Recognizing Images -Deep Learning

**White Paper**

DSC, Bellevue University

Prashant Raghuwanshi

DSC680-T301 Applied Data Science (2225-1)

 Professor Catie Williams

05/01/2022

Recognizing object Images by using Deep Learning:

The current state-of-the-art Computer Vision (CV) and Machine Learning (ML) allow the detection and tracking of single objects classes (such as faces, pedestrians, or cars) in an unconstrained setting at a level that allows the realization of smart cameras that recognize smiling persons, driver assistance (pedestrian detection), surveillance applications and image-based web search.

Here I am going to build an image-driven application that leverages computer vision to classify or categorize an image file based on its visual content. In essence, these applications pull valuable information and insights out of visual content.

**Business Problem :**

The use cases for image classification are virtually unlimited. Across a wide range of industries, organizations can put these AI-driven capabilities to work to streamline business processes, automate tedious manual tasks and gain insights in real-time.

Please find below a few of the listed use cases for image classification :

* Colleges & universities can deploy an image verification access system to keep check of unauthorized person entry into their campus.
* A social media company might use video classification to categorize and annotate content.
* A retailer might use image classification to enable a checkout-free store.

The list of potential applications of image classification goes on and on.

A commonality these applications share is deep learning, one of the key building blocks for AI solutions.

**Background :**

To solve the above-mentioned critical one of the Business use cases for identifying unauthorized persons or objects inside the university campus. This project is going to develop the computer vision ML software which is going to integrate with security cameras’ IOT feed data, and in return, this model will identify the images and send a notification to the responsible security personnel’s mobile device. Image classification is a supervised learning problem: define a set of target classes (objects to identify in images), and train a model to recognize them using labeled example photos.

In this project I am going to use the leading architecture used for image recognition and detection tasks is Convolutional Neural Networks (CNNs).

**Data Explanation:**

The CIFAR-10 are labeled subsets of the 80 million tiny images dataset. It is a freely available dataset on the web and can be used for education and learning purposes.

They were collected by Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton.

The CIFAR-10 dataset consists of 60000 32x32 color images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.

The dataset is divided into five training batches and one test batch, each with 10000 images. The test batch contains exactly 1000 randomly-selected images from each class.

The training batches contain the remaining images in random order, but some training batches may contain more images from one class than another.

Between them, the training batches contain exactly 5000 images from each class.

Link: <https://www.cs.toronto.edu/~kriz/cifar-10-python.tar.gz>

**Describe Data**.

Here are the classes in the dataset, as well as 10 random images from each:

A picture containing text

Description automatically generated

Graphical user interface, application, Word

Description automatically generated

**Methods:**

Computer vision joins two technologies, deep learning and Convolutional Neural Networks (CNNs), to analyze and identify images

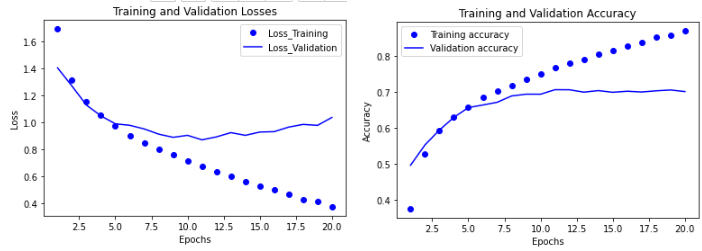
In this project, models were created by using **Convolutional Neural Networks** are used to extract features from images (and videos), employing convolutions as their primary operator. Convolutional neural networks consist of several layers with small neuron collections, each of them perceiving small parts of an image. The results from all the collections in a layer partially overlap in a way to create the entire image representation. The layer below then repeats this process on the new image representation, allowing the system to learn about the image composition

Diagram

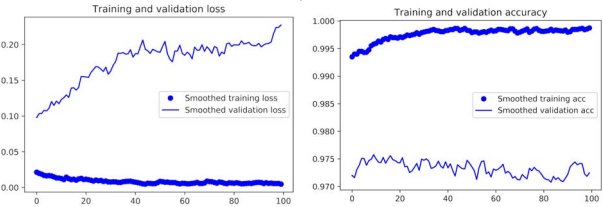
Description automatically generated

**Analysis :**

The training accuracy increases linearly over time, until it reaches nearly 100%, whereas the validation accuracy stalls at 70–72%. The validation loss reaches its minimum after only ten epochs and then stalls, whereas the training loss keeps decreasing linearly until it reaches nearly 0

****

**After fine-tuning** a network, The validation accuracy curve looks much cleaner. You’re seeing a nice 1% absolute improvement in accuracy, from about 96% to above 97%.

****

Seems our model had predicted accurately the image of a school bus.

****

**Conclusion:** The generated fine tunned CNN model is the best fit for object predictions. We can use this classification model to develop Automatic object identification systems that can be designed for many processes such as university security systems.

**Assumption**: Here I am assuming the Preprocessing data extraction operations is not having any limitations and the images were extracted & preprocessed successfully by the image data conversion tool or Keras image preprocessing libraries.

**Limitations:** This Model is limited to identifying the selected few varieties of objects only. To make this model more useful we need to train it with real-time images and to work with added new images this model required regular training.

**Challenges:** Major Constraints are related to used datasets and processed images, here the used datasets contain a total of 10 thousand object images, including 1 thousand for each object. However, the rapid addition of new objects in image lists would require regular re-training and deployment of the model. and this model is not trained for night vision cameras.

**Future Uses:**

* Elimination of the manual comparison of images, via the introduction of an automated, quicker, and more accurate and efficient way of comparing images and UI screens.
* Capability to spot minute differences that might otherwise be missed by manual inspection, including identification of pixel-level differences for fonts, colors, etc.
* Solution delivery as an API service over the internet.
* Does not require extensive processing capabilities or any additional expensive software or hardware. Because the AI/ML solution resides on the internet as an API service, all processing happens on the internet rather than locally.

**Recommendations:**

Modern software can recognize a large number of everyday objects, human faces, printed and handwritten text in images, and other entities. and we’ll continue witnessing how more and more businesses and organizations implement image recognition and other computer vision tasks to stand out from competitors and optimize operations. The developed image identification mode is best suited to recognize multiple individual objects and it is required to add more images in training datasets to make it more compatible with the real-time need.

**Implementation Plan:**

To develop ML applications, we have below implementation plan:

IDE’s to use: Anaconda Jupyter

Coding languages: Python

Platform: AWS and local PC

Resource Data Science Engineer

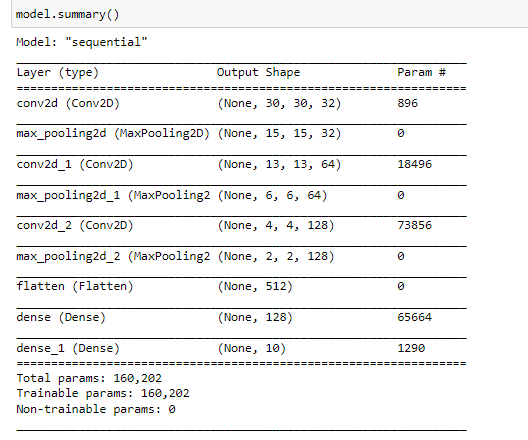
Timeline: 04/20/2022 to 5/16/2022

**Ethical Considerations:**

* This Data contains processed physical image information related to multiple verities of objects like multiple animals, cars, bus and does not contain any images of human individuals’.
* In the feature we definitely need government authorities & individual students’ approval to use the images of individual students to train the model.
* Datasets and information on data were extracted from the public websites 🡪 UCL machine learning repositories.
* This data research is not going to harm any privacy.

**Appendix :**

Model build screenprint:

****

**Questions May be asked by the audience:**

1. In the Real world how the data science team is going to capture the required student biometric data
2. What type of legal considerations do we need to deal with during scaling out this project?
3. Do you have any best practices for handling sensitive data?
4. Here we are going to use human PIIS so what type of approval is required from the government agencies
5. Do the data processing in-house and procure the required equipment & Software for capturing & process the human image data.
6. For procuring data from vendors or government agencies what ethical considerations do we need to take care of.
7. What would be our revenue model?
8. The presented solution or model is meant for identifying few types of objects How you are going to make it fit for other varieties?
9. How do you conclude the CNN model is the best fit?
10. How accurate is your Model?
11. With the increase in counts of varieties of objects and human faces images, are you expecting any degradation in, accuracy?
12. How the Model is designed to behave with unknown object images.
13. please elaborate more on provided Recommendations:
14. Do you have a recommendation for making the model prediction more consistent:

**References :**

<https://www.altexsoft.com/blog/image-recognition-neural-networks-use-cases/>

<https://data-flair.training/blogs/convolutional-neural-networks-tutorial/>

<https://www.cs.toronto.edu/~kriz/cifar.html>

<https://www.delltechnologies.com/asset/en-us/products/ready-solutions/industry-market/ai-driven-image-video-classification-white-paper.pdf>

<https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html>

Deep Learning with Python 1st Edition

Andrej Karpathy, The Unreasonable Electiveness of Recurrent Neural Networks