



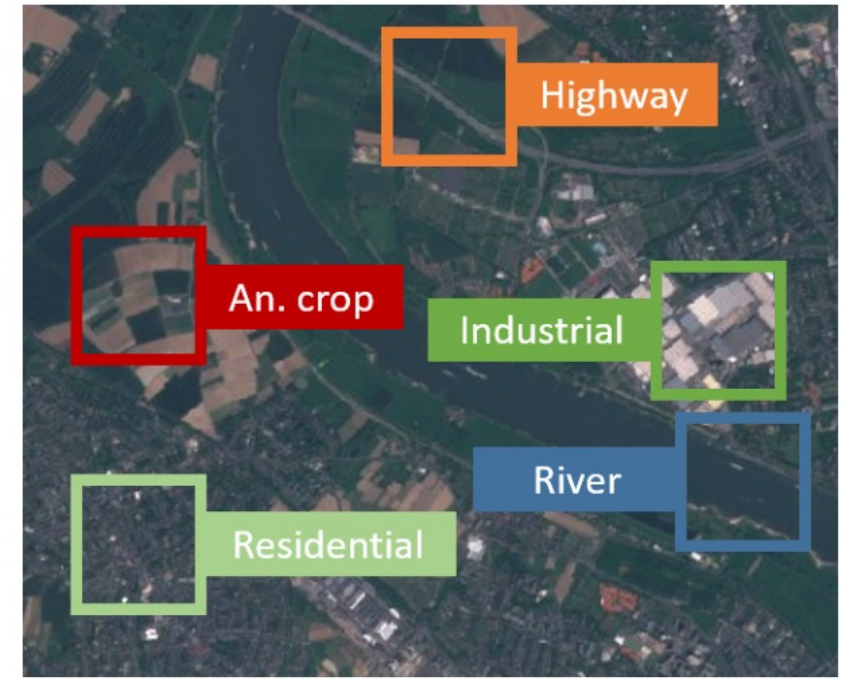
Multiclass Image Classification of Satellite Images

Presentation by Lucas Hoefft

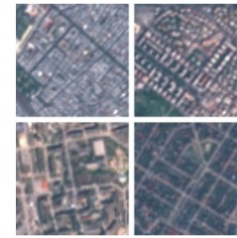
Problem Setting

- 27k labeled images of size 64x64 with Channels RGB
- 10 labels

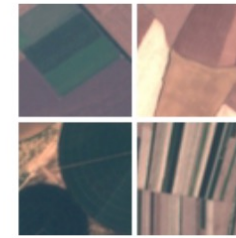
→ Multiclass Image Classification



(a) Industrial Buildings



(b) Residential Buildings



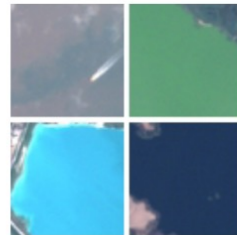
(c) Annual Crop



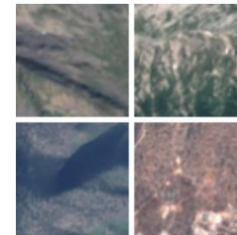
(d) Permanent Crop



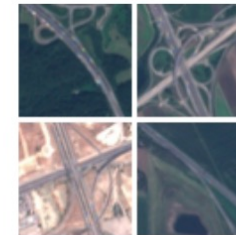
(e) River



(f) Sea & Lake



(g) Herbaceous Vegetation



(h) Highway



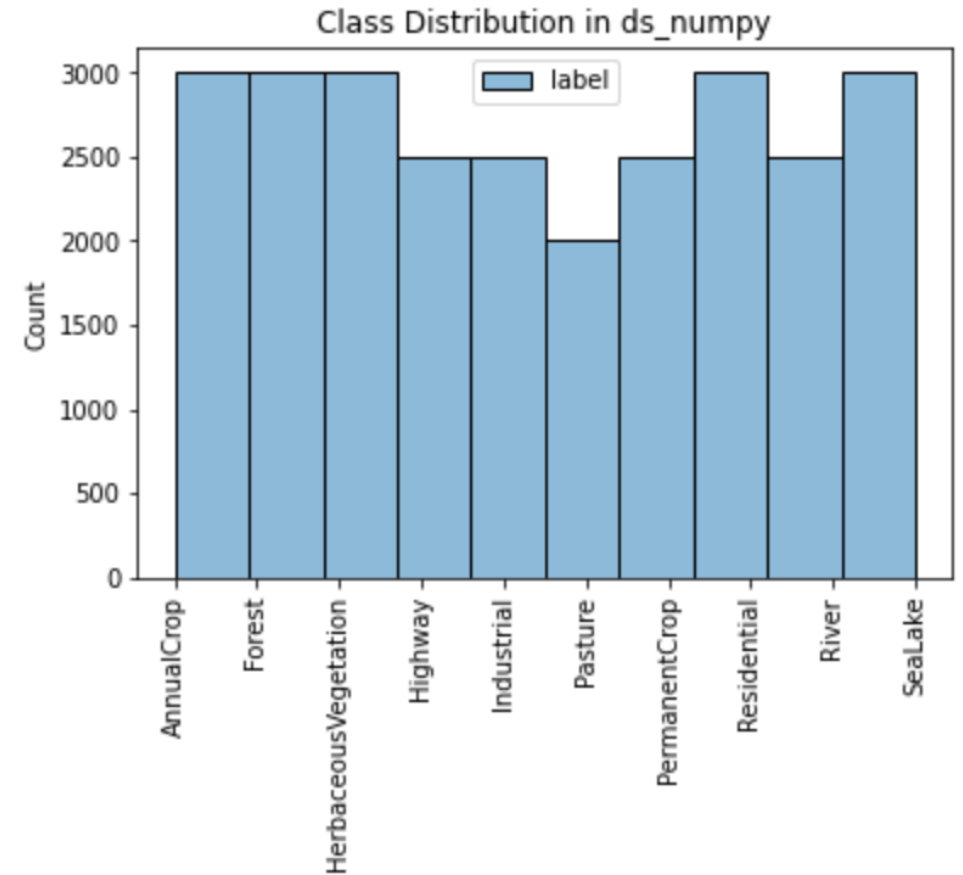
(i) Pasture



(j) Forest

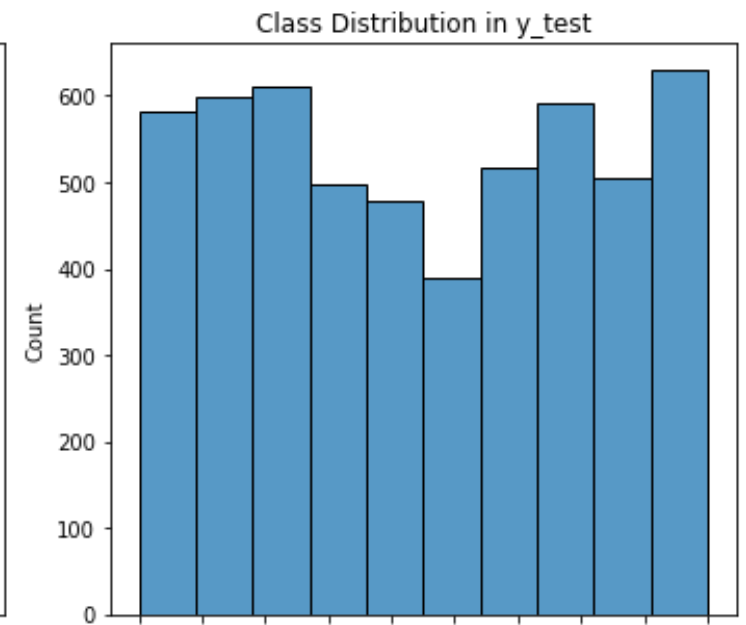
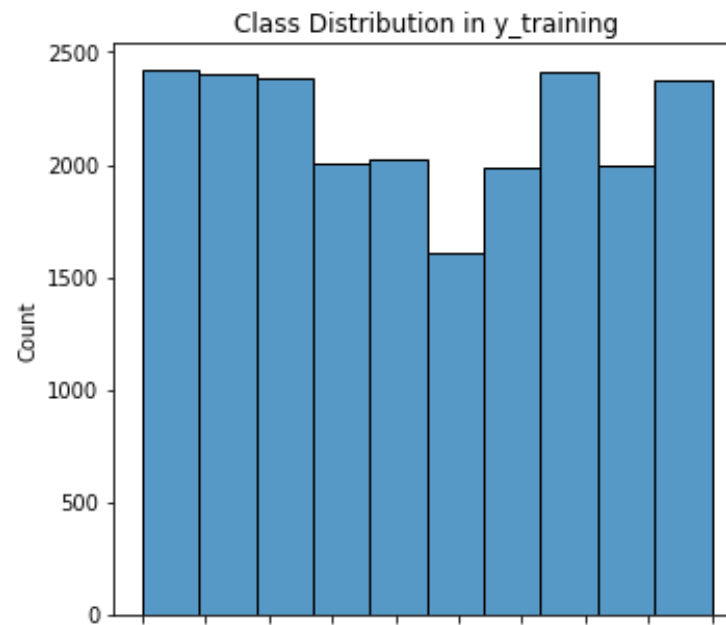
Explorative Data Analysis

- Balanced Classes
- Edges more present in Industrial, Residential
- Similarities between Forest and SeaLake



Preprocessing

- Train/Test-Split (80/20)
- Augmentation (Rotation, Flip)
- Normalization (from $[0,255]$ to $[0,1]$)
- One-Hot Encode Label



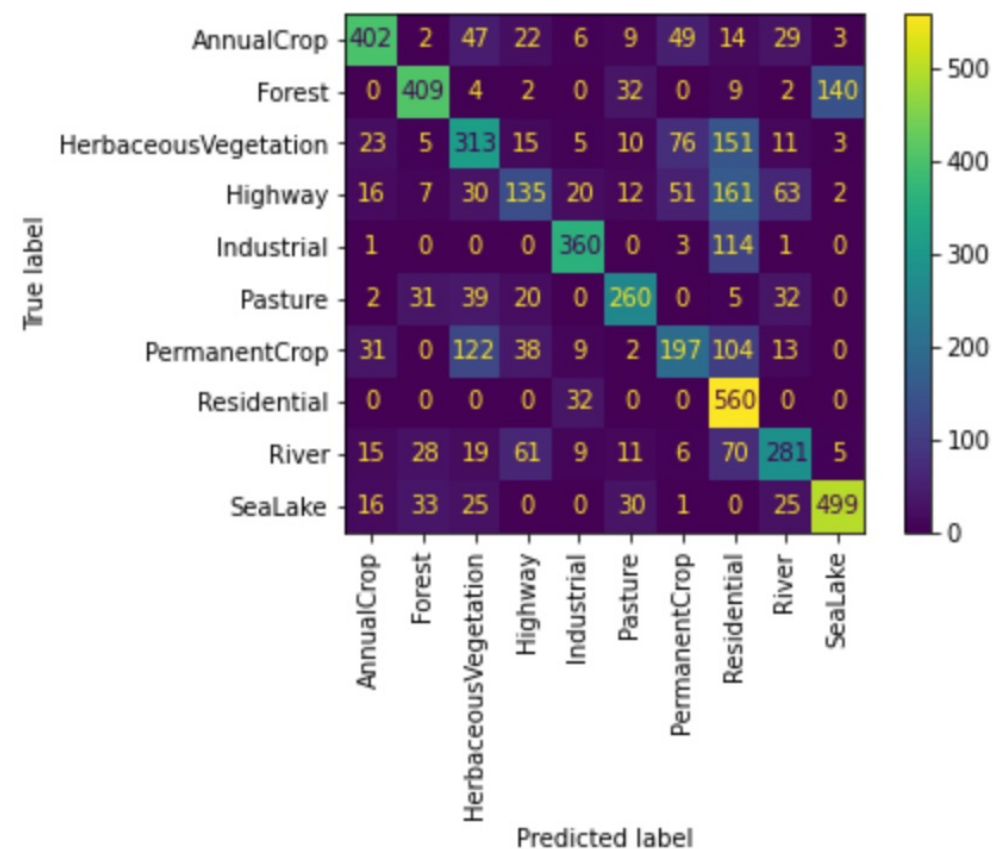
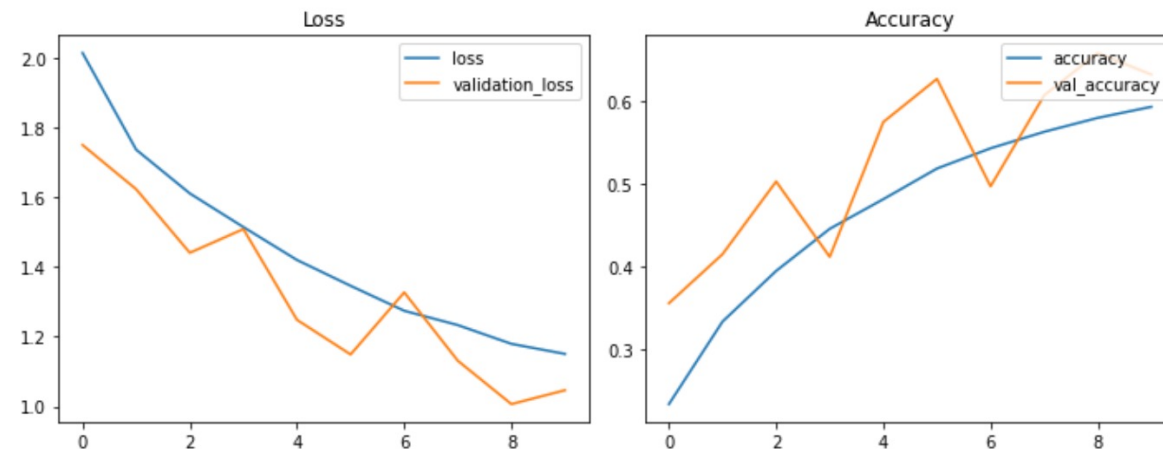
Method Application

- Baseline CNN
- Advanced CNN
- ResNet50 (Transfer Learning)
- Hyperparameter-Tuned Baseline CNN

Baseline CNN

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 62, 62, 32)	896
max_pooling2d (MaxPooling2D)	(None, 31, 31, 32)	0
dropout (Dropout)	(None, 31, 31, 32)	0
flatten (Flatten)	(None, 30752)	0
dense (Dense)	(None, 128)	3936384
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 10)	1290
Total params: 3,938,570		
Trainable params: 3,938,570		
Non-trainable params: 0		



→ Validation Accuracy: 0.6326

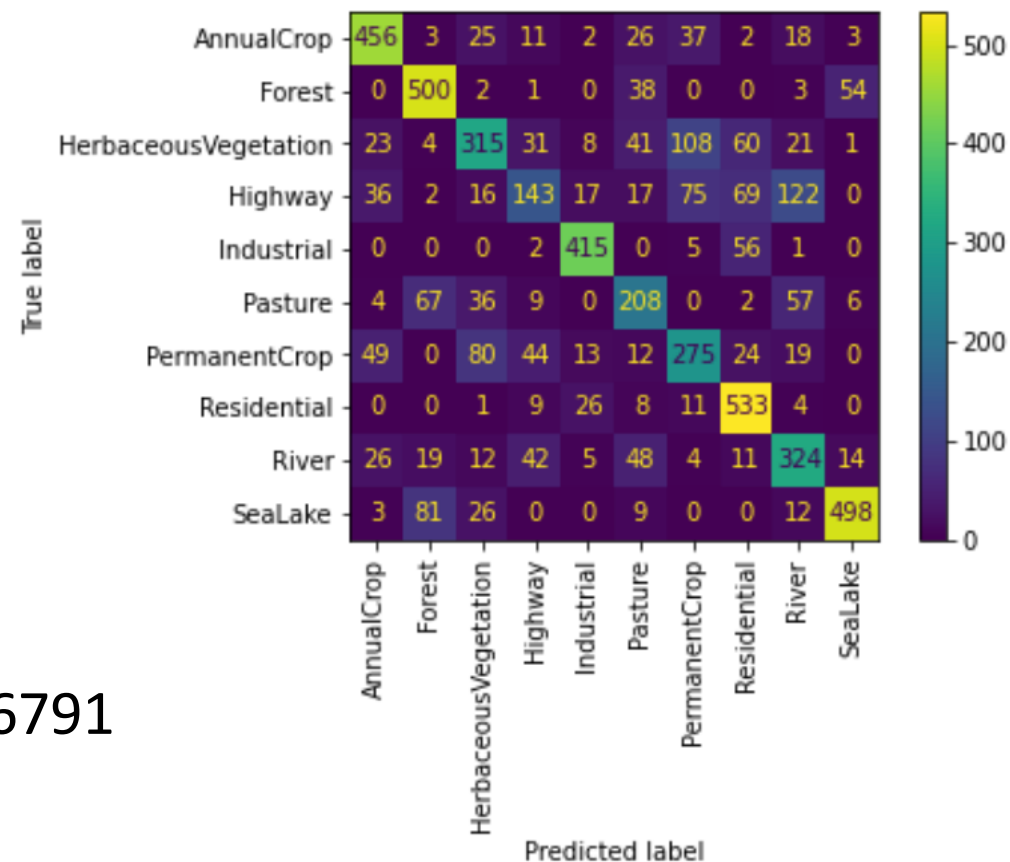
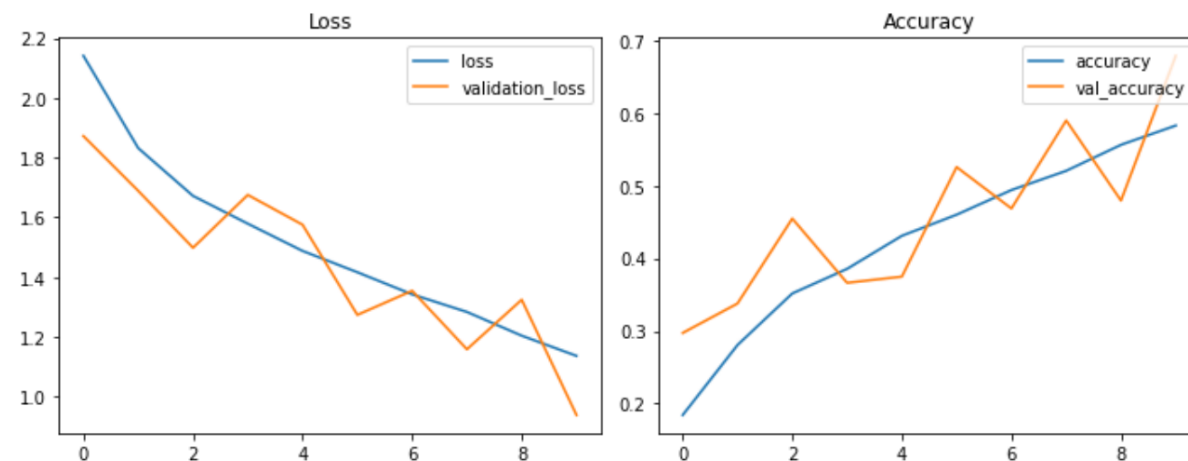
Advanced CNN

Model: "sequential_4"

Layer (type)	Output Shape	Param #
conv2d_8 (Conv2D)	(None, 62, 62, 32)	896
max_pooling2d_8 (MaxPooling 2D)	(None, 31, 31, 32)	0
conv2d_9 (Conv2D)	(None, 29, 29, 64)	18496
max_pooling2d_9 (MaxPooling 2D)	(None, 14, 14, 64)	0
conv2d_10 (Conv2D)	(None, 12, 12, 128)	73856
max_pooling2d_10 (MaxPooling 2D)	(None, 6, 6, 128)	0
dropout_9 (Dropout)	(None, 6, 6, 128)	0
flatten_5 (Flatten)	(None, 4608)	0
dense_10 (Dense)	(None, 256)	1179904
dropout_10 (Dropout)	(None, 256)	0
dense_11 (Dense)	(None, 10)	2570

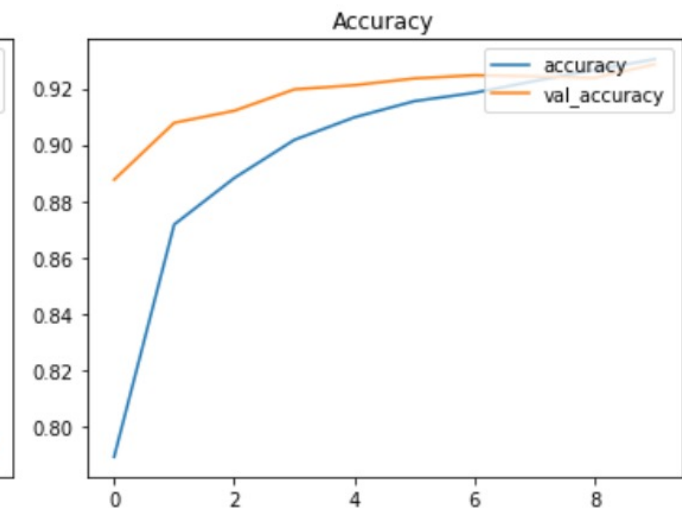
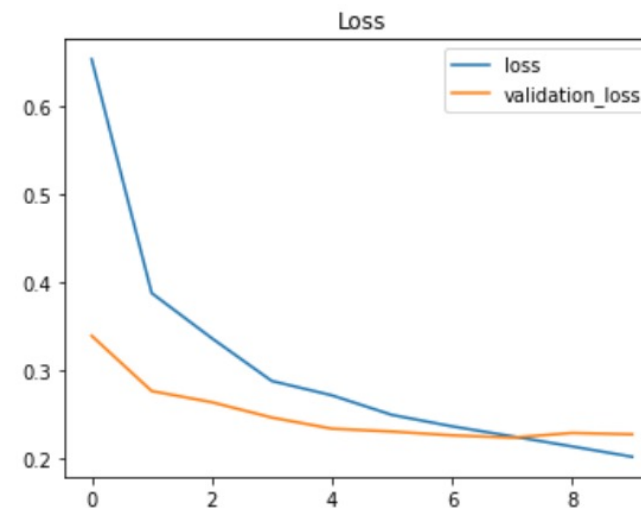
=====
 Total params: 1,275,722
 Trainable params: 1,275,722
 Non-trainable params: 0

→ Validation Accuracy: 0.6791



ResNet50

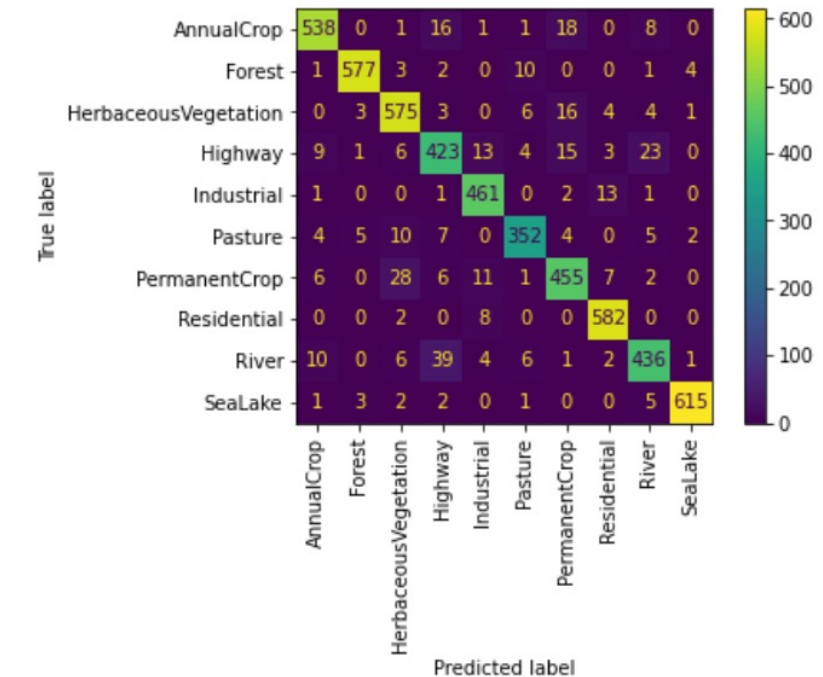
- ResNet50 architecture
- Pretrained weights of ImageNet (Transfer Learning)



Model: "model"

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 64, 64, 3)]	0	[]
⋮			
dense_7 (Dense)	(None, 10)	1290	['dropout_6[0][0]']

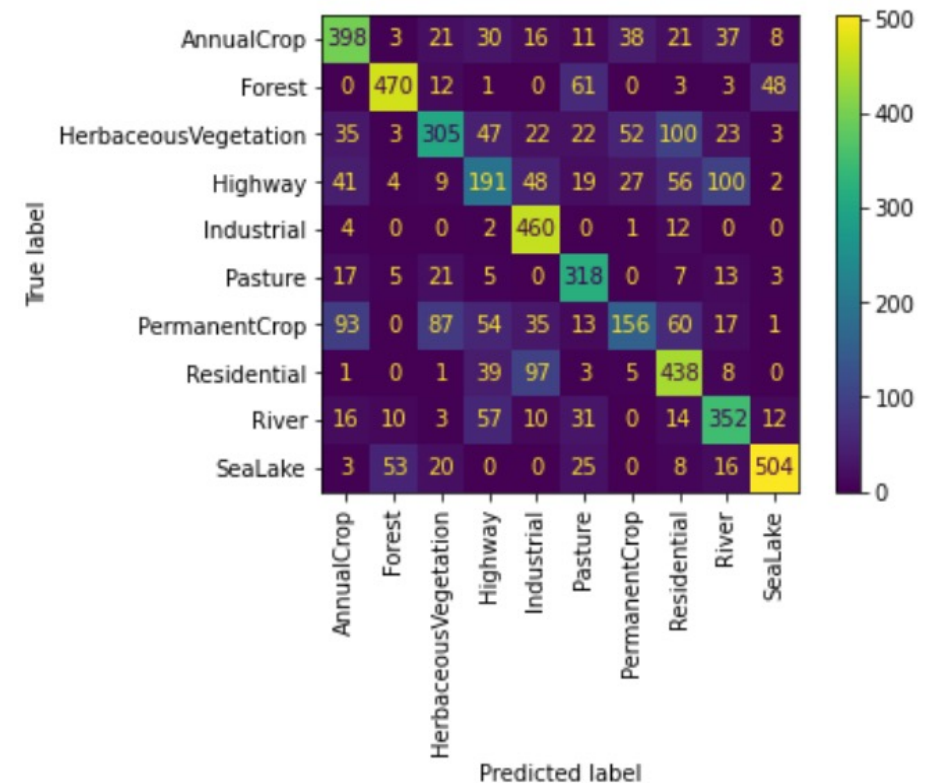
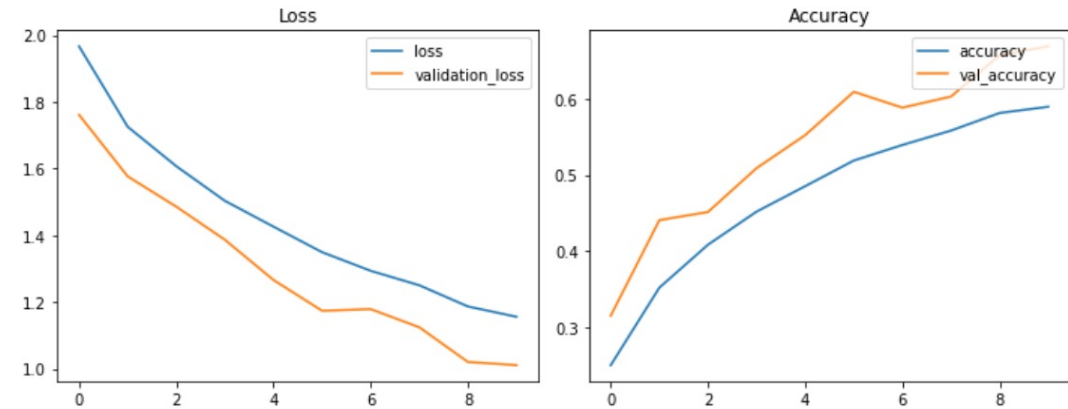
Total params: 23,851,274
 Trainable params: 263,562
 Non-trainable params: 23,587,712



→ Validation Accuracy: 0.9285

Hyperparameter-Tuned CNN

- Use architecture of Baseline CNN
- Hyperparameters (best model):
 - convolution size (7x7)
 - dense size (96)
 - learning rate (0.01)
- Random Search

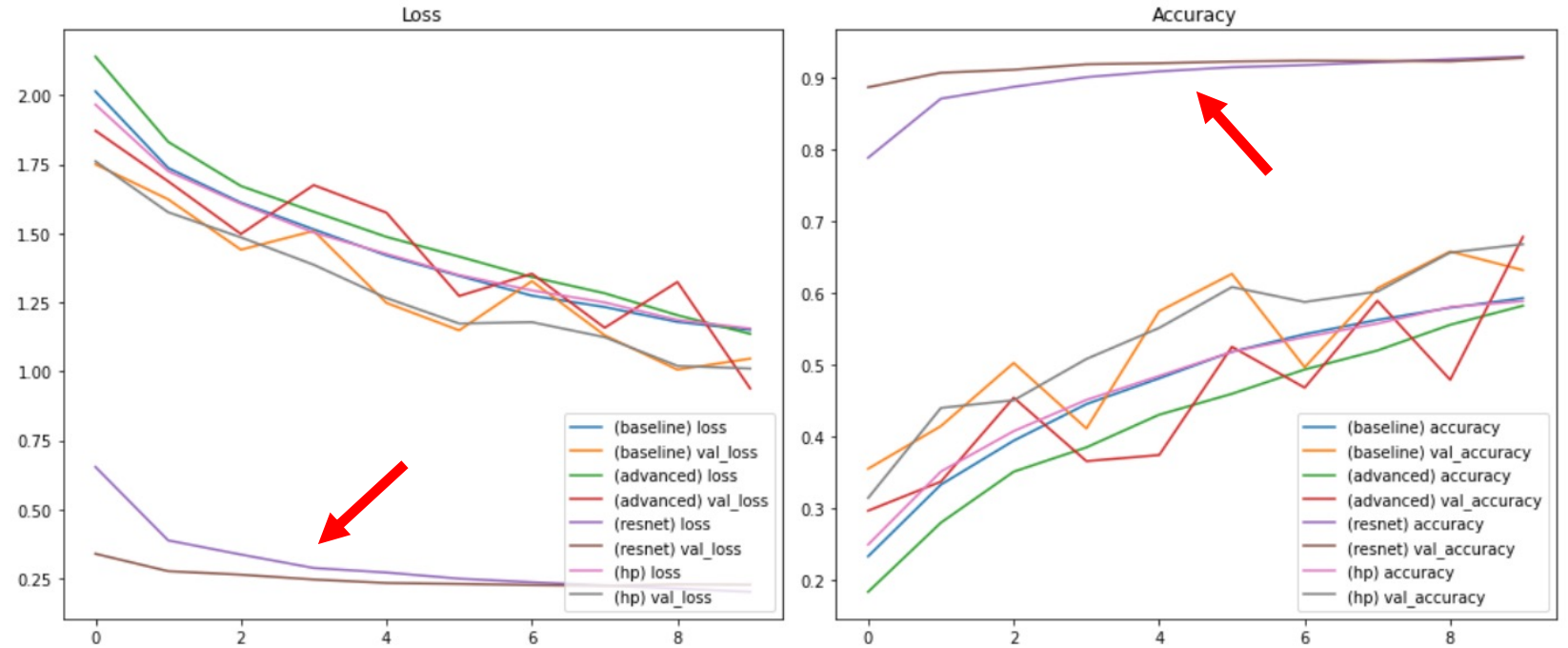


=====
Total params: 2,589,354
Trainable params: 2,589,354
Non-trainable params: 0

→ Validation Accuracy: 0.6687

Conclusions

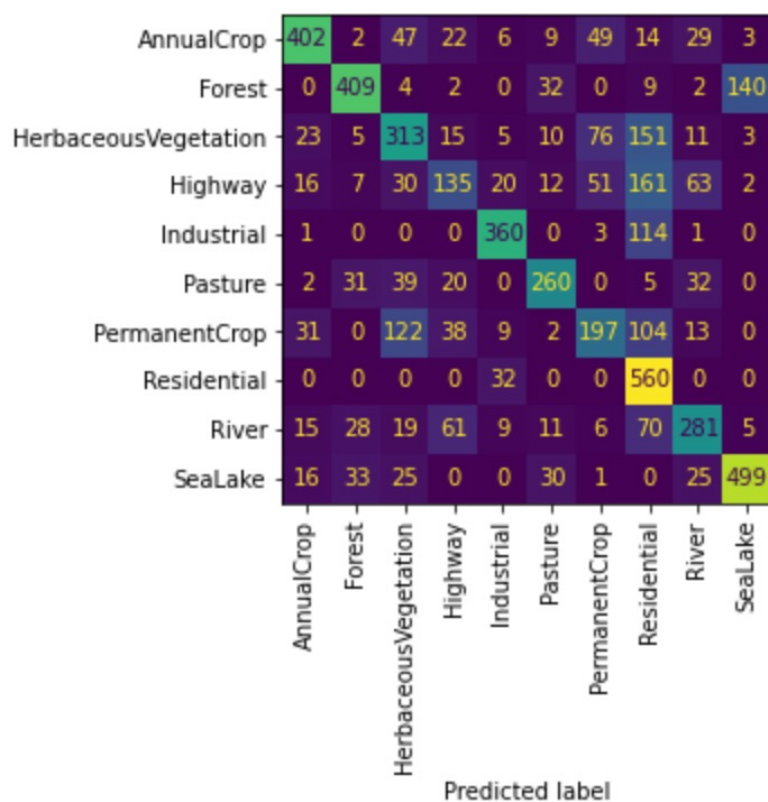
→ ResNet50
outperforms all other
models



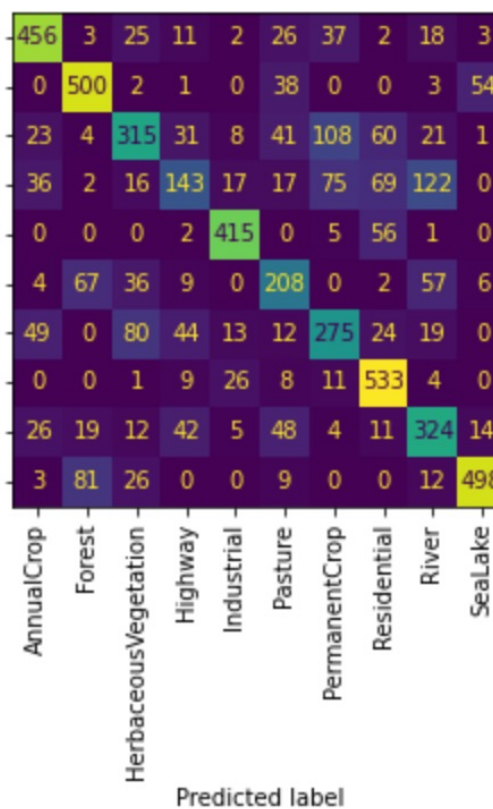
	Name	Total Params	Trainable Params	Non-Trainable Params	Loss	Val_Loss	Accuracy	Val_Accuracy
0	Baseline CNN	3938570	3938570	0	1.1507	1.0464	0.5936	0.6326
1	Advanced CNN	1275722	1275722	0	1.1362	0.9383	0.5830	0.6791
2	ResNet	23851274	263562	23587712	0.2017	0.2273	0.9303	0.9285
3	Baseline CNN (HP)	2589354	2589354	0	1.1557	1.0106	0.5894	0.6687

Conclusions

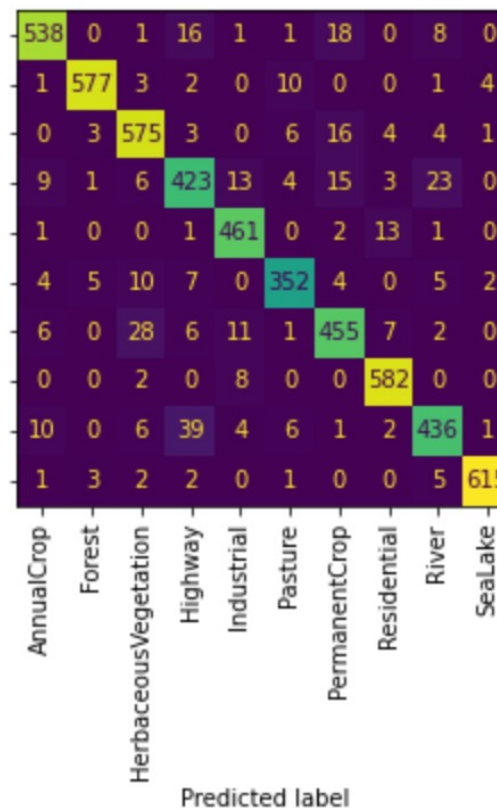
Baseline CNN



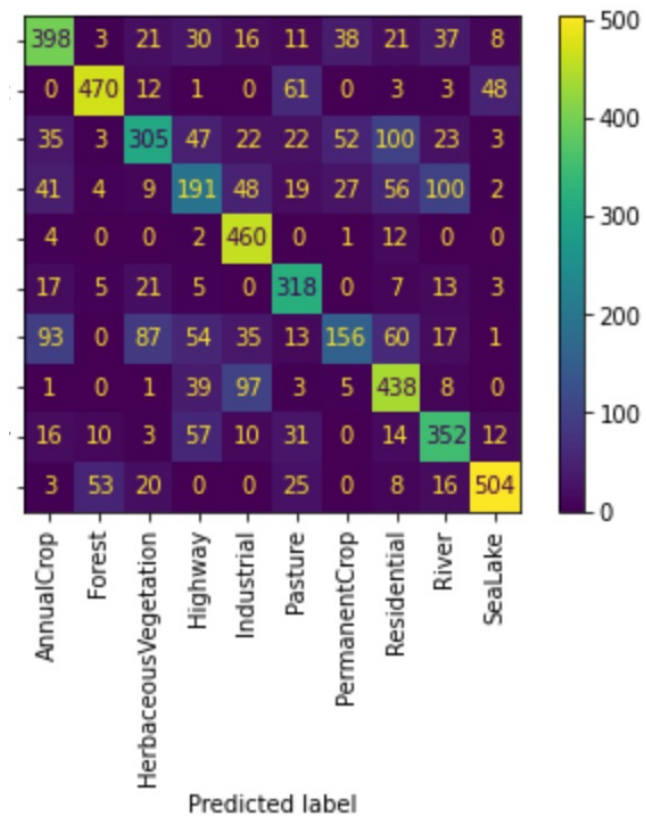
Advanced CNN



ResNet50



HP-Tuned CNN



References

- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778). doi: 10.1109/CVPR.2016.90
- Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(7), 2217-2226. <https://doi.org/10.5281/zenodo.7711810>