Group 67 - Requirements Document

Deep Learning for Object Recognition on a Mobile Robot

Julian Weisbord, Michael Rodriguez, Miles McCall

Oregon State University

CS 461 Fall 2017

———————————— ◆ ————————————

Abstract

This paper goes in-depth into the requirements that we as a team must accomplish. We will be working to build an image classifier to work with our autonomous robot to help recognize basic objects in different environmental settings.

CONTENTS

1 INTRODUCTION

1.1 Purpose & Scope

This capstone project is in partnership with the Personal Robotics Lab (PRL) at Oregon State University (OSU). The overall goal of our research is to take a bare-bones robot and train it to recognize the objects in any given environment. We will do this by utilizing several different machine learning algorithms, in an effort to make the robot into an Intelligent Agent. Currently their autonomous robots lack the ability to efficiently understand, learn, and classify objects in their environment with a high degree of certainty. By leveraging the ROS (Robot Operating System) libraries and the existing mobile robot platform, we aim to sequentially train a Convolutional Neural Network (CNN) with different online learning techniques. This CNN will then be capable of using robots as a data-collecting network to make itself smarter over time.

This is a research oriented project, which means the specific requirements are based on how we will both find and implement these machine learning algorithms. As a starting point, we will research algorithms that have worked well for similar projects. Once we have a substantial knowledge of common algorithms, we will apply them to the robots in the PRL of OSU. Moreover, this document proposes a plan of action for how robots like the PR2 and Fetch can be used to classify everyday objects in a wide variety of settings.

1.2 Definitions, Acronyms, & Abbreviations

Machine Learning: This evolved from the study of pattern recognition and computational learning theory in artificial intelligence,[4] machine learning explores the study and construction of algorithms that can learn from and make predictions on data

Overfitting: Trained model only makes predictions for data set that you have trained on and not new data.

Back-propagation: Used to calculate the error contribution of each neuron after a batch of data is processed. Calculates gradient of loss function (gradient descent/ synthetic gradients). Back-propagation is a generalization of the Delta Rule.

Forward Propagation: We apply a set of weights to the input data and calculate an output. For the first forward propagation, the set of weights is selected randomly.

Neural Network Training: Forward Prop (start with random weights) sum products of the inputs with their corresponding set of weights to arrive at first values of hidden layer; apply activation function to hidden layers; sum product of hidden layer results with the second set of weights to determine output sum; take activation function at that output sum to get the final output result; Back Propagation [1]

- Output sum margin of error = target - calculated

Online Machine Learning (sequential learning): A method of machine learning in which data becomes available sequentially. The opposite of batch learning, Online learning updates its predictor for future data continuously with each new piece of data. Ex: If you were to show a child a ripe banana for the first time, they would learn what a banana looks like. But maybe next time they see a banana, it is actually green and not ripe. The child would still know it is a banana but now they would know that bananas can be different colors. This is Online learning. It is more common for neural networks to use batch learning. With batch learning, if we apply the same analogy to the robot, if it were to see a green banana, its definition would completely forget that bananas can also be yellow. So we need Online Learning so that our intelligent robot can keep learning new information while remembering old information.

Intelligent Agent: In artificial intelligence, an intelligent agent (IA) is an autonomous entity which observes through sensors and acts upon an environment using actuators (i.e. it is an agent) and directs its activity towards achieving goals.

Intelligent agents may also learn or use knowledge to achieve their goals. They may be very simple or very complex: a reflex machine such as a thermostat is considered an example of an intelligent agent. [2]

Catastrophic Interference: Catastrophic Interference becomes present when learning is sequential. Sequential training involves the network learning an input-output pattern until the error is reduced below a specific criterion, then training the network on another set of input-output patterns. A backpropagation network will forget information A if it learns input A and then input B. Catastrophic forgetting occurs because when many of the weights, where knowledge is stored, are changed, it is impossible for prior knowledge about past data to be kept intact. During sequential learning, the inputs become mixed with the new input being superimposed over top of the old input. To recognize multiple sets of patterns, the network must find a place in weight space that can represent both the new and the old output.One way to do this is by connecting a hidden unit to only a subset of the input units. [3] [4]

Elastic Weight Consolidation (EWC) Elastic Weight Consolidation is a technique used to prevent a neural network from forgetting the past data its learned while training on new results. To do this, the algorithm locks the weights generated by previous runs in a way where the values converge on low error rates for both old and new data. EWC uses a different loss function than a standard neural network architecture to accomplish this.

1.3 References

REFERENCES

[1] S. Miller, "Neural nets," https://stevenmiller888.github.io/mind-how-to-build-a-neural-network/, 08 2015, (Accessed on 10/25/2017).

[2] S. J. R. Peter Norvig, "Artificial intelligence a modern approach," Prentice Hall.

[3] A. Karpathy, "Image recognition lectures," http://cs231n.github.io/, (Accessed on 10/26/2017).

[4] F. P. Ron Kohavi, "Glossary of terms," http://robotics.stanford.edu/~ronnyk/glossary.html, 1998, glossary.

1.4 Overview

To build a robust classifier for the robots, we will pursue three different methods to improve on the current system at the Personal Robotics Lab of OSU. These are: improving the data capture process, overfitting to data in different environmental contexts using multiple classifiers, and testing different online learning models to achieve the best classification rate. Project Goals:

- We will analyze the benefits of at least 4 different lifelong learning techniques added to the pipeline such as: Dual Memory Architecture, Functional Approximation, etc. This will allow the intelligent agent to continue to learn more about its environment, and thus be able to classify objects at higher and higher rates as time goes on.

- We will research different ways to improve the current image capture system in place at the OSU Personal Robotics Lab. A possible bottleneck of image recognition can occur when an image dataset contains too many different objects in one image. The team can improve this with more accurate image data capturing.

- The intelligent agent will be able to classify at least 3 different object classes(i.e: Dog, Mug, table) with a minimum of 80% accuracy. The intelligent agent should also be able to identify the differences between objects within the same class(i.e his mug vs. my mug).

2 OVERALL DESCRIPTION

2.1 Product perspective

The product we aim to deliver consists of a code base to be ran on a PR2 and/or Fetch robot with the ROS operating system. As a component of this larger existing system, our code base will append to and in some cases, replace current models running on the robots.

2.2 Product functions

Our software package encapsulates a few key functionalities. It will augment the robot's data capturing system with modern image identification algorithms, image tagging, and image recognition software. Next, the robot will take the image from its sensors and query the neural network classifiers that we have built in order to accurately classify the image. Finally, once the image has been successfully classified, the robot will update its classifiers with this image so that it can use it as an example when trying to classify other images. This is known as Sequential Learning or Online Learning. All of these functionalities currently exist within the running system, but we will improve upon these functions with algorithms that can classify at a higher rate.

2.3 User characteristics

The intended user of our software package primarily consists of individuals already involved in the robotics lab. These users have a high degree of knowledge on the major topics covered within our project, and already have training on how to interact with the robot and integrated systems. Beyond the lab, however, users expand into a much broader range of individuals. Its entirely possible that or models could be deployed on a robot in house settings, labs, or workplace environments. The end goal user will have a large spectrum of knowledge regarding the robot and encompassing software and use the robot in a supporting role to increase their productivity.

2.4 Constraints

Our project is largely framed by the technologies provided to us. We are given, the Fetch and the PR2 mobile robots, servers with NVDIA GPU's, and lab access. We are very much constrained by the quality of the camera/sensors and the overall movement capabilities. In regards to training, testing, and designing our machine learning algorithms, we will be utilizing the robotics lab's server environments. Our local/personal machines will likely lack the performance needed to effectively train our models.

2.5 Assumptions and dependencies

Our overarching project depends entirely on the described and agreed upon robot hardware and operating system. The robots have different functionality and features which will affect the resulting software we create. The operating system, however, should remain more consistent across different platforms. Luckily we can assume the ROS platform handles all communications with the underlying hardware and robot accessories, so API's will make our code more portable and modular.

3 SPECIFIC REQUIREMENTS

- Researchers will research three different ways to improve the current image capture system in place at the OSU Personal Robotics Lab.

- Researchers will analyze the benefits of at least three different lifelong/online learning techniques (defined above) such as: Dual Memory Architecture, Functional Approximation, etc. They will use the most accurate Online Learning technique (one that remembers the most information) for the final classifiers.

- Researchers will implement two Convolutional Neural Networks to showcase the benefits of certain models for image classification.

- Researchers will evaluate the success of the different algorithms by how much the average rate of classification increases compared to the current classifier.

- The intelligent agent will classify at least 3 different object classes with a minimum average of 80% accuracy.

- Researchers will compile a comprehensive write up detailing the results of their findings.

4 GANTT CHART

Gantt Chart - Task Planning							
	Sep	Oct	Nov	Dec	Jan	Feb	March
Prep	████						
Research CNN Algorithms		▭					
Data Collection			███				
Improve Data Collection			▭				
Create Image Datasets			▭				
Code Implementation				███			
Implement CNNs & Classifiers				▭			
Implement Online Learning				▭▭			
Data Processing					██		
Classify Created Datasets with Models					▭		
Analysis						███	
Analyze Output Results						▭	
Technical Write Up						▭	
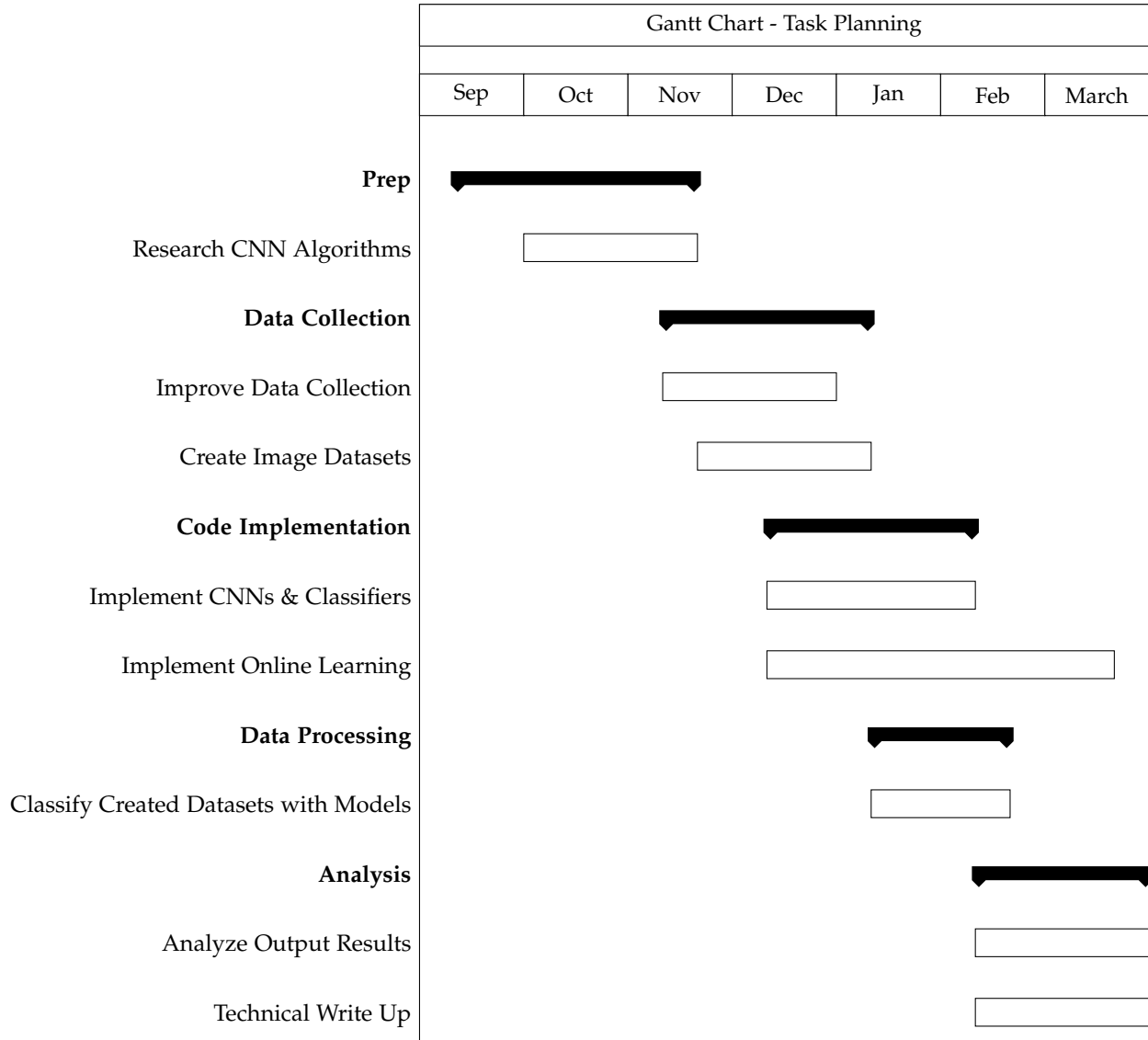
5 REVISION CHANGE LOG

- Definitions, Acronyms, &Abbreviations Section

    - Added: Elastic Weight Consolidation (EWC)

- Specific Requirements Section

    - Original

        * Researchers will research different ways to improve the current image capture system in place at the OSU Personal Robotics Lab.

            · They use a data capture system that utilizes AR tagging to show the robot where the image that it wants to capture is located. Researchers will research and implement several common methods to see which yields the most accurate classification rate in the end.
            · Researchers will document and report the benefits of each methodology in a technical writeup.

        * Researchers will analyze the benefits of at least 3 different lifelong/online learning techniques(defined above) added to the pipeline such as: Dual Memory Architecture, Functional Approximation, etc. Use the most accurate (remembers the most information) Online Learning technique in the final classifiers.
        * Researchers will implement at least 3 Convolutional Neural Networks that will be responsible for different aspects of image classification.
        * Researchers will evaluate the success of the different algorithms by how much the average rate of classification increases compared to the current working classifier.
        * The intelligent agent will classify at least 3 different object classes with a minimum average of 80% accuracy.

    - Revised

        * Researchers will analyze the benefits of at least three different lifelong/online learning techniques (defined above) such as: Dual Memory Architecture, Functional Approximation, etc. They will use the most accurate Online Learning technique (one that remembers the most information) for the final classifiers.
        * Researchers will implement two Convolutional Neural Networks to showcase the benefits of certain models for image classification.
        * Researchers will evaluate the success of the different algorithms by how much the average rate of classification increases compared to the current classifier.
        * The intelligent agent will classify at least 3 different object classes with a minimum average of 80% accuracy.