Image classification

When a computer analyses a picture, it may determine the "class" that the image belongs to. (Or a likelihood that the image belongs to a "class"). A class is basically a term, like "vehicle," "animal," "structure," and so on. (HuggingFace, 2022)

You may enter a picture of a sheep, for instance. The method of having a computer analyse an image and inform you that it is a sheep is known as image classification. (Or the likelihood that a sheep is present.)

Raw pixel data was the foundation of early picture classification. This implied that computers would dissect pictures into their component pixels. The issue is that the same subject might seem substantially differently in two distinct photographs. They may have various backdrops, perspectives, postures, etc. This makes it very difficult for computers to accurately "see" and classify images. (TensorFlow, 2022)

Giving a complete image a name or class is the work of image classification. It is assumed that each picture will only have one class. The class to which an image belongs is predicted by image classification models when they receive an image as input. (HuggingFace, 2022)

Computer Vision

Is a kind of deep learning where people train computers to see and understand the environment. Its advancement accelerated along with advancements in imaging and scanning technology. With computer vision's primary objective being:

- Facial recognition
- Image classification
- Image segmentation
- Object identification

(HuggingFace, 2022)

How does computer vision work?

1. Acquiring an image: Through video or picture analysis, images—even massive sets—can be gathered in real-time.

2. Image processing: Deep learning models automate a large portion of this process, although the models are frequently initially trained on a large number of tagged, previously recognised photos.

3. Understanding the image: The last stage, which identifies and categorises an item, is the interpretation step.

Use cases

When we are not interested in particular instances of objects with position information or their shape, image classification models can be applied. (TensorFlow, 2022)

Image Classification Metrics

The performance of different models is compared and evaluated using the metrics listed below:

1.Precision

Each class has a definition for precision, which tells us what percentage of the data that the model predicted would belong to that class in the validation data really did.

2. Recall

It provides information on the percentage of data from the validation set that were properly recognised as belonging to the class, similar to how precision is described for each class.

3. F1

Is a statistic that shows how accuracy and recall are balanced.

(HuggingFace, 2022)

The dataset

This is a collection of global natural scene images.

Link: https://www.kaggle.com/datasets/puneet6060/intel-image-classification

The dataset contains:

- Training set of 7000 images
- Test set of 3000 images
- Each example:
    - Six categories include 150 x 150 image distributions.
    - Associated to a label from 6 classes.
- Pictures:
    - 'buildings' -> 0,
    - 'forest' -> 1,
    - 'glacier' -> 2,
    - 'mountain' -> 3,
    - 'sea' -> 4,
    - 'street' -> 5

Why is the dataset appropriate

A carefully managed collection of photos is called an image classification dataset, and it is used to train, test, and assess how well machine learning algorithms work. The dataset must contain high-quality, varied, and multi-dimensional pictures since the algorithms learn from the samples of images they are given. (HuggingFace, 2022)

This is accomplished through the worldwide natural scene dataset:

• The dataset's simplicity makes it possible for a straightforward picture classification model to be used and understood.

• The data has already been processed and sanitised.

• Less intensive hardware processing is possible because to the tiny size.

• A lot of different examples.

• It is suitable for machine learning (deep learning) models due to the total volume of data (25000 instances).

• Excellent training data.

• The classes and pixel values for the photos are present.

• The abundance of data also helps the model avoid being underfit.

(ArcMap, 2022)

Analysis that will be conducted on the dataset

In order to classify photos, I will attempt to implement CNN using Keras in code.

1. Importing useful packages will come first.
2. Next, we'll load the data and pre-process it before visualising it.
3. Then, we'll assess how well the straightforward CNN model performed.
4. To overcome this difficulty as well, we will then employ a pre-trained model.

The model evaluation was to be done using TensorFlow

What is TensorFlow?

Is a group of tools that makes it possible for beginners to build machine learning models. It provides a machine learning end-to-end platform with numerous tools for the following:

• Common datasets for early validation and training.

• Data loading pipelines that are extremely scalable.

• Layers of pre-processing for typical input manipulations.

• Instruments for transforming and validating huge datasets.

• AI technologies to assist in identifying and removing bias in your data.

(TensorFlow, 2022)

Imports and Libraries used:

- Pandas
- NumPy
- Seaborn
- Matplotlib
- Cv2
- Os
- Sklearn
- Tensor Flow
- Tqdm

Exploratory Data Analysis

In our exploratory data analysis, we will try to answer these questions:

- How many training and testing examples do we have?
- What is the size of the images?
- What is the proportion of each observed category?

Overfitting and Underfitting

Ways to deal with underfitting:

- Remove noise from data.
- Increasing model complexity.
- Increase duration of training.

Ways to deal with overfitting:

- Increase training data.
- Reduce model complexity.

Model Creation and Training

Steps are:

1. Create the model,
2. Compile the model,
3. Train the model and fit the data to it,
4. The model's performance on the test set,
5. Run our model through an error analysis.

We can create a simple model with various layers, such as:

1. Conv2D: (32 filters of size 3 by 3) (32 filters of size 3 by 3) From the image, the features will be "extracted."

2. Images are halved in size thanks to MaxPooling2D.

3. Flattening changes, the photos' format from a two-dimensional array to a one-dimensional array with 150 150 3 pixel values.

4. Relu: returns the maximum value given x. (x, 0).

5. Softmax: 6 neurons, class membership probability for the image.

Interpret the results

- Using accuracy score
- Using a confusion matrix
- Using a classification report

Conclusion

• A complicated sequential model with numerous convolutional layers and 20 training epochs was able to predict the test with an accuracy of 0.76.

• To reduce overfitting and underfitting, the model was trained using an optimizer and a loss function.

• The validation accuracy and loss were examined, and it was found that the model was slightly underfit.

• The model has issues with buildings and streets. Well, given that the roadway is lined with buildings, that is somewhat understandable. Additionally, it struggles with the sea, glaciers, and mountains. I find it difficult to adequately discern between them. However, it has a very good sense of forest detection!

References

ArcMap (2022) *What is image classification?*, *What is image classification?-ArcMap | Documentation*. ArcMap. Available at: https://desktop.arcgis.com/en/arcmap/latest/extensions/spatial-analyst/image-classification/what-is-image-classification-.htm (Accessed: November 23, 2022).

HuggingFace (2022) *What is image classification? - hugging face*, *What is Image Classification? - Hugging Face*. HuggingFace. Available at: https://huggingface.co/tasks/image-classification (Accessed: November 23, 2022).

TensorFlow (2022) *Image classification : Tensorflow Core*, *Image classification*. TensorFlow. Available at: https://www.tensorflow.org/tutorials/images/classification (Accessed: November 23, 2022).