Exploring Object Classification with Artificial Intelligence

Weijie Cai

University of Sussex, Brighton, BN1 9RH, United Kingdom

Abstract. Intelligence has significantly advanced in recent years and has become a convenient technique that has influenced our daily lives like artifical intelligence applying in medicine making it a popular research topic. My project is focused on object classification, which is a computer-based technique of artificial intelligence that enables the location and identification of objects. This technique involves using advanced algorithms and machine learning models that can process large amounts of data to accurately recognize and classify objects based on their visual features. Object classification has a wide range of applications in various industries such as manufacturing, retail, and security. It can be used to automate processes, improve efficiency, and enhance safety.

Keywords: Artificial Intelligence, Object Classification, Medicine.

1. Introduction

The purpose of this paper is to present the project on object detection using artificial intelligence. Object detection is primarily based on machine learning, and it requires an enormous amount of data for training and testing to achieve high accuracy in recognizing objects. In this paper, we have developed and implemented a method for detecting objects that involves deep learning for object classification. However, object classification is dependent on computer vision, and we require tools to help us analyze images. Therefore, Edge Impulse is an excellent platform for assisting us in using artificial intelligence, where we can upload data and sources.

Once the data is uploaded successfully, Edge Impulse can train it to help artificial intelligence become familiar with the objects. After training, we can upload additional data to test our project. Once the testing and training are complete, Edge Impulse generates the features of the data. With enough experience, Edge Impulse can classify objects, and we can upload pictures for artificial intelligence to classify the objects.

Our experiments have shown that artificial intelligence needs an enormous amount of data to train and test. Therefore, to maintain a high classification precision rate, we need to train on even more data to assist artificial intelligence classify objects more easily. Experiment has retained recall and f1 score in a high accuracy.

2. Methodology

To classify objects, I chose to use Edge Impulse, an artificial intelligence platform that can recognize sounds, images, and objects. My project involves classifying Bluetooth audio devices and cups using Edge Impulse. To start the project, I collected 50 images of Bluetooth audio devices and cups respectively, which provides enough data for Edge Impulse to train. I divided the images into training and testing data with rate 6:4. For training data, I used 30 images of cups and Bluetooth audio devices. The more images in the training data, the better performance the artificial intelligence can recognize the objects. The remaining 40 images were used as testing data, with an equal number of cup and Bluetooth audio device images. This balanced distribution of objects in the testing data is conducive to the artificial intelligence classify the data more accurately.

Here are some examples of the training and testing images, as shown in Figure 1 and Figure 2 respectively.

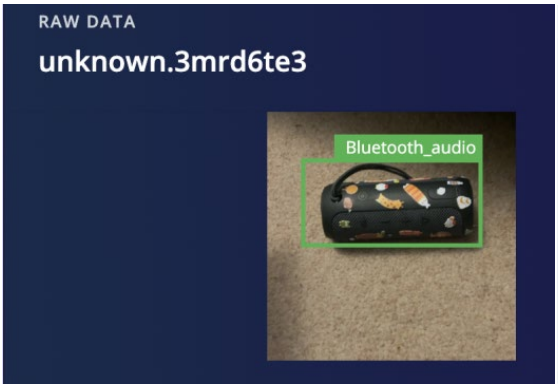

Figure 1 the example of speaker 1

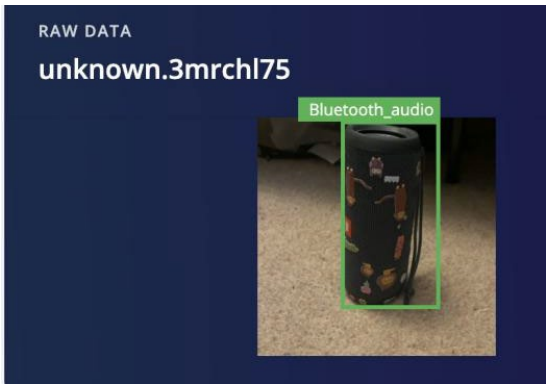

Figure 2 the example of speaker 2

Here are some examples of the training and testing images of figure3 and figure4:

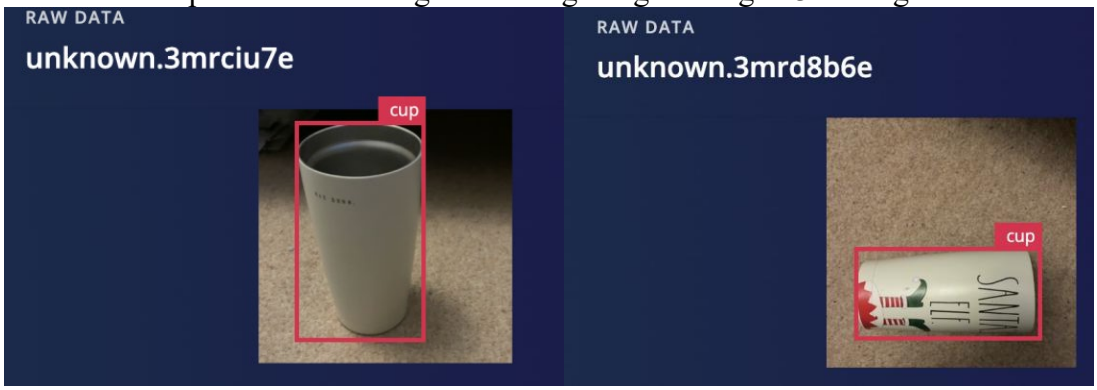

Figure 3 the example of cup 1

Figure 4 the example of cup 2

After uploading the photos, I labeled them into two categories, "Bluetooth_audio" and "cup," which is displayed in the pictures. Labeling is important for object classification, as the artificial intelligence needs to be trained on the labeled data to recognize the objects. Object classification differs from image recognition, which only detects one thing based on the image name. Object classification detects several objects in a single picture. Once the data was labeled, I started the training process. The training process involves using Edge Impulse to design features for the pictures. To create features, I added a processing block for the images and a learning block for object detection. Before generating features, I adjust the image parameters to set the color depth to RGB. Then, I perform generating features. This step aims to simplify and speed up data transformations while enhancing model accuracy, and can produce new features for both supervised and unsupervised learning.

After generating features, I obtained a features picture figure 5:

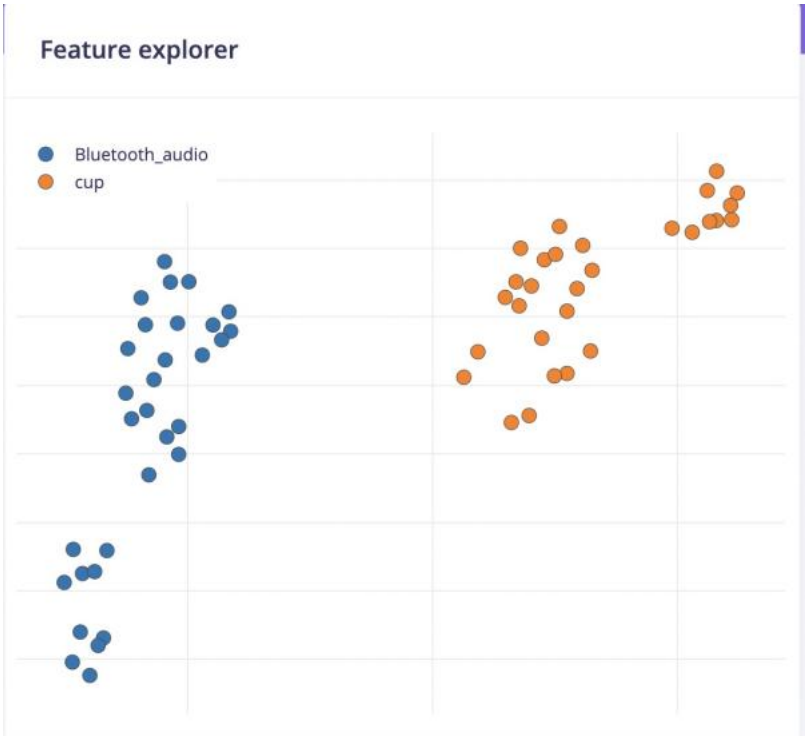

Figure 5 the example of feature

The generated figure is shown in Figure 1 according to DSP result[1]. The final step was to decide on the data for the training data set, and then I could start training the model to generate features. Overall, the process of object classification involves labeling the data, designing features, generating features, and training the model. Each step is critical for the model to accurately recognize and classify objects.

In this picture, the features are split into separate parts, with Bluetooth audio represented by blue and the cup represented by orange. This division generates two parts, each with its own color-coded features.

Before object detection, it's important to establish training settings such as the number of training cycles, learning rate, validation set size, and data augmentation. These settings create a training standard to ensure accurate detection. Object detection also requires a layer dedicated to image detection, making it a complex model.

Object detection can calculate a confusion matrix[2] to precisely evaluate its performance. Additionally, the F1 score[3] accuracy is a reliable metric, and in this case, it is 95.7%. The confusion matrix graph illustrates the results:

Table 1 Confusion matrix(1)

	Background	BlueTooth_Audio	CUP
Background	100%	0%	0%
Bluetooth_Audio	0%	100%	0%
Cup	14.3%	0%	85.7%
F1 Score	1.00	1.00	0.92

In the Confusion Matrix results, the tag for Bluetooth audio background had 0% detection, whereas cups had 14.3% detection in the background, with only 85.7% detection for cups themselves. This indicates that cups are more difficult to detect than Bluetooth audio. Therefore, we need to gather more pictures of cups to train the artificial intelligence model, while the data of Bluetooth audio can be classified with 100% accuracy. Moreover, I also took pictures of different versions of cups.

However, the model may not detect cups directly as the pictures contain a proportion of background elements that are higher than that of the cups. Therefore, cups were still classified with 14.3% background elements in the results. The F1 score for Bluetooth audio and cups were 1.0 and 0.92, respectively, with the F1 score for cups being lower than that for Bluetooth audio. In my opinion, the reason for the higher F1 score of Bluetooth audio is because the results showed 0% detection for its background element, while cups had a proportion of 14.3% background elements.

After training all elements and generating features for the training data, I proceeded to test my model. The testing model achieved an accuracy of 80%, meaning that 80% of the testing data can be correctly detected by the artificial intelligence. However, the results showed some errors in detecting Bluetooth audio compared to cups. Bluetooth audio could not be detected as well as cups.

The picture figure 6 displays the results:

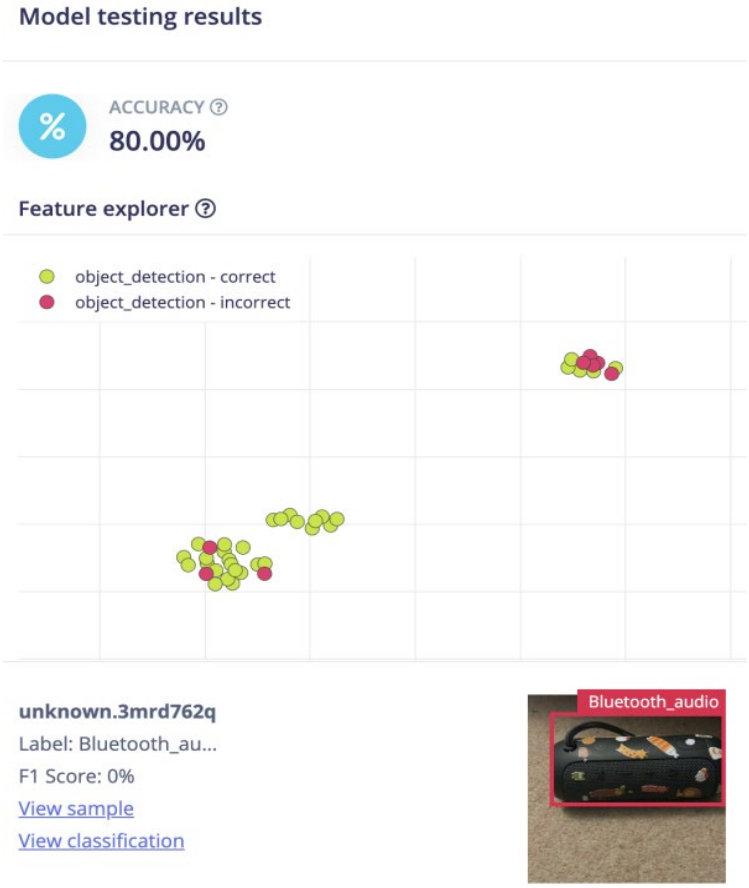

Figure 6 the example of accuracy 1

From my perspective, Bluetooth audio contains less background noise compared to a cup. Additionally, the image of Bluetooth audio has not undergone various versions, resulting in more errors than with cups. One reason for the inability to test a model is the insufficient quantity of data. Therefore, it is essential to increase the amount of data. Apart from this reason, errors can also occur due to unclear images.

As figure 7 display that:

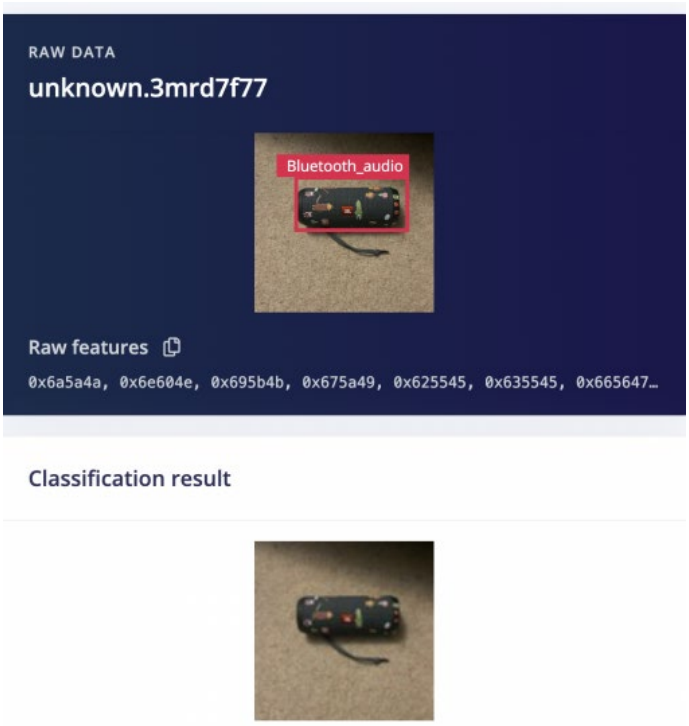

Figure 7 the example of classification 1

There have a picture of the testing model figure 8:

Figure 8 the example of testing 1

As shown in the image, there are still errors in object detection. To improve this, I plan to enhanced the quantity of data for training model.

Towards the end of the project, I constract a online classification feature which can recognize the similarity between pictures and identify the objects they depict. To test this feature, I uploaded images of cups and Bluetooth audio devices, and the model automatically detected these objects and assigned them appropriate tags. Online classification is a valuable tool for refining the model accuracy.

Here are figure 9 and figure 10 of the online classification feature:

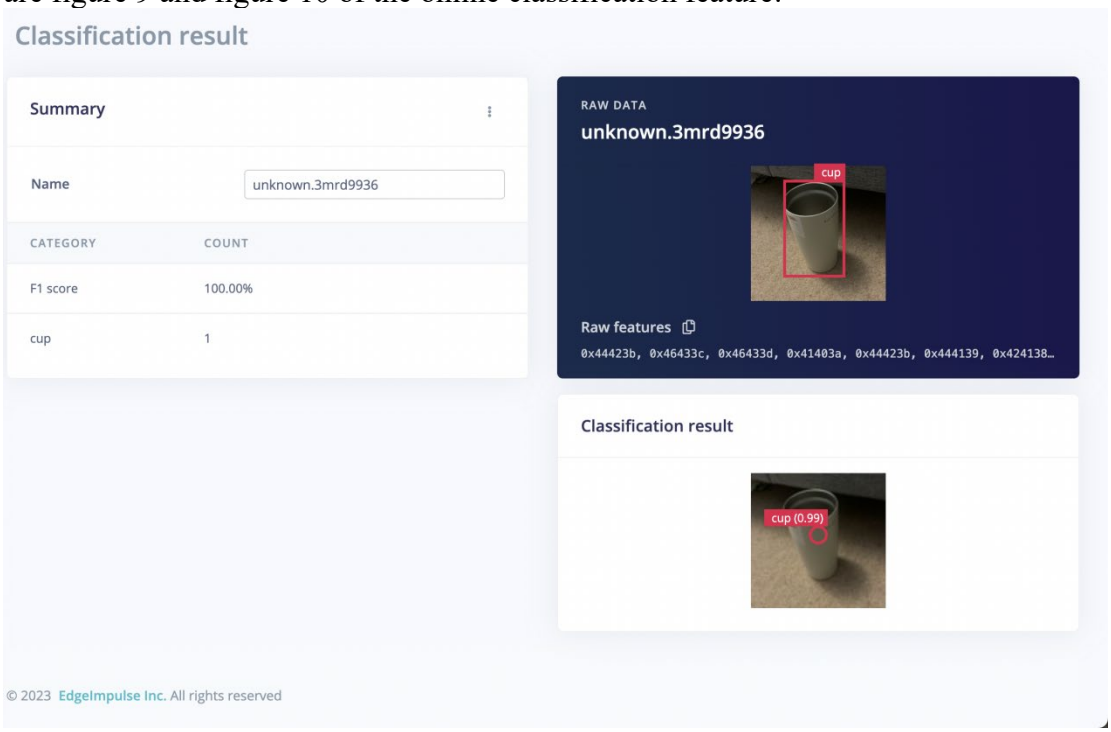

Figure 9 the example of cup_classification(1)

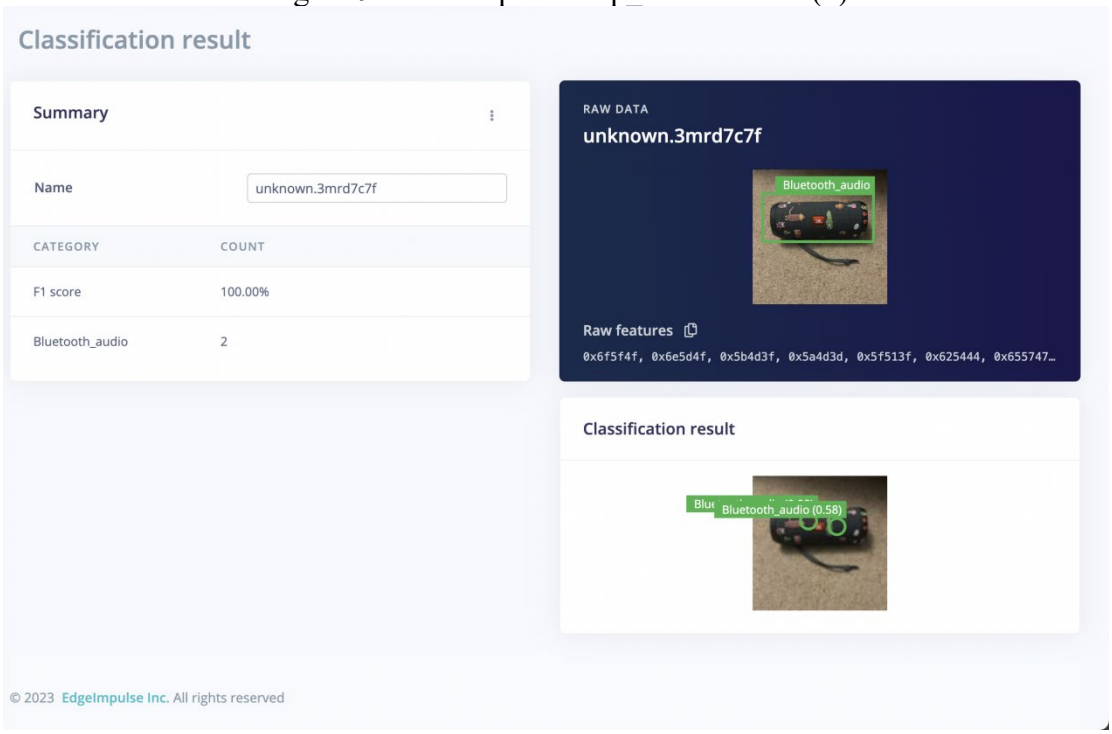

Figure 10 the example of Bluetooth classification(1)

3. Conclusion

The report outlines the process of using artificial intelligence to achieve object detection. The first step is to prepare the training and testing data. Feature generation is a critical step in training data

analysis, which helps to identify sources of information. The training data will be evaluated for accuracy through object detection, represented by the Confusion matrix, which reflects the performance of the previous training. Once the training data is complete, the testing model can be developed. It is important to use different parts of dataset for testing and training, and more training data will increase the learning performance. The testing model will evaluate all the testing data, providing an accuracy score for object classification. Object classification is a complex deep learning activity that requires extensive data to support it. Deep machine learning algorithms need a large amount of data to train and test to maintain high accuracy. Object classification is widely applied in our daily lives, and its popularity is increasing rapidly. In addition, the online classification feature allows for object classification by uploading photos. This technology has enormous potential and can significantly impact our lives in many positive ways. As Derrick Mwiti[4]says:"Object detection has been applied widely in video surveillance, self-driving cars, and object/people tracking. In this piece, we'll look at the basics of object detection and review some of the most commonly-used algorithms and a few brand new approaches, as well. Object detection locates the presence of an object in an image and draws a bounding box around that object. This usually involves two processes; classifying the object's type, and then drawing a box around that object."

Reference

[1] Pascual, C., Song, Z., Krein, P. T., Sarwate, D. V., Midya, P., & Roeckner, W. J. (2003). High-fidelity PWM inverter for digital audio amplification: Spectral analysis, real-time DSP implementation, and results. IEEE Transactions on Power Electronics, 18(1), 473-485.

[2] Pascual, C., Song, Z., Krein, P. T., Sarwate, D. V., Midya, P., & Roeckner, W. J. (2003). High-fidelity PWM inverter for digital audio amplification: Spectral analysis, real-time DSP implementation, and results. IEEE Transactions on Power Electronics, 18(1), 473-485.

[3] Room, C. (2019). Confusion Matrix. Mach. Learn, 6, 27.

[4] Goutte, C., & Gaussier, E. (2005). A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In Advances in Information Retrieval: 27th European Conference on IR Research, ECIR 2005, Santiago de Compostela, Spain, March 21-23, 2005. Proceedings 27 (pp. 345-359). Springer Berlin Heidelberg.