

VISUAL QUESTION ANSWERING MODEL FOR PLANT DISEASE IDENTIFICATION

Balasooriya Arachchilage Gayani Indunil Chandrarathne

208524C

Master of Science in Artificial Intelligence

Department of Computational Mathematics
Faculty of Information Technology

University of Moratuwa
Sri Lanka

April 2023

VISUAL QUESTION ANSWERING MODEL FOR PLANT DISEASE IDENTIFICATION

Balasooriya Arachchilage Gayani Indunil Chandrarathne

208524C

Thesis/Dissertation submitted in partial fulfillment of the requirements for the degree
Master of Science in Artificial Intelligence

Department of Computational Mathematics
Faculty of Information Technology

University of Moratuwa
Sri Lanka

April 2023

DECLARATION

I declare that this is my own work and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature:

Date:

The above candidate has carried out research for the PhD/MPhil/Masters thesis/dissertation under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor:

Signature of the Supervisor:

Date:

DEDICATION

I dedicate this work to my loving family, who has always been my source of strength, inspiration, and motivation throughout my academic journey.

ACKNOWLEDGEMENT

I would like to express our sincere gratitude to all those who have contributed to this project.

ABSTRACT

The notable achievements in AI tasks owe their success to the natural language processing (NLP) domain with Large Language Models (LLM) and led to the emergence of new research directions in Deep Learning. The Visual Question Answering (VQA) task has garnered considerable attention owing to its promising results obtained through the use of pre trained LLMs. Here we are investigated a VQA as a domain specific expert system for domain specific knowledge representation and extraction. We have implemented a novel approach for plant disease identification, an expert-level task, utilizing fine-tuning LLM. The VQA technique has been utilized as a means of knowledge extraction, making it more accessible to non-expert users. We proposed a new VQA architecture that employs a fine-tuned GPT2 model for domain-specific knowledge representation, with the aim of enhancing both explicit and implicit reasoning in the context of plant disease question answering.

Keywords: VQA, Large language Models, Plant Disease Identification, Knowledge Representation

TABLE OF CONTENTS

Declaration	i
Dedication	ii
Acknowledgement.....	iii
Abstract	iv
Table of Contents	v
List of Figures	ix
List of Tables.....	x
List of Abbreviations.....	xi
List of Appendices	xii
Chapter 1	1
Introduction	1
1.1 Prolegomena.....	1
1.2 Objectives	2
1.3 Background Study and Motivation.....	4
1.4 Problem Definition	5
1.5 PlantDOC: VQA model for Plant Disease identification using LLM fine tuning for domain knowledge representation.....	6
1.6 Structure of the Thesis.....	6
1.7 Summary	7
Chapter 2	8
Background of AI Based Plant Disease Identification Systems and Visual Question Answering Models	8
2.1 Introduction	8
2.2 Plant Disease Identification using AI.....	8
2.3 Overview of Deep Learning based VQA Systems	9
2.3.1 Vanila VQA models.....	10
2.3.2 Transformer based VQA	11
2.3.3 Reasoning in VQA	12
2.4 Analysis of Literature	14

2.5	Problem Definition	16
2.6	Summary	17
Chapter 3		18
Technology Adopted		18
3.1	Introduction	18
3.2	VQA Model Architecture	18
3.2.1	Image Feature Representation	18
3.2.2	Language Feature Representation	19
3.2.3	Transformer Architecture	20
3.2.4	Large Language Models	22
3.2.5	Fine-tuning LLM	22
3.2.6	Transformer Model for Image Classification	23
3.3	Evaluation Methods	24
3.4	Summary	25
Chapter 4		26
LLM based Approach to Implement VQA system for Plant Diseases Identification		26
4.1	Introduction	26
4.2	Hypothesis	26
4.3	Input	27
4.4	Output	27
4.5	Process	27
4.6	Features	29
4.7	Users	29
4.8	Summary	29
Chapter 5		30
Design of PlantDOc VQA model		30
5.1	Introduction	30
5.2	Dataset Creation	30
5.2.1	Dataset for Knowledge Base Training	33
5.2.2	Plant Disease VQA dataset	33
5.3	Design VQA Architecture for PlantDOC	34
5.3.1	Image Feature Representation	34

5.3.2	Text Embedding	35
5.3.3	Knowledge Base Training.....	35
5.3.4	Knowledge Base Data Extraction	35
5.3.5	VQA Model.....	36
5.4	Model Training.....	36
5.5	Model Evaluation	36
5.6	Summary	36
Chapter 6.....		37
Implementation		37
6.1	Introduction	37
6.2	Development Frameworks	37
6.2.1	PyTorch DL Framework	37
6.2.2	HuggingFace NLP Framework	37
6.2.3	Anaconda Platform.....	38
6.3	Training Setup	38
6.4	Implementation.....	38
6.4.1	Fine-Tuning GPT2 for Plant Disease Knowledge	38
6.4.2	Training PlantDOC VQA.....	40
6.5	Model Parameters and Training	41
6.6	Summary	41
Chapter 7.....		42
Evaluation		42
7.1	Introduction	42
7.2	Results	42
7.2.1	Plant Disease Classification Model Results.....	42
7.2.2	PlantDOC_LLM Model Training Results.....	42
7.2.3	PlantDOC VQA Model Training Results	43
7.3	Summary	44
Chapter 8.....		45
Conclusion and future work		45
8.1	Introduction	45
8.2	Conclusion.....	45

8.3	Future Work	46
8.4	Summary	47
	References	48

LIST OF FIGURES

Figure	Description	Page
Figure 2.1	The bench marked VQA architecture introduced in [3]	11
Figure 3.1	Typical VQA architecture [54]	18
Figure 3.2	VGG and ResNet architecture comparison [22]	19
Figure 3.3	Transformer architecture [54]	22
Figure 3.4	Original ViT architecture diagram presented in [52]	23
Figure 4.1	Flow Diagram of the system	28
Figure 5.1	sample data element of PlantDOC VQA dataset	33
Figure 5.2	Proposed Deep Learning architecture	34
Figure 5.2	PlantDOC LLM prediction results	44

LIST OF TABLES

Table	Description	Page
Table 2.1	Example scenarios with reasoning in VQA	13
Table 2.2	Analysis of DL based plant disease classification	14
Table 2.3	Comparison of available methodologies for VQA systems	15
Table 5.1	Statistics of Plant Disease Image Dataset	29
Table 5.2	Sample Question data used for training	31
Table 7.1	Plant Disease Classification model results	42
Table 7.2	GPT2 finetune model results	44
Table 7.3	PlantDOC VQA model results	45

LIST OF ABBREVIATIONS

Abbreviation	Description
DL	Deep Learning
AI	Artificial Intelligence
NLP	Natural Language Processing
LLM	Large Language Model
VQA	Visual Question Answering
CV	Computer Vision
MLLM	Multimodal Large Language Model
GDP	Gross Domestic Product
ML	Machine Learning
RNN	Recurrent Neural Network
CNN	Convolutional Neural Network
NN	Neural Network
ViT	Vision Transformer
UAS	Unmanned Aerial Systems
VD	Visual Dialog
SOTA	State-of-the-art

LIST OF APPENDICES

Appendix	Description	Page
Appendix - A	Plant VQA base questions	41