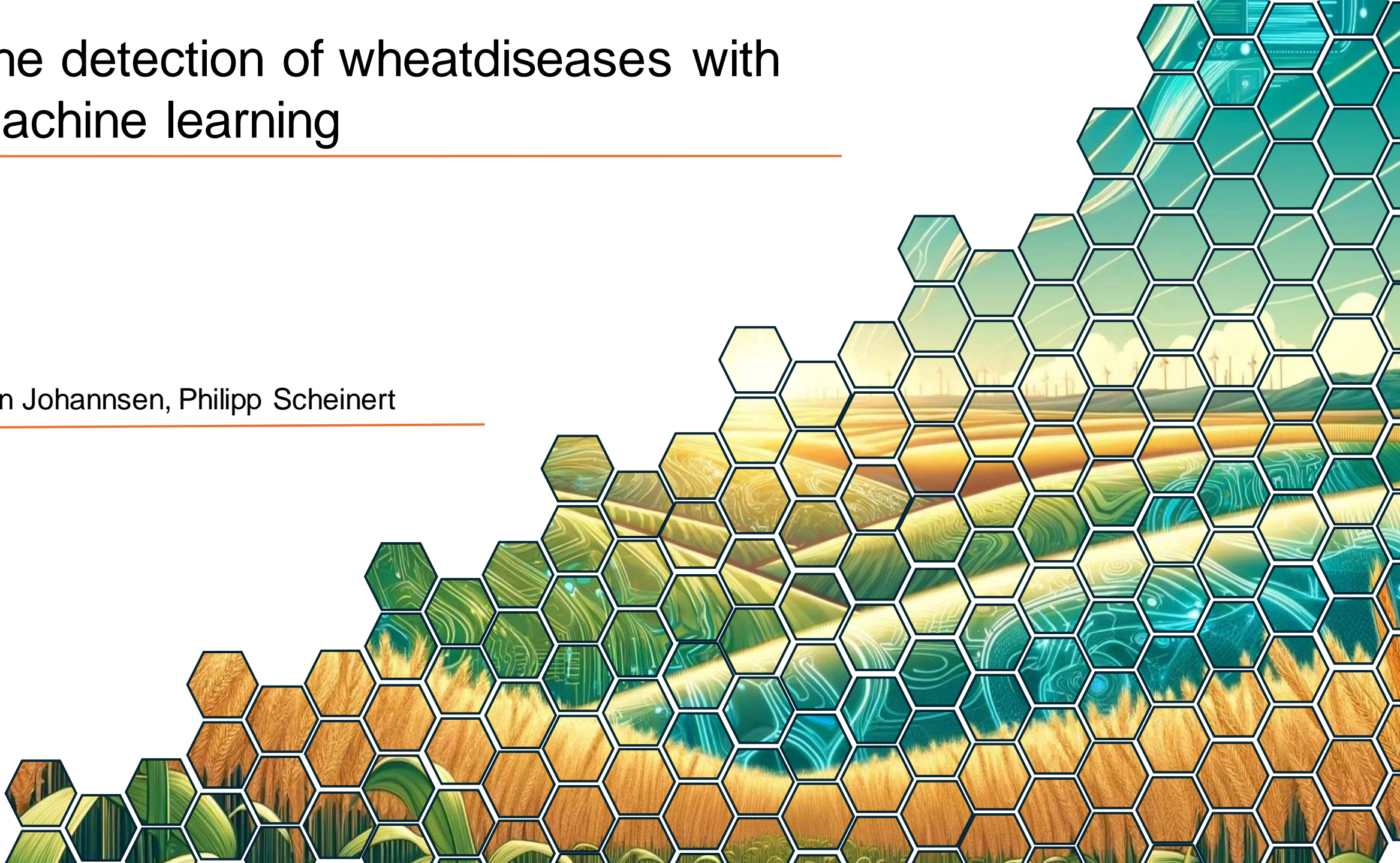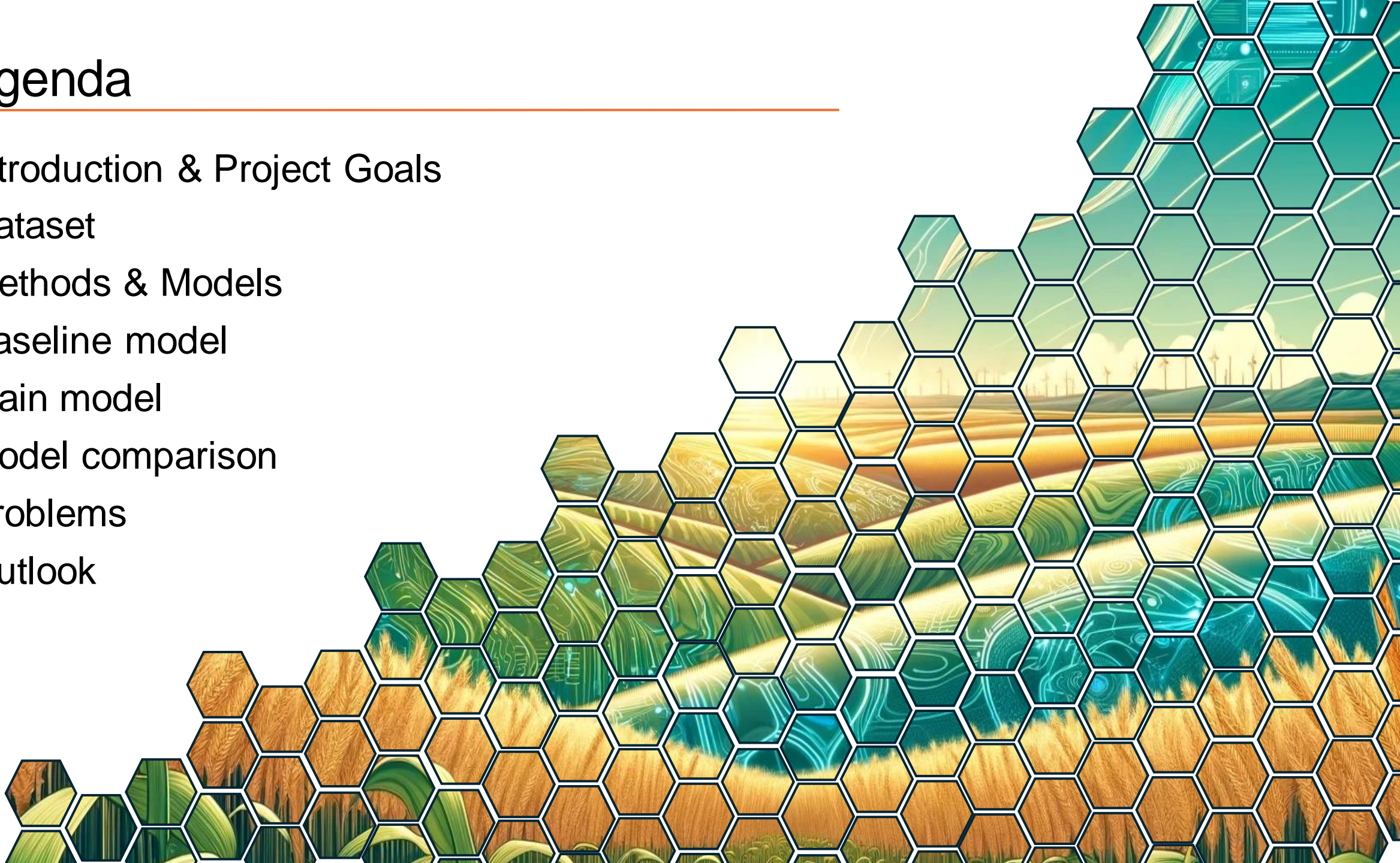# The detection of wheatdiseases with machine learning

Finn Johannsen, Philipp Scheinert

Agenda

- Introduction & Project Goals

- Dataset

- Methods & Models

- Baseline model

- Main model

- Model comparison

- Problems

- Outlook

Introduction & Project Goals

- We wanted to do something that is related to real life

- Therefore, we ended up searching for projects in the STEM –field

- After a long discussion, we decided to develop a model for recognizing diseases in wheat plants
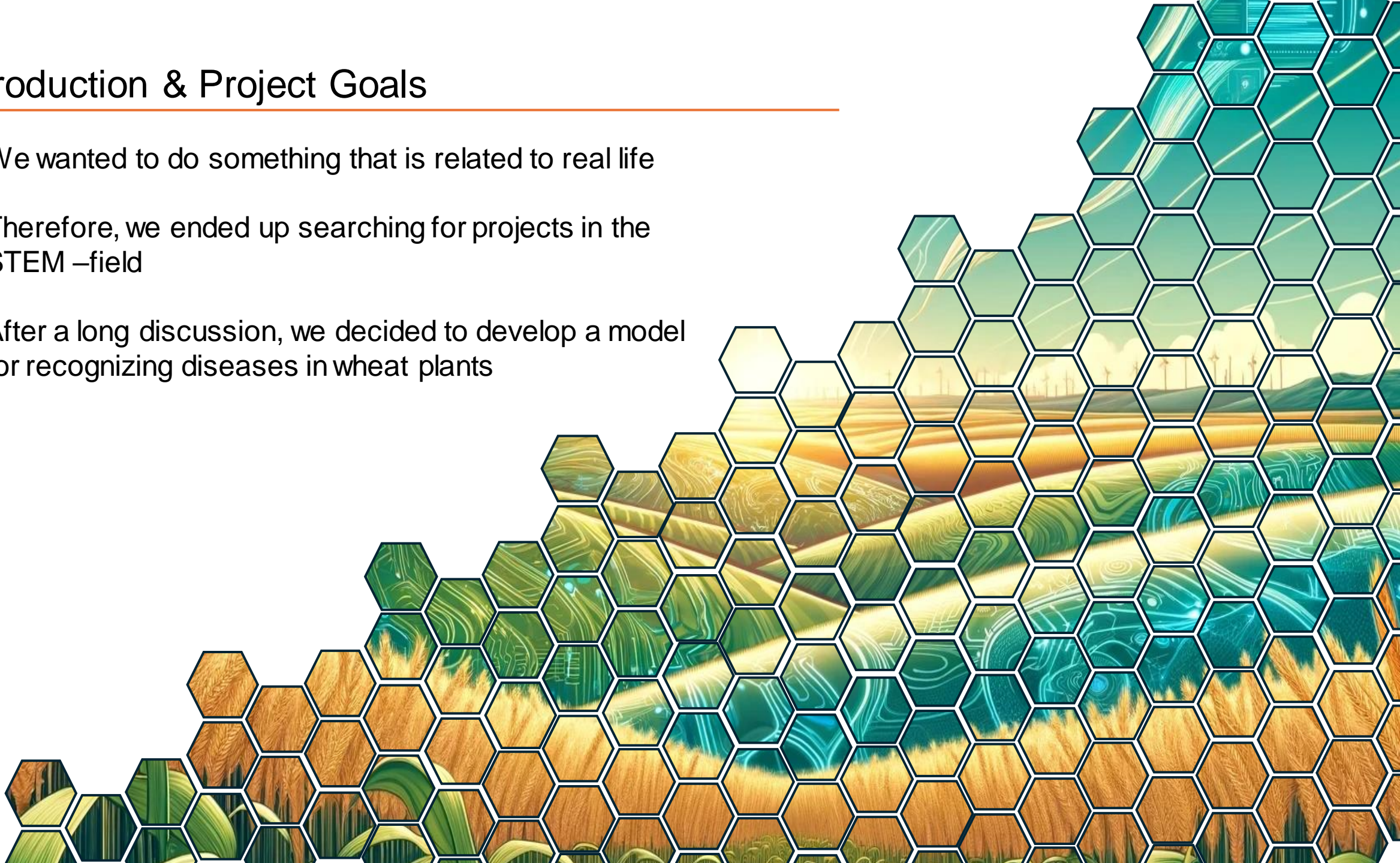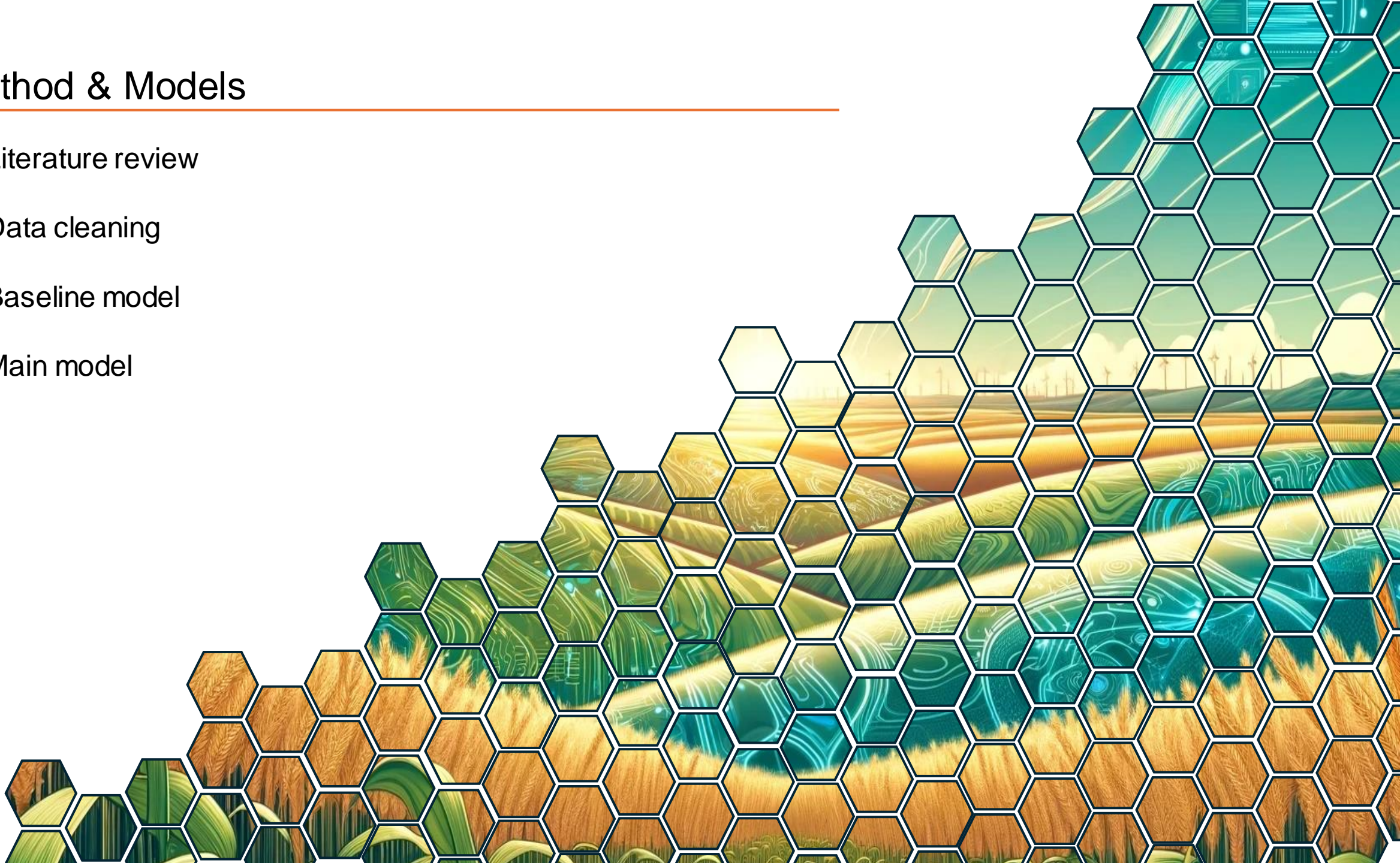
Dataset

- We used a data set from Kaggle as a basis

- Over 14,000 pictures

- 16 different wheat diseases

- Problem: The dataset is partly of low quality

- In the end we just selected 10 different diseases with 2,800 pictures
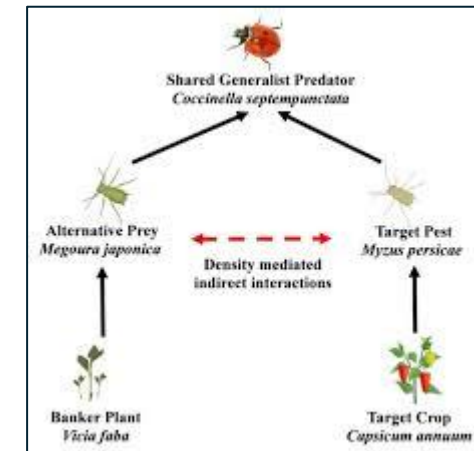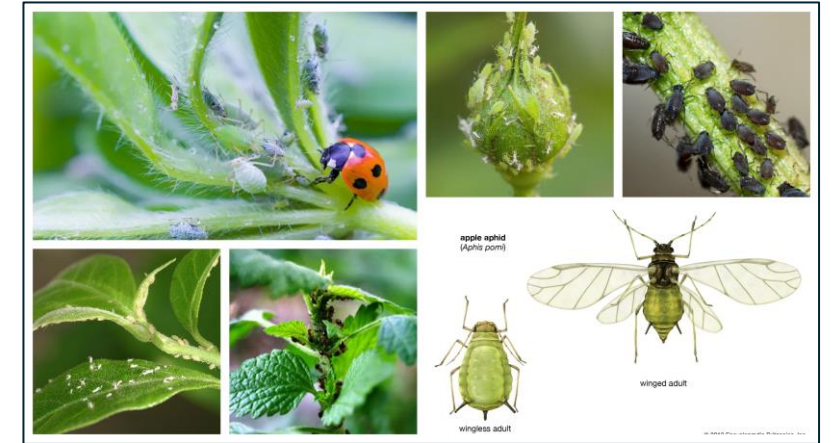
Method & Models

- Literature review

- Data cleaning

- Baseline model

- Main model

Baseline model

3x3 cernel size of the convolution

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 254, 254, 32)	896
max_pooling2d (MaxPooling2D)	(None, 127, 127, 32)	0
conv2d_1 (Conv2D)	(None, 125, 125, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 62, 62, 64)	0
conv2d_2 (Conv2D)	(None, 60, 60, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(None, 30, 30, 128)	0
flatten (Flatten)	(None, 115200)	0
dense (Dense)	(None, 64)	7,372,864
dense_1 (Dense)	(None, 10)	650

Accuracy of the baseline model on the test data for each class
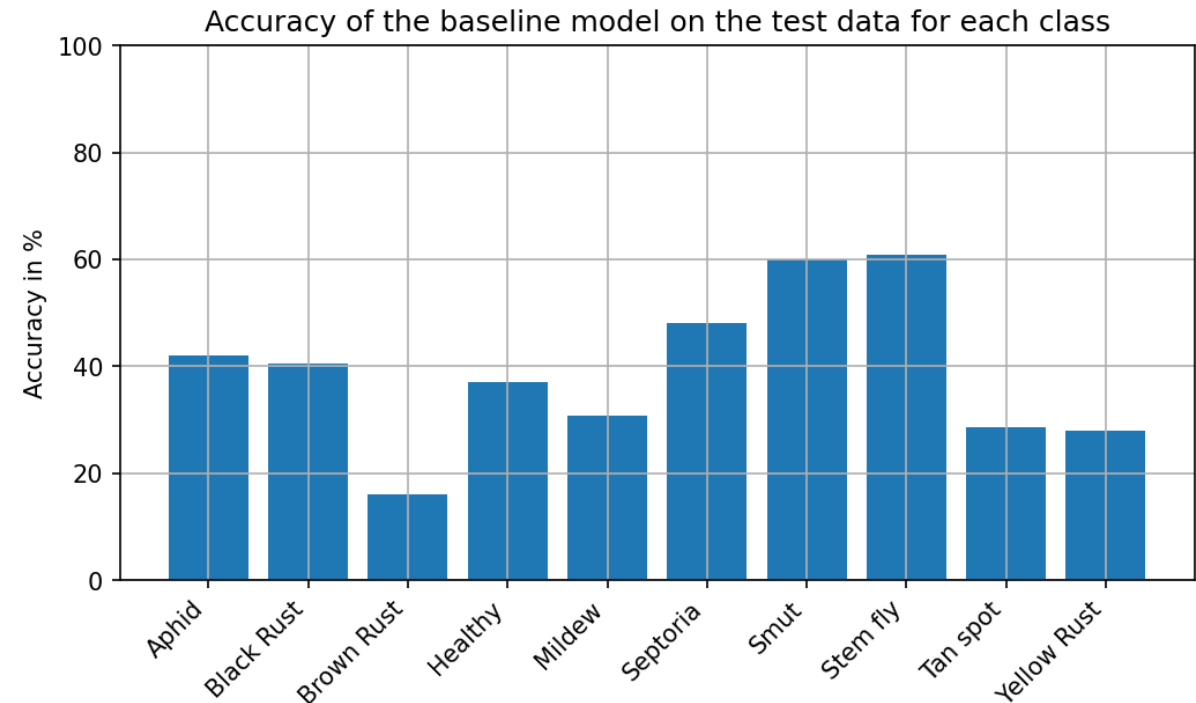
Main model

Layer (type)	Output Shape	Param #
resnet50 (Functional)	(None, 8, 8, 2048)	23,587,712
global_average_pooling2d (GlobalAveragePooling2D)	(None, 2048)	0
batch_normalization (BatchNormalization)	(None, 2048)	8,192
dense (Dense)	(None, 256)	524,544
batch_normalization_1 (BatchNormalization)	(None, 256)	1,024
dense_1 (Dense)	(None, 128)	32,896
dropout (Dropout)	(None, 128)	0
batch_normalization_2 (BatchNormalization)	(None, 128)	512
dense_2 (Dense)	(None, 64)	8,256
dropout_1 (Dropout)	(None, 64)	0
batch_normalization_3 (BatchNormalization)	(None, 64)	256
dense_3 (Dense)	(None, 10)	650

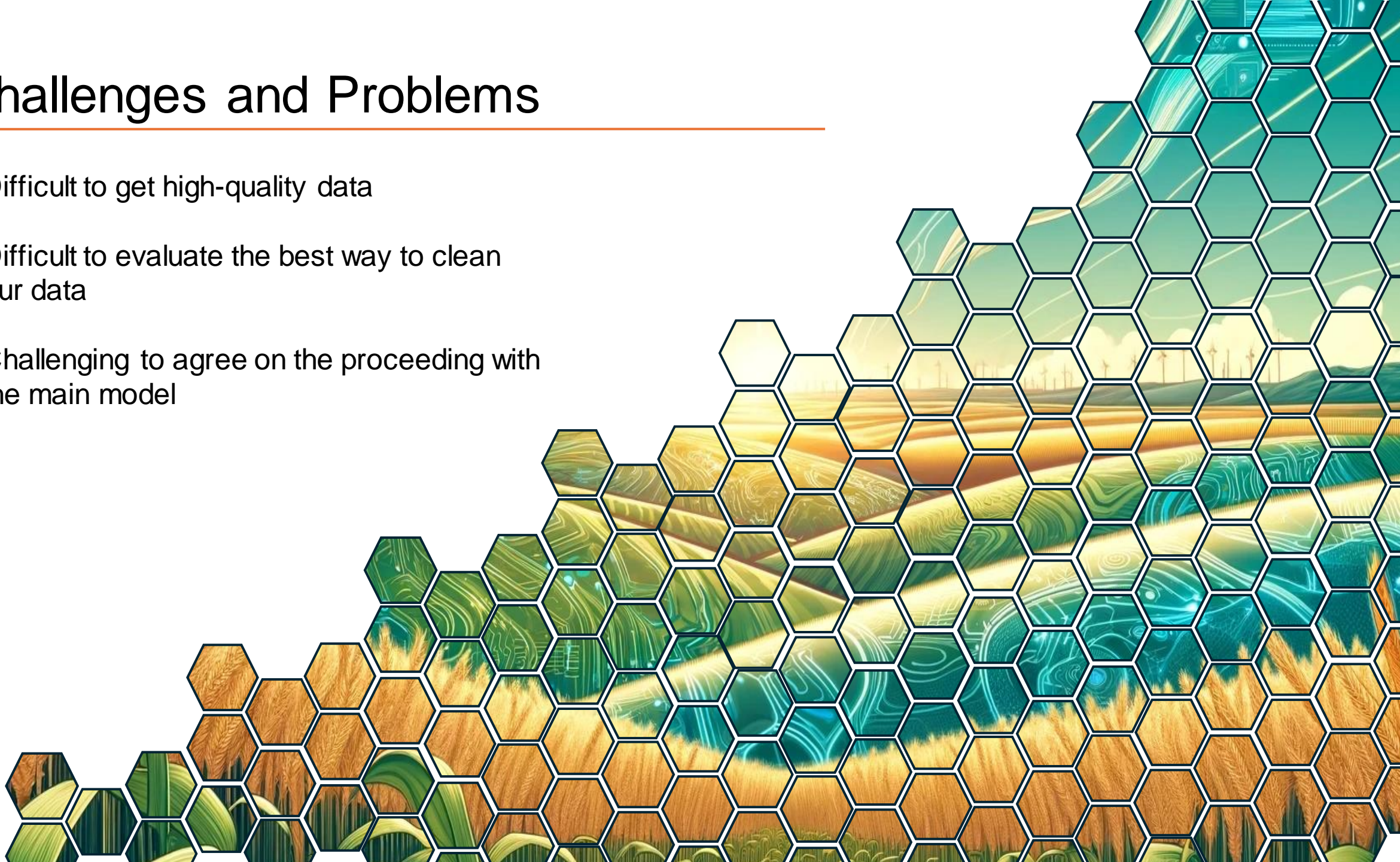Accuracy of the ResNet50 based model on the test data for each class
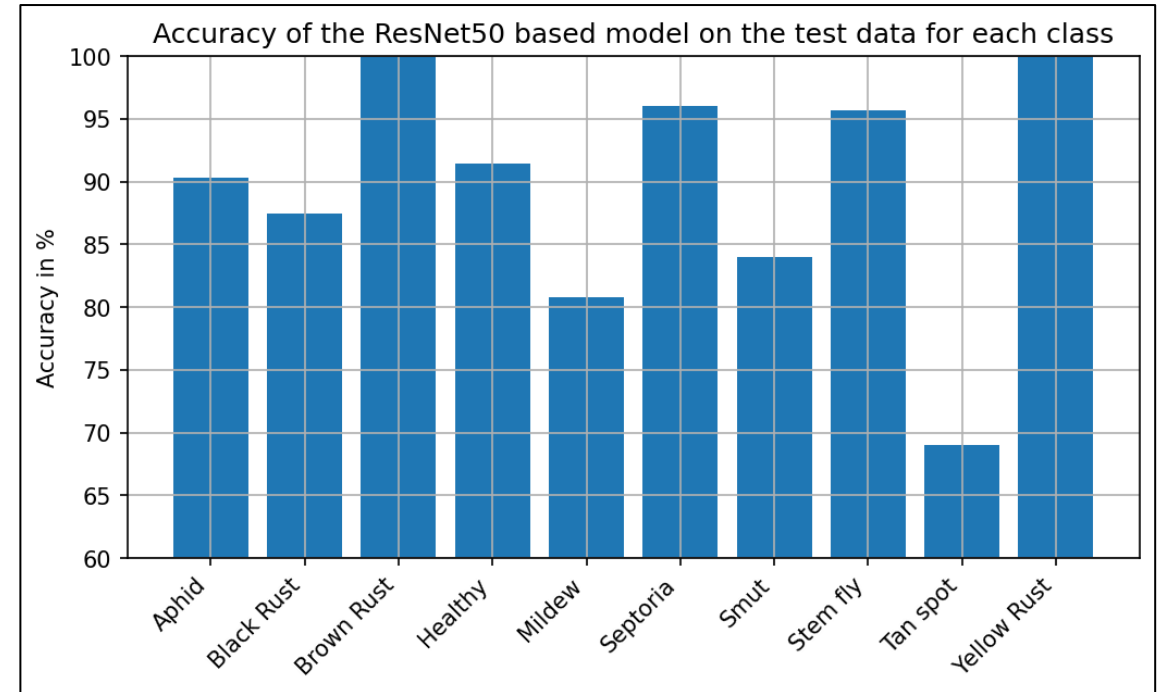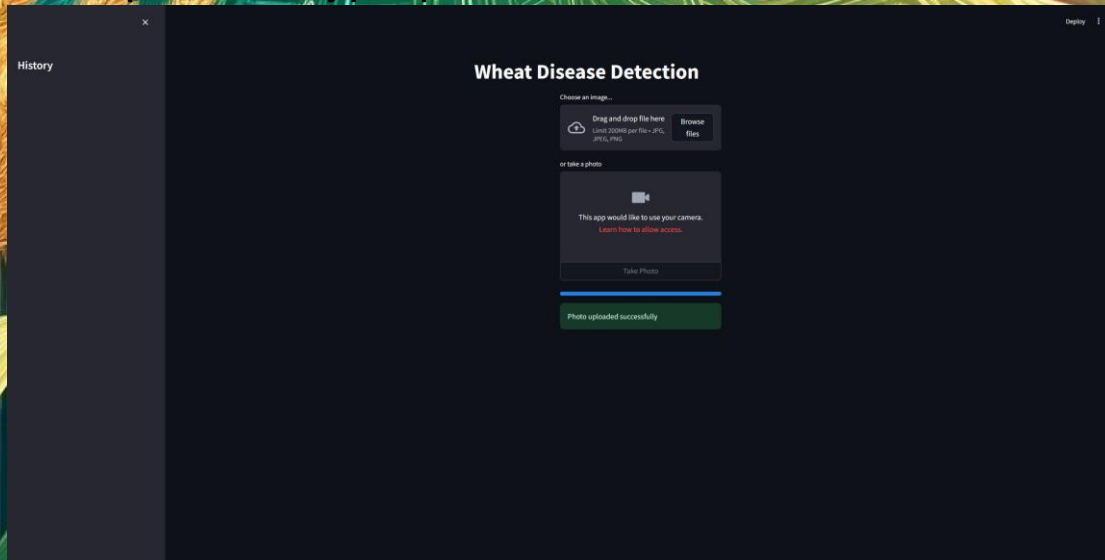
Model Comparison
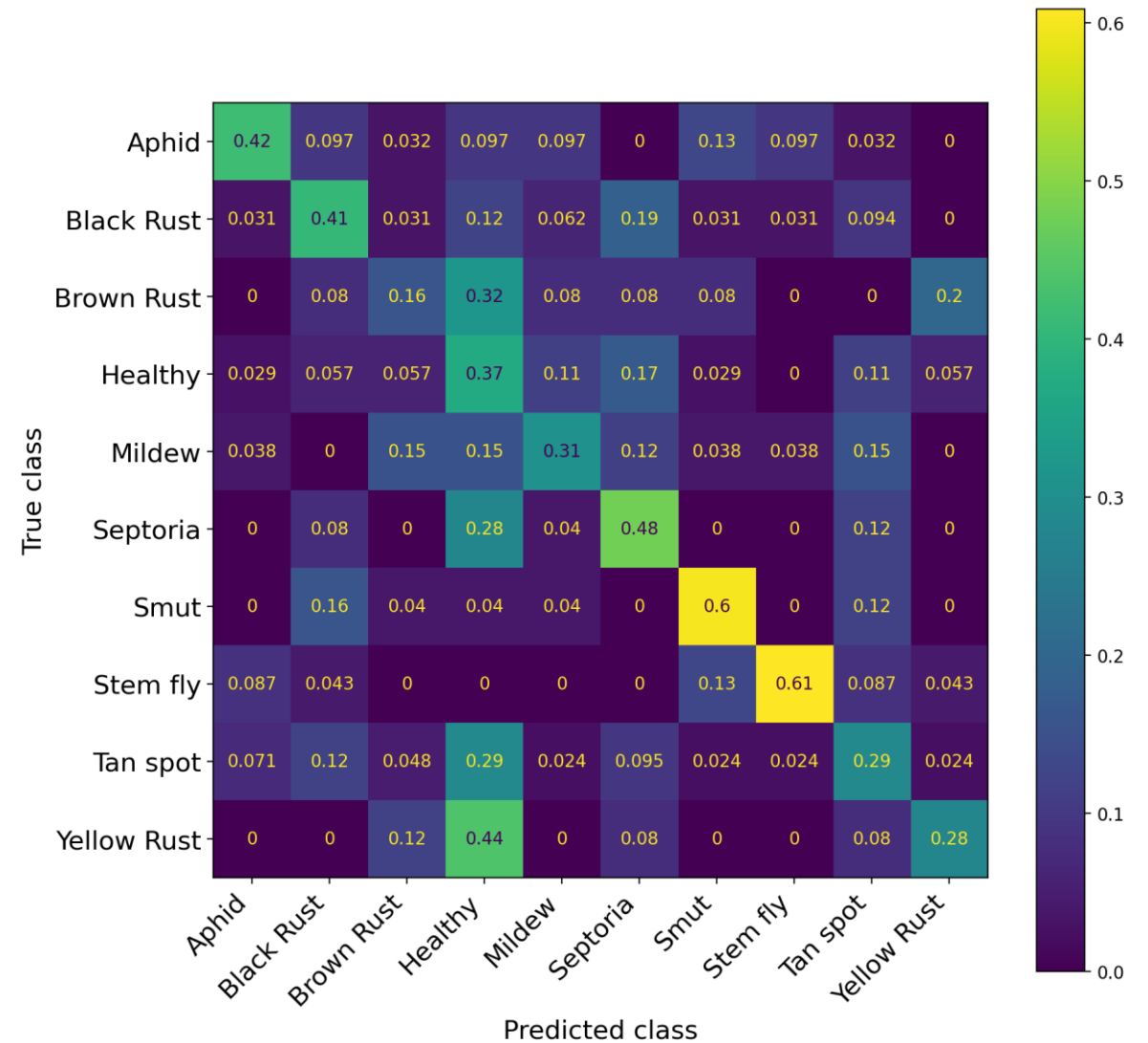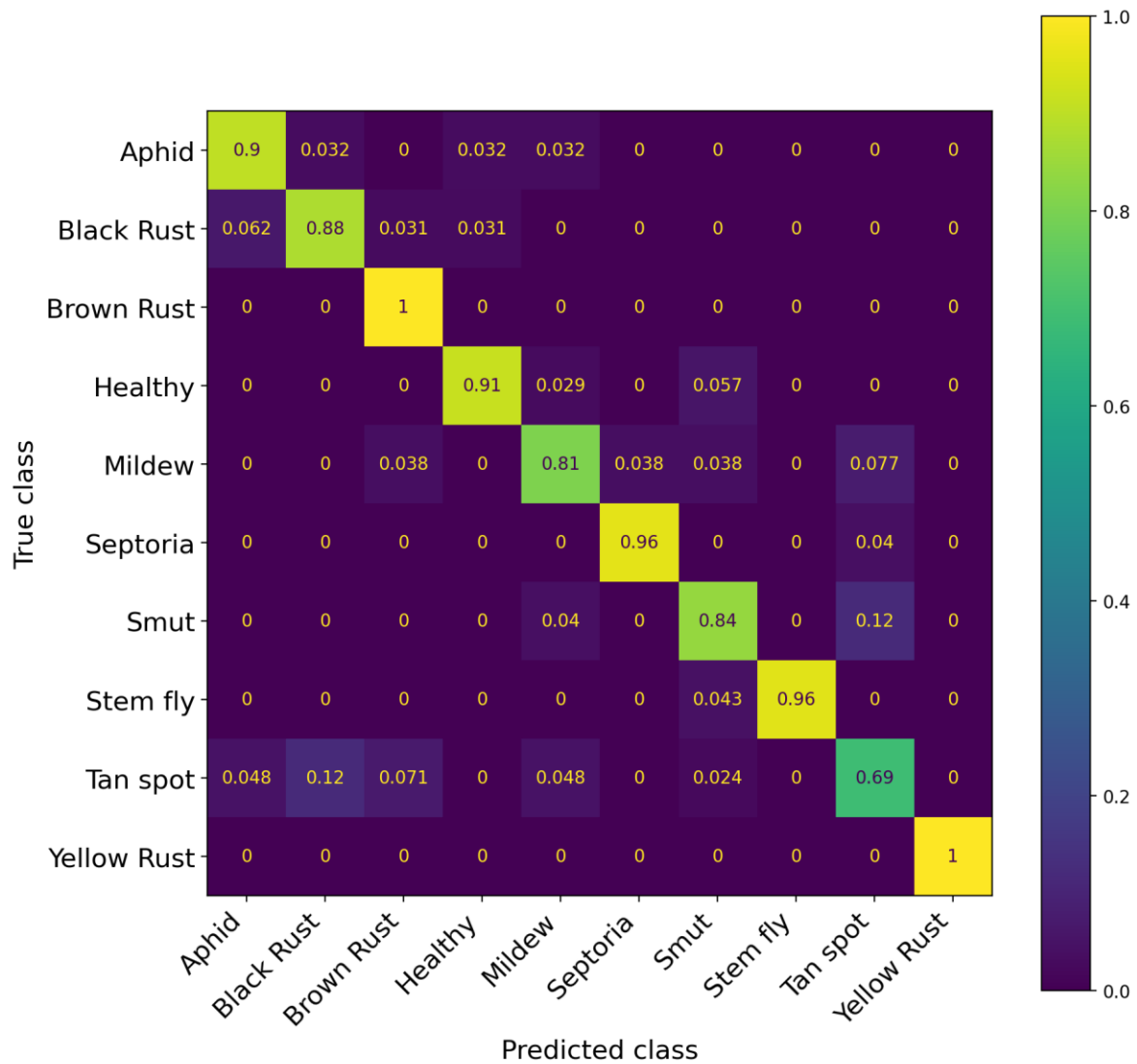
Challenges and Problems

- Difficult to get high-quality data

- Difficult to evaluate the best way to clean our data

- Challenging to agree on the proceeding with the main model

Outlook

- Real-time detection of wheat diseases through video surveillance

- Linking the recorded diseases with GIS data to determine spatial patterns in fields

- Maybe transferlearning with a different dataset

- Optimize Hyperparameter

Thank You!