



Actua's AI Activities Series

Activity 4

Sort Things Out: Exploring Image Classification

Sort Things Out: Exploring Image Classification

If you're accessing this activity directly, did you know there are eight other activities in this series up on our website? If you find yourself unfamiliar with any of the AI concepts and terminology introduced in these activities, please refer to our [AI Glossary](#). These activities also follow a space exploration narrative when done in order. It is recommended to complete the activities in order but they can also be done on their own.

You and your group-mates are astronauts and scientists aboard the Actua Orbital Station. Unfortunately, your station just got bombarded by magnetic rays and your electronics have begun to shut down! The only one who can save you is the station's AI, DANN. DANN stands for Dedicated Actua Neural Network, and its gone a little loopy. Brush up on your technical skills, learn about AI, and save yourself and your crewmates! So far, we've restored DANN's basic thinking abilities and its mathematical skills. Now we need to make sure we can communicate with it!

DANN has finished its diagnostic, and we have finished our study from "[Regression Analysis: Making Predictions using Data](#)" Now, we can begin to fix DANN! When it was functioning, DANN was a state-of-the-art AI system that you could interact with using your voice. Because of the damage, however, DANN's audio core was knocked offline and, unfortunately, you don't seem to be able to access the audio core without DANN's help. Our mission specialists think that you might be able to use DANN's visual core to communicate, but the visual core isn't currently set up to recognize our requests. They suggest training an image classification model to recognize our poses or hand shapes so that we can get access to the audio core. Once we do that, we can bring DANN's other senses back online in "[Hand Commands: Training Image Classification Models](#)"!



Activity Summary

In this activity, participants will experiment with machine vision's application to classification tasks. They will learn to identify classification schemes and classes. Participants can then explore and compare two pre-trained machine vision models: COCO-SSD and MobileNet. Participants will evaluate these models and test their suitability for a defined task. This activity will lay the foundation for participants to train their own image classification model in a following activity.

Developed by Actua, 2022.

Delivery Environment	Activity Duration	Intended Audience
Classroom with computer access	45 minutes	Grades 9-12 (Ages 13-18)

Achievement Goals

Learning Goals

Learning goals are statements referring to the understanding, knowledge, skills or application participants acquire during the activity. **Following this activity, participants will:**

- Identify and describe classification, one of the major applications of artificial intelligence.
- Interpret the output information of a machine learning program (label, confidence) in general terms (e.g. very confident, confident, not very confident).
- Assess the suitability of a trained artificial intelligence model for a specific task.



Logistics (Timing, Group Size, Materials)

Section Title	Time	Group Size	Materials
Opening Hook: What is classification?	10 minutes	<i>Whole Group</i>	Whole Group <ul style="list-style-type: none"> • Pen/pencil • Paper
Section 1: Playing with Machine Vision	15 minutes	<i>Individual; Small Groups</i>	Small Group <ul style="list-style-type: none"> • Computer with functioning webcam • Pen/pencil • Paper • Small objects (stapler, cup, silverware, etc)
Section 2: Can you Break It?	15 minutes	<i>Individual; Small Groups</i>	Small Group <ul style="list-style-type: none"> • Computer with functioning webcam • Pen/pencil • Paper • Small objects (stapler, cup, silverware, etc)
Reflection & Debrief	5 minutes	<i>Whole Group</i>	



Safety Considerations

Safety considerations have been provided below to support safety during this activity, however they are not necessarily comprehensive. It is important that you review the activity and your delivery environment to determine any additional safety considerations that you should be implementing for the delivery of these activities.

Online Safety

Some components of this activity require the use of devices connected to the internet.

- Facilitators should review the provided videos and read/explore provided websites and materials to determine if they are suitable for their participants.
- Where applicable, facilitators should remind participants to stay on task and only use links provided within this activity.

Activity Procedure

To Do in Advance

SECTION 1 & 2

- For an in-person group, gather small objects to use for the visual AI.
 - Good objects for this purpose are things like a stapler, scissors, a cup, silverware, or writing utensils.)

Opening Hook: What is image classification?

What is classification?

Classification is the process of separating data into distinct categories, also known as classes, based on certain features of the data. For example, you could classify your friends or family members by eye color (e.g. brown, blue, hazel), hair colour (e.g. blonde, brown, black, red), or height (e.g. tall, average, short). You can also classify objects! Cars can be classified by type (e.g. compact, sedan, SUV), colour, or manufacturer (e.g. Honda, Ford, Toyota).

{Individually / In small groups / As a large group}, come up with 1-3 groups of objects that can be classified, and what features you could use to classify them.



Think about broad groups of objects that can have lots of differences, like cars or books!

Facilitator's note: Many possible answers. For any suggested classification scheme, ask for examples of classes that might make it up if they are not volunteered.

Possible example: foods for different meals of the day (e.g. breakfast foods, such as cereal, pancakes, or eggs; lunch and dinner foods, such as pizza or pasta; or snacks, such as chips or candy).

Having come up with examples, consider these two questions: What's the point of classifying things (e.g., objects, items, data)? What does classification do?

- Classification is a way of understanding, sorting, and simplifying data. Computers in general are very good at tasks that give them clear, well-defined data. Through classification, data can be converted into a format that computers can more easily work with.
- Some classes can also be thought of as “shorthand” for a collection of features. For example, classes of pants might include sweatpants, jeans, and pajamas. Each of these classes might tell us something about the features of the pants: leg length, whether or not they're made of denim, or whether they have a zipper. Sweatpants will likely have full-length legs, not be made of denim, and have no zipper, so we can classify them!

Image classification is the task of determining what's shown in an image (e.g. photo, video frame). This might mean identifying what an object is, or what action is taking place, or who specifically is in the image. This is the kind of classification task that we will be exploring in this activity, since Mission Control hopes that we will be able to use it to develop a solution to move DANN's repairs forward.

To get us started, specialists at Mission Control have set up two AI models (Model α and Model β , below) for us to explore and test. These models are pre-trained, which means that they should be ready for us to use. They've included some instructions to help with our testing and exploration, which starts with Model α .



Section 1: Playing With Machine Learning

1. Open [Model \$\alpha\$ - COCO-SSD AI Program](#) to load the first AI model and begin detecting.
 - a. You may be asked for permission to use your webcam. No data from your webcam will be stored—the whole process happens on your computer.
 - b. After a few moments, the model window should say that it's currently "detecting".
2. You should see a live, but somewhat choppy, feed from your webcam, but you may also see a pink rectangle around parts of the image, with other information there as well. In small groups, jot down your answers to the following questions about the webcam feed:
 - a. What is happening in the video window? Make 2-3 observations that describe what you are seeing in the video window (e.g. "I see a rectangle that moves when I move.").
 - b. What do you think the shapes, words, and numbers that appear on camera mean? What is their purpose?
 - i. The rectangle is a frame that indicates the program has recognized an object. The word is what the program thinks the object is, and the number is how confident the program is that it's correct (from 0.00 to 1.00).

The notes from Mission Control refer to Model α as "Common Objects in Context, Single Shot Multibox Detector", or COCO-SSD for short. They include a few extra details about the name:

- COCO is the name of the collection of images (the dataset) that was used to train this model to recognize different types of objects.
- SSD refers to what the model has been trained to do, which is detect (i.e. "are they there at all?") then classify (i.e. "what are they?") multiple objects in an image or video.



With that in mind, you will need to conduct a short test of the detection and classification capabilities of this model to see if it could be used for your purposes. To do so:

1. Gather a few handheld objects from the space around you. These can be things like writing utensils, paper, silverware, a phone, or anything like that.
2. Create a data table to record your observations. The table needs to have one row for each of the objects that you gathered, so if you gathered 5 objects, there should be at least 5 rows. A data table template has been included to help you out but you are free to create your own.

Object	Recognized?	Model output (words and numbers)	Other observations (e.g. orientation, distance)
<i>[Add a row for each object]</i>			

3. Show each object to your webcam.
4. What happens? Does the model react in any way?
5. Does moving the object change the result? Try changing how close or far the object is from the camera and changing the orientation (e.g. whether it's vertical or horizontal).
6. A rectangle will be drawn around each object that is detected by the model. Within each rectangle, there should be a word or a few words followed by a number. Record these words and numbers.
7. If you show the AI an object and it is not recognized (i.e. no rectangle, no words, and no numbers), write "Not recognized" or "Not detected" or "No", in the "Recognized?" column, if you're using the data table template.
8. Record any other information you wish to observe.



9. COCO-SSD looks for potential objects in a given image or video frame and draws a rectangle, called a bounding box, around them. It also attempts to classify each object that it thinks it found: the words you see within the bounding box are the model's guess at what the object is.
 - a. The number next to the words is how confident the model is in identifying the object. The model reports its confidence as a number from 0 to 1, where 0 is 0% confident and 1 is 100% confident (Meaning that it is 100% sure that it knows what the object is). Most models will only report guesses above a certain confidence threshold (for example, 0.5 or 50%).
 - b. According to Mission Control's notes, the overlay on the video also has a threshold of 50%, and won't display anything below that threshold, since it would likely be wrong at that point.

Section 2: Can You Break It?

The second model that Mission Control provided, Model β , is a simpler machine vision model. Instead of searching for separate objects within an image, Model β tries to classify the image or video frame as a whole. Similar to COCO-SSD, Model β reports how confident it is in its classification. The instructions for testing Model β are similar to those for COCO-SSD:

Technical example: [Model \$\beta\$ - MobileNet AI Program](#)

1. Gather a few objects from the space around you, similar to the ones you used for the last activity. If you can find a few new objects, that's great!
2. Create a data table to record your observations. The table needs to have one row for each of the objects that you gathered. If you have 5 objects to test with, there should be at least 5 rows in your table. You can use the table structure from the last activity for this.



3. Show each object to your webcam. What happens? Does the model react in any way?
 - a. Does moving the object change anything? Try changing how close or far the object is from the camera and changing the orientation (e.g. whether it's vertical or horizontal).
 - b. The model's guess for classification (i.e. the label for the image) and its confidence in its guess are reported at the bottom of the video window. Record the model's guesses (classification and confidence).
 - c. If you show the AI an object and it is not recognized (i.e. no rectangle, no words, and no numbers), write "Not recognized" or "Not detected".
4. Record any other information you wish to observe.

Mission Control refers to Model β as "MobileNet" and notes that this model was trained using the ImageNet dataset. Imagenet is a popular dataset that contains millions of labelled images of all sorts of things, meaning our second model here is very extensively trained.

Reflection & Debrief

With your results and observations from the quick tests of both AI models in hand, { in small groups / as a large group } discuss the following questions:

1. Did the AI models successfully recognize all of the objects that you showed them? How many of the objects did they successfully recognize?
2. Why can the AI models recognize some objects, but not others?
3. Did either of the AI models confidently (any confidence over 70%) label some objects that they were wrong about?
4. Can you think of a reason why either of the AI models might have thought that the object that you showed it was something else?
5. Do you think that either of these AI models could be used without additional training or other modifications to accomplish a hand shape classification task? How might either the model or the task be changed to suit the other?



6. How can visual recognition AI programs like these be used in society? What would a bigger version of this program look like, and how could it help people?

Many answers are possible here, though they should be backed up by evidence and observations from the tables.

As you look through your notes from Mission Control to see what the next step is, you come across a section with a heading in all caps. It has one word that's been double-underlined. It reads: TRAINING.

In the next activity, we will learn about model training and how to train our own image classifier using Google's Teachable Machine platform. We can take the model we produce with Teachable Machine and upload it to DANN's visual core so that we can continue with repairs.

Delivery Adaptations

How might you adapt the time, space, materials, group sizes, or instructions to make this activity more approachable or more challenging? **Modifications** are ways to make the activity more accessible, **extensions** are ways to make the activity last longer or more challenging.

Modifications

- In a virtual environment, ensure each participant has a selection of objects to use for the classification activities.
- If participants do not have access to writing supplies, create a shared Google Doc with observation tables to fill out.

Extensions

- Begin finding larger or more complicated objects that the AI models might not recognize, and see what they classify those objects as. Is there a pattern in the model's incorrect answers? Do either of the models have "default" answers that show up more often?



- Begin testing the model with multiple objects at once. Can it recognize them separately? Does it lump them together into another object? How does this affect the accuracy of the model?

References & Gratitude

The Imagenet dataset helped make this module possible. It can be found at <http://www.image-net.org/>.

P5.js is a javascript library that provided support for the code found within this activity. It can be found at <https://p5js.org/>.



Terms of Use

Prior to using this activity or parts thereof, you agree and understand that:

- It is your responsibility to review all aspects of this activity and ensure safety measures are in place for the protection of all involved parties.
- Any safety precautions contained in the “Safety Considerations” section of this write-up are not intended as a complete list or to replace your own safety review process.
- Actua shall not be responsible or liable for any damage that may occur due to your use of this content.
- This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. For more information, please see <https://creativecommons.org/licenses/by-nc-sa/4.0/>.
- You may adapt the content for your program (remix, transform, and build upon the material), providing appropriate credit to Actua and indicating if changes were made. No sharing of content with third parties without written permission from Actua.

About Actua

Actua is Canada’s leading science, technology, engineering and mathematics (STEM) youth outreach network, representing a growing network of over 40 universities and colleges across the country. Each year 350,000 young Canadians in over 500 communities nationwide are inspired through hands-on educational workshops, camps and community outreach initiatives. Actua focuses on the engagement of underrepresented youth through specialized programs for Indigenous youth, girls and young women, at-risk youth and youth living in Northern and remote communities. For more information, please visit us online at www.actua.ca and on social media: [Twitter](#), [Facebook](#), [Instagram](#) and [YouTube](#)!



Appendices

Appendix A: Career & Mentor Connections

- Machine learning researcher
- Computer Programmer
- Data Analyst
- Statistician

Appendix B: Extension Questions

Possible responses and discussion points for debrief questions

1. Did the AI models successfully recognize all of the objects that you showed them? How many of the objects did they successfully recognize?
 - *The likely answer is no, not all objects would have been recognized. Both AIs should only recognize a few of the objects presented to them unless they're shown objects of the same type that they recognize.*
2. Why might the AI models be able to recognize some objects and not others?
 - *What each model can recognize is based on the training data that each model has learned from. These models can't recognize objects that they haven't seen before and even then, may only recognize objects if they're presented in a certain way (again based on the data they've been trained with).*
 - *AI does not think of objects the same way that we each do because it doesn't have the benefit of being able to draw on our senses and experiences to help identify objects. These AI models, for example, look for specific patterns in an image to be able to guess what the object might be.*
3. Did either of the AI models incorrectly but confidently (say, confidence >0.7) label some objects?
 - *This is a possibility, but if it didn't happen, you could ask this as a hypothetical: "Why might an AI model incorrectly, but confidently, label some objects?"*



4. Can you think of a reason why either of the AI models might have thought that the object that you showed it was something else?
- *There isn't a definitive answer to this question, but a good way to approach it is thinking about what connections may exist between the actual object and the AI guess. Since these models only work on images, focus on whether the objects could have similar forms, sizes, colors, or other visible features in common.*
5. Do you think that either of these AI models could be used **without additional training or other modifications** to accomplish a hand shape classification task? How might either the model or the task be changed to suit the other?
- *The bold and underlined phrase is key here: No, without additional training or other modifications, these models wouldn't be able to accomplish a hand shape classification task: arbitrary (made up) hand shapes are likely not part of either of the datasets that they've been trained on and these models can only recognize something they've seen before in their training data.*
 - *For the second part of this:*
 - The models could be trained on the hand shapes that you want to use so that they can recognize them.*
 - If you didn't want to additionally train the models, you could change the task to showing the models objects that you know they can recognize instead of using self-defined hand shapes.*

Appendix C: Supporting Resources

SECTION 1

- [Model \$\alpha\$ - COCO-SSD AI Program](#)

SECTION 2

- [Model \$\beta\$ - MobileNet AI Program](#)

