Potato Disease Classification

Leveraging Convolutional Neural Networks to identify and categorize potato diseases from images of potato tubers

Artificial Intelligence
Hendrik Lüdemann, Linus von Maltzan

Wirtschaftsingenieurwesen, TWE 22
Duale Hochschule Baden-Württemberg

Friedrichshafen 2025

Abb 1. Lars Blankers: Potatoes, https://unsplash.com/photos/brown-potato-lot-B0s3Xndk6tw, 18.02.2018

Use Case

- Potato farmers sort multiple million tubers manually
- Shortage of workers raises wages and reduces capacities

→ Optical potato sorter

Abb 2. Wevano: Sortieren, https://wevano.nl/de/agrarisch/sortieren/, 18.02.2018

Agenda

1. Problem Statement
2. Dataset Description
3. Exploratory Data Analysis
4. Preprocessing
5. Models Used
6. Evaluation Metrics
7. Conclusion & Future Work
8. References

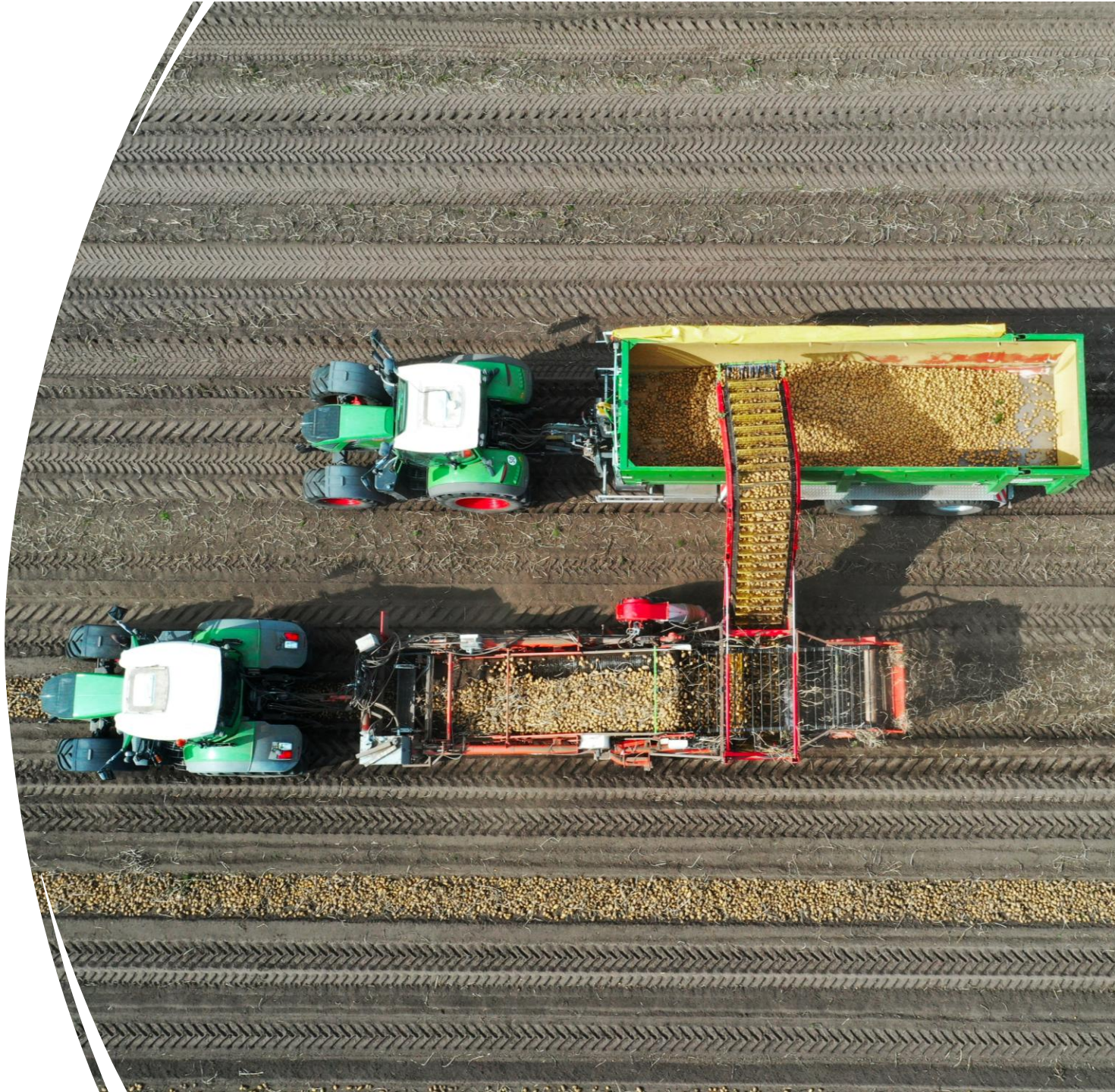


Abb 3. Hendrik Lüdemann, Potato Harvest, 2022

Problem Statement

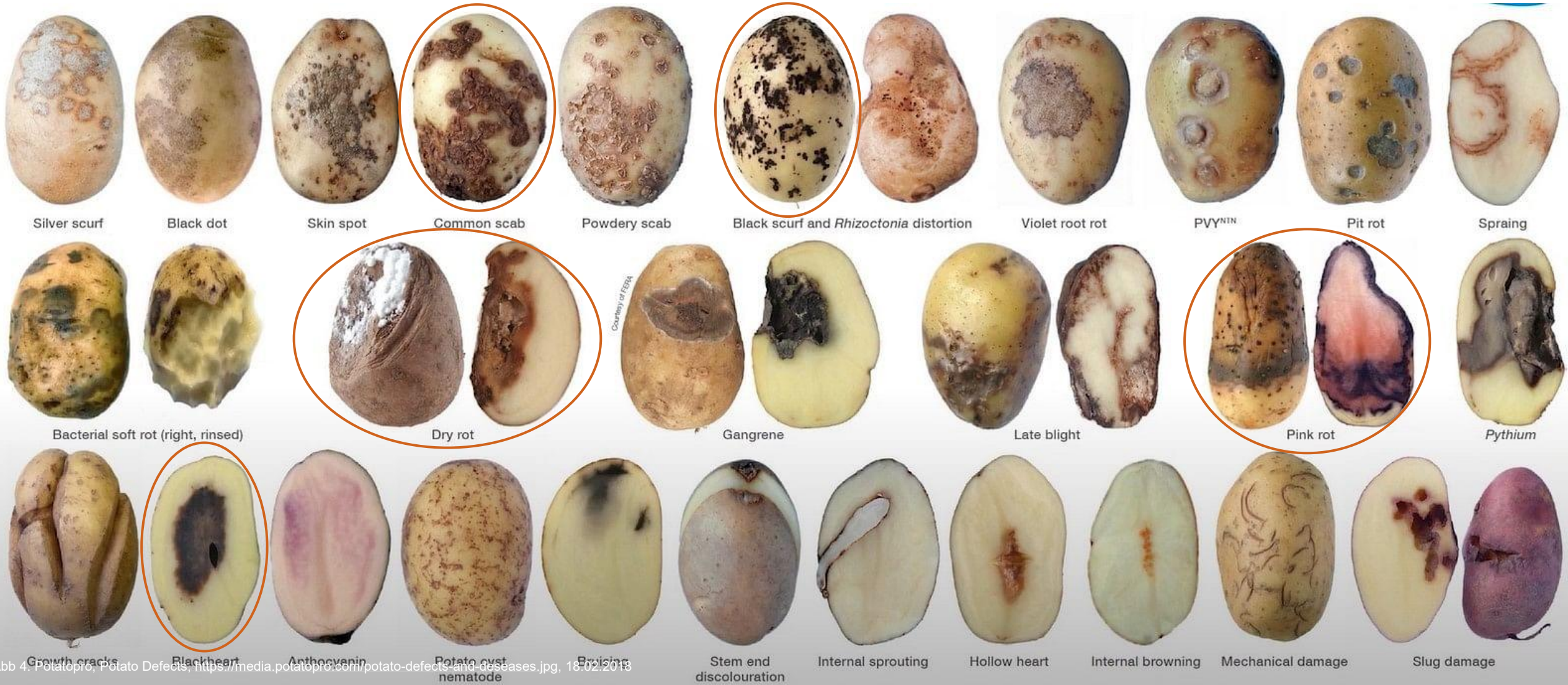
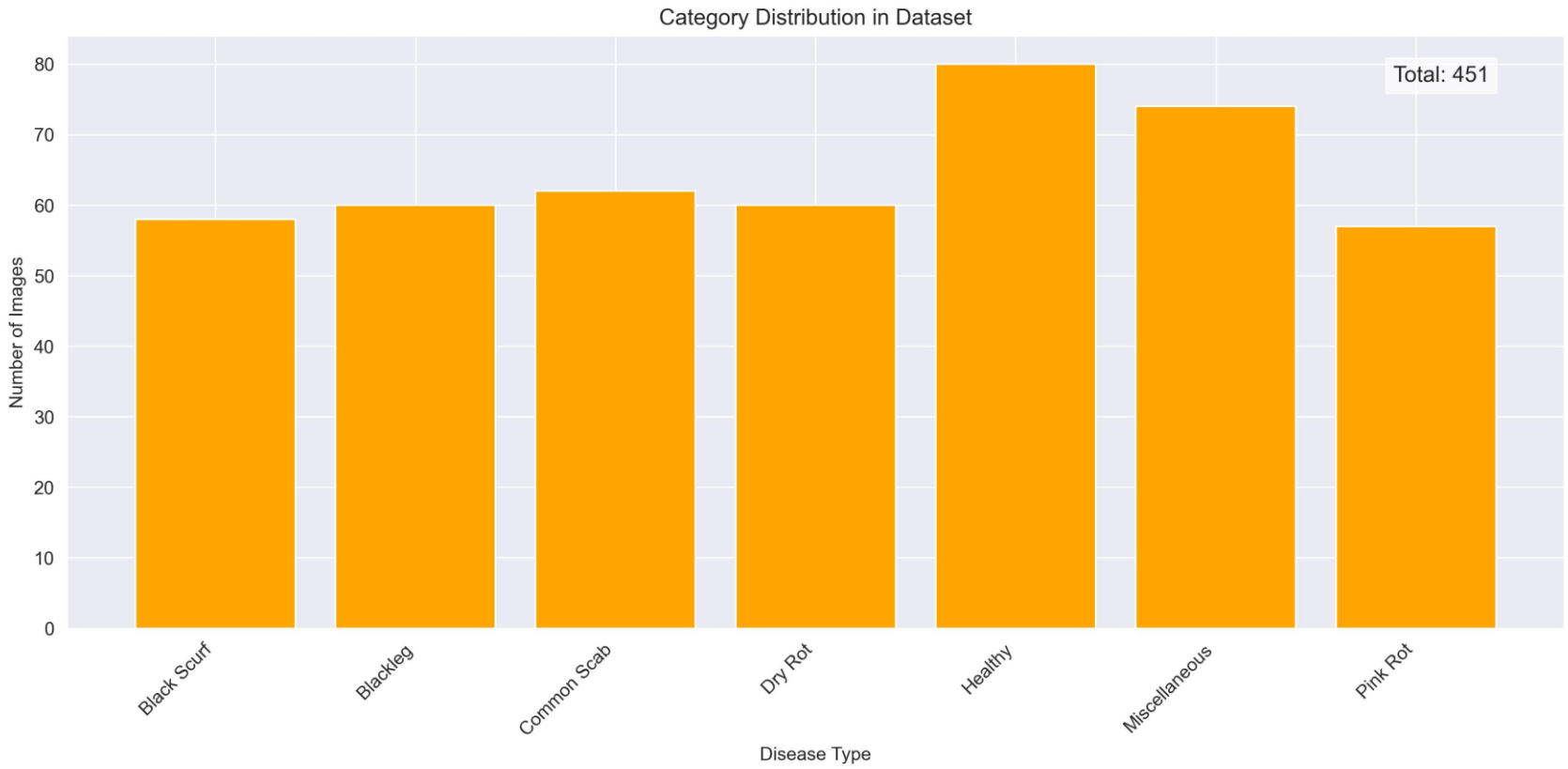
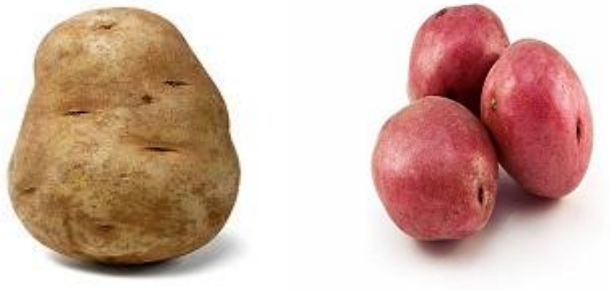


Abb 4: Potatopro, Potato Defects, https://media.potatopro.com/potato-defects-and-deseases.jpg, 18.02.2018

Dataset Description



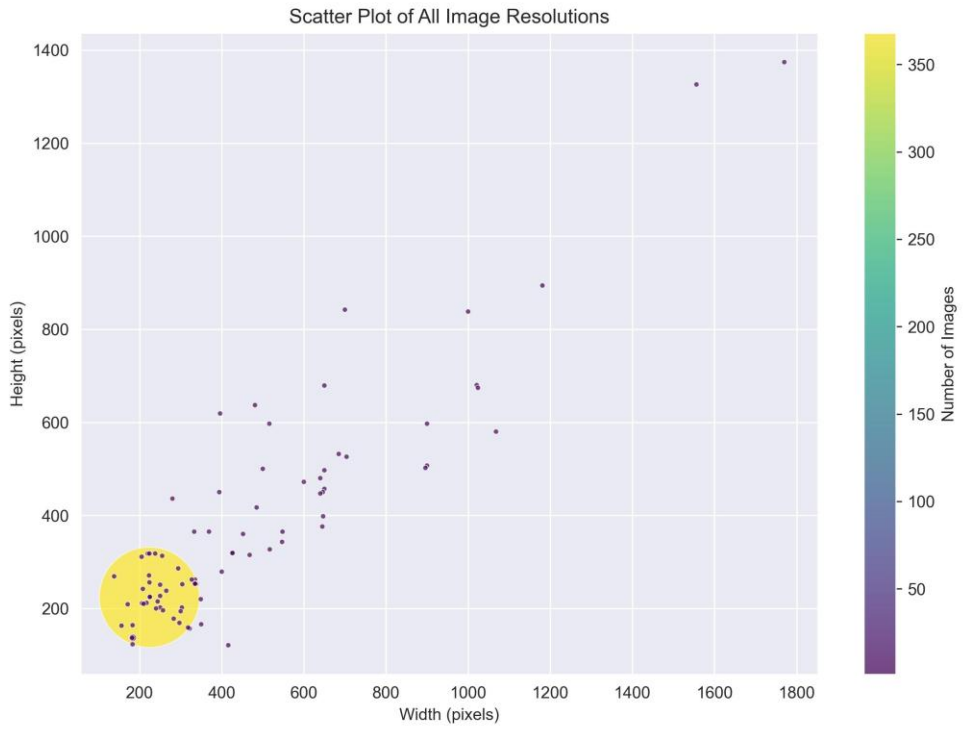


Healthy

Black Scurf

Abb 5. Mukaffi: Potatoes, https://www.kaggle.com/datasets/mukaffimoin/potato-diseases-datasets, 18.02.2018

Resolution Analysis



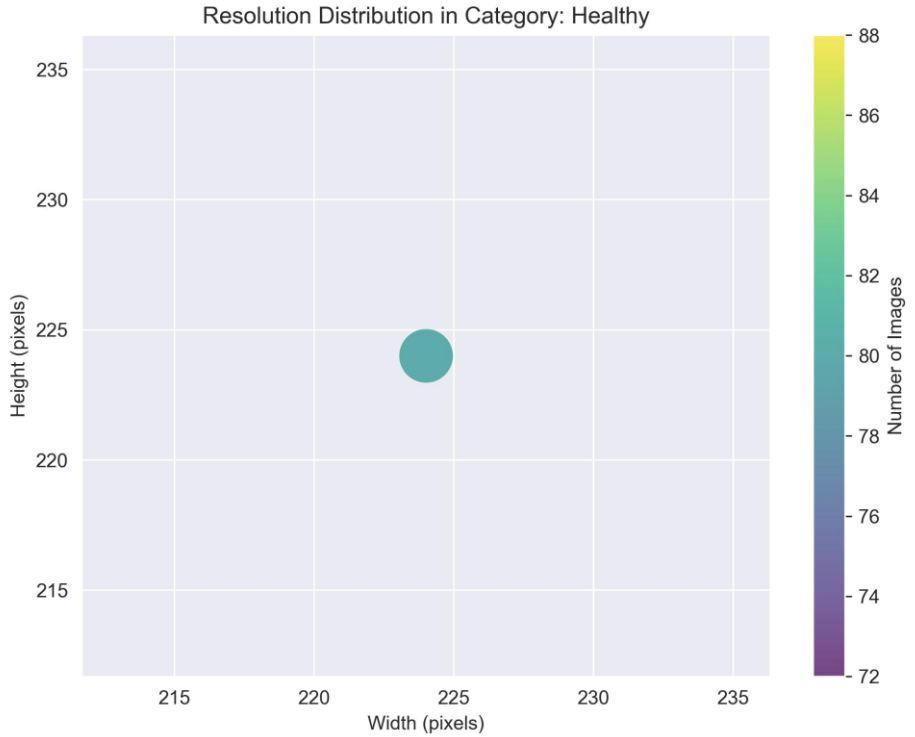
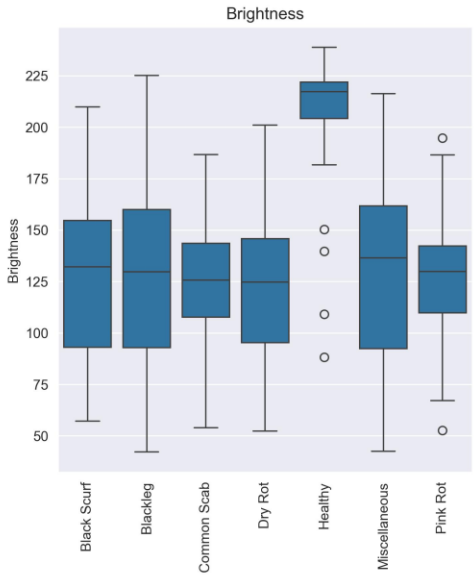
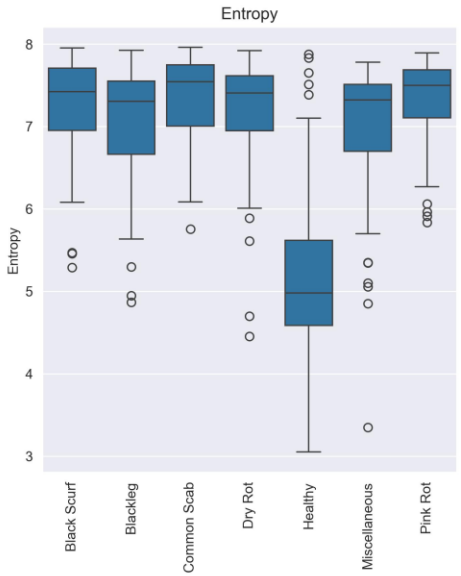
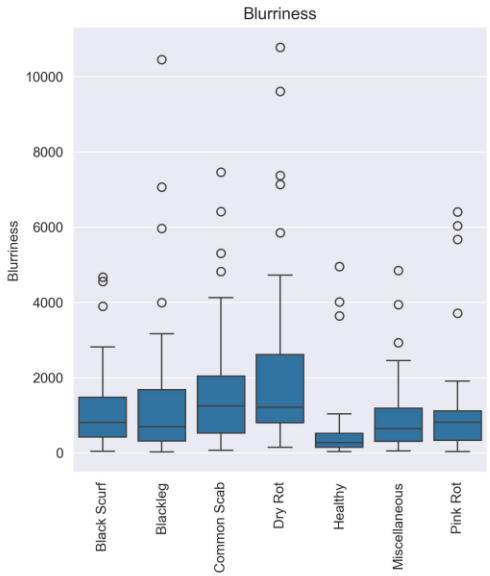
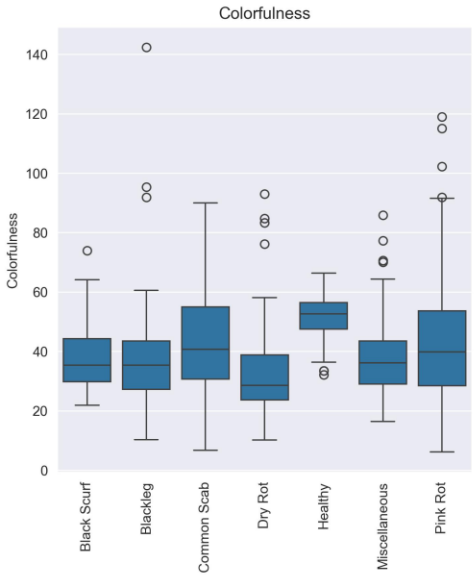


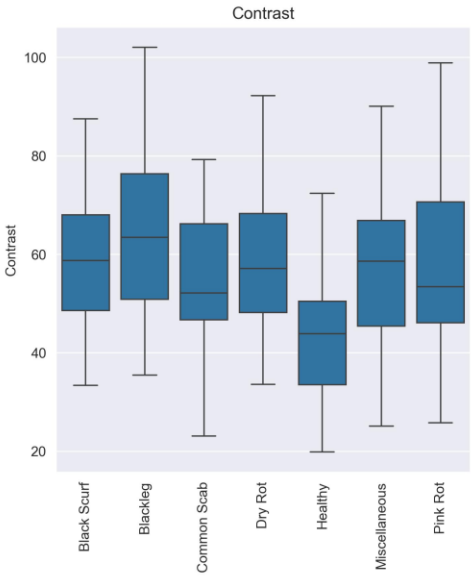
Image Property Analysis



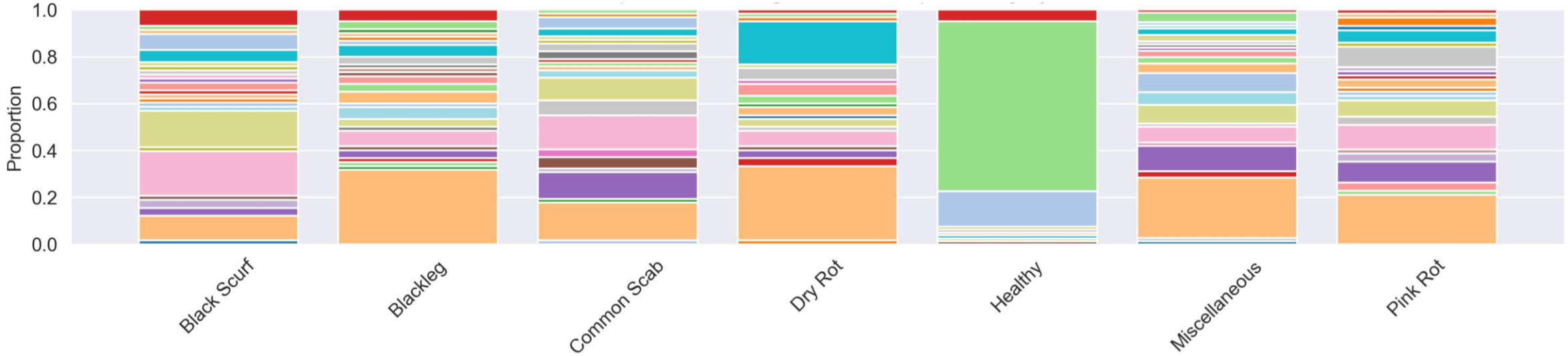




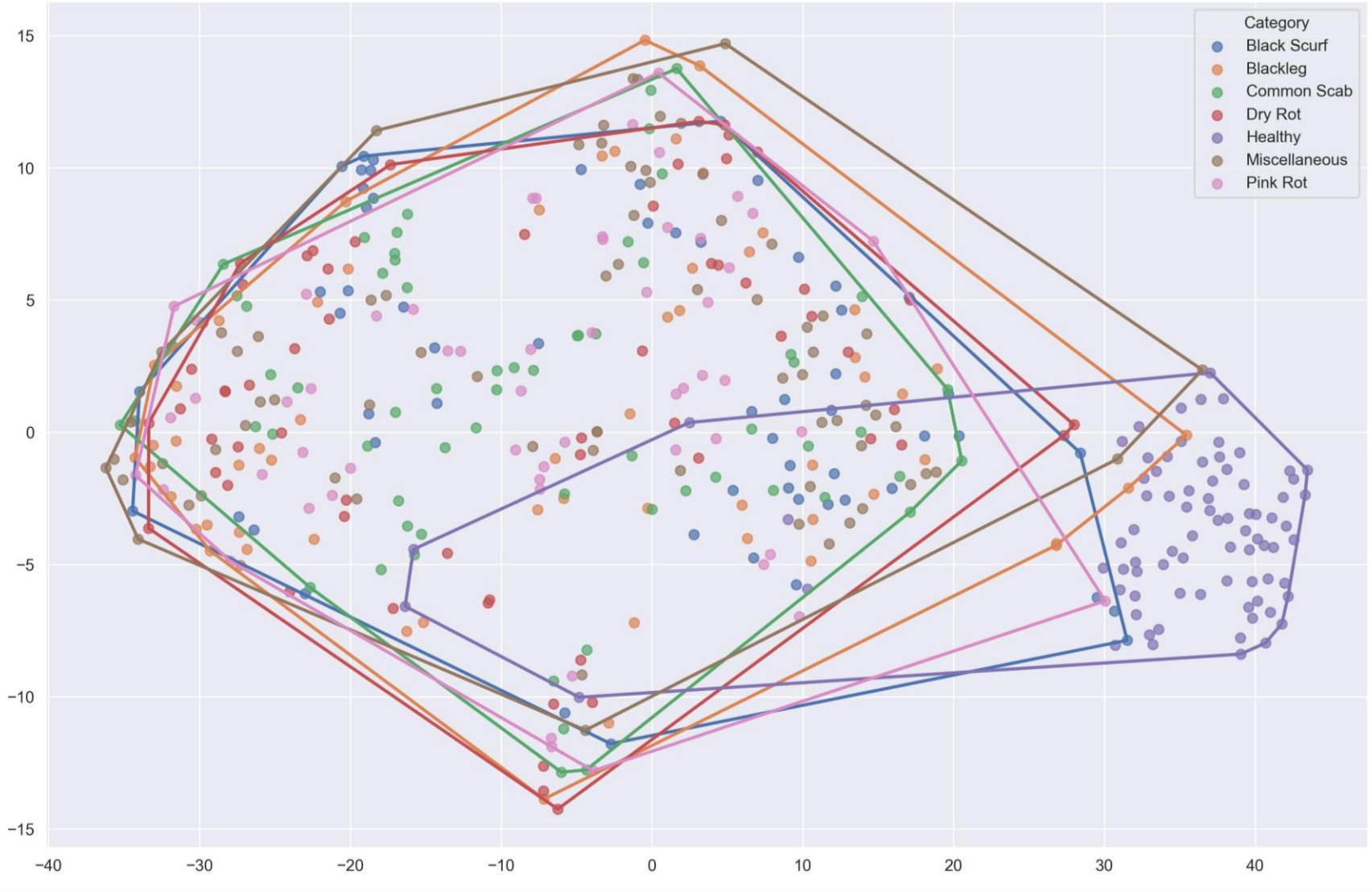




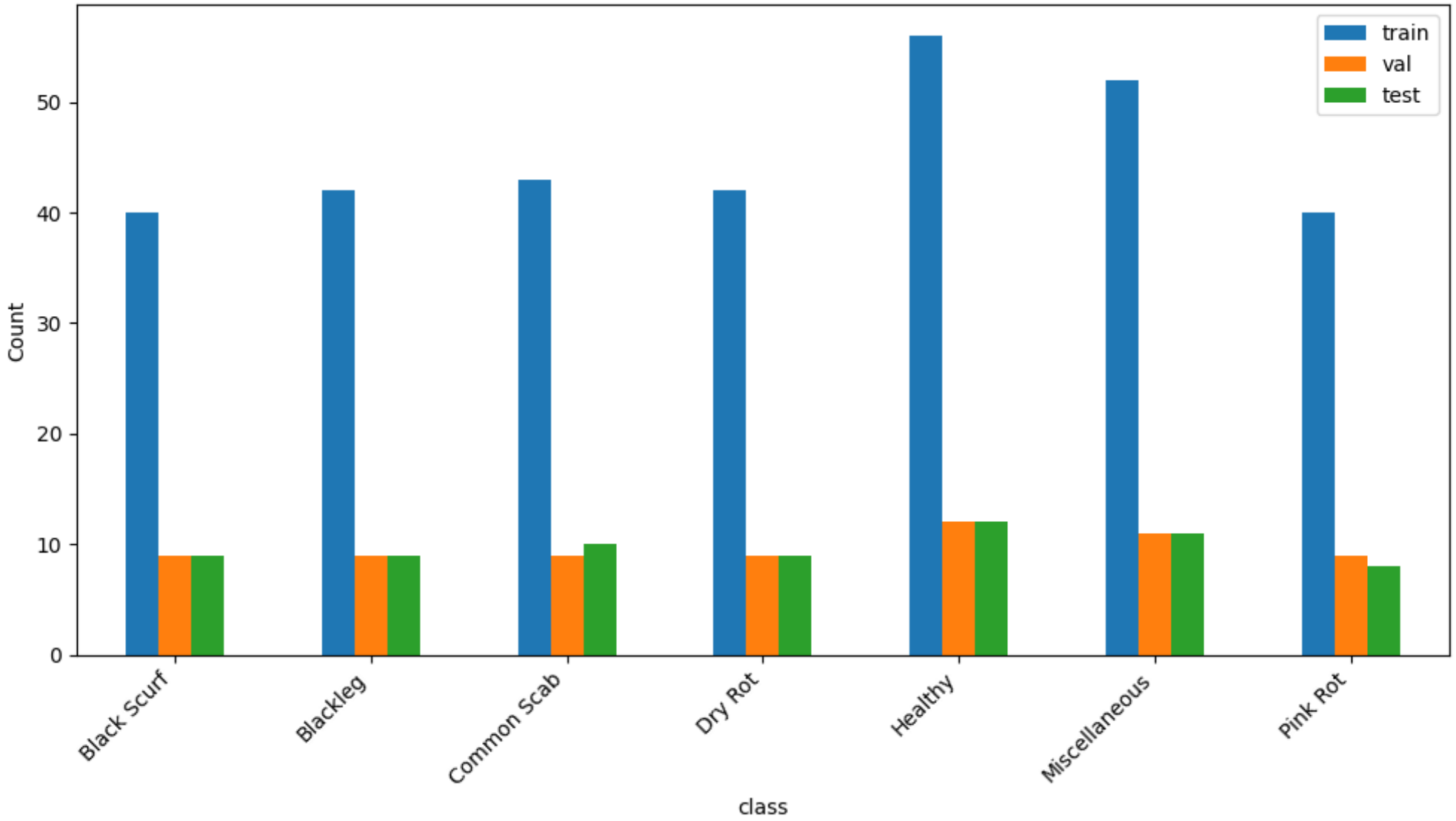
Background Color Analysis



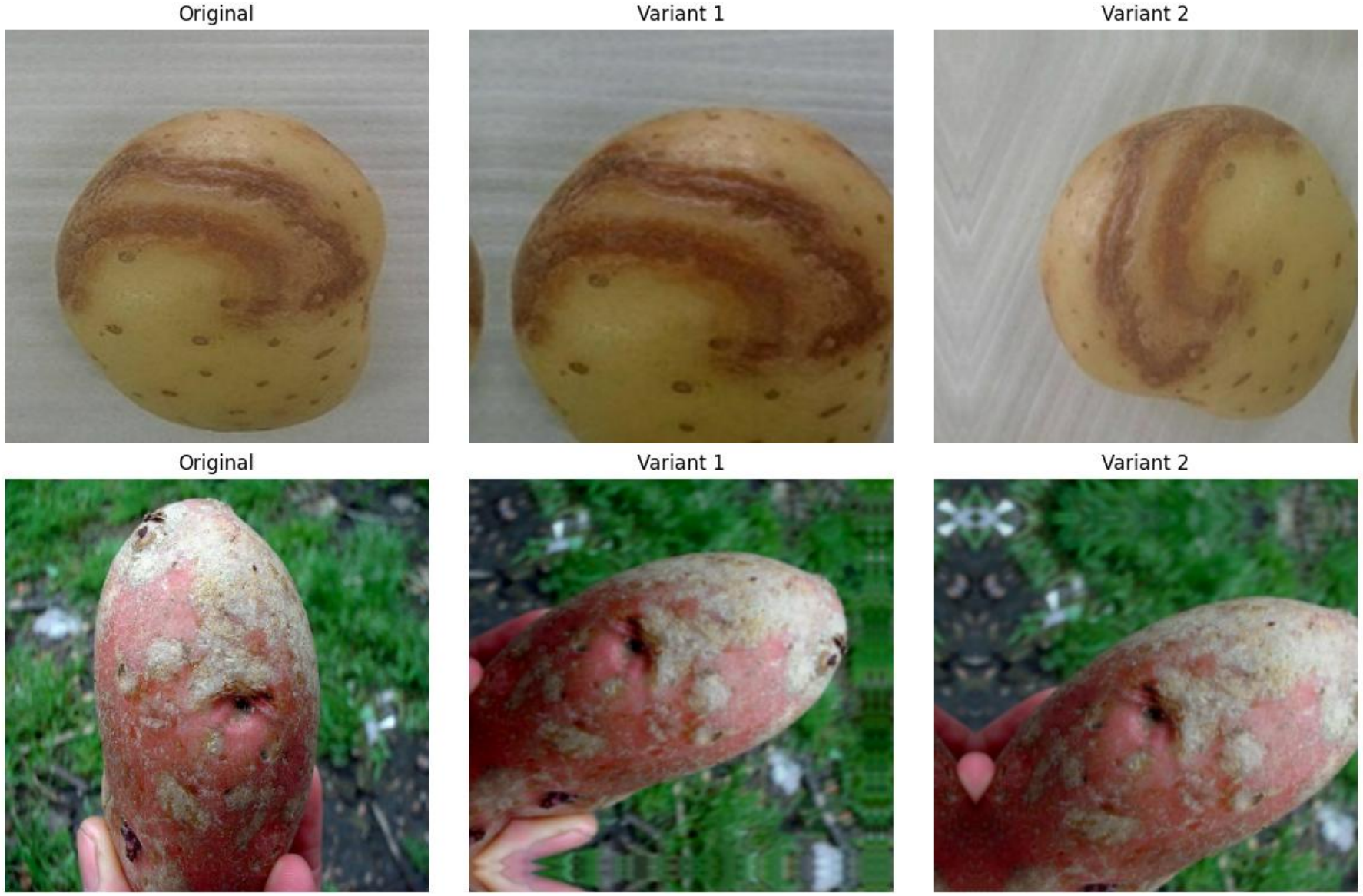
t-SNE Projection



Dataset Splitting using Stratified Sampling



Preprocessing



Machine Learning Approach

Model Architecture

Input Image

Takes an RGB image (224×224) as input to the network.

Pre-processing

Applies Efficient-Net specific normalization to prepare pixel values.

Efficient-NetV2

Extracts rich visual features using a pretrained convolutional backbone.

Global Avg. Pooling

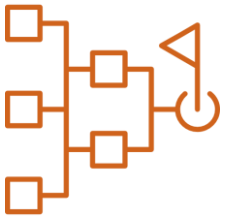
Compresses each feature map to a single value, reducing spatial dimensions.

Dropout

Randomly drops 30% of features to reduce overfitting and improve generalization.

Dense Softmax

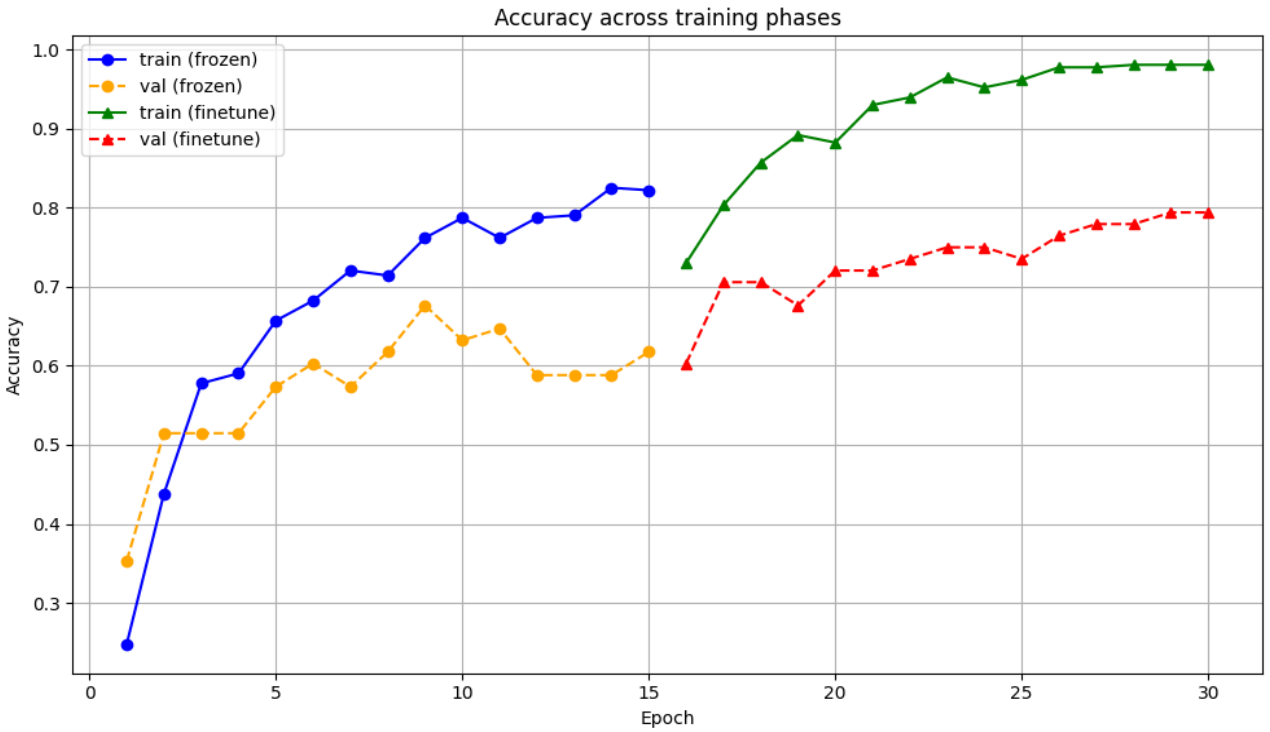
Maps features to class probabilities using a fully connected softmax layer.

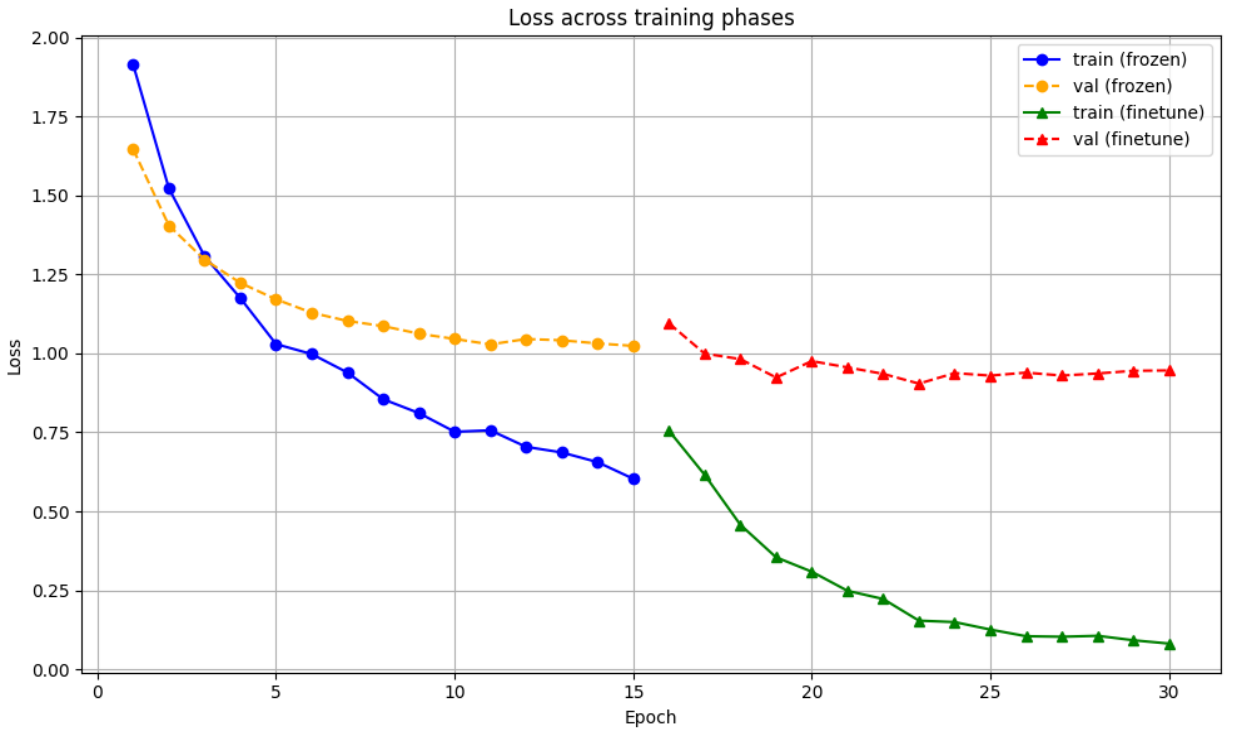


Output

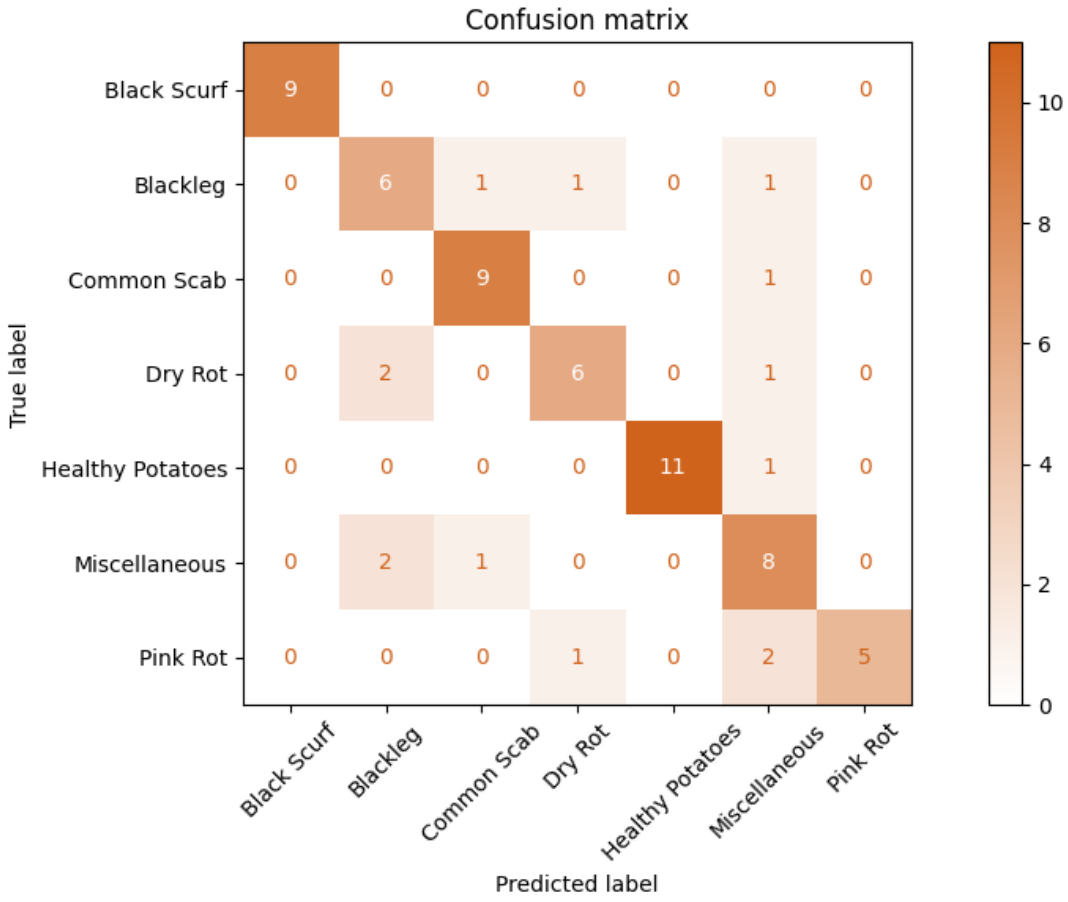
Outputs a probability distribution over target classes for final prediction.

Evaluation Metrics





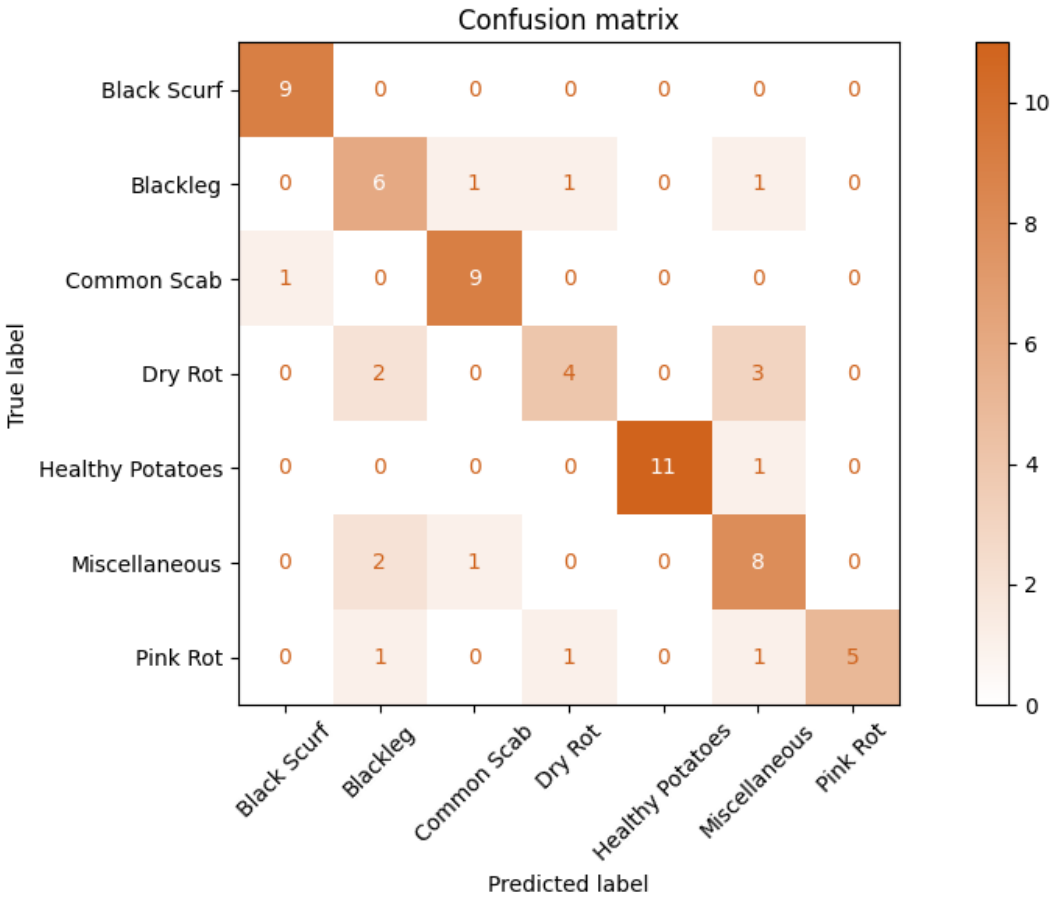
Evaluation Metrics



Class	Precision	Recall	F1-Score	Support
Black Scurf	1.00	1.00	1.00	9
Blackleg	0.60	0.67	0.63	9
Common Scab	0.82	0.90	0.86	10
Dry Rot	0.75	0.67	0.71	9
Healthy Potatoes	1.00	0.92	0.96	12
Miscellaneous	0.57	0.73	0.64	11
Pink Rot	1.00	0.62	0.77	8
Overall Accuracy			**0.79**	**68**

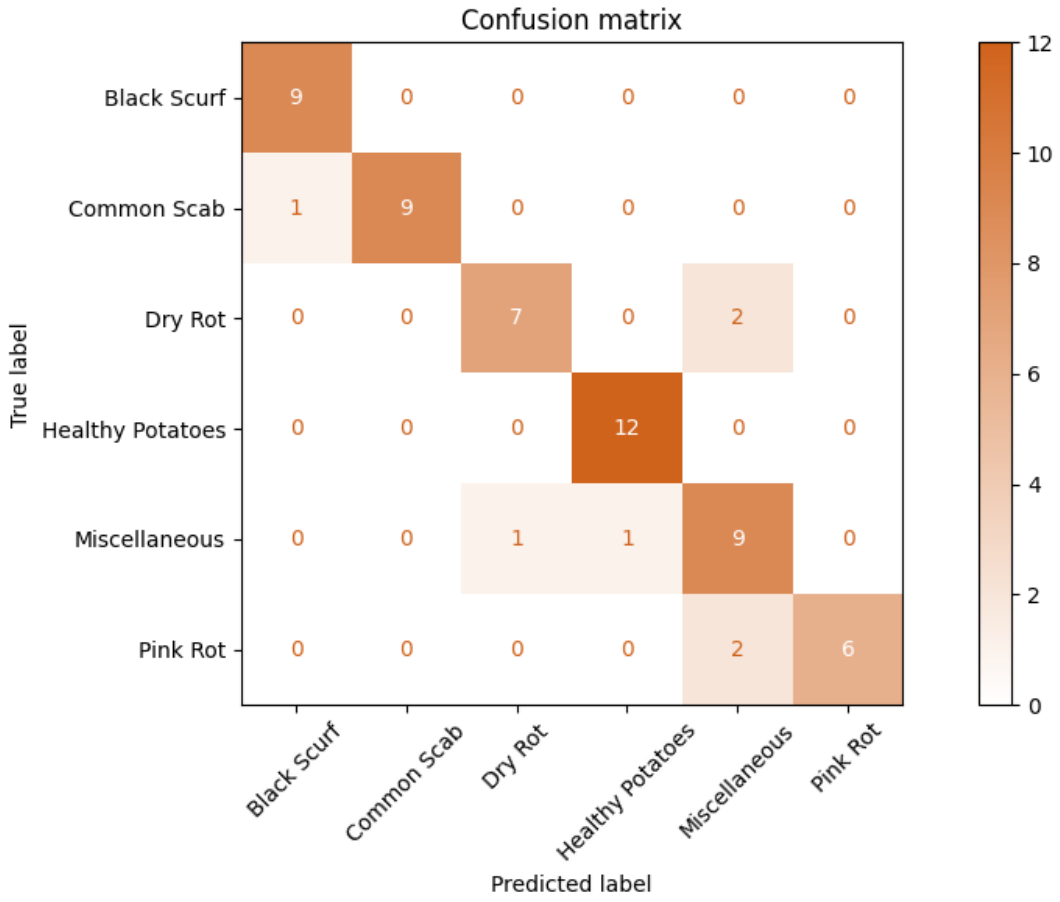
Hyperparameter-Tuning

Training with blackleg



Class	Precision	Recall	F1-Score	Support
Black Scurf	0.90	1.00	0.95	9
Blackleg	0.55	0.67	0.60	9
Common Scab	0.82	0.90	0.86	10
Dry Rot	0.67	0.44	0.53	9
Healthy Potatoes	1.00	0.92	0.96	12
Miscellaneous	0.57	0.73	0.64	11
Pink Rot	1.00	0.62	0.77	8
Overall Accuracy			**0.76**	**68**

Training without blackleg



Class	Precision	Recall	F1-Score	Support
Black Scurf	0.90	1.00	0.95	9
Common Scab	0.82	0.90	0.95	10
Dry Rot	0.88	0.78	0.82	9
Healthy Potatoes	0.92	1.00	0.96	12
Miscellaneous	0.69	0.82	0.75	11
Pink Rot	1.00	0.75	0.86	8
Overall Accuracy			**0.88**	**59**

Conclusion & Future Work

Expand Dataset

Increase the volume and diversity of training images by sourcing potatoes from multiple regions, varieties and growth stages to boost model robustness.

Real-World Imagery

Gather real-world photos directly from the production line's conveyor belt, reflecting true lighting, motion blur and occlusions for reliable in-line classification.

Broaden Disease Coverage

Add images of additional potato diseases to ensure the algorithm detects and differentiates every relevant disease category.

Integrate diverse imagery and expanded disease categories to achieve robust, potato disease classification in real-world production environments.

References

A Complete Guide to Data Augmentation, Accessed 10 June 2025. https://www.datacamp.com/tutorial/complete-guide-data-augmentation.

Erdem (burnpiro), Kemal. 'T-SNE Clearly Explained'. TDS Archive (blog), 21 July 2022. https://medium.com/data-science/t-sne-clearly-explained-d84c537f53a.

Gulli, Antonio, and Sujit Pal, Deep Learning with Keras. Packt Publishing Ltd, 2017. https://books.google.com/books?hl=en&lr=&id=20EwDwAAQBAJ&oi=fnd&pg=PP1&dq=info:dtP3ug-T4p4J:scholar.google.com&ots=lJcE9meUO0&sig=VsR7KO3VLkEqTE78kRGPW2uQaGk.

Kaggle, Potato Diseases Datasets, Accessed 10 June 2025. https://www.kaggle.com/datasets/mukaffimoin/potato-diseases-datasets

'**K-Means Clustering,** In Wikipedia, 13 March 2025. https://en.wikipedia.org/w/index.php?title=K-means_clustering&oldid=1280238914.

Potato, Identifying Diseases, Vegetable : Center for Agriculture, Food, and the Environment at UMass Amherst', 6 September 2022. https://www.umass.edu/agriculture-food-environment/vegetable/fact-sheets/potato-identifying-diseases.

Rastovski, A., and Aren Jacob Huibert van Es. Storage of Potatoes: Post-Harvest Behaviour, Store Design, Storage Practice, Handling. Wageningen: Centre for Agriculture Publishing and Documentation, 1981.

Regina, 'Potato Diseases & Common Pests'. Plantura (blog), 18 October 2021. https://plantura.garden/uk/vegetables/potatoes/potato-diseases.

Tan, Mingxing, and Quoc V. Le. 'EfficientNetV2: Smaller Models and Faster Training'. arXiv, 23 June 2021. https://doi.org/10.48550/arXiv.2104.00298.

T-Distributed Stochastic Neighbor Embedding (t-SNE), Dimensionality Reduction Techniques (4/5), 2024. https://www.youtube.com/watch?v=U-s8q6HshZw.

TensorFlow. 'Introduction to TensorFlow'. Accessed 10 June 2025. https://www.tensorflow.org/learn.