

A Project Report on

Real Time Bare Skinned Image Filtering Using CNN

Submitted in partial fulfillment of the requirements for the award
of the degree of

Bachelor of Engineering

in

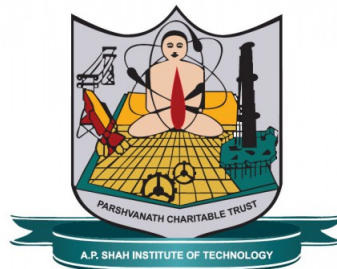
Computers

by

Pratik Jain(16102060)
Tina Shah(16102019)
Kshitija Shah(16102027)

Under the Guidance of

Prof.Jaya Gupta



Department of Computers

A.P. Shah Institute of Technology
G.B.Road,Kasarvadavli, Thane(W), Mumbai-400615
UNIVERSITY OF MUMBAI

Academic Year 2019-2020

Approval Sheet

This Project Report entitled “ *Real Time Bare Skinned Image Filtering Using CNN*” Submitted by “*Pratik Jain*” (16102060), “*Tina Shah*” (16102019), “*Kshitija Shah*” (16102027) is approved for the partial fulfillment of the requirement for the award of the degree of *Bachelor of Engineering* in *Computers* from *University of Mumbai*.

Prof. Jaya Gupta
Guide

Prof. Sachin Malve
Head Department of Computers

Place: A.P. Shah Institute of Technology, Thane
Date:

CERTIFICATE

This is to certify that the project entitled “*Real Time Bare Skinned Image Filtering Using CNN*” submitted by “*Pratik Jain*” (16102060), “*Tina Shah*” (16102019), “*Kshitija Shah*” (16102027) for the partial fulfillment of the requirement for award of a degree *Bachelor of Engineering* in *Computers*, to the University of Mumbai, is a bonafide work carried out during academic year 2019-2020.

Prof. Jaya Gupta
Guide

Prof. Sachin Malve
Head Department of Computers

Dr. Uttam D. Kolekar
Principal

External Examiner(s)

1.

2.

Place: A.P. Shah Institute of Technology, Thane

Date:

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, We have adequately cited and referenced the original sources. We also declare that We have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

(Pratik Jain 16102060)

(Tina Shah 16102019)

(Kshitija Shah 16102027)

Date:

Abstract

Image classification is critical and significant research problems in computer vision applications such as facial expression classification, satellite image classification, and plant classification based on images. This project proposes the image classification model applied for identifying the display of daunting pictures on the internet. The proposed model uses Convolutional Neural Network (CNN) to identify these images and filter them through different blocks of the network, so that it can be classified accurately. The model works on TensorFlow a user friendly platform, which provides different high levelled APIs (in Keras) which are used to build any basic model. It also permits us to add our own libraries. The input data for the model are images we collected online through various resources. Our model will work as an extension to the web browser and will work on all websites when activated. The output of the proposed model is blurring the images and deactivating the links. This means that it will scan the entire web page and find all the daunting images present on that page. Then we will blur those images before they are loaded and the children could see them. Apart from it, we will also disable any clickable links present. This ensures protection from disturbing images and links, to the children.

Contents

1	Introduction	1
1.1	Objective	1
2	Literature Review	2
3	Problem Definition	4
4	Scope	5
5	Technology Stack	6
6	Benefits for the environment and society	7
7	Applications	8
8	Proposed System Architecture	9
9	Flow Of Models	10
10	Activity Diagram	11
11	Use Case Diagram	12
12	Modules	13
12.1	Module 1 : Classification Of Images Using CNN	13
12.2	Module 2 : Optimizers and Activation Function Analysis	14
12.2.1	Activation Function:	14
12.2.2	Optimizers	15
12.3	Module 3 : Testing of Trained Model	17
12.4	Module 4 : Creating the required extension	17
13	Future Scope	18
	Annexture A	19
	Bibliography	20

List of Figures

10.1 Activity Diagram	11
11.1 Use Case Diagram	12
13.1 Gantt Chart	19

Chapter 1

Introduction

Skin colour detection has been used in numerous computer vision applications like face detection, nudity recognition, hand gesture detection and person identification. Skin colour detection is often used as a preliminary step in these applications. Colour is the most robust and useful clue for skin detection and also allows fast processing of the skin patterns. Other cues like shape and geometry can be used to build accurate face detection systems. Skin colour detection is a challenging task as the skin colour in an image is sensitive to various factors like illumination, camera characteristics, ethnicity, individual characteristics such as age, sex and body parts and other factors like makeup, hairstyle and glasses. All these factors affect appearance of skin colour. Another problem is that there is a significant overlap between the skin and non-skin pixels. Most of the skin detection techniques discussed in literature are used as preprocessor for face detection and tracking systems. However when these techniques are used in real-time, it is crucial to follow time deadlines and memory constraints. Sometimes, accuracy may need to be sacrificed when the skin detection strategy is used only as a preprocessing step to face detection, particularly in real time applications. So in our model we are using classification using the CNN which uses curves, edges and other features to identify and classify an image. We are training our model on both nude and non-nude images. In this study we have focused on the problem of developing an accurate and robust model for the nude images.

1.1 Objective

The model will offer child proof surfing on the internet without parent intervention. So that parents do not have to worry about their children coming across nude images at such an early age without knowing the actual meaning of it. That is it will monitor every page and filter all the images, hiding their details from the children and disabling their activation upon any click. Even if the people in the image is not completely nude, may it be just the upper or lower half of the persons body; thus ensuring guaranteed protection.

Chapter 2

Literature Review

The Convolutional Neural Network (CNN) classifies the image regions as a collection of either skin or non-skin regions, it also classifies it based on the curves and edges it encounters in the image. Various approaches to skin modelling are used in the literature. Here we give a brief review of the neural network models for skin detection. The approach presented by Lee et al. employed a learning scheme based on the skin colour distribution of the image, using a neural network to learn and classify whether the input image contains skin exposure. Earlier work on nude identification focused on human skin detection, in which the idea is that greater amounts of detected skin would lead to higher probabilities of nudity within the image, hence characterizing the content as nude. Nevertheless, these approaches suffer with a high rate of false positives, especially in the context of beaches or practice of aquatic sports. An Tien Vo, Hai Son Tran, Thai Hoang Le suggested a method which involved ConvNetJS an open library which implements Deep Convolutional Neural Network. It is written in JavaScript and HTML with various supporting features for research purpose. It is an effective Neural Network, and is able to use CPU instead of GPU. The model proposed by them is an advertisement image classification system based on CNN consisting of 4 steps. It takes the image as an input from the webpage. It takes the image on the web browser by skin capturing and then saving that image. Then at the end classifying it by passing it through the CNN model. Our model is based on real time image disabling, so using this method would not work in our case.

Methods of data mining such as k-means were as well utilized for detection of skin and worthwhile results were obtained as demonstrated in the study by Hamid A. Jalab using a cluster pixel model designed segmentation of skin under specifiable environmental conditions. This proposed model can overcome the changes in complex backgrounds and sensitivity to brightness conditions. Kakumanu et al. presented modeling and detection of skin and this is recommended for further study. In conclusion, associated with the current issue, the most outstanding results is obtained using fuzzy system combined with support vector machine (FSFCSVM).

A residual learning framework to ease the training of networks that are substantially deeper than those used previously was put forth by Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. They explicitly reformulated the layers as learning residual functions with reference to the layer inputs, instead of learning unreferenced functions. They provided comprehensive empirical evidence showing that these residual networks are easier to optimize, and can gain accuracy from considerably increased depth.

Apart from this the model proposed by Brown et al. contained two types of skin models

used in the literature (Brown et al., 2001) viz., symmetric and asymmetric. Symmetric model uses a single classifier for both the classes whereas asymmetric model uses two separate classifiers for skin and non-skin pixels that are separately trained using respective features. Advantage of asymmetric skin classifier is that it increases the distances in certain skin related features between a positive (skin) and a negative (non-skin) image, with disadvantage of increased time complexity for training two classifiers. The neural classifiers used in the literature either uses a symmetric model with single neuron in the output layer or uses two separate neural networks (asymmetric model) for each of the skin and non-skin classes. The novelty of our approach is that it has multiple convolutional that is hidden layer working on only skin coloured images containing humans. Thus by using a CNN classifier

we have improved the separability between these two classes, eliminating additional time complexity that is needed in asymmetric classifier.

Chapter 3

Problem Definition

The main problem faced with CNN is training the model with the images it needs to identify. So in this study we train the model only on one type of image, which contains of exposed skin of humans, using a classifier. The classifier will differentiate the image based on its type. An image classified as positive by the classifier equals a nude or daunting image. Any other type of image will be filtered right through it. Thus this classifier makes the work easier, as we dont have to make two different classifiers for identification and filtering.

Chapter 4

Scope

This project proposes the image classification model applied for identifying the display of daunting pictures on the internet. In order to differentiate them from images containing humans from other images of any type we train our model on those images. Our scope is to write our own CNN model for training the images to classify whether the daunting images. Model will be able to classify whether the image is nude or non-nude for further processing. Even if the people in the image is not completely nude, may it be just the upper or lower half of the person's body, the model should also be classify such type of images as nude images. Also CNN Model should classify images with good accuracy for better results.

Chapter 5

Technology Stack

Our technology is built on different APIs which help us detect the nude pictures on the web pages. There is an API for face detection and different APIs for nude body detection. Apart from this our model mainly uses Convolutional Neural Network (CNN) for classifying the images correctly. Images containing nudity should be separated from all other images in order to disable them. All this code is written in Colab, a platform which allows importing python libraries and uploading data from our computer or any other source. We have also used OpenCV-Python which makes use of Numpy, which is a highly optimized library for numerical operations. It is an open source computer vision and machine learning software library built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception. We are also using NVIDIA a graphics processing units for the gaming and professional markets, as well as system. It is a chip unit for the mobile computing and automotive market. We are using this NVIDIA embedded chip computer for our training purpose. Our model consists of 25000 training images. Normal computers without a graphic card cannot handle this amount of training data. For this reason the NVIDIA system is an important aspect for our project.

Chapter 6

Benefits for the environment and society

This project would work as an excellent child proof surfing on the internet without us monitoring their usage 24*7. Not only the child would not come across extreme pictures and videos at such an early age where their minds are growing and has the most grasping power. But it will also help the parents to not waste their time and worry about their child getting exposed to nudity at such an early age. Thus benefitting the society as a whole in multiple ways the project mainly ensures protection of the children from a number of things while surfing on the internet.

Chapter 7

Applications

We will try to put this technology on the web browser itself so that it works on all the websites. Thus being said we will try to provide it in the form of an extension which when activated would work for all the websites you visit while it is activated. So it has a wide range of applications, though it cannot work on all websites (e.g. porn websites) where there are multiple images and videos or links to them which cannot be blurred and disabled all at once. Our technology will still try to cover the maximum websites possible. The extension would work in Google Chrome as any other extension you may use or see. You will have to download and add it to the Google Chrome and give it permissions to scan the pages you visit while surfing on the internet.

Chapter 8

Proposed System Architecture

Convolution Neural Network (CNN) is one of the most popular and effective technology for image recognition and image classification. The CNN classifier takes an image as a input, processes it and classifies it under certain categories. Our main aim is to disable the view of daunting images and get the link inactivated so that children do not get redirected to some irrelevant content which they are not ready to get exposed to at such an early age. We will first be training the CNN model (VGG16/VGG19) on bare skinned images as well as advertisements of any other type. During the prediction process the bare images will be giving an output of one, while all other type of images will have a value assigned to them greater than one. Once the model is trained with a good accuracy of prediction and filtering, a chrome extension will be created for the same, which will capture all the images on the websites and if it comes across any such image, the image will be blurred, also automatically disabling any redirect able links. Hence this model would work as an excellent child proof surfing on the internet without us monitoring their usage 24*7.

Chapter 9

Flow Of Models

First Step is the collection of dataset of both kind of nude and non-nude images. Images will be scrapped from the internet using the links of that particular image using concept of BeautifulSoup in python. Approximately 25000 images will be used in dataset for more accurate results. Next step is the classification of images, done by using Convolution neural network. The training of CNN would be done on two kinds of images positive and negative; positive being the nude images and negative being the images of any other type. Testing will be done for the same, so that we come to know how accurate the prediction is. Testing will also consists of both kind of images that is nude images and the images of advertisements. Extension will be created just like any other Google extension, which will blur and disable the images on the website visited. The user just needs to download this Chrome extension and enable it, after that all the working will be taken care of by the extension itself without any kind of parental control. This extension will work 24*7 and provide the safest browsing experience on the internet to the kids.

Chapter 10

Activity Diagram

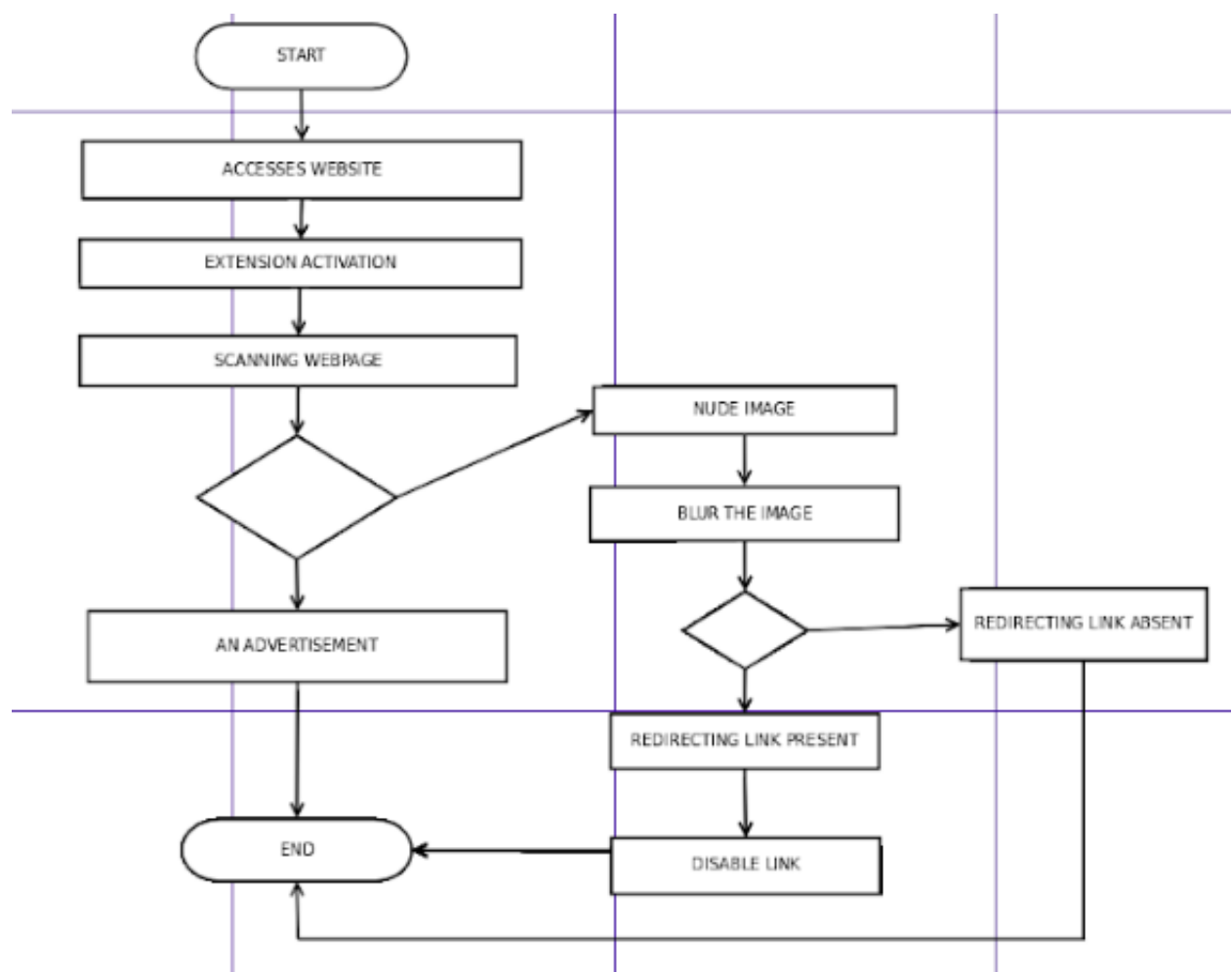


Figure 10.1: Activity Diagram

Chapter 11

Use Case Diagram

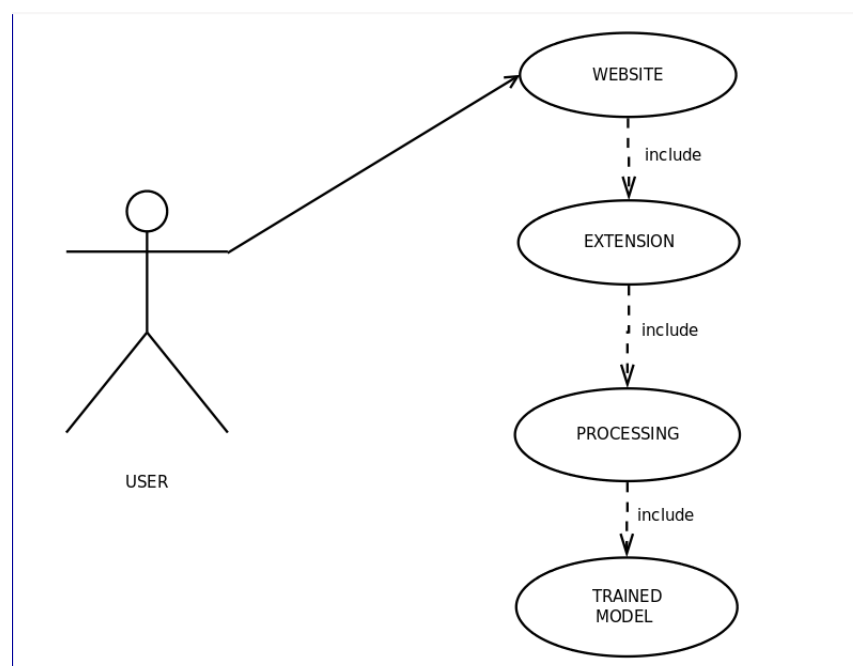


Figure 11.1: Use Case Diagram

Description of Use Case diagram

The above representation of our project is in the form of a Use Case diagram. This diagram represents the flow and the working of each model in the diagram. The diagram depicts that when the user accesses the website through the Google Chrome, the extensions working comes into the picture. The extension processes the entire webpage differentiating the images and blurring the positive images (nude images). This work is done on the basis of a trained model which the extension works upon to complete the desired action.

Chapter 12

Modules

The project is divided into 4 models which are as follows:

12.1 Module 1 : Classification Of Images Using CNN

In this module, training is done using the CNN model to identify the bare images and the images of advertisement. Convolution neural networks have different models. Few of them are enlisted below:

- LeNet-5
- AlexNet
- VGG16 and VGG19
- Inception-v1 and Inception-v3 and Inception-v4
- ResNet-50
- Xception

Our model proposes the use of VGG16/VGG19 model. Training is done on a dataset of approximately 25000 images for more accurate and correct results. The dataset of 25000 images was not easy to find for one reason; it was not readily available on any website like Kaggle for example which is a source of various types of datasets. So now we had to prepare our own dataset. For this purpose we used a concept known as BeautifulSoup or more commonly known as Scrapping. Scrapping basically is downloading all the images from a page and saving it into our computer. The name scrapping suggests that we do not download each and every content from the web page, but only scrap out the content we need. To achieve the mentioned task we wrote a code which runs in a loop format. It scans each and every link from a text file, visits that link and download whatever image is present on that link. We found this list of links from a GitHub account. This collection of data takes several weeks as, the amount of data is huge. Also one problem you might face is the blank links. As these links are random, it is possible that the image that was previously present is not available any more. When the program encounters such a problem it stops over there and we have to remove that specific link from the text folder in order to scan the rest of the image links. So it requires your attention on a periodic basis even though it is a self-running

program. Now that the data is collected we had to train the model from the same. For this we cannot use our regular computers because their configuration cannot handle such a large amount of data. So you will have to use a computer with a GPU. We will use a computer containing NVIDIA system in order to train the model.

12.2 Module 2 : Optimizers and Activation Function Analysis

12.2.1 Activation Function:

Activation functions are really important for an Artificial Neural Network to learn and make sense of things which are really complicated. They introduce non-linear properties to our Neural Network. Their main purpose is to convert an input signal of a node to an output signal. That output signal now is used as an input in the next layer in the stack. Specifically in an Artificial Neural Network we do the sum of products of inputs (X) and their corresponding Weights (W) and apply an Activation function $f(x)$ to it to get the output of that layer and feed it as an input to the next layer. A Neural Network without Activation function would simply be a Linear regression Model, which has limited power and without activation function our Neural network would not be able to learn and model other complicated kinds of data such as images, videos, audio, speech etc. That is why we use Artificial Neural network techniques such as Deep learning to make sense of something complicated, high dimensional, non-linear -big datasets, where the model has lots and lots of hidden layers in between and has a very complicated architecture which helps us to make sense and extract knowledge from such complicated big datasets.

Types of Activation functions -

- Sigmoid or Logistic
- Tanh Hyperbolic tangent
- ReLu -Rectified linear units
- Elu
- Hard sigmoid
- Selu
- Softsign
- Softmax
- Softplus
- Exponential
- Linear

Out of the 11 types of activation functions mentioned above 3 are the most popular and most widely used functions.

1. **Tanh :**

Its mathematical formula is $f(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)}$. Now its output is zero centred because its range is between -1 to 1 i.e. -1 < output < 1. Hence optimization is easier in this method hence in practice it is always preferred over Sigmoid function. But still it suffers from Vanishing gradient problem.

2. **ReLU :**

Rectified Linear units: It has become very popular in the past couple of years. It was recently proved that it had 6 times improvement in convergence from Tanh function. Its just $R(x) = \max(0, x)$ i.e. if $x < 0$, $R(x) = 0$ and if $x \geq 0$, $R(x) = x$. Hence as seeing the mathematical form of this function we can see that it is very simple and efficient. A lot of times in Machine learning and computer science we notice that most simple and consistent techniques and methods are only preferred and are best. Hence it avoids and rectifies vanishing gradient problem. Almost all deep learning Models use ReLU nowadays. But its limitation is that it should only be used within Hidden layers of a Neural Network Model. Another problem with ReLU is that some gradients can be fragile during training and can die. It can cause a weight update which might result in no activation of any data point, resulting into Dead Neurons.

3. **Sigmoid Activation function:**

It is an activation function of form $f(x) = \frac{1}{1 + \exp(-x)}$. Its Range is between 0 and 1. It is an S shaped curve. For our model we are specifically using Sigmoid function as it gives the best accuracy of the test cases when it works along with the optimizer named SGD.

12.2.2 Optimizers

Optimizers tie together the loss function and model parameters by updating the model in response to the output of the loss function. In simpler terms, optimizers shape and mould your model into its most accurate possible form by editing the weights. The loss function is the guide to the terrain, telling the optimizer when its moving in the right or wrong direction. It is impossible to know what your models weights should be right from the start. But with some trial and error based on the loss function, you can end up getting the right and the best values eventually. Here are few of the optimizers that are present in today's Machine Learning world:

1. Gradient Descent:

This algorithm is used across all types of Machine Learning (and other math problems) to optimize. Its fast, robust, and flexible. Here's how it works:

- 1] Calculate what a small change in each individual weight would do to the loss function (i.e. which direction should the hiker walk in)
- 2] Adjust each individual weight based on its gradient (i.e. take a small step in the determined direction)
- 3] Keep doing steps 1 and 2 until the loss function gets as low as possible.

2. Stochastic Gradient Descent:

Instead of calculating the gradients for all of your training examples on every pass of gradient descent, it's sometimes more efficient to only use a subset of the training examples each time. Stochastic gradient descent is an implementation that either uses batches of examples at a time or random examples on each pass. We specifically haven't included the formal functions for the concepts in this post because we're trying to explain things intuitively. For more insight into the math involved and a more technical analysis, this walkthrough guide with Excel examples is helpful.

3. Adagrad:

Adagrad adapts the learning rate specifically to individual features: that means that some of the weights in your dataset will have different learning rates than others. This works really well for sparse datasets where a lot of input examples are missing. Adagrad has a major issue though: the adaptive learning rate tends to get really small over time. Some other optimizers below seek to eliminate this problem.

4. RMSprop:

RMSprop is a special version of Adagrad developed by Professor Geoffrey Hinton in his neural nets class. Instead of letting all of the gradients accumulate for momentum, it only accumulates gradients in a fixed window. RMSprop is similar to Adagrad, which is another optimizer that seeks to solve some of the issues that Adagrad leaves open.

5. Adam:

Adam stands for adaptive moment estimation, and is another way of using past gradients to calculate current gradients. Adam also utilizes the concept of momentum by adding fractions of previous gradients to the current one. This optimizer has become pretty widespread, and is practically accepted for use in training neural nets.

It's easy to get lost in the complexity of some of these new optimizers. Just remember that they all have the same goal: minimizing our loss function. Even the most complex ways of doing that are simple at their core. But for our model the best that works is SGD along with sigmoid. So we have used that in our model as well. This gives us an accuracy of 87.29

12.3 Module 3 : Testing of Trained Model

Once we have trained the model on the bare skinned images, the next step is to test the model that is to see whether the model is providing the appropriate output we are hoping for or not. We will be passing both a set of bare skinned images and a set of any image which does not pose nudity of any kind. If the model detects any bare skinned image it will assign a value of one to that image, if not then the value greater than one will be assigned to it. For testing we do not require any special computer, as the testing data is comparatively very small in front of the training data and our personal computers can handle it.

12.4 Module 4 : Creating the required extension

We will try to put our trained model on the web browser itself so that it works on all the websites. Thus being said we will provide it in the form of an extension which when activated would work for all the websites you visit while it is activated. Our technology will try to cover the maximum websites possible. The extension would work in Google Chrome as any other extension you may use or see. You will have to download and add it to the Google Chrome and give it the required permissions to scan the pages you visit while surfing on the internet. This extension will consist of an already trained model and would not require training it again. So once it is installed and given permissions, it is ready to be used. Disabling the link is the next step that follows in our process. Just blurring the images does not serve the entire purpose, as clicking on image may redirect us to some other website. So the extension created will not only blur the image but also disable the links.

Chapter 13

Future Scope

Our scope is to blur all the extreme images and deactivating their links. This means that it will scan the entire web page and find all the extreme images present on that page. Then we will blur those images before they are loaded, so that the children would not be able to see them. This will be done in a 2 step process first it will identify the image as nude or non-nude image. If identified as nude it needs to be blurred, if not keep the image as it is and display it. Apart from it, we will also disable any clickable links present on or below those images. This ensures double protection of children from the disturbing images and links that are easily available on en-number of websites we visit from day to day. Our model will try to do this in the most minimum amount of time possible.

Annexture

Gantt Chart

GANTT CHART TEMPLATE

Smartsheet Tip → A Gantt chart's visual timeline allows you to see details about each task as well as project dependencies.

PROJECT TITLE Real Time Bare Skinned Image Filtering using CNN
PROJECT GUIDE Prof. Jaya Gupta

COMPANY NAME A.P.Shah Institute of Technology
DATE 8/27/19

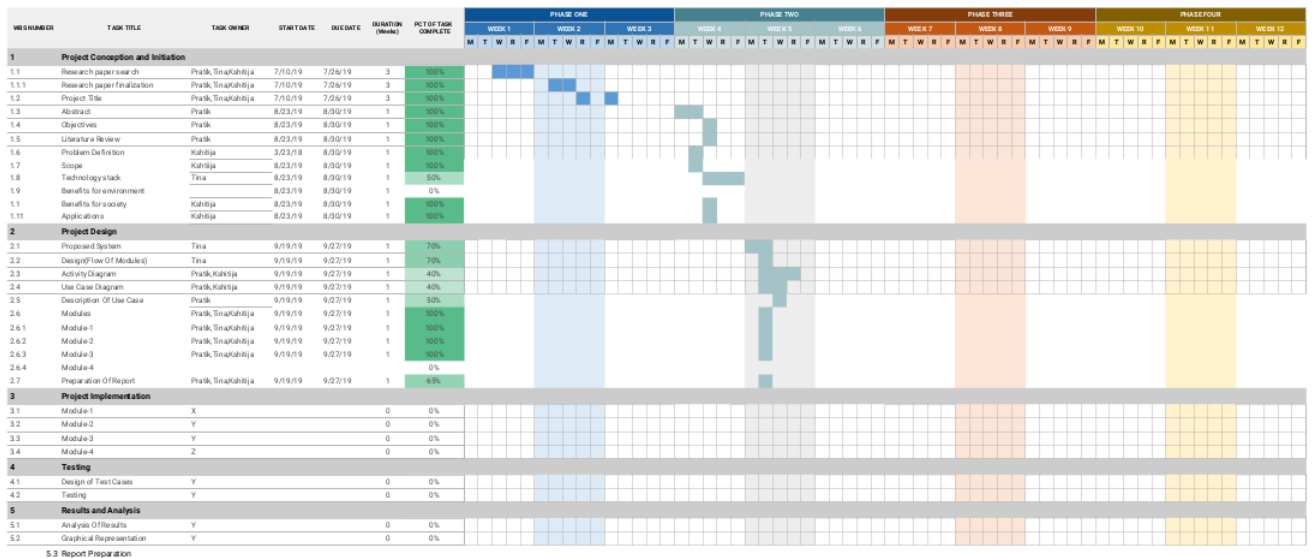


Figure 13.1: Gantt Chart

Bibliography

- [1] 2017 9th International Conference on Knowledge and Systems Engineering(KSE) Advertisement Image Classification Using Convolutional Neural Network
- [2] <https://machinelearningmastery.com/how-to-develop-a-convolutional-neural-network-to-classify-photos-of-dogs-and-cats/>
- [3] <https://medium.com/@alkeshab/face-detection-using-opencv-in-google-colaboratory-a7529a2bb921>
- [4] <https://www.analyticsvidya.com/blog/2017/06/architecture-of-convolution-neural-networks-simplified-demystified/>
- [5] <https://neurohive.io/en/popular-networks/vgg16/>
- [6] <https://www.pyimagesearch.com/2018/09/24/open-cv-face-recognition/>
- [7] <https://www.willberger.org/cascade-haar-explained/>

Acknowledgement

We have great pleasure in presenting the report on "**Real Time Bare Skinned Image Classification Using CNN**". We take this opportunity to express our sincere thanks towards our guide **Prof.Jaya Gupta** & Co-Guide **Prof.Amol Kalugade** Department of Computers, APSIT thane for providing the technical guidelines and suggestions regarding line of work. We would like to express our gratitude towards his constant encouragement, support and guidance through the development of project.

We thank **Prof.Sachin Malve** Head of Department,Computers, APSIT for his encouragement during progress meeting and providing guidelines to write this report.

We thank **Prof.Amol Kalugade** BE project co-ordinator, Department of Computers, APSIT for being encouraging throughout the course and for guidance.

We also thank the entire staff of APSIT for their invaluable help rendered during the course of this work. We wish to express our deep gratitude towards all our colleagues of APSIT for their encouragement.

Student Name1:Pratik Jain
Student ID1:16102060

Student Name2:Tina Shah
Student ID2:16102019

Student Name3:Kshitija Shah
Student ID3:16102027