KLE Society's

KLE Technological University

September 27, 2022

Course Project
On
Birthday Celebration Scenario Image Classification

Machine Learning(17ECSC306)

Submitted by

Name	Roll No	USN
Sharad S Shettar	106	01FE19BCS008
Sairajath R Nayak	116	01FE19BCS020
Ramakrishna M Desai	120	01FE19BCS024
Sharath S Shanbhag	124	01FE19BCS029

Team Number:A08

Under The Guidance Of
Dr. Meena S M
Mr. Uday Nagraj Kulkarni

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
HUBLI-580031

Academic year 2021-2022

Abstract

We live in the era of data. With the Internet of Things (IoT) and Artificial Intelligence (AI) becoming ubiquitous technologies, we now have huge volumes of data being generated. Differing in form, data could be speech, text, image, or a mix of any of these. In the form of photos or videos, images make up for a significant share of global data creation. Image classification is probably the most important part of digital image analysis. It uses AI-based deep learning models to analyze images with results that for specific tasks already surpass human-level accuracy .Since the vast amount of image data we obtain from cameras and sensors is unstructured, we depend on advanced techniques such as machine learning algorithms to analyze the images efficiently. Image classification is probably the most important part of digital image analysis. It uses AI-based deep learning models to analyze images with results that for specific tasks already surpass human-level accuracy (for example, in face recognition).

Since AI is computationally very intensive and involves the transmission of huge amounts of potentially sensitive visual information, processing image data in the cloud comes with severe limitations. Therefore, there is a big emerging trend called Edge AI that aims to move machine learning (ML) tasks from the cloud to the edge. This allows moving ML computing close to the source of data, specifically to edge devices (computers) that are connected to cameras.

Acknowledgement

Every project is successful largely due to the effort of a number of people who have always given their valuable advice or lent a helping hand. We sincerely appreciate the inspiration, support and guidance of all those people who have been instrumental in making this project a success.

We take this opportunity to thank Dr.Ashok Shettar, Vice-chancellor,KLE Technological University and to Dr. Prakash Tewari, Principal, B V Bhoomaraddi College of Engineering and Technology, Hubli.

We take this opportunity to thank Dr.Meena S M, Head, School of Computer Science and Engineering for having provided us with an academic environment which nurtured our practical skills contributing to the success of our project.

We also take this opportunity to thank Prof. Uday Kulkarni for the guidance and constant supervision as well as for providing necessary information regarding the project & also for supporting us throughout the project. We are thankful and fortunate enough to get constant encouragement, support and guidance from all teaching and non teaching staff of SoCSE which helped us in successfully completing our project.

We owe our gratitude to the School of Computer Science Engineering, KLE Technological University for providing us with the resources necessary for the completion of this project. Our gratitude will not complete without thanking the Almighty, our beloved parents, our seniors and our friends who have been a constant source of blessings and aspirations.

Contents

1 Introduction

1.1 Overview of the project

Image classification is the task of categorizing and assigning labels to groups of pixels or vectors within an image dependent on particular rules. The categorization law can be applied through one or multiple spectral or textural characterizations.

With billions of smartphones around the globe, wouldn't it be great if the smartphone could be turned into a tool for recognizing the various Birthday Celebration scenarios from images it captures with its camera?Performing on-device image recognition makes it possible to overcome the limitations of the cloud in terms of privacy, real-time performance, efficacy, robustness, and more. Hence, the use of Edge AI for computer vision makes it possible to scale image recognition applications in real-world scenarios.

2 Objectives

1) To pre-process the images.
2) To categorise the Images into different categories of Celebration Scenarios.
3) Build an application using the trained model to predict the class of the real-time celebration scenario image.

2.1 Problem definition

We are provided with two tasks:

*Our first task is to categorise the images into three sub-categories of the Birthday Celebration Scenario from the collected dataset.

*Our second task is to classify different categories of celebration considering 5 categories and given the image of a celebration scenario, our task is to predict what category the provided image belongs to.

3 Approach

3.1 Dataset

The aim was to collect different scenarios of Birthday celebrations.We collected a total of 685 images.

Figure 1: All The Classes of Birthday Celebration Scenarios in Dataset

One image from each category is shown in Figure 1.The three categories are-GreenTheme,CharacterTheme and SurpriseParties .

3.2 Methodology

Our Methodology consists of 3 phases, Pre-processing, Model Training, App Building. The images in the dataset are Pre-processed i.e.

Also the height and width of the images is kept 224 itself. The images are split into train and test datasets using splitfolders library in python. In second phase i.e. model training the model is trained on the pre-processed data the model with best test accuracy is selected. These model weights are saved and used for building the app in the third phase.

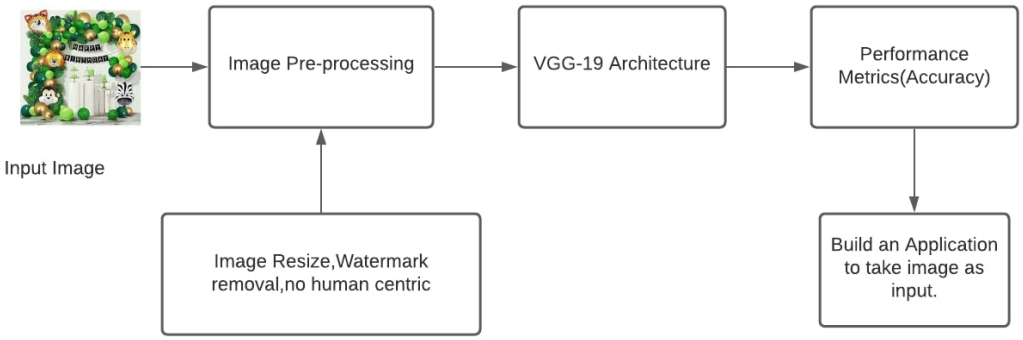

Figure 2: Methodology

4 A Deep Walk Through The Pipeline

4.1 Pre-processing

All the collected images were converted to an array.The images were then converted to a default size(224x224) and with no watermarks.

4.2 Loss Function

Here we are using categorical cross-entropy as loss function to minimize. It is also called as Soft-max Loss. It is a Soft-max activation plus a Categorical-Entropy loss. If we use this loss, we will train a CNN to output a probability over the classes for each image. It is used for multi-class classification.

Figure 3: categorical Entropy

4.3 Architecture Design

VGG19 is a variant of VGG model which in short consists of 19 layers (16 convolution layers, 3 Fully connected layer, 5 MaxPool layers and 1 SoftMax layer). There are other variants of VGG like VGG11, VGG16 and others. VGG19 has 19.6 billion FLOPs.

A fixed size of (224 * 224) RGB image was given as input to this network which means that the matrix was of shape (224,224,3).

The only preprocessing that was done is that they subtracted the mean RGB value from each pixel, computed over the whole training set.

Used kernels of (3 * 3) size with a stride size of 1 pixel, this enabled them to cover the whole notion of the image.

spatial padding was used to preserve the spatial resolution of the image. max pooling was performed over a 2 * 2 pixel windows with sride 2.

This was followed by Rectified linear unit(ReLu) to introduce non-linearity to make the model classify better and to improve computational time as the previous models used tanh or sigmoid functions this proved much better than those.

5 Results

5.1 VGG-19 Architecture

Total params: 20,099,651 Trainable params: 75,267 Non-trainable params: 20,024,384

Model	Accuracy
VGG-19(BirthdayCelebration)	98.00
VGG-19(CelebrationScenario)	97.4

input_1 (InputLayer)	[(None, 224, 224, 3)]	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv4 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv4 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv4 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0

Figure 4: VGG-19

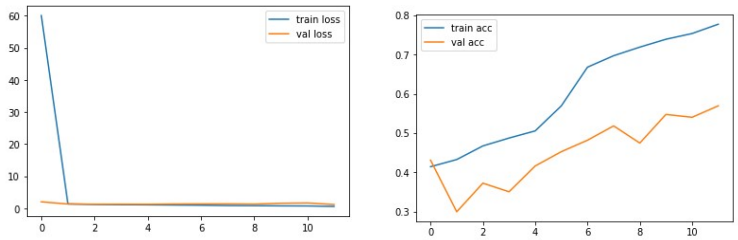

Figure 5: Birthday Celebration Loss And Accuracy Graphs

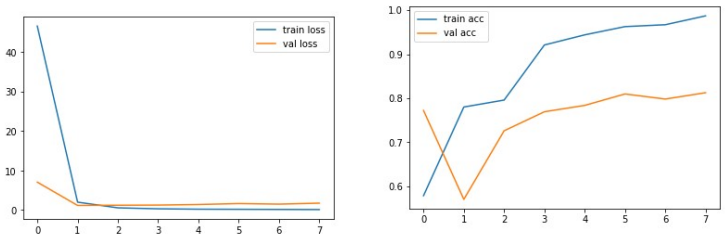

Figure 6: Celebration Loss And Accuracy Graphs

6 Application

6.1 User Interface

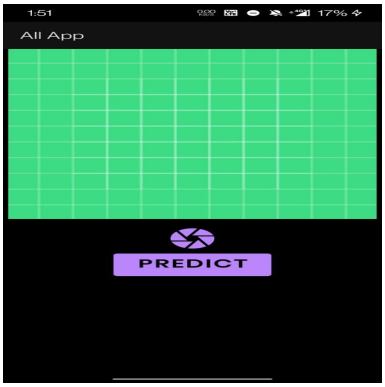

Figure 7: Main Screen of Application

6.2 Screenshots of Results

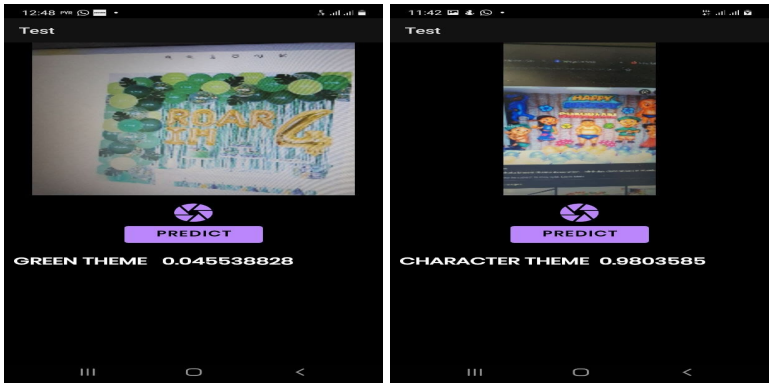

Figure 8: Result for Birthday Celebration scenario

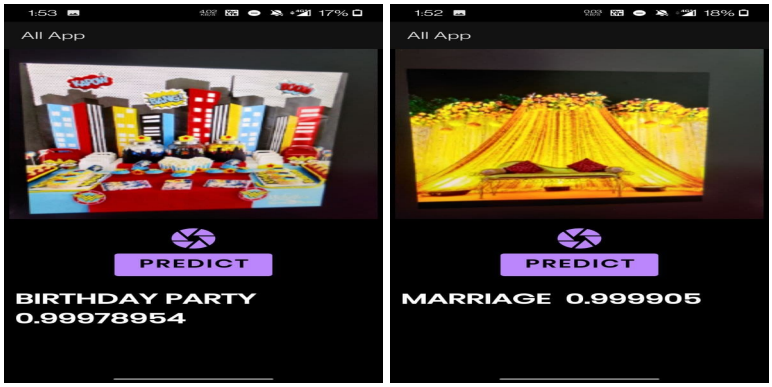

Figure 9: Result for Celebration scenario

7 Conclusion

With respect to the context of proposed learning model to categorise celebration scenarios, we proceeded with VGG-19 Architecture and it was trained to:

- To predict the category label of the given image.

GreenTheme,CharacterTheme and SurpriseParties are the 3 Categories on which proposed method is tested. Therefore, related scenario images were taken for identification which spans over 3 classes.
Five Celebration scenarios namely marriage ,Birthday,Festival,Baby shower and naming ceremony were also classified using VGG-19 architecture.

8 Future Scope

- The images in our custom dataset are sourced from Google and Pinterest and not in real conditions of Celebration scenarios, this motivates to build a public dataset which captures the images in real life conditions which would enable us to achieve a better efficacy in solving this problem.
- Few more techniques like image segmentation can be tried along with different CNN models to help in categorising the images of the celebration scenarios.
- To improve recognition rate in classification process, some other CNN architectures and hybrid algorithms can also be used.

9 References

[1]- https://medium.com/analytics-vidhya/python-based-project-covid-19-detector-with-vgg-19-convolutional-neural-network-f9602fc40b81

[2]- https://github.com/Sharathsshanbhag/CustomImage-CelebrationScenari
.git