Exploring and investigating hands-on AI projects: How image and pattern recognition, assisted by artificial intelligence and machine learning functions and impacts our world.

Ethan Yang

About A.I.

Artificial intelligence (AI) is the development of computer systems to be able to perform tasks emulating human intelligence. AI has developed increasing popularity for its accuracy, efficiency, and potential to foster exponential economic growth. It is changing the world with intelligent assistance like Alexa and Siri, self-driving cars, and email spam filters which have become an indispensable part of our daily lives. More advanced AI systems such as manufacturing robots, automated financial investing, and marketing chatbots continue to facilitate breakthroughs in our lives and provide efficiency, stability, and reliability to our world.

However, though machines and AI can be quick and precise, some of their products are considered controversial. In September 2022, a winning artwork made by AI at the Colorado State Fair's fine art competition sparked an AI-generated art backlash. John Allen created the piece "Théâtre D'Opéra Spatial" using the AI text-to-image program Midjourney and won a $300 cash prize[1]. Since then, criticism regarding how AI-generated arts are a form of plagiarism and hurt the feelings of human artists stormed social media, as well as parodies mocking AI's inability to comprehend simple objects in their creation. Poor implementation of AI as above could negatively impact an industry, even if unintended. If AI were to take over the art industry, creativity and uniqueness would start to cease in illustrations. Thus it is crucial for engineers to responsibly utilize AI effectively and rightfully, for security and privacy purposes as well.

For example, computer vision utilizes machine learning and neural networks to enable and train computer systems to recognize images and patterns. In past First Robotics Competitions, I have used limelight, a smart camera that deflects reflective tape, to automatically aim the robot at the goal and throw a perfect curve into the hoops. In a broader context, computer vision is also used in medicine (radiology, orthopedics, cardiology,

---

[1]"Controversy Erupts over Prize Awarded to Ai-Generated Art." Arts and Culture News | Al Jazeera, 5 Sept. 2022, www.aljazeera.com/news/2022/9/4/controversy-erupts-over-prize-awarded-to-ai-generated-piece.

ophthalmology) to aid computer inspection of X-Ray, CT, and MRI scans, monitor the progression of heart diseases, detect heart or eye development anomalies, etc. Of course, there are disruptive computer vision applications as well. It is known that companies are tracking facial features and eye movements of passersby for offline advertisements to research how people react to advertisements. Though it could improve and help personalize online video content, this computer vision application surely violates people's privacy[2].

In this project, I aim to investigate other usages of computer vision, such as handwriting and object recognition, through the use of computers and single-board computers, including Arduinos and the Raspberry Pi, while avoiding harmful applications.

Development of A.I.

Machine learning is required in creating AI. It is how computer systems develop their intelligence, by using algorithms, statistical models, and neural networks to analyze and draw inferences from patterns in data.

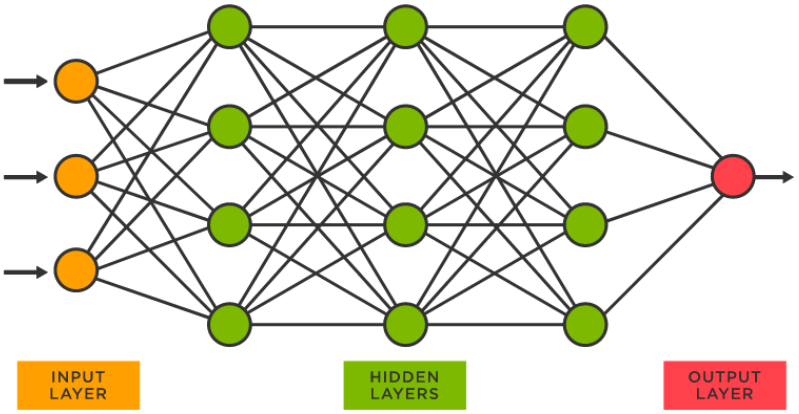
For many AIs to function, neural networks are used to develop machine learning. Consisting of an input layer, hidden layers, and an output layer, the neural network is crucial to practicing machine learning, as it is the collection of methods used in machine learning to model data using graphs of neurons. Neurons were inputted and sent to the nodes of hidden layers, where calculations and algorithms such as sigmoid, tanh, and ReLu functions operate, resulting in different weights and biases ready to be outputted.[3]

To operate a neural network, the user must set up by deciding the architecture of the neurons, the number of neurons in each layer, and the number of layers. In the process of training a neural network, known data are fed into the neural network with known outputs, seeding the network with known input and output information. The hidden layers would alter the input data by multiplying by the weights, adding the bias, and running through sigmoid/tanh/ReLu functions to keep the value of each neuron between 0 and 1. The values from the input layer would be passed down to the next layer and ultimately to the output layer, and the value of the output layer would relate to a specific result or identification.

---

[2]"Top 5 Disruptive Computer Vision Applications." Logikk, 27 Feb. 2019, www.logikk.com/articles/disruptive-computer-vision-applications/.

[3] 3Blue1Brown. "But What Is a Neural Network? | Chapter 1, Deep Learning." *YouTube*, YouTube, 5 Oct. 2017, https://www.youtube.com/watch?v=aircAruvnKk.

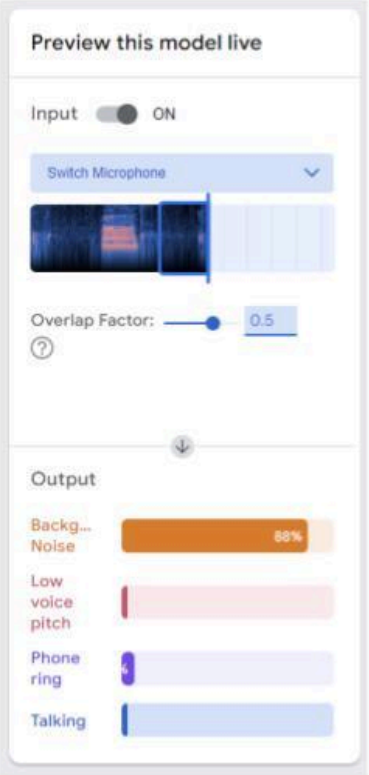


The concept of a neural network[4]

Let's use Imaginary Soundscape[5], a website that pairs appropriate audio with an inputted image, as an example. Its input layer receives neurons containing the data of the pixel of the image, then, through hidden layers and training, the network would recognize the content of the image and pair it with associated audio. Specifically, when a picture with greens and trees is inputted, the neural network would familiarize the image with the theme of forest and play audio of birds chirping and winds blowing leaves.

My Investigations

In my first investigation, I tried out audio recognition with Teachable Machine and object recognition with Tensorflow. Tensorflow is an open-source software library for machine learning and artificial intelligence used across a range of tasks but has a particular focus on the training and inference of deep neural networks. In Teachable Machine, I was able to create samples of background noises and three other types of audio and adjust the model's sample sizes, epochs, and the overlap factors to test artificial intelligence's ability to recognize different categories of sounds. Epochs indicate the least amount of training samples fed through the model. According to my model, the more epochs, the more accurate the model is, and vice versa. On the other hand, overlap factor describes the amount of mode overlap with the



An overview of Teachable Machine audio recognition

[4] https://www.tibco.com/reference-center/what-is-a-neural-network
[5] http://www.imaginarysoundscape.net/#/street

doped core/an aspect shared by two or more tests. The higher the overlap factor, the model would try to only output one kind of audio it recognizes, whereas with a lower overlap factor an inputted sound may output different types of audio with different percentages. Composing and training a neural network, as evident in the investigation, requires careful examination and modification of data.

Secondly, I experimented with basic TensorFlow object recognition. The result came out both expected and unexpected. It was anticipated that the computer would have a lower level of confidence in obstructed and rotated objects. However, particular items such as cell phones, tissue paper boxes, and human faces are unrecognizable. Though the lack of data on these particular items was expected, I came to question why it is difficult for computers to interpret them, especially human faces. The answer is, the quality of delivery, illuminations, poses, occlusions, expressions, and low resolution, for instance, affects the accuracy of facial recognition systems.[6] According to professor Emily Mower Provost from the University of Michigan, complex emotions are "notoriously difficult to interpret" due to mismatches between the speaker's expression and the observer's perception of the emotion[7]. In addition, despite technical issues, facial recognition systems also face ethical challenges such as identity fraud, racial discrimination in their algorithms, data privacy, and more[8]. After all, the need for cautious usage of artificial intelligence is also evident in facial recognition systems.

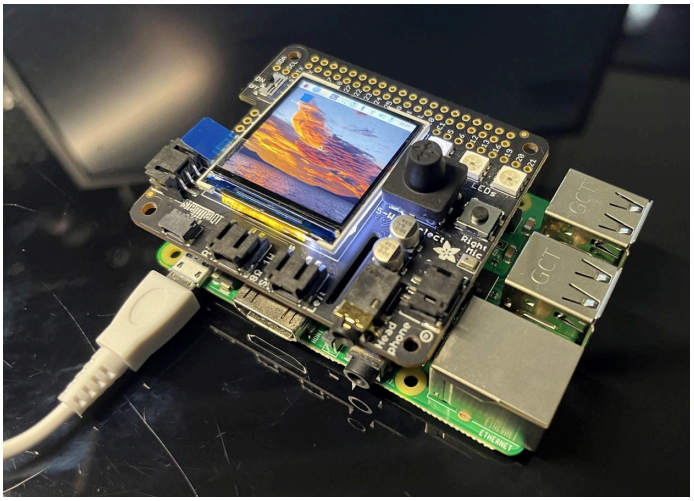
The Tensorflow Object Recognition: the model recognizes the Extra gum's pack as a packet with a 76.5% confidence rate

[6]"Challenges Faced by Facial Recognition System." PathPartnerTech, 20 Oct. 2021, www.pathpartnertech.com/challenges-faced-by-facial-recognition-system/.
[7]"Prof. Emily Mower Provost." Electrical Engineering and Computer Science, emp.engin.umich.edu/. Accessed 11 May 2023.
[8]Gangarapu, Katam Raju. "Ethics of Facial Recognition: Key Issues and Solutions." Learn Hub, learn.g2.com/ethics-of-facial-recognition. Accessed 11 May 2023.

I then became interested in handwriting text recognition since it was used as an example in a neural network introductory video I watched. I was tempted to recreate Harald Scheidl's simple HTR. However, I encountered various technically challenging errors, version incapability, for say. Under my mentor's recommendation, I decided to halt the project and move on for the moment. Realizing that not everything operates the way I expected couldn't be a more valuable lesson from this obstacle so my determination to continue research in other ways on AI continued.

My Projects

I purchased a Raspberry Pi 3 Model B+ and an Adafruit Braincraft HAT, a device built with microcontrollers and microcomputers designed for machine learning purposes.



An image of the Raspberry Pi and the Adafruit Braincraft HAT

The Adafruit Braincraft HAT has a 240×240 TFT IPS display for inference output, slots for camera connector cable for imaging projects, a 5-way joystick, a button for UI input, left and right microphones, stereo headphone out, stereo 1 W speaker out, three RGB DotStar LEDs, two 3 pin STEMMA connectors on PWM pins so they can drive NeoPixels or servos, and Grove/STEMMA/Qwiic I2C port.[9]

I initiated establishing a remote connection to my laptop through the secure shell (SSH) and configuring the Raspberry Pi.

The procedures are as follows:

1. Burning Raspbian Lite to an SD card using Raspberry Pi Imager with SSH enabled.

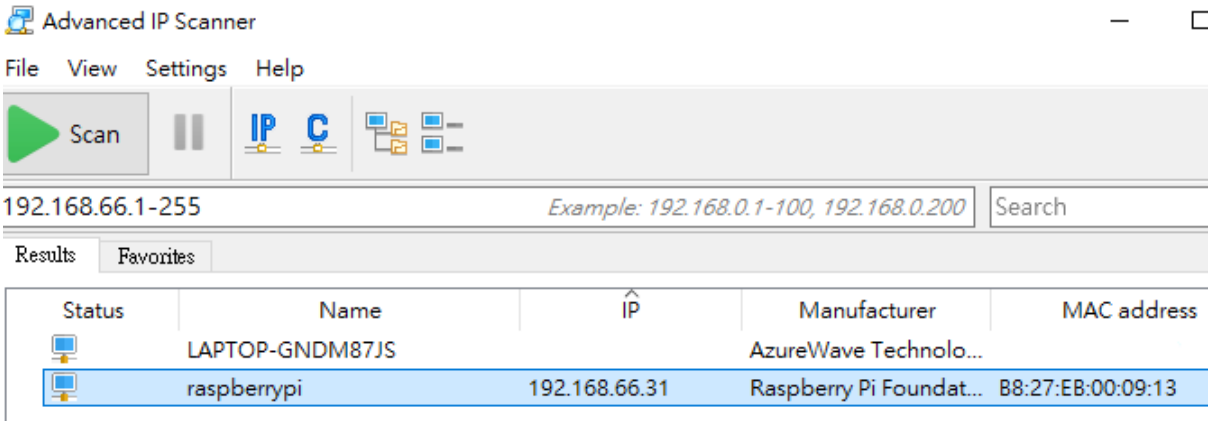
[9]Industries, Adafruit. "Adafruit Braincraft Hat - Machine Learning for Raspberry Pi 4." Adafruit Industries Blog RSS, www.adafruit.com/product/4374. Accessed 11 May 2023.

2. Create a “wpa_supplicant.conf” text file and input the wifi’s basic information, including SSID, password, country, and the network security key. Plugging in an Ethernet cable is also plausible.
3. Create a null file named “ssh” to enable SSH.
4. Search for Raspberry Pi’s IP address.

I struggled in searching for the Pi’s IP address. One tutorial introduced me to ping IP addresses with python-

```
ping -c 3 raspberrypi.local
```

However, I not only wasn’t able to distinguish which IP address belongs to the Raspberry Pi from other devices, but I also couldn’t connect or access any device from the listed IP addresses to my laptop’s terminal. This became extremely troublesome since whether or not the Pi was connected to the internet was ambiguous. Later I came upon another tutorial that recommends using Advanced IP Scanner to look for the Pi’s address. The scanner was more efficient as it shows the name and manufacturer of the corresponding device, making the Pi identifiable.



The Advanced IP Scanner showing the IP address of my Raspberry Pi

Though I was able to connect to the Pi through my laptop, subsequent attempts to connect failed as “Connection timed out”. I was frustrated until I came to my realization that I was simply having weak connections after seeing the wifi symbol on the Pi’s display. My laptop and Pi are located far away from my router, causing a slow and unstable connection. I resolved it by plugging the Raspberry Pi into a power bank so I was able to bring the Pi to the router for a stronger connection.

```
PS C:\Windows\system32> ssh pi@192.168.66.25
ssh: connect to host 192.168.66.25 port 22: Connection timed out
```

Error message for SSH connection failure

Later I installed Blinka, an API that brings CircuitPython libraries to single board computers, to the Raspberry Pi using pip, the package installer for Python to set up and test the display module, audio, buttons, fan, LED, and joysticks. Due to the tutorial's lack of explanation, I was creating python files and codes on the SD card instead using the Raspberry Pi terminal and the python library "nano". Thus the files were not found in the first place.

I looked for projects available on the Braincraft HAT and came across a tutorial on using Google Assistant on the HAT. It utilizes the Raspberry Pi's speaker to listen to the user's request after a button was pressed to initiate one. The LED lights on the HAT would turn from red to green to indicate that the assistant is recording an audio request. The Google Assistant would recognize the speech and respond accordingly using the Google Assistant API and through a connected headphone.

```
Press button to initiate a new request
INFO:root:Recording audio request.
INFO:root:Transcript of user request: "what".
INFO:root:Transcript of user request: "what is".
INFO:root:Transcript of user request: "what is the".
INFO:root:Transcript of user request: "what is the weather".
INFO:root:Transcript of user request: "what  is the weather".
INFO:root:Transcript of user request: "what is  the weather".
INFO:root:Transcript of user request: "what is the  weather".
INFO:root:End of audio request detected.
INFO:root:Stopping recording.
INFO:root:Transcript of user request: "what is the weather".
INFO:root:Playing assistant response.
INFO:root:Finished playing assistant response.
Press button to initiate a new request
INFO:root:Recording audio request.
INFO:root:Transcript of user request: "what".
INFO:root:Transcript of user request: "what do".
INFO:root:Transcript of user request: "what is".
INFO:root:Transcript of user request: "what is a".
INFO:root:Transcript of user request: "what is your".
INFO:root:Transcript of user request: "what  is your".
INFO:root:Transcript of user request: "what  is your favorite".
INFO:root:Transcript of user request: "what is  your favorite".
INFO:root:Transcript of user request: "what is  your favorite color".
INFO:root:Transcript of user request: "what is your  favorite color".
INFO:root:End of audio request detected.
INFO:root:Stopping recording.
INFO:root:Transcript of user request: "what is your favorite  color".
INFO:root:Transcript of user request: "what is your favorite color".
INFO:root:Expecting follow-on query from user.
INFO:root:Playing assistant response.
INFO:root:Finished playing assistant response.
```

Google Assistance listens to my requests, recognizing and answering them

Lastly, to be consistent with previous works, I decided to train a Teachable Machine object recognition model on the Braincraft HAT instead of on my laptop. Since I wasn't able to get my hands on the Raspberry Pi Camera Module, I used the Logitech C270 Webcam as an alternative, and thus I am required to use "fswebcam" instead of the picamera software.

Challenge: Unable to use OpenCV to capture images

Error: The Raspberry crashes and disconnects from SSH when the code compiles.

Solution: Give up on OpenCV and look for an alternative way to capture images, which is the "fswebcam" library.

Challenge: Saving images to a specified directory
Error: Error message "Error opening file for output"
Solution: adding "sudo" in front to enable access to write in specified directories.

In the limited amount of time, I was able to accomplish a total of 2 investigations with online resources and 2 projects on the Braincraft HAT and the Raspberry Pi. Given more time, I would merge the aforementioned investigations and projects into one smart mirror that has Google Assistant incorporated and is capable of recognizing objects and audio, and possibly even more functions!

Impact of A.I.

Challenges are often inevitable, but they're helpful for individual growth. Throughout my projects, I faced several obstacles in utilizing libraries, establishing the Raspberry Pi's SSH, etc. Yet, these technological challenges act as opportunities to build resilience capacity and skills in troubleshooting and debugging.

Furthermore, besides building skills throughout the way, the projects and investigations ultimately lead to answering one question: why is studying AI crucial? Nowadays, with every company looking to introduce AI to their domain, AI engineering is highly demanded from all fields. Due to the wide application of AI from pharmacies to stock markets, studying AI could guarantee a career or living as it is the "skill of the century". AI engineers are capable of analyzing large quantities of data, improving user experiences, and more to benefit society. Hence, investigating AI in the project could not only enhance my understanding but also prepare me for the unknown future. And in this unknown future, I am motivated to create a neural network from scratch as my ability and knowledge in AI matures.

Works Cited

3Blue1Brown. “But What Is a Neural Network? | Chapter 1, Deep Learning.” *YouTube*, YouTube, 5 Oct. 2017, https://www.youtube.com/watch?v=aircAruvnKk.

Allan, Alasdair. “A Magic Mirror with Added Tensorflow.” Medium, Medium, 25 Apr. 2018, https://aallan.medium.com/a-magic-mirror-with-added-tensorflow-b8fcc5528a6.

“Controversy Erupts over Prize Awarded to Ai-Generated Art.” Arts and Culture News | Al Jazeera, 5 Sept. 2022, www.aljazeera.com/news/2022/9/4/controversy-erupts-over-prize-awarded-to-ai-generated-piece.

“Challenges Faced by Facial Recognition System.” PathPartnerTech, 20 Oct. 2021, www.pathpartnertech.com/challenges-faced-by-facial-recognition-system/.

CrashCourse. “Neural Networks and Deep Learning: Crash Course AI #3.” YouTube, YouTube, 23 Aug. 2019, https://www.youtube.com/watch?v=oV3ZY6tJiA0.

Gangarapu, Katam Raju. “Ethics of Facial Recognition: Key Issues and Solutions.” Learn Hub, learn.g2.com/ethics-of-facial-recognition. Accessed 11 May 2023.

Githubharald. “Githubharald/Simplehtr: Handwritten Text Recognition (HTR) System Implemented with Tensorflow.” GitHub, https://github.com/githubharald/SimpleHTR.

Googlecreativelab. “Teachablemachine-Community/Gettingstarted.md at Master · Googlecreativelab/Teachablemachine-Community.” GitHub, 8 June 2021, https://github.com/googlecreativelab/teachablemachine-community/blob/master/snippets/markdown/tiny_image/GettingStarted.md.

“Guides for Product: Adafruit Braincraft Hat - Machine Learning for Raspberry Pi 4.” Adafruit Learning System, https://learn.adafruit.com/products/4374/guides.

“Hand Gesture Recognition Sensor (PAJ7620).” Arduino Project Hub, https://create.arduino.cc/projecthub/SurtrTech/hand-gesture-recognition-sensor-paj7620-9be62f.

Industries, Adafruit. “Adafruit Braincraft Hat - Machine Learning for Raspberry Pi 4.” Adafruit Industries Blog RSS, www.adafruit.com/product/4374. Accessed 11 May 2023.

LearnCode.academy. “Neural Networks Explained - Machine Learning Tutorial for Beginners.” YouTube, YouTube, 11 June 2018, https://www.youtube.com/watch?v=GvQwE2OhL8I.

Levine, Zach. “Build Your Own Google Home-Enabled Smart Mirror in about Two Hours.” Howchoo, 4 May 2021, https://howchoo.com/pi/build-your-own-google-home-enabled-smart-mirror.

“Prof. Emily Mower Provost.” Electrical Engineering and Computer Science, emp.engin.umich.edu/. Accessed 11 May 2023.

Reusch, Andrew. “Basic Tensorflow Object Recognition on Any Computer or IOS Device with Google Colab.” Adafruit Learning System, https://learn.adafruit.com/basic-tensorflow-object-recognition-in-the-cloud-google-colab.

“Top 5 Disruptive Computer Vision Applications.” Logikk, 27 Feb. 2019, www.logikk.com/articles/disruptive-computer-vision-applications/.