IMAGE RECOGNITION

TO

DETECT MULTIPLE ATTRIBUTES

Statement of Work

Course: AIDI 1002-01 : Final Project

Course Facilitator: Marcos Bittencourt

Presented by:

Tarun Kumar Punna

PROJECT OVERVIEW:

Image recognition is one of the many applications of Machine Learning, it can solve problems for security purposes, object detection, face detection, healthcare, entertainment, among others. This application has an enormous potential to help our society1 , so it is important to find new uses for this tool, improve the current methods and get more accurate and useful insights from it.

In this project, the researcher creates three models to predict three attributes out of the same picture, it uses the Inception-V3 as based model for image recognition , and the data set uses is the CelebA dataset , which contains over 200,000 celebrity images and 40 binary attribute annotations per image.

PROBLEM STATEMENT:

Many machine learning algorithms used in face recognition or object detection are built to detect a specific attribute, for example, the mood of a person in a picture, the algorithm will predict if the person is happy, sad, neutral, surprised, etc. What if we want to explore many non-exclusive attributes that are of a picture at the same time? As to classify if the person is smiling or not and at the same time know if is wearing eyeglasses or if is wearing a hat. Most probably the same algorithm will not perform well predicting two or more attributes. This is a classification problem for each of the selected attributes, one model per attribute, but using the same image. The output will be the predicted attribute of each model.

The use of different algorithms, one of them predicting a specific attribute will give us better insights of the picture to be analyzed, results will be measure against the real target and the algorithms will be replicable to pictures outside of the dataset used in this project.

The goal is to build three classification models based on image recognition, the attributes to be predicted are:

- If the subject is smiling or not.

- If the subject is female or male.

- If the subject is young or not.

DATA SOURCES

A popular component of computer vision and deep learning revolves around identifying faces for various applications from logging into your phone with your face or searching through surveillance images for a particular suspect. This dataset is great for training and testing models for face detection, particularly for recognising facial attributes such as finding people with brown hair, are smiling, or wearing glasses. Images cover large pose variations, background clutter, diverse people, supported by a large quantity of images and rich annotations.

This data was originally collected by researchers at MMLAB, The Chinese University of Hong Kong (specific reference in Acknowledgment section).

Contents:

Overall

- 202,599 number of face images of various celebrities
- 10,177 unique identities, but names of identities are not given
- 40 binary attribute annotations per image
- 5 landmark locations

Data Files:

- imgalignceleba.zip: All the face images, cropped and aligned
- listevalpartition.csv: Recommended partitioning of images into training, validation, testing sets. Images 1-162770 are training, 162771-182637 are validation, 182638-202599 are testing
- listbboxceleba.csv: Bounding box information for each image. "x1" and "y1" represent the upper left point coordinate of bounding box. "width" and "height" represent the width and height of bounding box
- listlandmarksalign_celeba.csv: Image landmarks and their respective coordinates. There are 5 landmarks: left eye, right eye, nose, left mouth, right mouth
- listattrceleba.csv: Attribute labels for each image. There are 40 attributes. "1" represents positive while "-1" represents negative

METRICS

The metrics to measure the performance of the models are:

Accuracy:

Accuracy is a common metric to measure binary classifiers, this adapts well to the problem to be solved, the corresponding formula to this metrics is:

$$accuracy = \frac{True\ Positives + True\ Negatives}{Total\ number\ of\ pictures}$$

Where:

- True Positives: Number of times the model predicted to the positive class and it was the positive class.
- True Negatives: Number of times the model predicted to the negative class and it was the negative class.
- Total number of pictures: Total number of pictures included in the test data set.

F1 Score:

F1 score is a metric that consider precision and recall for test accuracy. This measure gives relevant information about how the positive class is being predicted. The perfect model will have a value of 1 and the worst a value of 0.

$$F1\ score = 2 * \frac{precision * recall}{precision + recall}$$

Where:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

RESEARCH TECHNIQUES

Image Classification:

Contextual image classification, a topic of pattern recognition in computer vision, is an approach of classification based on contextual information in images. "Contextual" means this approach is focusing on the relationship of the nearby pixels, which is also called neighborhood.

Feature Engineering:

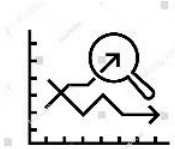
Discovery analysis collects data and consolidates it into a single source that can be easily and instantly evaluated. Once your raw data is converted, follow the train of thought by drilling down into the data. When a trend is identified, it unearths the contributing factors.

Deep Learning:

Deep learning is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised.

Predictive Analysis:

Predictive analysis helps to answer the question "What might happen?" by extracting information from existing data to determine patterns and then form assumptions of future outcomes and trends. With this, we will describe what is the best course of action.



DEVELOPMENT ENVIORNMENT

Python

Python is an interpreted, high-level, general-purpose programming language used for performing the statistical analysis. When applying the technique of Web Scraping, Python coding will scrap the internet for selected data.

Google Colab

Colab you can import an image dataset, train an image classifier on it, and evaluate the model, all in just a few lines of code. Colab notebooks execute code on Google's cloud servers, meaning you can leverage the power of Google hardware, including GPUs and TPUs. Colab is used extensively in the machine learning community.

TensorFlow

TensorFlow is a free and open-source software library for machine learning. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. TensorFlow is a symbolic math library based on dataflow and differentiable programming.

Keras

Keras is an open-source library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library. Up until version 2.3 Keras supported multiple backends, including TensorFlow, Microsoft Cognitive Toolkit, R, Theano, and PlaidML.

ANALYSIS PROCESS

- Collect the data
    - Images collected from various sources in all sections.
- Clean the data
    - Clean Secure storage of cleaned datasets
- Perform Feature Engineering
- Develop the Data Preprocessing Pipeline
- Develop the Model Architecture
- Split the data into train, test, and validation data.
- Train Model using Training Data Set.
- Perform Validation using validation dataset
- Test Model
- Evaluate the model
- Form and document conclusions

IMPLEMENTATION PLAN

Item	Phase	Major Tasks / Milestones	Dates
1.	Project Organization	Setup	Oct 28, 2020
2.	Business Understanding & Problem Discovery	Statement of Work	Nov 06, 2020
3.	Data Acquisition & Understanding	Data Collection & EDA	Nov 23, 2020
4.	ML Modeling & Evaluation	Evaluate learning algorithms and Prototype model architecture	Nov 23, 2020
5.	Delivery & Acceptance	Development & deployment of software pipeline & Solution endpoint	Dec 18, 2020