

Applied Computer Science

**Thesis: Dataset Investigation & Exploratory
Research**

Reid C. S. Lowdon

February 20, 2021

Contents

1	Literature Review	3
1.1	Introduction	3
1.2	Problem Definition	3
1.3	Existing Research	3
1.4	Contributions	3
2	Foundations	4
2.1	Artificial Intelligence	4
2.2	Neural Network	4
2.3	Convolutional Neural Network	4
2.4	Deep Convolutional Neural Network	4
3	Existing Architectures	5
3.1	DenseNet	5
3.2	ResNet	5
3.3	GoogleNet	5
3.4	VGG16 & VGG19	5
4	Implementation	6
4.1	Dataset Acquisition	6
4.2	Data Preparation and Augmentation	6
4.3	Experiment Setup	6
5	Results and Analysis	7
5.1	Evaluation Metrics	7
5.2	Quantitative Results and Comparisons	7
5.3	Qualitative Results and Comparisons	7
5.4	Summary	7
6	Conclusions	8

1 Literature Review

1.1 Introduction

Probably do this section last as that's generally how introductions are written

1.2 Problem Definition

- Major research questions:
 - How much data is enough?
 - What is the tipping point for “dirty data”?
- AI is new and emerging → not enough research done with the datasets that are utilized
- General rule is “more data is better”
 - There must be a point where more data does not matter (i.e. does not significantly improve model accuracy)
- With pumping out more data comes increased risk for biases in the data
 - Especially when this data comes from one source

1.3 Existing Research

- Information regarding current/past research on the topics listed in “1.2 Problem Definition”
- Probably won't find a ton of research as this is a fairly new/unexplored area

1.4 Contributions

- Why is my work valuable?
- What do I add to the existing knowledge base?

2 Foundations

2.1 Artificial Intelligence

- Starting with this section may be too broad, talk to Chris about this

2.2 Neural Network

- Information regarding what a “neural network” is and how they function

2.3 Convolutional Neural Network

- Information regarding what a “convolutional neural network” is, how they function, and differ from a standard neural network

2.4 Deep Convolutional Neural Network

- Information regarding what a “deep convolutional neural network”, how they function, and how they differ from a standard convolutional neural network

3 Existing Architectures

- In this section I will talk about:
 - The architectures I investigated (pros and cons)
 - Why I chose the architecture that I chose (probably DenseNet)
- Information about each architecture in detail
 - How it's built (what it's layers are, how they function, general information about the architecture)

3.1 DenseNet

3.2 ResNet

3.3 GoogleNet

3.4 VGG16 & VGG19

4 Implementation

4.1 Dataset Acquisition

- Talk about the gantry system here and how we have obtained the dataset that I am working with

4.2 Data Preparation and Augmentation

- What did I do to prepare the dataset in order to reduce bias, ensure that the data is clean, transformations, everything I did to manipulate the data for my experiments

4.3 Experiment Setup

- Specifics on the hardware that was utilized
 - Ask Chris if I should include my own hardware here as I am initializing my experiments on my personal machine
 - Include details about the server that I use at school
- Also mention the learning rates, number of epochs, which optimization algorithm

5 Results and Analysis

5.1 Evaluation Metrics

- Talk about how accuracy was calculated (provide formula)

5.2 Quantitative Results and Comparisons

- Include tables showing the accuracy of each architecture I trained and used for inference

5.3 Qualitative Results and Comparisons

- Discuss the results and if possible have samples from each network as a visual comparison

5.4 Summary

6 Conclusions