parcels for 1 tile*

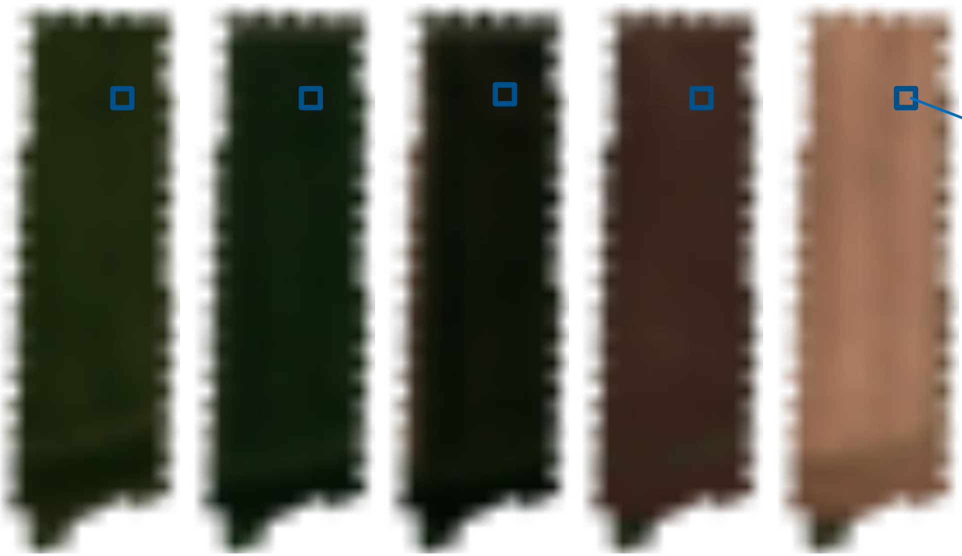
Wheat

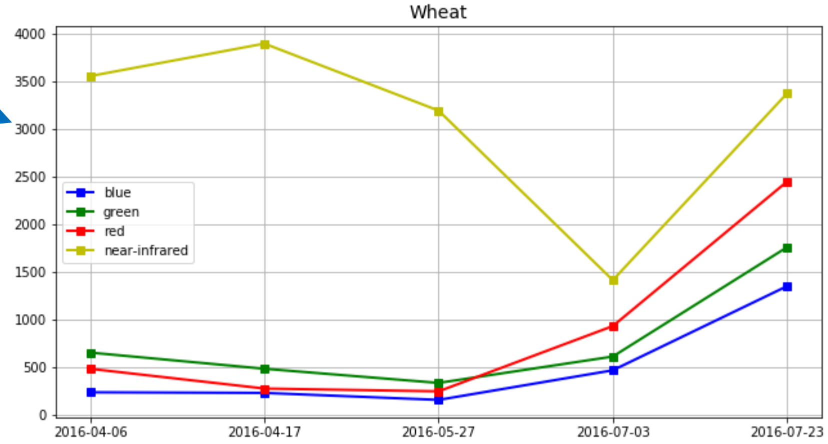
Maize

Legumes

Fruits

Satellite Image time series

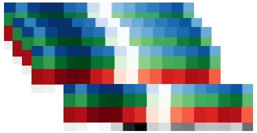




Time series of pixel values

Approach

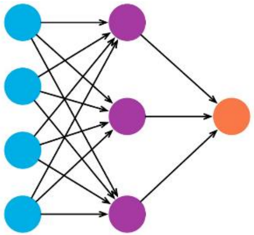
Input: *Time-series of pixel spectral values*



Preprocessing

Learning

Classifier



Add new samples

Experts Validation

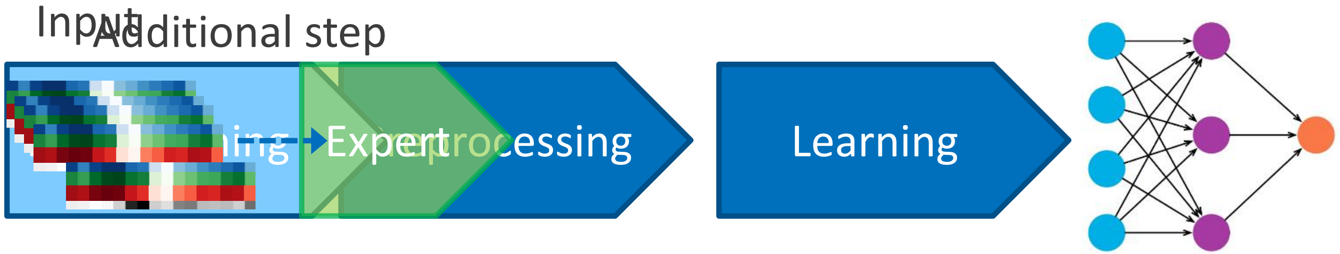
Classification

Inference

New Data

Quality of training data

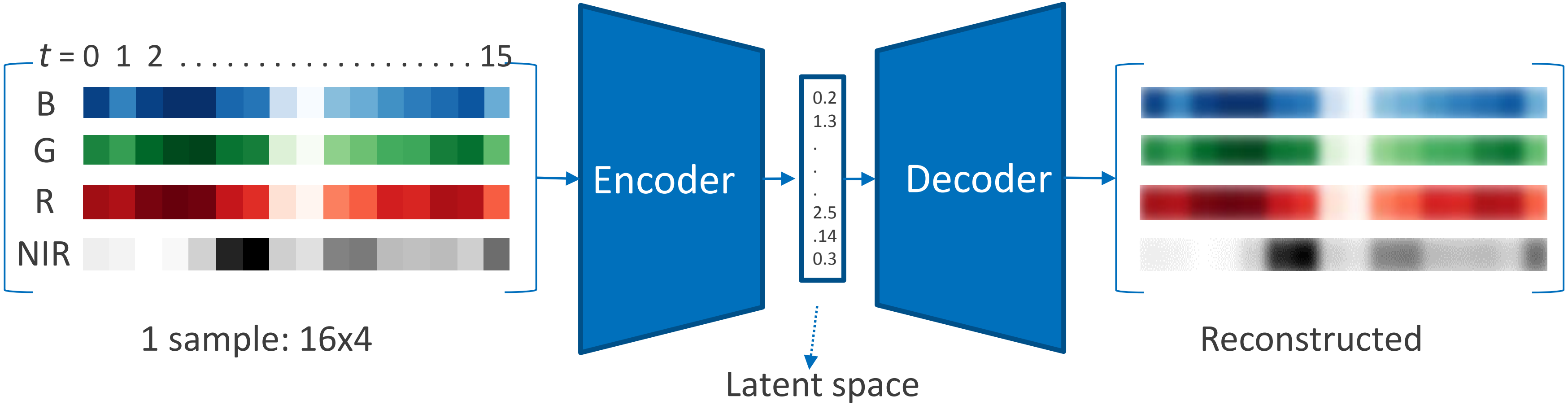
What if the parcels are incorrectly labelled?



- Identify outliers and use clean parcels to train classifier
- Traditional approach: statistically remove outliers, needs understanding of data

Data cleaning

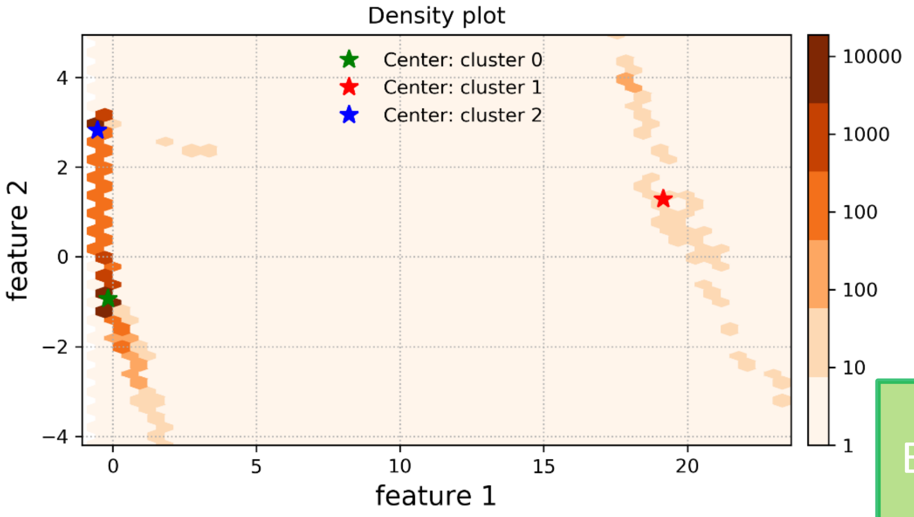
- AutoEncoder learns optimal filters for reconstructing the input signal

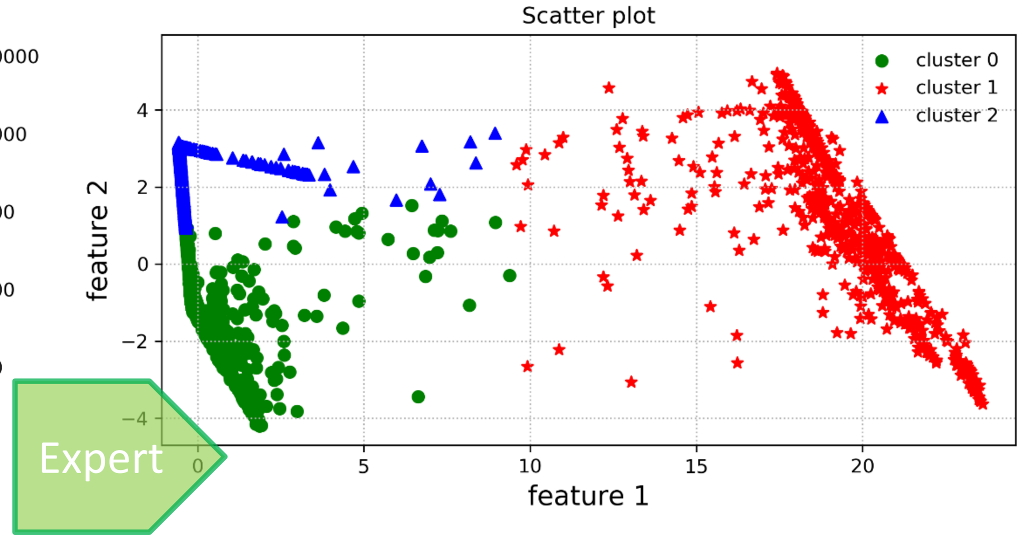


- RNN-AE encodes spectral temporal changes for a particular crop

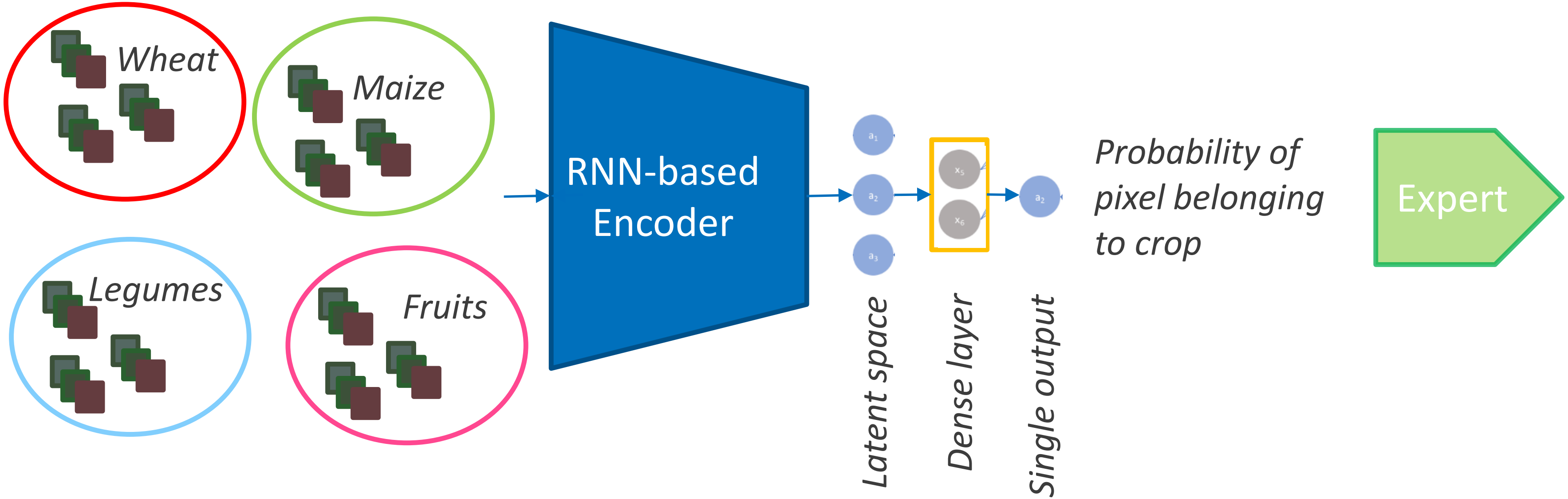
Data cleaning

- Wheat: latent space features for pixels from 1312 parcels
  - k-mean clustering (k=2) iterated 2 times, choose clean parcels (424)





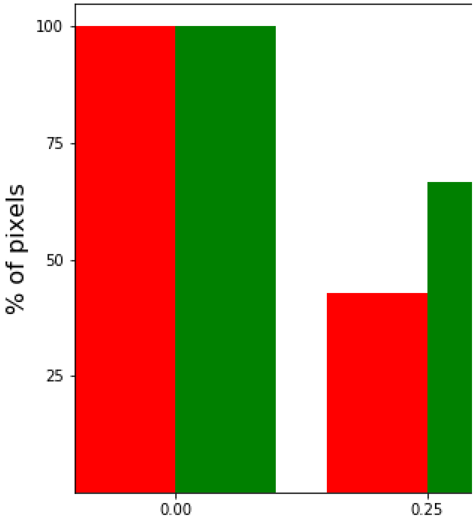
Binary Classifier

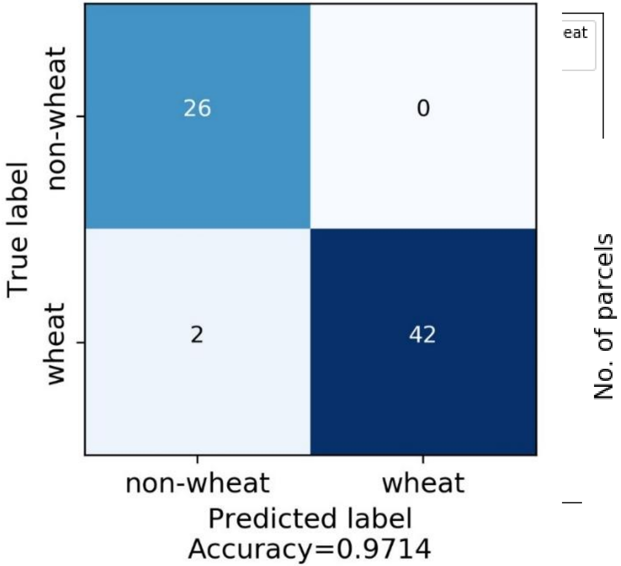


- Build a classifier for each crop using its pre-trained AE
- Non-crop classes selected from other crops (balanced classes)

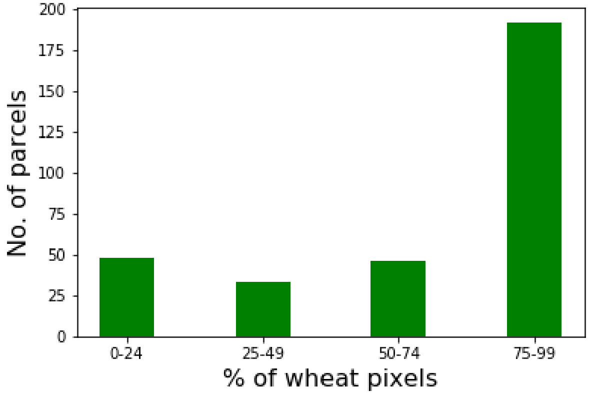
Parcel results

- Classifier generated probability for each pixel in a parcel
- Parcel classified using *“Majority Vote”*





Mixed pixel parcels:
802/888 as wheat



Detect non-wheat

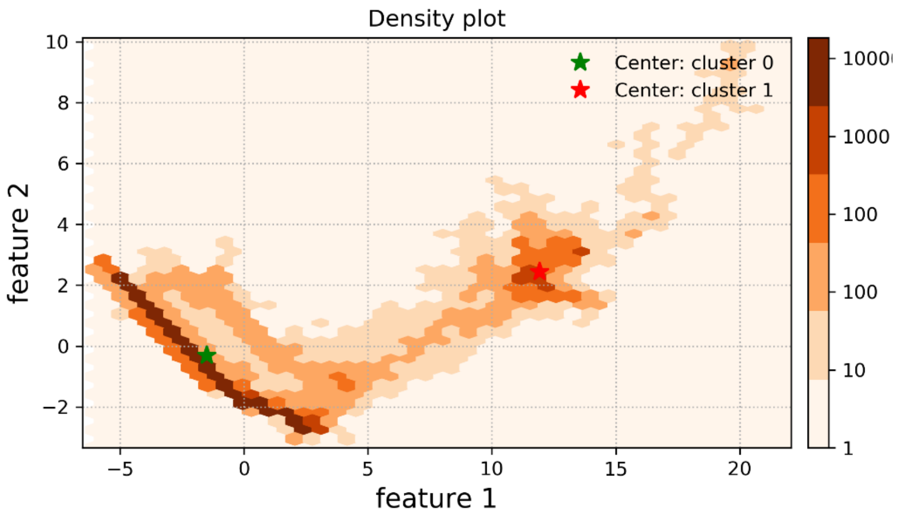
- Parcels from legumes and stone fruits: 516 parcels

Non-wheat	Wheat
509	**7**

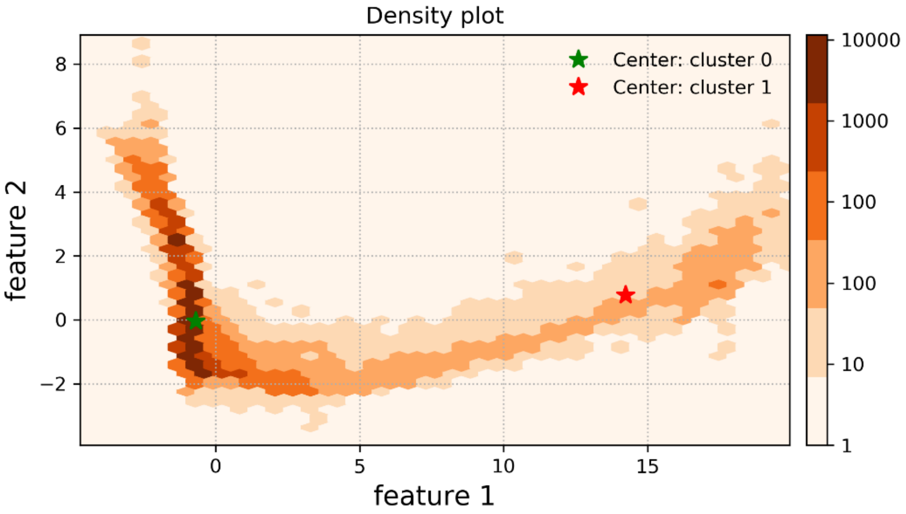
Predicted labels
Accuracy = 0.98

Results on other varieties

- Maize (971 clean from 2003 parcels)



- Legumes (1102 clean from 1437 parcels)



- Mixed pixel parcels: 818/1032 as maize
- Mixed pixel parcels: 253/335 as legumes

Results on untrained sub-varieties

- Binary classifier tried out-of-box on sub-varieties with more than 100 parcels

Wheat (7931*)

	Variety	#parcels	Accuracy(%)
1	7886	882	95.9
2	7927	239	92.8
3	99176	244	95.0
4	7937	380	91.8
5	9194	146	87.6
6	7876	335	86.2
7	7956	451	91.1
8	99171	107	95.3
9	7895	275	81.8
10	7964	107	94.3

Maize (8019*)

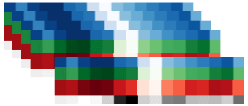
	Variety	#parcels	Accuracy(%)
1	8013	1846	81.9

Legumes (1043*)

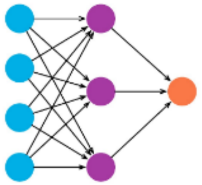
	Variety	#parcels	Accuracy(%)
1	1144	163	51.5
2	1045	160	25.0
3	1046	126	22.2

*** Crop Variety used for training AE and classifier**

Conclusion



Data cleaning → Expert → Preprocessing → Learning ← Expert



- Systematic approach
- Efficient and precise method
- Extendable to other sensors
- **Digitalization of knowledge**

- Future direction: Prediction