CAPSTONE PROJECT REPORT

A thesis submitted in partial fulfilment of the requirements for the degree of Bachelor of Technology in Data Science (Business Analytics).

“Artificial Intelligence in Research of Human Hair & Product Efficacy Determination”

By

Advait Thergaonkar (SAP ID: 70091018060)

under the guidance of

Prof. Pranav Nerurkar (Academic Mentor)

Department of Data Science

MPSTME, Mumbai

NMIMS University

October, 2021.

CERTIFICATE

This is to certify that the thesis entitled “Capstone Project Report” is a bonafide work of “Advait Thergaonkar (SAP ID: 70091018060)” submitted to the NMIMS University in partial fulfilment of the requirement for the award of the degree of “Bachelor of Technology” in “Data Science (Business Analytics)”.

Prof. Sarada Samantaray
Associate Dean
NMIMS University

ACKNOWLEDGEMENT

It is with a feeling of great pleasure that I express my sincere thanks to Prof. Pranav Nerurkar, for his endless support and consistent efforts in implementing this project. It is with utmost gratitude that I thank Prof. Sarada Samantaray (HOD, Data Science Department) and Dr. Siba Panda for providing their expertise and guidance throughout my academic journey. I am extremely thankful to all the other professors and staff members of the Data Science department for constantly inspiring me and enabling me to successfully complete this capstone project.

APPENDIX

INTRODUCTION

Quality testing of products is a boon for any industry that produces edible items, personal care products or any product that a human uses in his daily lifestyle. In the cosmetic or the personal care industry quality testing of products is one of the most important steps as these products would be applied on the human body and hence have to go through multiple testing rounds wherein their efficacy, long term effects, stability and their reactions to different application areas is tested for. After this testing is done reports are generated and if any changes are required, they are made and one the product meets the safety and regulatory standards the product is sent for mass manufacturing and retail.

In the hair care industry there are several time consuming steps involved in the testing of products and determination of their efficacy, one of the processes is, scientists when formulating and creating hair products like, shampoo's, bleaches, conditioner's and other products, they try these products on small hair samples wherein they check and evaluate certain parameters like volume, dryness, frizzyness, colour, shine and several other parameters, they rate these parameters on a fixed scale and then develop the product if certain changes are required based on the ratings found.

What our solution does is that it automates the rating process as it uses artificial intelligence to study the hair image provided and then give ratings depending on the reaction that hair shows when a particular product is applied, after the product is applied a report is generated with all the predicted ratings and the input image and it can be used for the scientists later usage to make any modifications to the product.

One of the main unique selling points of our product is its affordability, easy usage, lack of proprietary hardware requirements and mainly reducing the time taken in testing of products so that more productive work like innovation can be done.

CURRENT SCENARIO

Currently the cosmetic industry is a vast fast growing industry. There are firms spread around the globe ranging from developed nations like the United States of America to developing nations like India and other subcontinent nations. Although these firms are from different backgrounds, have different business ethics and different styles of working, what they have in common is that they need physical devices and software's for the testing of the products they make.

Currently there are proprietary devices which determine the efficacy of the product when applied on hair, these products are Cocam Evaluator, Gloss meter and Visual analysers. These devices range anywhere starting from a few lakhs of rupees to anywhere in multiples of tens of lakhs. There are other tests where Human trained panels are used for evaluating the sensory parameters. These evaluations again costs lakhs of rupees and ethics committee clearances many a times to start the studies.

Although these machines and tests are fairly accurate, they aren't as affordable for a newly found start-up, even though these firms don't want to compromise on the quality and testing of their products they aren't able to afford machines and devices which do quality testing for them. Although the big firms say they want to support the new entrants in the industry there are hardly any inexpensive quality testing devices that are in the market which are affordable for small firms.

For quality testing of products, in small to medium scale firms many processes are still a manual task undertaken by scientists or special workforce, this team of quality testers manually undertake the quality testing of products without the usage of state-of-the-art technology, not a fault of the firm but because such machines and devices are very expensive for such firms to afford the quality testing is done manually leading to a huge waste of time, resources, accuracy and efficiency. These methods, though inexpensive and light on the company finances, are huge time wasters and cost the firm a lot of quality work in the long term.

MOTIVATION

Since economical constraints have restricted start-ups, small and medium companies to evaluate the cosmetic products for their efficacy an alternative is imperative to provide them the database as well as help the consumers get properly evaluated products which will improve the quality of products in the market as well as reduce the risk and side effects of the products applied on the body.

The project aims at creating a quality and efficacy evaluation product which is affordable for any firm, individual or organization. They can use this software without having any specialized hardware and can be used using just a smartphone, iPad/tablet or any portable device having a camera.

Since this does not use any specialized hardware its cost price is extremely low and is in an affordable range for customers. As our product is an outcome of a CNN model, we have solidly defined accuracy metrics which we can show to our customers, and prove our case.

When customers use this product, it will lead them to reduce their manual labour as our product has batch image inputs and then generate reports accordingly for each image. This will lead to them not only saving time in generating reports but also improving the productivity in labs.

This is a serious problem as firm’s are looking for solutions to reduce time, improve efficiency, accuracy and quality of work, but aren’t able to come up with efficient solutions which can help them nullify the problems mentioned.

LITERATURE REVIEW

The volume of research in Artificial Intelligence and Machine Learning is ever increasing and its application in the cosmetic and personal care industry is increasing too. Currently there are a few applications of artificial intelligence in the cosmetic and personal care industry like, custom cosmetic products recommendation systems, reaction of the body after applying certain products and prediction of how a product would react on a certain skin or hair type, at the same time there are various attempts being made on making quality testing devices of cosmetic products using artificial intelligence.

The global hair care market is projected to grow from $77.15 billion in 2021 to $112.97 billion in 2028 at a CAGR of 5.6% in forecast period, 2021-2028[14].

Hair care involves the utilization of different products to cleanse, maintain, and improve the overall texture and appearance of the hair. These products generally include gels, oils, rinses, tonics, masks, serums, shampoos, dressings, conditioners, and sprays. Presently, several companies are introducing novel and innovative products that are silicone and sulphate-free to keep up with the continuously evolving hair trends. Moreover, inflating disposable incomes, improving living standards, and the increasing adoption of a modern approach to grooming practices are positively influencing the demand for hair care products worldwide. Due to the diversity and complexity of hair types and concerns, several companies are offering customized products to meet specialized consumer needs, which range from cleansing, styling, softening, strengthening to scalp care, damage repair, thermal protection, frizz control, colour retention, moisturization, and bond-building. This, along with the increasing awareness about hair-related issues due to climate change, air pollution, and other environmental factors, represents one of the significant factors fuelling the market growth. Apart from this, with a considerable shift towards sustainable manufacturing and chemical-free products, the leading manufacturers are

incorporating natural oils, aloe vera, cocoa butter, and extracts of other organic ingredients [15]

Generally, in the industry sensory evaluation of hair tresses is done by using different panels who evaluate the visual appearance of the shampoo, the texture of the foam and the touch of the hair after rinse are all criteria to be considered and evaluated. The most suitable method to evaluate hair care products sensory performance is the sensory evaluation which permits to score the different sensory attributes with the help of a panel of trained judges. However, this method is time consuming and sometime it is not really relevant for product development and so without all the data to assure safety. For this reason, some instrumental investigation is preferred and can be made as well in order to save time on formulation screening.[16]

Recently, CNN, one of the most commonly used deep learning methods, becomes popular because of its outstanding performance on object recognition [8], face information extraction [9], image retrieval [10] and lots of other tasks in computer vision areas. CNN is an end-to-end learning architecture to leverage collected information [11]. It trains the description features directly from raw data, which means decreasing the human effort of designing features. Meanwhile, CNN is able to deal with large scale data because of its sufficient capacity and reasonable model structure. Therefore, CNN solves two of the most important problems of traditional algorithms in computer vision, the large efforts of feature designing [12] and the weak ability to deal with big data [13]. These advantages of CNN ensure its excellent performances.

Gabriela Daniels et al used machine learning to calculate hair assembly features they used a typical hair fibre of individuals from European descent (Caucasian hair). The first objective of this study was to apply computer vision and machine learning techniques to quantify the effects of haircare treatments on hair assembly and to identify correctly whether unknown tresses were treated or not.

The outcomes of these analyses were compared with the expected theoretical behaviour (ETB) of the treatment and were envisaged to offer more accurate methods for the assessment of hair assembly.

The second objective was to explore and compare the performance of human assessment with that obtained by AI algorithms. In order to do that, after performing automatic image analysis, we carried out a sensory test with naive assessors and an online survey based on tress images from the employed data set[3].

This study differentiates from earlier studies therein that we undertake a holistic approach in creating a product that is in-expensive, widely accessible, usable on a portable device and does not need specialised hardware for it to function.

Additionally the study ultimately aims to make a simple to use user interface that can be installed on a device and be used by a professional and does not need any training to understand how to use the application.

Convolutional Neural network has been used for solving and creating most of the problems in the research papers referred so far.

- **Architecture of a Convolutional Neural Network**

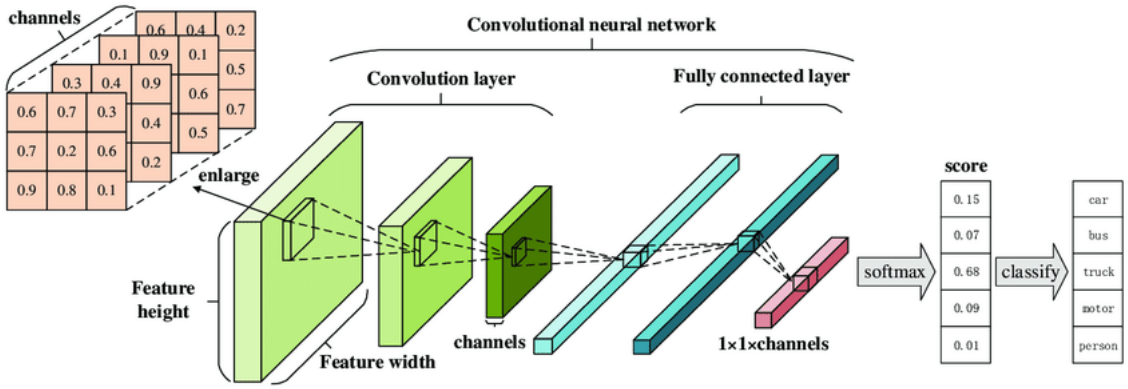


Fig 1: Architecture of a typical CNN Model [6]

The traditional CNN structure is mainly composed of convolution layers, pooling layers, fully connected layers, and some activation functions. Each convolution kernel is connected to the part of feature maps. The input is connected to functions. Each convolution kernel is connected to the part of feature maps. The input is connected to all of the output elements in the fully connected layer.[6]

- **Convolutional Layer**

  Convolution is the first layer to extract features from an input image. Convolution preserves the relationship between pixels by learning image features using small squares of input data. It is a mathematical operation that takes two inputs such as image matrix and a filter or kernel. [7]

- **Strides**

  Stride is the number of pixels shifts over the input matrix. When the stride is 1 then we move the filters to 1 pixel at a time. When the stride is 2 then we move the filters to 2 pixels at a time and so on. [7]

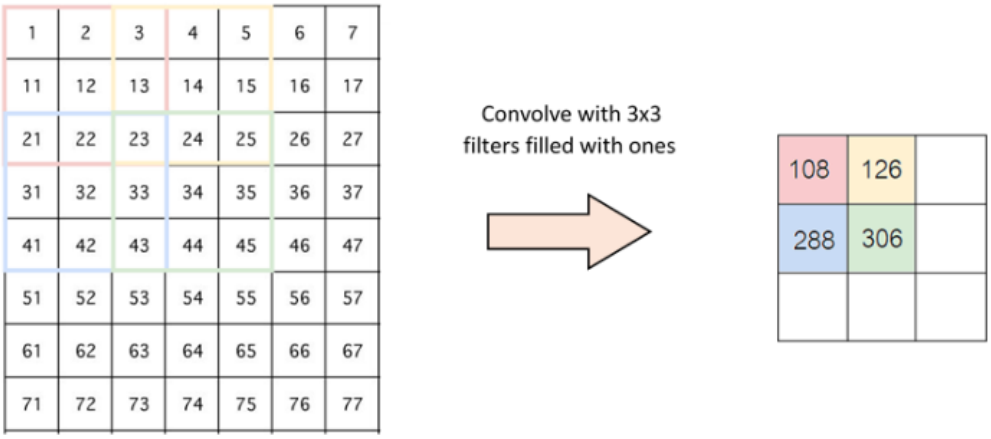


Fig 2: Stride of 2 Pixels (6)

- **Pooling Layer**

  Pooling layers section would reduce the number of parameters when the images are too large. Spatial pooling also called subsampling or down sampling which reduces the dimensionality of each map but retains important information. [7]

  There are 3 types of pooling; Max Pooling, Sum Pooling and Average Pooling.

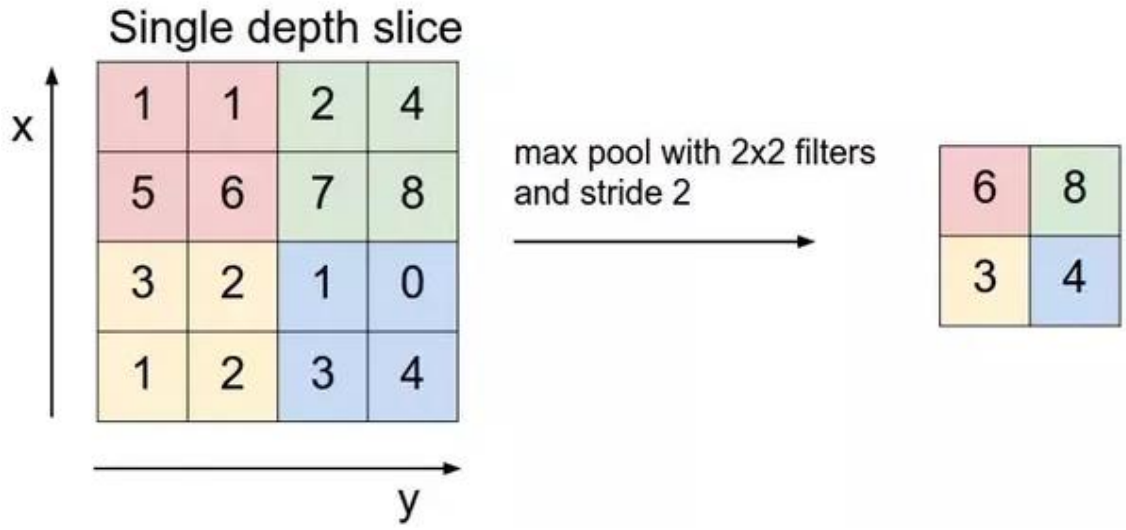


Fig 3: Max Pooling (6)

- **Fully Connected Layer:**

  The layer we call as FC layer, we flattened our matrix into vector and feed it into a fully connected layer like a neural network.[7]

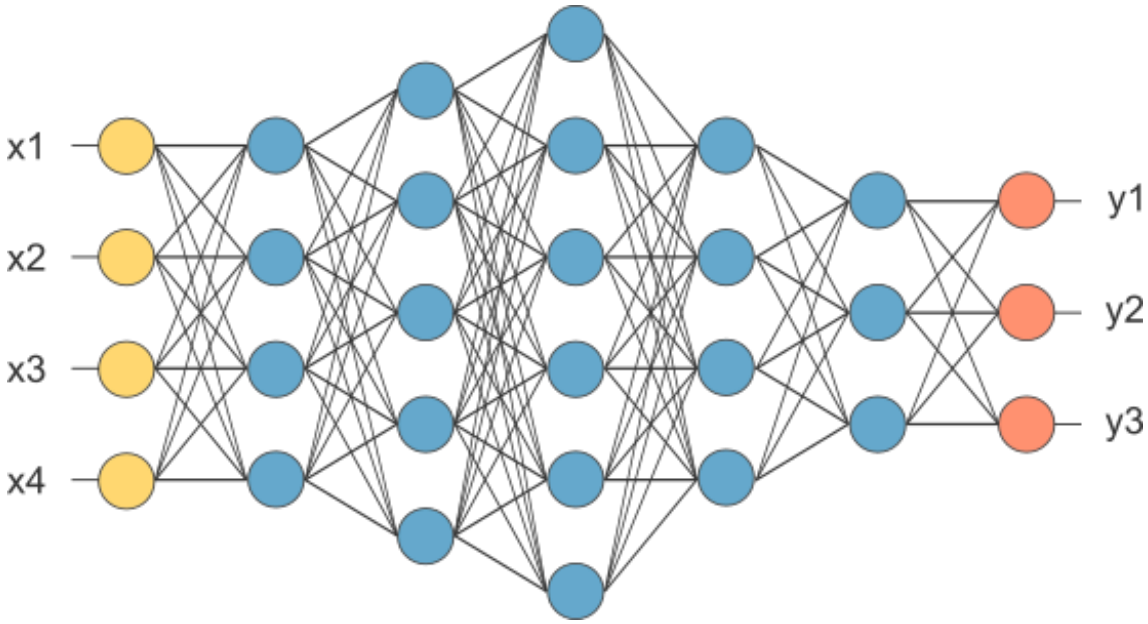
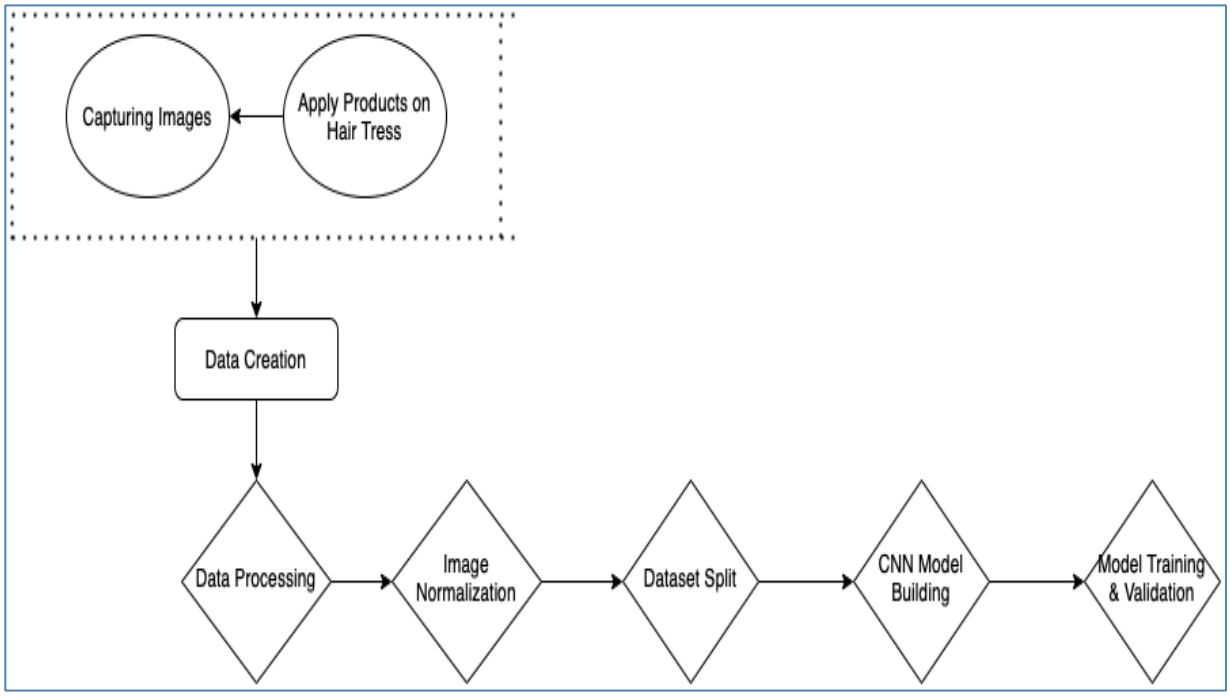


Fig 4: Example of a Fully Connected Layer (6)

Hence though there is dearth of proper systems for sensory evaluation of hair care products and much work is needed to create a cost efficient , time saving technology to evaluate these products.

METHODOLOGY

I. Algorithm of the Model Building



In building the model for the hair rating system there are several steps involved.

- The first step of the algorithm is creating the dataset, here images were captured of the hair tresses on which the hair care product was applied.
- The second step was to processes the data, here images were cropped and scaled to a size best suited for our CNN model.
- The third stem was to normalize all the pixel values of our images.
- The fourth step was to split the dataset into training and testing parts.
- The fifth was to build the CNN model and train.
- The last step was to evaluate the model and save it for future use.

II. Data Creation

A. Creation of the Image Dataset

Creating data or editing existing data according to the needs of a project is one of the most important steps in actually going ahead in the creation of a product.

Being an image-based project required us to take several input hair images for training, testing and validating purposes. Our dataset consisted of hair images of mainly two types, ‘Bleached Hair’, ‘Virgin Hair’ and a few images of ‘Conditioned Hair’.

Three images of each hair tress, from three different angles (−45°, 0°, +45°) were taken at each treatment point *t*0, *t*1 and *t*2, resulting in 9 images per tress (image size: 3264 pixels × 2448 pixels). Each tress was combed twice just before the photograph was taken. A total of 720 images were then used for analysis [2]

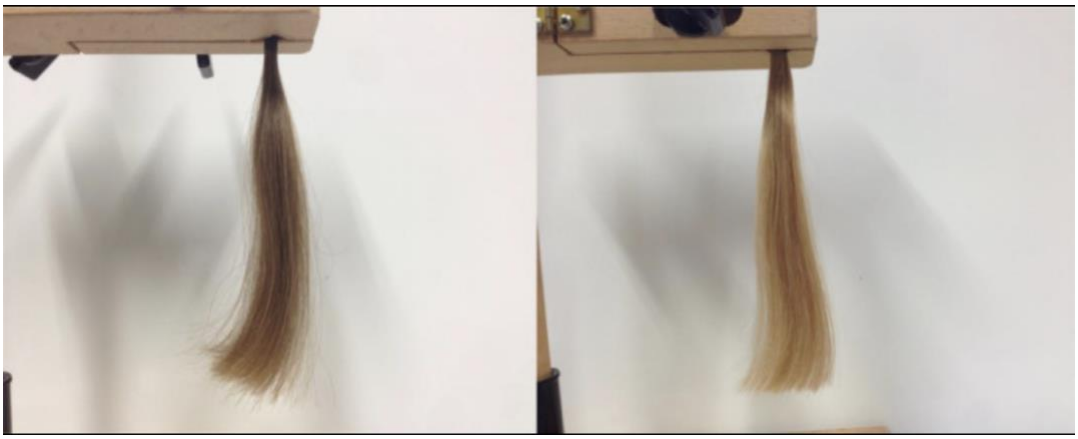


Fig 5: Example of images in the Data Set

For hair tresses applied with conditioner, serum and other hair products a similar approach was followed for creating more images.

Fig 6: Lab image for hair tress

B. Creation of the Ratings Dataset

After the creation of the image dataset, the next step was rating the selected images on various parameters.

The parameters chosen were in accordance with the requirements of the firm R&D Global Consultants Pvt Ltd. The image data set was evaluated by a trained panel of domain experts.

In this study, a first panel composed of 12 trained panellists of the company. The company has developed a lexicon for characterizing sensory characteristic of shampoo applied on hair tress.

This evaluation provides a total of 5 attributes divided as follows:

1. Hair Volume
2. Hair Shine
3. Hair Smoothness
4. Hair Frizzyness
5. Hair Flyaway

Parameters Volume, Shine and Smoothness are illustrated by references (low = 0 and high = 5 of the attribute).

Parameters Frizzyness and Flyaway are illustrated by (Low = 0 is a better rating and High = 5 is a bad rating).

C. Details of Parameters

The parameters were chosen keeping common hair problems in mind, these are caused by improper care, handling and different chemical treatments. Hair damage can occur in form of by any of the reasons below:

i. Split Ends: Excessive split ends near the hair tip shows damaged hairs due to excess shampooing and brushing. Split ends depict cuticle erosion and splitting of cortex in hair.[17]
ii. Breaking of Hair: Chemical treatment of hair causes surface modification, reduced hair strength and increased porosity which leads to breaking of hair intermittently. [18]
iii. Rough and Dull Hair: Ageing of hair and excessive hair treatments leads to deterioration of the cuticular region from root to tip thus giving roughness and dullness to hair. [19]
iv. Discolored Hair: Sun rays causes bleaching of brown hair and browning of black hair. [20]
v. Tangling: Reduces tensile strength and gloss, increases friction and tangling. [21]
vi. Thinning of Hair: Ageing causes reduction in the diameter of hair. [22]
vii. Static Charge: Can happen due to dryness, Chlorine from swimming pool also causes increased hair friction.[23]

All these damages are caused due to hair protein damage. Various amino acids are indicators for hair protein damage.

The use of Hair care products can damage hair indirectly as a result of fiber abrasion occurring when hair is rubbed against each other during cleaning.

Experiments indicate that exposure to shampoos can have a deleterious effect on hair structure over time. The deleterious effect is seen in the form of reduction in hair volume, loss of hair gloss, Loss of hair smoothness, increased fly away and fizziness of the products.

D. Final Dataset of Target Variables

After rating all the images on the set scale with the help of a domain expert a database was created with all the 888 images rated and those 5 chosen parameters.

While creating the dataset each images file name was linked to its corresponding row to prevent any mixing or corruption of data.

target_variables

Type	Washes	Image_Name	Volume	Shine	Smoothness	Frizzyness	Fly away	
Bleached	No Wash	101.JPG	4	2	4	3	3	
Bleached	No Wash	102.JPG	3	2	3	2	2	
Bleached	No Wash	103.JPG	4	2	3	3	4	
Bleached	No Wash	201.JPG	4	3	4	1	4	
Bleached	No Wash	202.JPG	4	4	3	3	4	
Bleached	No Wash	203.JPG	3	4	2	3	5	
Bleached	No Wash	301.JPG	3	3	4	1	2	
Bleached	No Wash	302.JPG	4	2	4	1	2	

E. Data Processing

After gathering and arranging all the required images for building the convolutional neural network (CNN) model we pre-process the data. This data pre-processing involved cropping out the parts of the image which did not contain the hair tress, the next step involved was reducing the size of the image to fit the constraints of our network.

- **Cropping the Images:**
  For cropping images and keeping only the interested areas so that the algorithm does not learn any noise an AI based tool named Rembg (Background Removal Tool) was used.
  Rembg is a python tool used to remove the backgrounds from images and videos.

- **Usage of Rembg:**
  Rembg is a very simple tool to use, it can be installed from PyPi using PIP. After installation of rembg we passed in all our required images and cropped them to our liking.
  Since rembg is an AI based image cropper there were some images which were not cropped to our liking and hence were discarded and finally we were left with 888 number of images.

```python
from rembg.bg import remove
import numpy as np
import io
from PIL import Image

input_path = 'input.png'
output_path = 'out.png'

f = np.fromfile(input_path)
result = remove(f)
img = Image.open(io.BytesIO(result)).convert("RGBA")
img.save(output_path)
```

Fig 7: A general usage of the REMBG

Fig 8: General Image Captured and Image after cropping and using REMBG

- **Reshaping and normalizing the Images**:

  After the images were cropped and the background removed the images were reduced to a size of (256, 256, 3), after reshaping, the image pixels were normalized to a value between 0-1.

- **Splitting the Dataset into training and testing**:

  Once the pre-processing was done the data was split into two sets of training (90%) and testing (10%) with a random state set to 42.

F. Convolutional Neural Networks (CNN) Model Building

For the model building of our proposed algorithm of the image-based evaluation system a convolutional neural network was built on the dataset of images created.

The convolutional neural network was built using the programming language python. Jupyter Notebook and Google Collaboratory were the editors used.

After the images were processed, the ratings dataset built and the data split into training and testing sets the model had to be built.

The libraries used in building the model were OpenCV, Numpy, Pandas, Keras, Tensorflow, SKlearn and OS.

```
import os
import cv2

import numpy as np
import pandas as pd

import matplotlib.pyplot as plt

import keras
import tensorflow as tf

from sklearn.model_selection import train_test_split
from tensorflow.keras.layers import Dense, Input, Dropout,Flatten, Conv2D
from tensorflow.keras.layers import BatchNormalization, Activation, MaxPooling2D

from tensorflow.keras.models import Model, Sequential
from tensorflow.keras.optimizers import Adam
from tensorflow.keras.callbacks import ModelCheckpoint
from keras.callbacks import ModelCheckpoint, EarlyStopping
from keras.models import load_model

from tensorflow.keras.utils import to_categorical

import time
```

Importing the necessary libraries

Once the libraries were imported and prior to the model being compiled some hyperparameters had to be set.

Setting the model checkpoints and call-backs, for the model checkpoint only the best model was saved where the validation tests, mean absolute error was checked and the epoch with the lowest MAE was set as the best model and eventually saved. The patience was set to 15 wherein if the validation MAE did not improve for 15 consecutive epochs, then the training would stop and the best model be saved.

- **Model Architecture**

  The CNN model had a total of 536,879,617 parameters. The hidden layers had an input shape of (256,256,3). Once the image is input its batch got normalized with an activation of ‘relu’ and max pooling of (2,2) once both the hidden layers were made the input was flattened and passed onto an output layer with an activation of ‘linear’ as our task was a regression task.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 256, 256, 32)      2432
_________________________________________________________________
batch_normalization (BatchNo (None, 256, 256, 32)      128
_________________________________________________________________
activation (Activation)      (None, 256, 256, 32)      0
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 128, 128, 32)      0
_________________________________________________________________
dropout (Dropout)            (None, 128, 128, 32)      0
_________________________________________________________________
flatten (Flatten)            (None, 524288)            0
_________________________________________________________________
dense (Dense)                (None, 1024)              536871936
_________________________________________________________________
batch_normalization_1 (Batch (None, 1024)              4096
_________________________________________________________________
activation_1 (Activation)    (None, 1024)              0
_________________________________________________________________
dropout_1 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_1 (Dense)              (None, 1)                 1025
=================================================================
Total params: 536,879,617
Trainable params: 536,877,505
Non-trainable params: 2,112
_________________________________________________________________
```

Model Architecture

This model was compiled using an optimizer ‘adam’ and a loss function of ‘mean absolute error’.

Once the model was compiled, the next task was to input all the five parameters and get a model for each parameter respectively.

- Parameter Volume
  A model was built for the volume parameter and a validation MAE was achieved of **0.5496** on the test set.
- Parameter Shine
  A model was built for the shine parameter and a validation MAE was achieved of **0.5033** on the test set.
- Parameter Smoothness
  A model was built for the smoothness parameter and a validation MAE was achieved of **0.3470** on the test set.
- Parameter Frizzyness
  A model was built for the frizzyness parameter and a validation MAE was achieved of **0.4645** on the test set.
- Parameter Fly Away
  A model was built for the flyaway parameter and a validation MAE was achieved of **0.4645** on the test set.

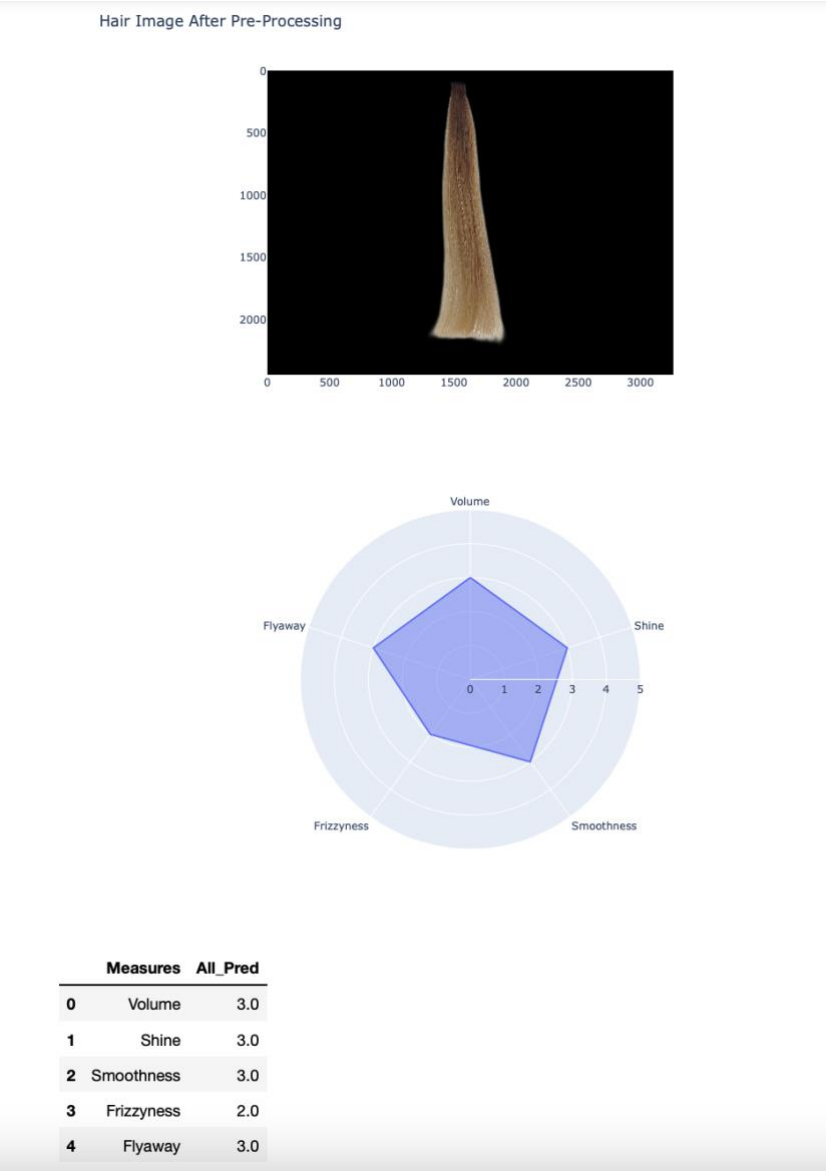
  Hence the respective five Models were created for each of the parameters respectively using the CNN techniques for the built dataset.

G. End User Application

After the models were built and were ready to be deployed, the next step was to create a user interface for the customers to use.

For the web application a website was made wherein you could upload and capture images and then get the ratings and a report generated.

The report consists of the input image and a web graph which is the norm of the industry.



	Measures	All_Pred
0	Volume	3.0
1	Shine	3.0
2	Smoothness	3.0
3	Frizzyness	2.0
4	Flyaway	3.0

Output as seen for the User

RESULTS AND DISCUSSION

Cosmetic Industry is the most fast-growing industry in the FMCG sector. The industry is divided into hair care and skin care industry. The global hair care market is projected to grow from $77.15 billion in 2021 to $112.97 billion in 2028 at a CAGR of 5.6% in forecast period, 2021-2028. Various Research & Development and manufacturing companies in cosmetic sector are today trying to find several cheaper and time saving evaluation methods to help in quick creation of formulations to cater the consumers. Hence keeping with this demand from the research and development company the opportunity was provided to create an application for quick sensory evaluation of hair care products.

To make the model for the application a dataset was created by clicking images of hair tresses. The images clicked were in two batches first was virgin hair which was combed twice and then bleached hair which were combed twice and washed for multiple times. After the images were clicked a panel of 12 experts from the company rated the images on a set scale of 1-5 on 5 primary parameters namely, 'Volume', 'Shine', 'Smoothness', 'Frizzyness' and 'Fly Away'. Parameters Volume, Shine and Smoothness are illustrated by references (low = 0 and high = 5 of the attribute). Parameters Frizzyness and Flyaway are illustrated by (Low = 0 is a better rating and High = 5 is a bad rating).

These images were scaled to a resolution of (256,256), cropped to remove any noise and then normalized between a pixel value of 0-1.

After the image data set processing a. CNN model was built for the five parameters and trained on the dataset with these 5 parameters. This convolutional neural network model had of 536,877,505 trainable parameters and 2112 of non-trainable parameters. A total of 5 models were built one for each of the parameters and then deployed for usage.

The model error rate was calculated using the mean absolute error on the testing set. The values for each of the models were as follows:

- Parameter Volume

  A model was built for the volume parameter and a validation MAE was achieved of **0.5496** on the test set.
- Parameter Shine

  A model was built for the shine parameter and a validation MAE was achieved of **0.5033** on the test set.
- Parameter Smoothness

  A model was built for the smoothness parameter and a validation MAE was achieved of **0.3470** on the test set.
- Parameter Frizzyness

  A model was built for the frizzyness parameter and a validation MAE was achieved of **0.4645** on the test set.
- Parameter Fly Away

  A model was built for the flyaway parameter and a validation MAE was achieved of **0.7086** on the test set.

In the created models the principal used for evaluation was Mean Absolute error wherein lower the value of error the better the model is. The ideal MAE value should be below 1.

After the model creation a front-end web application was built wherein the user can upload all the images required for testing and get the reports generated.

The model was then tested in the R&D lab of the company by the scientists and it was observed that it was giving good result matching with their estimation. It was observed that the above model helped in reduction of 3 days of manual labour and gave the results in 10 minutes.

CONTRIBUTIONS

The project dealt in this dissertation was from the aspect of creating a cheaper evaluation platform for the cosmetic industry to help the sensory evaluation of the hair care products. Since this project was not done by anyone in the world the company R &D Global came up with the challenge of providing a solution to this problem. The project helped the cosmetic community in providing the following:

- Creation of datasets for future references and applications
- Providing time saving solutions for tedious tasks
- Providing cost efficient solutions for firms of all economic backgrounds
- Sample model for further improvements in the sensory evaluation field
- Help the scientists evaluate several samples in a small duration
- Provide results with minimum error and evaluation results with no human bias which is very important for real evaluation of the products.

NOVELTY

While creating the application for quality testing of hair care products, one of the most important aspects kept in mind was the uniqueness or novelty of the product created. Currently in the cosmetic industry there exists no such product which can do sensory evaluation using the camera of a mobile device and also do the evaluation on the same device using advance AI techniques like CNN.

At the same time the application created is extremely cost efficient which is a boon for small and medium sized firms which cannot afford the expensive machines currently existing in the market.

The application also has a unique selling point of being extremely time efficient as it can do hours of evaluation work in just a matter of few minutes and get results with a minimum error.

FUTURE SCOPE

The application made currently does sensory evaluation on 5 parameters, in the future these parameters can be extended to whatever the company proposes.

Another future improvement can be to build a single model that gives ratings for all the parameters.

This application can be improved to not only make sensory evaluations but also recommend the haircare products that the company makes based on the hair type that the human possess.

CONCLUSION

The project aims at creating a quality and efficacy evaluation product which is affordable for any firm, individual or organization. They can use this software without having any specialized hardware and can be used using just a smartphone, iPad/tablet or any portable device having a camera.

Currently the most widely used method to evaluate hair care products sensory performance is the sensory evaluation which permits to score the different sensory attributes with the help of a panel of trained judges. However, this method is time consuming and sometime it is not really relevant for product development and so without all the data to assure safety. For this reason, some instrumental investigation is preferred and can be made as well in order to save time on formulation screening.

This study differentiates from earlier studies therein that we undertake a holistic approach in creating a product that is in-expensive, widely accessible, usable on a portable device and does not need specialised hardware for it to function.

To make the model for the application a dataset was created by clicking images of hair tresses. The images clicked were in two batches first was virgin hair which was combed twice and then bleached hair which were combed twice and washed for multiple times. After the images were clicked a panel of 12 experts from the company rated the images on a set scale of 1-5 on 5 primary parameters namely, ‘Volume’, ‘Shine’, ‘Smoothness’, ‘Frizzyness’ and ‘Fly Away’. These images were scaled to a resolution of (256,256), cropped to remove any noise and then normalized between a pixel value of 0-1.

After the curation of the dataset a Convolutional Neural Network Model was built and trained on the dataset with these 5 parameters. This convolutional neural network model had of 536,877,505 trainable parameters and 2112 non-trainable

parameters. A total of 5 models were built one for each of the parameters and then deployed for usage.

For the end user a front-end web application was built wherein the user can upload all the images required for testing and get the reports generated. The model when used was giving good result matching with the estimation of the company scientists and helped in reduction of 18 manhours gave the results in 10 minutes.

REFERENCES

[1] Isik, Selin & Abba, Sani & Usman, Abdullahi & Mericli, Filiz. (2020). Artificial intelligence-based models for qualitative and quantitative prediction of phytochemical compound using HPLC method. Turkish Journal of Chemistry. 44. 10.3906/kim-2003-6.

[2] Savardi, M., Daniels, G., Tamburic, S., Muhammad, U. R., & Benini, S. (2020). Figaro-tresses: A dataset for evaluating hair assembly features before and after cosmetic treatment. *Data in brief*, *31*, 105964. https://doi.org/10.1016/j.dib.2020.105964

[3] Daniels, Gabriela & Tamburic, Slobodanka & Benini, Sergio & Randall, Jane & Sanderson, Tracey & Savardi, Mattia. (2021). Artificial Intelligence in hair research: A proof-of-concept study on evaluating hair assembly features. International Journal of Cosmetic Science. 43. 10.1111/ics.12706.

[4] https://github.com/danielgatis/rembg

[5] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), 2017, pp. 1-6, doi; 10.1109/ICEngTechnol.2017.8308186.

[6] Kang, Xu & Song, Bin & Sun, Fengyao. (2019). A Deep Similarity Metric Method Based on Incomplete Data for Traffic Anomaly Detection in IoT. Applied Sciences. 9. 135. 10.3390/app9010135.

[7] CNN Medium

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in neural information pro-

cessing systems, 2012, pp. 1097–1105.

[9] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” arXiv preprint arXiv:1409.4842, 2014.

[10] D. Yi, Z. Lei, and S. Z. Li, “Age estimation by multi-scale convolutional network,” in Computer Vision–ACCV 2014. Springer, 2015, pp. 144–158.

[11] K. Kavukcuoglu, P. Sermanet, Y.-L. Boureau, K. Gregor, M. Mathieu, and Y. L. Cun, “Learning convolutional feature hierarchies for visual recognition,” in Advances in neural information processing systems, 2010, pp. 1090–1098.

[12] R. Rautkorpi and J. Iivarinen, “A novel shape feature for image classification and retrieval,” in Image Analysis and Recognition. Springer, 2004, pp. 753–760.

[13] S.-C. Zhu, R. Zhang, and Z. Tu, “Integrating bottom-up/top-down for object recognition by data driven markov chain monte carlo,” in Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on, vol. 1. IEEE, 2000, pp. 738–745.

[14] https://www.fortunebusinessinsights.com/hair-care-market-102555

[15] Hair Care Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2021-2026, IMarc

[16] Bouton, Francois & Lheritier, Anne & Lavarde, Marc. Sensory Assessment Of Shampoo Containing New Conditioner In Comparison With Shampoos Containing Silicone. (2018). Conference IFSCC, Munich, October 2018.

[17] Garcia ML, Epps JA, Yare RS; Normal cuticle-wear pattern in human hair;*J Soc Cos Chem*; 1978; 29:155–175.

[18] Kelley S, Robinson VNE; The effect of grooming on the hair surface; *J Soc Cos Chem*; 1982; 33:203–215.

[19] Jachowicz J.; Hair damage and attempts to its repair;*J Soc Cos Chem*; 1987; 38:263–286.

[20] Pande C, Jachowicz J.; Hair photo-damage – measurement and prevention;*J Soc Cos Chem*; 1993; 44:109.

[21] Jachowicz J.; Hair damage and attempts to its repair;*J Soc Cos Chem*; 1987; 38:263–286

[22] Trüeb RM; Aging of hair;*J Cosm Derma*; 2005;4:60-

[23] Lunn AC, Evans RE; The electrostatic properties of human hair;*J Soc Cos Chem*; 1977; 28:549–569.