Plant Disease Detection using Deep Learning

Nathan Abela

Institute of Information & Communication Technology

MCAST, Paola, Malta

Problem Definition

- **Plant diseases** are **responsible** for **20-40%** of global crop **loss** per year [1].
    - **Amounting** to over **$220 billion annually** [1].
- Detection systems aid in **early disease identification**.
- Led to the **P**recision **A**griculture (**PA**) technology creation [2].

Literature - Datasets

Dataset	No. of Images	No. of Classes	Resolution
PlantVillage	54,303	38	256x256
PlantDoc	2,598	17	Various
PlantLeaves	4,503	17	6000x4000
PlantNet-300K	306,146	1,081	600x600

Table 1: Dataset Comparison

Literature – Key Papers

- **Dataset** – PlantVillage – all 38 classes
- **Models** – EfficientNet – VGG16

Study	Classifier	Accuracy
Atila et al. [3]	EfficientNet	99.97%
Mohameth et al. [4]	VGG16	97.82%

Table 2: Results Comparison between two key papers

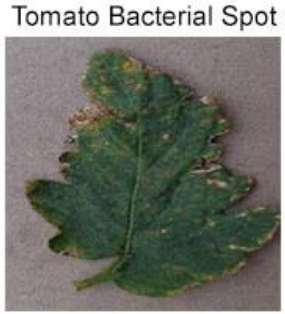
Literature – Local Contributions

Aspect	Galdes [5]	Busutill [6]	Mercieca [7]	Xuereb [8]
Detection Type	Disease	Disease	Crop & Disease	Disease
Datasets	PlantVillage	PlantVillage	PlantVillage & Custom	PlantVillage
Plants	Tomato	Potato	Apples, Grapes & Potatoes	Various
Classes	10	3	11	15
Algorithm	XGBoost + VGG16	YOLOv5	VGG16	GoogLeNet & AlexNet
Accuracy	91%	98.5%	98.67%	95% & 78%

Table 3: Summary comparison of local research contribution

Methodology – Hypothesis

This research **hypothesises** that **recent advancements** in deep learning models, particularly CNNs, permit a more **holistic** and **accurate detection** and **classification** of various plant diseases.



Methodology – Research Questions

1. How well can **established** deep learning **methods** for plant disease detection be **replicated**, and what **strengths** and **weaknesses** can be gained?

2. Can computer vision techniques differentiate between **healthy** and diseased plants?

3. How can an **experimentation plan** help **advance** plant disease **detection**?

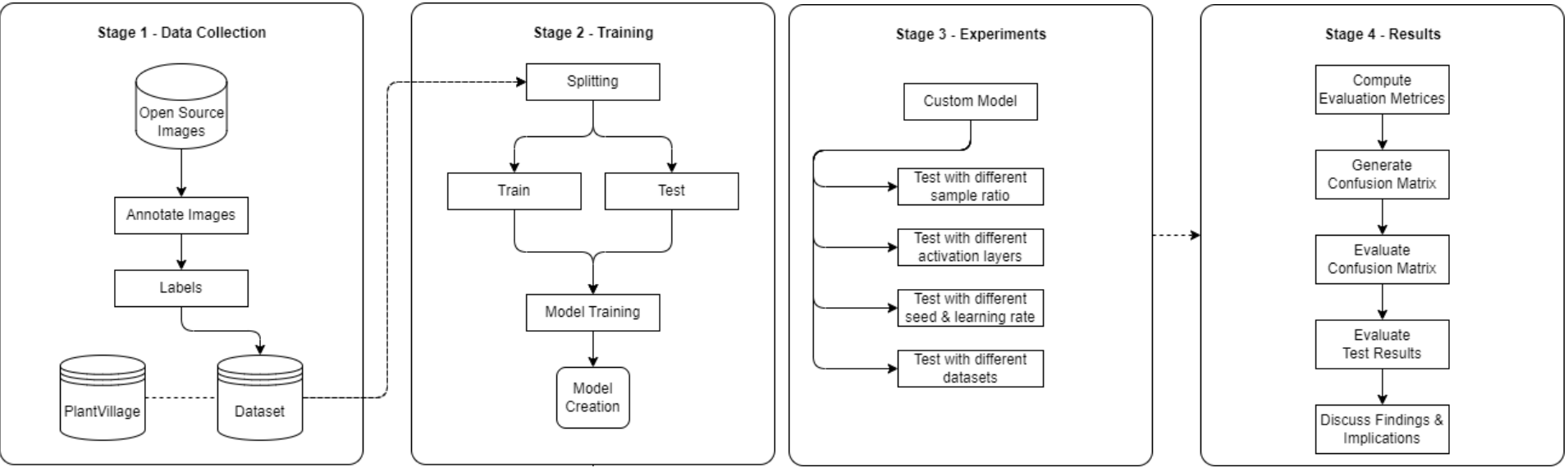
4. How do **variations** in CNN architecture **affect** the **efficiency** of the model?

Methodology – Research Questions

5. What **impact** do different **hyperparameters** have on the **performance** of **deep-learning models** for plant disease detection?

6. How does the **proposed solution** perform with key **local crops**?

Methodology – Novel Pipeline



Methodology - Experiments

Phase 1:

Calibration

- **Replicating** key **methodologies** from two key papers **Atila et al.** [3] and **Mohameth et al.** [4]

Phase 2:

Novel

- **Custom** deep-learning **algorithm** using **insights** gained from **calibration phase**

Phase 3:

Real-World Application

- Testing **custom dataset** using **custom algorithm**
- Use of **Grad-CAM** to visualise model under **real-world conditions**

Results – Calibration Phase

Calibration Experiments – Replication of Methodologies from Key Papers [RQ 1]

Study	Atila et al. [3]	Atila et al. Replication	Mohameth et al. [4]	Mohameth et al. Replication
Accuracy	99.84%	95.86%	97.82%	95.64%
Precision	97.24%	96.33%	N/A	96.64%
Recall	96.82%	95.55%	N/A	94.71%
F1-Score	N/A	95.94%	96.42%	95.66%

Table 4: Results comparison of the replicated methodologies and the original studies using the PlantVillage dataset

Results – Novel Phase

Novel Experiments – Custom Algorithm [RQ 3, 4 & 5]

Study	Custom Algorithm
Accuracy	81.01%
Precision	85.21%
Recall	78.27%
F1-Score	81.59%

Table 5: Results summary of the Custom model using the PlantVillage dataset

Results – Real-World Application Phase

Real-World Application Experiments – Custom Dataset [RQ 2, 3 & 6]

Study	Custom Dataset
Accuracy	85.13%
Precision	91.94%
Recall	81.82%
F1-Score	86.13%

Table 6: Results summary of the Custom model using the Custom dataset

Custom Dataset Overview

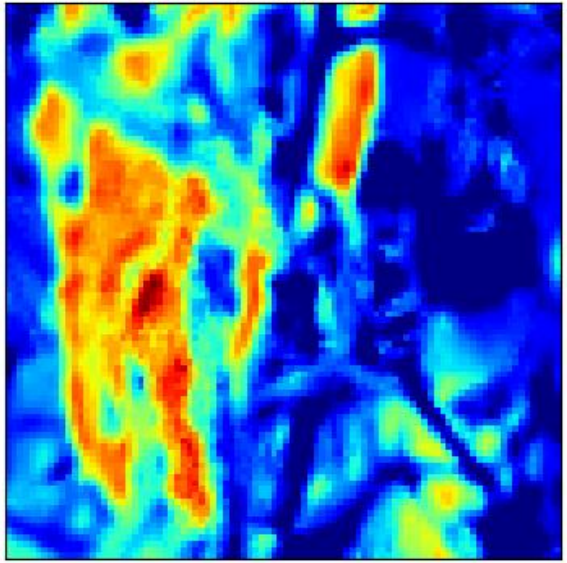
Class	Plant	Healthy or Diseased	Disease Name	Type of Disease	No. of Images
C_1	Potato	Diseased	Late Blight	Fungal	80
C_2	Potato	Healthy	/	/	80
C_3	Strawberry	Diseased	Leaf Scorch	Fungal	80
C_4	Strawberry	Healthy	/	/	80
C_5	Tomato	Diseased	Yellow Leaf	Viral	80
C_6	Tomato	Healthy	/	/	80

Table 7: Class breakdown of the Custom dataset

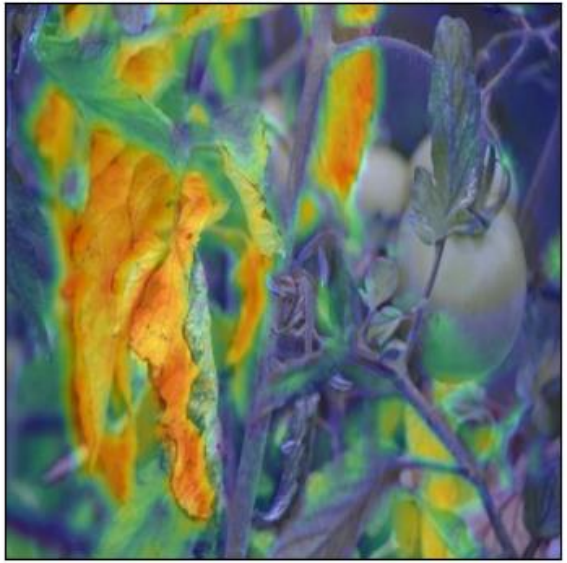
Custom Dataset Evaluation

Original Image

Grad CAM Heatmap



Superimposed Image



Research Question Evaluation

- **How** well can **established** deep learning **methods** for plant disease detection be **replicated**, and what **strengths** and **weaknesses** can be gained?
    - **Effective** and **Successful replication**.
- **Can computer vision techniques differentiate** between **healthy** and diseased plants?
    - CV **techniques** can distinguish between **healthy** and **diseased** crops.
- **How** does the **proposed solution** perform with key **local crops**?
    - Despite lack of diversity in **custom dataset**, the **outcome** does address such.

Research Question Evaluation

- How can an **experimentation plan** help **advance** plant disease **detection**?
- How do **variations** in CNN architecture **affect** the **efficiency** of the model?
- What **impact** do different **hyperparameters** have on the **performance** of **deep-learning models** for plant disease detection?
    - With an experimentation plan **insights** were gathered on how **hyperparameters** and **various CNN** architecture **contribute** and their **effect**.

Conclusions - Limitations

- **Absence** of **locally** sourced **images** – limits study validity
- Custom dataset sourced from **online sources** rather than **local fields**.
- **Seasonal availability** of the **diseased leaves**.

Conclusions - Recommendations

- Prioritise the **development** of a **more** representative **local** dataset.
- **Collaboration** with **local farmers** and **agricultural organisations**.
- More **advanced neural network architectures** that require **less computational power** to be used in **real-time applications**.

References

- [1] H. Canton, "*Food and agriculture organization of the united nations-fao*", in The Europa directory of international organizations 2021. Routledge, 2021, pp. 297–305.
- [2] S. Candiago, F. Remondino, M. De Giglio, M. Dubbini, and M. Gattelli, "Evaluating multispectral images and vegetation indices for precision farming applications from uav images," Remote Sensing, vol. 7, no. 4, pp. 4026–4047, 04 2015.
- [3] Umit Atila, Murat Uc¸ar, Kemal Akyol, and Emine Uc¸ar. Plant leaf disease classification using efficientnet deep learning model. Ecological Informatics, 61:101182, 03 2021.

References

- [4] Faye Mohameth, Chen Bingcai, and Kane Amath Sada. Plant disease detection with deep learning and feature extraction using plant village. Journal of Computer and Communications, 08:10–22, 01 2020.
- [5] Dylan Galdes. Comparing image classification methods for tomato plant disease detection. Bachelor of Science (Honours) in Software Development, 06 2021.
- [6] Jasmine Busuttil. Identifying plant diseases using yolov5. Bachelor of Science (Honours) in Software Development, 06 2022.

References

- [7] Robert Mercieca. Using convolutional neural networks to identify different plants and common diseases associated with them. Bachelor of Science (Honours) in Software Development, 05 2022.
- [8] Melchior Xuereb. Plant diseases: Investigating how they can be detected and identified using ai. Bachelor of Science (Honours) in Software Development, 06 2020.

Thank you

Nathan Abela

nathan.abela.e25765@mcast.edu.mt

MCAST