

Kavan Mehta

Mrs. Secord

Independent Study & Mentorship Program 1 (ISM 1)

April 6, 2022

Product Set-up and Completion Summary for Hand Gestures Recognition Project

Introduction and Statement of Purpose

For my final product, I planned to create a research project which has helped me gain an excellent understanding of deep learning and implement my knowledge of machine learning, computer vision, and natural language processing (NLP) to solve real-world problems. I learned and demonstrated practical implementations of machine learning and integrated machine learning models into software applications. For my Final Product, my aspiration was to learn and use different machine learning algorithms to expand and enhance my knowledge. My plan was to implement a hand gestures project that recognizes our hand gestures to perform basic functions such as opening applications and basic search for websites etc. To accomplish this, I conducted research on articles, books, and online courses to learn more about how I could implement these concepts. This solution has addressed many of our day to day situations as we often carry objects and need our device to execute certain functions hands-free. In this way, I tried to apply my research skills and knowledge of machine learning algorithms and programming to create a machine learning model that has utilized computer vision to help people in their daily life and save time. I gained practical knowledge on how to actually implement machine learning models on real-world problems.

Review of skills and Research

Research topics that were instrumental for my project include learning about different machine learning algorithms using deep learning and specifically, computer vision. Furthermore, I learned the practical applications of specific algorithms such as convolutional neural networks, and K-nearest neighbors classifiers for a simpler approach to the model. I also learned specific applications, methods, and functions in some essential libraries such as OpenCV, Keras, Matplotlib, Glob, OS, and Numpy that were really important to my project.

Since I learned and used python in my entire project, I also gained an extensive understanding of the python language and how it can be used to implement machine learning. I was also able to significantly improve my problem-solving skills while debugging different issues for developing my project. I also used my knowledge of different programming languages such as Swift, Javascript & Java and understanding of several machine learning algorithms to implement the model using programming skills.

Methodology

Materials:

This project was conducted digitally to create a machine learning model, its preprocessing, and a video live stream, the main materials for the project were following:

- Laptop.
- Digital Python Libraries, virtual environments, and Platforms such as TensorFlow, Anaconda, Jupyter Notebooks, VS Code, OpenCV, Keras, Matplotlib, Glob, OS, Re, and Numpy.
- Machine Learning e-books, video tutorials, article tutorials, and research articles.

Procedure:

1. I started my project by researching and gaining a deeper understanding of computer vision, deep learning concepts, different algorithms, libraries, and models such as VGG16, OpenCV library, and preprocessing concepts such as background subtraction and skin detection which have helped me create my final project ,“Hand Gesture Recognition Project.”
2. As a next step, I took a deep dive into the most critical elements of my project which are “Computer Vision” and “Convolutional Neural Network” algorithm and the theory behind these algorithms through my research by using scholarly articles, ebooks, and tutorials.
3. Then, I actually implemented a “Convolutional Neural Network” algorithm and different methods to create my final project that would refine and increase my practical knowledge/skills.
 - a. As a first step, I identified a clear problem statement for my final project, and designed an approach to solve the problem using a convolutional neural network with the guidance of my mentor.
 - b. Then, I prepared and collected my own data set required for implementing the computer vision algorithm for hand gestures by deciding 10 specific hand gestures such as Thumbs up, Victory Sign, Full Hand, Index finger, fist, “L” sign, “Rock” Sign, “Cool” Sign, “Gun” Sign and Only Fingers.
 - c. Then, I created the preprocessing routine in my program to resize each image to 64*64 pixel images and cleaned any erroneous files. I also created an entire

function to map my image labels to their correct images by using a program dictionary while reading the files in numerical order.

- d. I decided to use the convolutional neural network model for actually performing analysis on the images due to its efficient way for finding features that are important to predict the image.
- e. I trained and fine-tuned the model in an agile way to evaluate the model by changing the number of images in a certain batch and other categories like epochs, number of convolutions, and number of dense neural network layers.
- f. I fine tuned the machine learning model implementation to optimize the outcomes to get a strong accuracy.
- g. Finally, I launched this model to make predictions on live-stream video by analyzing each frame, predicting the hand gestures, and then performing specific functions such as opening up a variety of applications using hand gestures.

Utilization of Higher-Level Thinking Skills

During this project, I used multiple higher-level thinking skills such as problem-solving, designing, evaluating, analyzing, and synthesizing information. I used my analyzing skills to conduct successful research that helped me gain knowledge for the field of machine learning. I synthesized information from my research to create my final project by relating my knowledge from several topics such as algorithms, frameworks, libraries, and data management that I learned. After that, I designed the entire project with the guidance of my mentor and solved many programming issues that came up during the project and research. Whether it was from simply not finding online sources to learn about my topic or concerning the actual implementation of the project, I used my problem-solving skills in every area. Lastly, I evaluated

my model's performance and transformed my project into a usable form for the community to be implemented in the real world by creating a live video stream that can perform computer functions through different hand gestures.

Results

My results were very different from my initial predictions on the model's performance and its applications. I faced many challenges in this project while practically implementing the concepts. However, these challenges were extremely helpful in developing a strong understanding and foundation of machine learning and how we can apply machine learning models in day to day problems. Overall, when using 6200 training images and 2800 testing images from a dataset, I was able to achieve an accuracy of 87% on the test set. However, when using the same model on a live video stream or images of my hand with different hand gestures, its predictions weren't as accurate. Due to this, I tried using my own gesture dataset and essentially used Google's framework MediaPipe to use its pipelines for mapping 21 3D keypoints of the hand and predict the hand gesture based on its position and orientation. Furthermore, I found that I had much better recognition after using MediaPipe and I was even able to perform certain tasks such as opening up different applications based on certain hand gestures.

Conclusions

I was able to learn and get a good understanding about machine learning and many Python libraries such as OpenCV during this entire project. This project has incredibly increased my knowledge of machine learning, its practical applications, possible complications, and constraints in the implementation of these models. Due to the challenges I faced, I was able to learn the importance of having consistent, matching datasets for training and testing the data and

learned the core understanding of machine learning. I also learned about different tools such as MediaPipe and its functions and was really able to explore the practical implementation of machine learning. Although my project did not conclude the way I had originally imagined, I was able to create my hand gestures recognition project by changing my approach to create the hand gestures recognition model independent of generic machine learning dataset through Google's machine learning pipelines. However, I was able to train and learn about machine learning through my own machine learning model as I was able to create, train, and test a model based on the 8,000 images from the Kaggle Dataset.

Application/Meaning

My project was aimed to help the community by making computer tasks easier and hands-free for their devices which could help create better time efficiency, increase productivity, and improve some daily functions on devices by using hand gestures to accomplish tasks. The concept of an application/website platform that implements my machine learning model in a usable form could be extremely beneficial in the future and serve as a very helpful tool for people who use sign language, improve accessibility for all by automating processes hands-free, and can help others work more efficiently through advanced technology. Furthermore, if the model is expanded to recognize a vast variety of hand gestures accurately and can perform more functions, it could fundamentally change the way people interact with computers in the future.