Lucerne University of
Applied Sciences and Arts

HOCHSCHULE
LUZERN

Wirtschaft

Influence of facial characteristics in brand selfies on social media engagement

(Master Thesis)

Author: Carole Mattmann
Hinterer Steinacker 40
CH-4600 Olten
carole.mattmann@stud.hslu.ch

Lecturer: Dr. Marek Gorny
Universität Luzern
marek.gorny@unilu.ch

Co-Lecturer: Prof. Dr. Mirko Birbaumer
Hochschule Luzern
mirko.birbaumer@hslu.ch

Lucerne University of Applied Sciences and Arts
Master in Applied Information and Data Science (MScIDS)
Spring Semester 2021

Lucerne, 29th May 2021

Abstract

User generated content on social media has become a relevant part of marketing. One way of promoting a brand on social media is users posting a selfie with the branded product. It is already known that this is a successful way to market a brand. This thesis aims to analyse the brand selfies with an emphasis on the facial characteristics of the selfie takers.

Humans react to human faces stronger than to most other objects. We look at faces to get information, to categorize and to form relationships. Advancements in computer vision and machine learning have make it possible to analyse faces in a variety of ways. In this thesis, the faces on the selfies are analysed with three different state-of-the-art analysis techniques. First, convolutional neural networks are trained to label the faces with the attributes *age*, *gender*, *emotion*, *ethnicity* and *attractiveness*. In a second analysis, facial landmarks are detected and face proportions using the *golden ratio* standards are calculated. The last analysis decomposes images into 30 components using the *eigenface* approach. The results of the three analyses are then put into relation to the engagement a selfie recieves on social media with regression models.

The results suggest that more attractive people get more engagement. The face proportions and the image decompositions provide some insights in how attractiveness is measured in this thesis. Also gender and ethnicity may have a small impact on the engagement. Other than facial characteristics, it may be of importance in what angle and light a selfie is photographed.

Table of Contents

List of Tables

List of Figures

1. Introduction and research question

Social media has become an important tool for a brands marketing. Especially social media posts with user generated content, oftentimes selfies, have proven to be effective. Therefore it can be of interest for brands and content producers to know what makes one selfie more successful than another. [1]

An important element in a selfie is the persons face. It has already been proven that a picture with a face receives more attention on social media than one without. This thesis aims to analyse different characteristics of the face in more detail and bring it into relation with the engagement a selfie receives.

With the advancements in computer vision, rather new techniques are available to analyse faces. Convolutional neural networks are trained in this thesis to detect several characteristics of a face. This approach is complemented by other proven approaches for face analysis, namely the calculation of facial proportions and an eigendecomposition of the image into principal components. Those different face analyses are then tested for their impact on the engagement for a post.

The reseach question of this thesis is : *How do facial characteristics in brand selfies influence the engagement rate for the selfie on social media?*

In other contexts it has been proven that humans are interested in faces to detect emotions that a person expresses or to identify the person for example by age, gender or ethnic background (Bakhshi et al., 2014). This leads to the formulation of the following hypotheses:

H1 : The emotion of a face on a brand selfie influences the engagement rate
H2 : The age of a face on a brand selfie influences the engagement rate
H3 : The ethnicity of a face on a brand selfie influences the engagement rate
H4 : The gender of a face on a brand selfie influences the engagement rate

It could be shown that people that are percieved as more attractive have certain advantages like preferential treatment and better social acceptance (Liu et al., 2016). In advertising it is common to employ models that are usually percieved as attractive to promote brands. In this thesis, the relation between attractiveness and the engagement will be analyzed. As attractiveness is a complex concept, two different methods will be used. One will assess attractiveness as a whole, the other will use face proportions to express attractiveness. Those are the two hypotheses :

H5 : The attractiveness of a face on a brand selfie influences the engagement rate
H6 : The face proportions of a face on a brand selfie influence the engagement rate

It is likely that not only the facial characteristics and proportions of a person on a brand selfie will have an impact on its engagement rate. For example, how an image is photographed may have an effect as well. So the last analysis uses a more holistic approach for image analysis. This leads to the formulation of the last hypothesis :

H7 : Images can be broken down into components which have an influence on the engagement rate

[1] More details and sources of all the information presented in this chapter can be found in chapter 2.

2. Background

This thesis will study people's faces on selfies. It will gain insights into how facial characteristics of the selfie taker are percieved by the audience. This will be measured by looking at the engagement rate with the selfie on social media. In the following subchapters, the most important pieces of the topic and the underlying literarture are described.

2.1. Images on social media

Li&Xie (2019, pp. 9–19) analyze different characteristics of posts on Twitter and Instagram. They find that the presence of an image has a positive impact on the posts engagement rate. If there is an image included in the post, professional shots, presence of a human face and the text-image fit also partly have an impact on the engagement.

The social media platform Instagram was designed as image sharing platform. With more than 1 billion users worldwide, Caliandro and Graham (2020, p. 1) state that it «has become one of the most important social networking sites globally and in the process has transformed the role of photographs and photography in visual culture». They describe that Instagram is designed as a peer-to-peer platform where «users […] share their amateur pictures with friends» (p. 2). Because of Instagrams importance and their focus on images and user generated content the platform will be used as data source in this thesis.

2.2. Self expression through user generated content on social media

Social media platforms are used for self-expression. Facebook for example has started as a platform where users could make a profile about themselves. Twitter provides a platform to share thoughts as short tweets. Instagram's way of expression is through images. In this thesis the focus is on images. Specifically, the focus is on *selfies*, meaning a «photograph that one has taken of oneself, typically one taken with a smartphone or webcam and shared via social media» ('Oxford English Dictionary - Selfie', n.d.).

Selfies are self expressions in the sense that they can «clearly communicate an individual's passion or interest that reinforces her or his social identity and serve as an artistic expression of that individual's fashion, beauty and/or possessions. […] For example, some consumers post selfies with brands/products they own with a brand-related hashtag. » (Sung et al., 2018, p. 15).

2.3. Marketing perspective on user generated content on social media

Social media is not only privately used for self-expression, but has become an essential part of marketing (Husain et al., 2016, p. 1). Hartmann et al. (2021) estimate based on a sample that around 1% of all social media images are brand images. They conclude that « this magnitude is likely to exceed any other advertising channel […], generating an immense level of reach » (p. 2).

User generated content featuring brand promotion was investigated by Lou, Tan and Chen (2019). They found that brands are more successfully promoted through user generated content than brand generated content, measured by engagement rate.

Influencer marketing and brand selfies are two typical examples of user generated content. The influencer marketing industry is already a $10 billion industry (Haenlein et al., 2020). *Influencer marketing* means collaborating with a popular user on social media who will promote your brand. Often this includes posting a picture of the influencer using the product. It is proven that influencers promoting brands will help making other people more aware of the brand (Brooks, 2018).

Marketers also use *selfie campaigns*. They encourage non-paid users to post a selfie promoting their brand. In exchange the user gets a prize or exposure in the form of a repost in the brand's profile. Bharti (2020, p. ix) describes that «the growing popularity of selfie campaigns has taken the world by storm». For example, the cosmetics company Lacome introduced a selfie campaign *#bareselfie* to encourage users to post a selfie without makeup (Sung et al., 2018, p. 17) or Lay's potato chips produced a special packaging with a smile to encourage users to post a selfie under the hashtag *#smilewithlays* (Hartmann et al., 2021, p. 2).

A tactic to position yourself as an influencer and to gain followers is also to post content with a brand and hope for a repost of the brand itself or being found by hashtag searches. This will get the user more exposure and potentially more followers. The brand in return can profit from being able to use the image on their channel (Kavanagh, 2020, p. 12). This is a tactic mostly used by lesser known or aspiring influencers, as more popular influencers would mostly not post branded content for free. It blends with the topic of self-expression.

Several studies prove that posting selfies mentioning a brand can also strengthen relationships between the selfie producer and the brand (Hofstetter et al., 2020). In this thesis however, the focus lies more on the connection of the selfie viewer to the brand than from the selfie producer to the brand.

2.4. Brand selfies

In the last years, the term *brand selfie* has been established. Hofstetter, Kunath and John (2020) define brand selfies as «self-photos with a brand logo or product from that brand in tow». The images can be produced by any user on a social media platform, which can mean influencers with millions of followers but also people who only use social media privately. The context of the production can be different, sometimes a paid job, a selfie challenge or just a user wanting to share a brand, either for self-expression or to get more exposure (see use cases described in chapter 2.2 and 2.3).

In this thesis, images are studied depicting a visible consumer face together with a branded product (see example in Figure 1). Even though the term *selfie* is used, it is not specifically required that the picture has been taken by the person on the picture itself. However, the picture should be user generated.

It can either be that a person is holding a branded product in their hands, wears the product (for example in the case of clothes, jewelery, makeup) or displays the brand in another manner. The brand logo does not need to be visible. Not studied are pictures without faces or without a product.

Figure 1

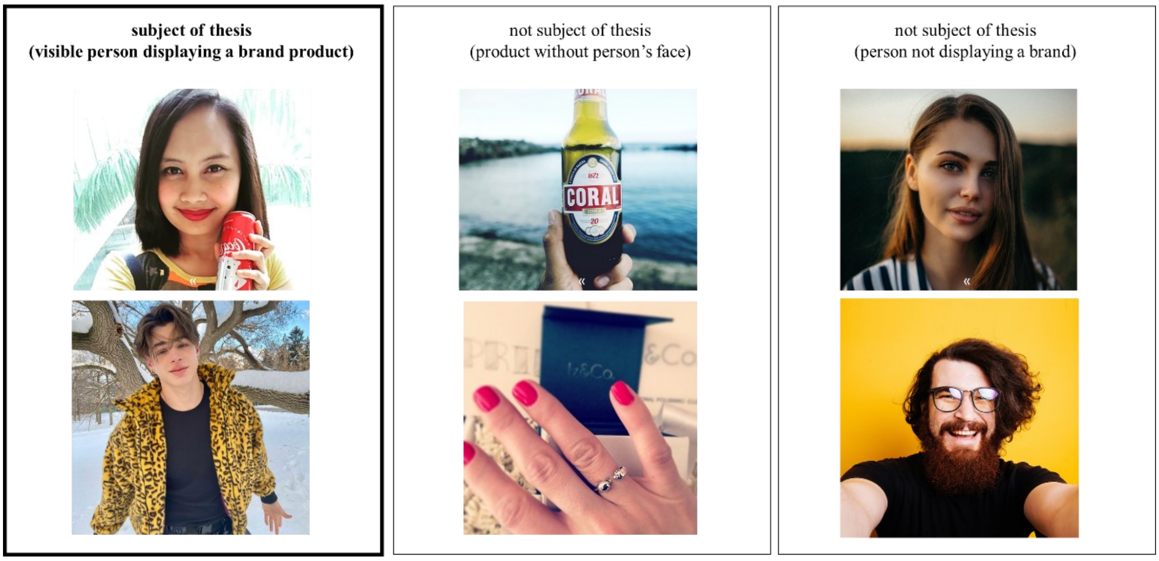


Figure 1. Topic definition «consumer brand selfies» (Authors compilation)

2.5. User Engagement

According to Li&Xie (2019) it is industry practise to categorize social media engagement into direct responses and sharing/reposting. They suggest that the direct response is «more private and directed because it gives a direct affirmation», whereas «sharing is more socially visible and undirected». Both types are seen as important.

Sharing and reposting is hard to measure on Instagram, as the official API does not allow such queries. This is why this thesis will only analyse direct engagement. Instagram offers the options to respond directly to a post by liking or commenting. In this thesis the examined engagement is limited to the number of *likes* for a post. The *likes* are «a commonly adopted metric, which enables readers to show enjoyment, appreciation, or endorsement of the content » (Li & Xie, 2019). Details on how engagement rate is calculated in this thesis can be found in chapter 3.3.2.

A variety of reasons can impact the engagement rate of a post. It can depend for example on how much emotion a post evokes (Berger & Milkman, 2012). Also the presence of a person seems to have an impact (see chapter 2.6). There are also reasons that depend more on the profile that posts it than on the post. It is common to speak about a *Like-Follower Ratio* when talking about Instagram engagement, which implies that the more followers a profile has, the more likes a post on that profile gets (Huynh, 2021). It also seems to be the case that women get more engagement than men (Fowler, 2017).

2.6. Faces in marketing and social media

It has been known for a long time that humans react to human faces (Morton & Johnson, 1991). This is also true on social media. Users on social media are more likely to react to a picture with a face on it than one without (Bakhshi et al., 2014, p. 965; Li&Xie, 2019, pp. 9-19). Bakshi, Shamma and Gilbert (2014) quantify the difference with 32% more likes. They also try to gain more insights by looking at number of faces on the picture or age and gender of a person but do not find a significant impact on the likes.

Not only the presence of a face but also its characteristics are becoming more relevant. In the last years, face recognition and face filters have been implemented on several social media platforms. On

Snapchat, users can add items to their face or transform the faces when taking a selfie. TikTok supposedly uses image recognition to define the attractiveness of a person and takes that into account when promoting the content (*How The Tiktok Algorithm Rates Your Face | Beauty Culture Ep. 6*, 2020).

2.7. Facial characteristics and attractiveness

Bakshi, Shamma and Gilbert (2014) claim that «face perception is perhaps the most highly developed human visual skill». People «monitor faces because they provide vital clues in an impressive variety of contexts: attraction, the complexity of emotions, identity, age, humor, and a person's regional and national background» (p. 965).

It is being studied extensively what attracts people to one face more that to another, in other words what makes it attractive. It is a complex topic that is looked at from different fields such as art, psychology, medicine and marketing. The emphasis in this thesis lies on gaining a better understanding for why a person engages with a post. Liu et al. (2016, p. 16634) summarize that «compared to people with unattractive faces, people with attractive faces usually receive higher-level of social acceptance, possess more positive personalities, receive preferential employment and professional advancement, and attain better outcomes in real life situations ». The hypothesis in this thesis is that the preferential treatment of more attractive faces will also apply to social media engagement.

Although describing or quantifying attractiveness is complicated, some traits have been identified that seem to have a universal impact on the perception of attractiveness. Examples are averageness (Halberstadt et al., 2013), certain face porportions (Hong et al., 2017), being young and having good skin (Etcoff, 2011, p. 8).

2.8. Face analysis methods

Faces can be analyzed using different machine learning methods. Liu at al. (2016) provide an overview over computational methods used on faces. They propose categorizing the methods into feature-based approaches, holistic approaches and hybrids.

Feature-based approaches try to extract certain traits of a face. Their advantage is that they are easy to interpret for humans. The disadvantage is that the features are often not robust because they change with face rotations or expressions. Types of feature-based approaches are geometric (analysis of geometry by localizing facial regions and landmarks and taking some distances or ratios like for example to «golden ratio» of face height to face width), texture (analysis of skin texture, for example skin smoothness) or color (analysis of colors in the face, for example hair and eye color).

Holistic approaches look at the face as a whole entity. This usually means performing an analysis on the pixel values of the whole picture. Two common methods are the « Eigenface » method where dimensionality reduction is performed on a vector of the pixel values or the application of convolutional neural networks.

Hybrid approaches combine different approaches together to compensate for weaknesses in the individual models. An example is Eisenthal et al. (2006) who work with Eigenfaces and facial ratios combined and reached better results by using this combination as a model for attractiveness.

This thesis combines the two most common holistic approaches and the most common feature-based approach: convolutional neural networks, eigenfaces and geometry (face proportions). More details on the used approaches are described chapter 3 (Methodology).

3. Methodology

In this chapter, the methodology from the data collection up to the final analysis of the engagement is presented. The methodology consists of four steps that are shown in Figure 2: collection the selfie data through scraping Instagram, preprocessing the data, performing three analyses on the images (CNN, face proportions, image decomposition) and finally modelling the results of those three analyses versus the engagement to identify variables that influence the engagement.

To train the CNN models, additional training data was collected from existing training datasets and preprocessed.

Figure 2

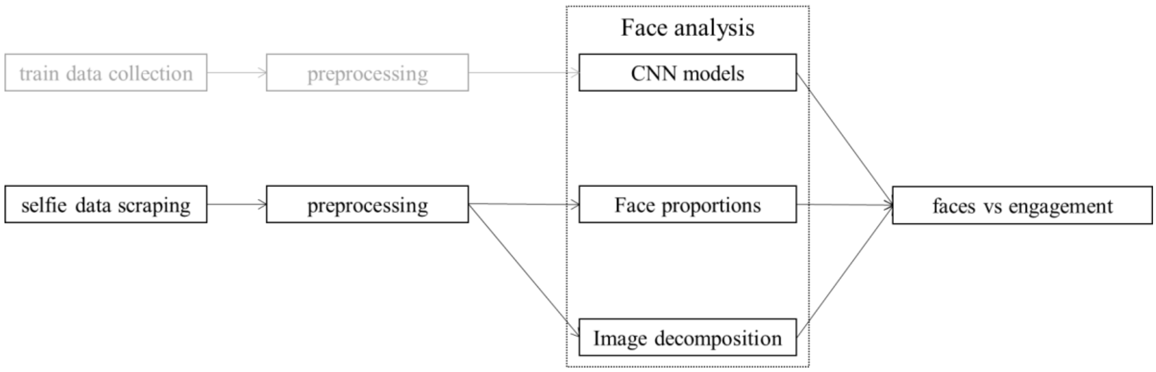


Figure 2. Steps in the methodology

3.1. Software, storage, applications

For all steps, Python is used as programming language and Tensorflow/Keras as machine learning framework for the CNN. Several python packages are used, for example *opencv* for face detection. All packages are referenced in the code in the appendix. All coding and modeling is done in Jupityer Notebooks on a local machine or on Google Colab. Google Golab Pro offers 25 GB RAM and GPU. The images are stored in folders and the metadata in csv files.

3.2. Selfie data scraping

3.2.1. Choice of data source

The data source is the social media platform Instagram. A total of over 42'000 images and their metadata were scraped with a python script. The scipt accesses the offial Instagram API (Facebook Graph API). The scraping was done in two steps, first the scraping of the metadata including a link to the image and as secondly downloading the corresponding picture via the link provided.

The selected posts are in the of the product categories food, beverages, clothing, cosmetics and jewelery. It is necessary to narrow down the product categories because depending on the product, an image of a person with that product looks different (for example a selfie with a car and a selfie with a beverage). An overview of the Instagram profiles scraped is provided in Table 1.

Table 1
Table of images scraped from instagram by profile

Profile	Product category	Brand	Number of images scraped
@beerselfie	Beverages	Craft beer	12'892
@dunkin	Food	Dunkin Donuts	1'590
@lululemon	Clothing	Lululemon	2'906
@asos	Clothing	Asos	7'272
@starbucks	Beverages	Starbucks	2'070
@summerfridays	Cosmetics	Summerfridays	872
@izandco	Jewelry	izandco	2'483
@lushcosmetics	Cosmetics	Lush	3'619
@birchbox	Cosmetics	Birchbox	7'628
@americaneagle	Clothing	American Eagle	1'490
Total images			**42'762**

As established in chapter 2.5 the number of likes a post gets is influenced by how many followers the user who posts an image has. On the official instagram API, it is not possible to get that information directly when searching for hashtags (for example *#selfie*). This was implemented following the Cambridge Analytica Scandal to prevent misuse of the data (Turcanu, 2018). There are also other issues with hashtag searches like the number of posts that is allowed to scrape, use of face filters or masks during covid times. In this thesis, posts are collected from 10 manually chosen brand or community profiles (profiles that promote a certain community or interest like « beerselfie). See Figure 3 for two examples of scraped profiles.

Figure 3

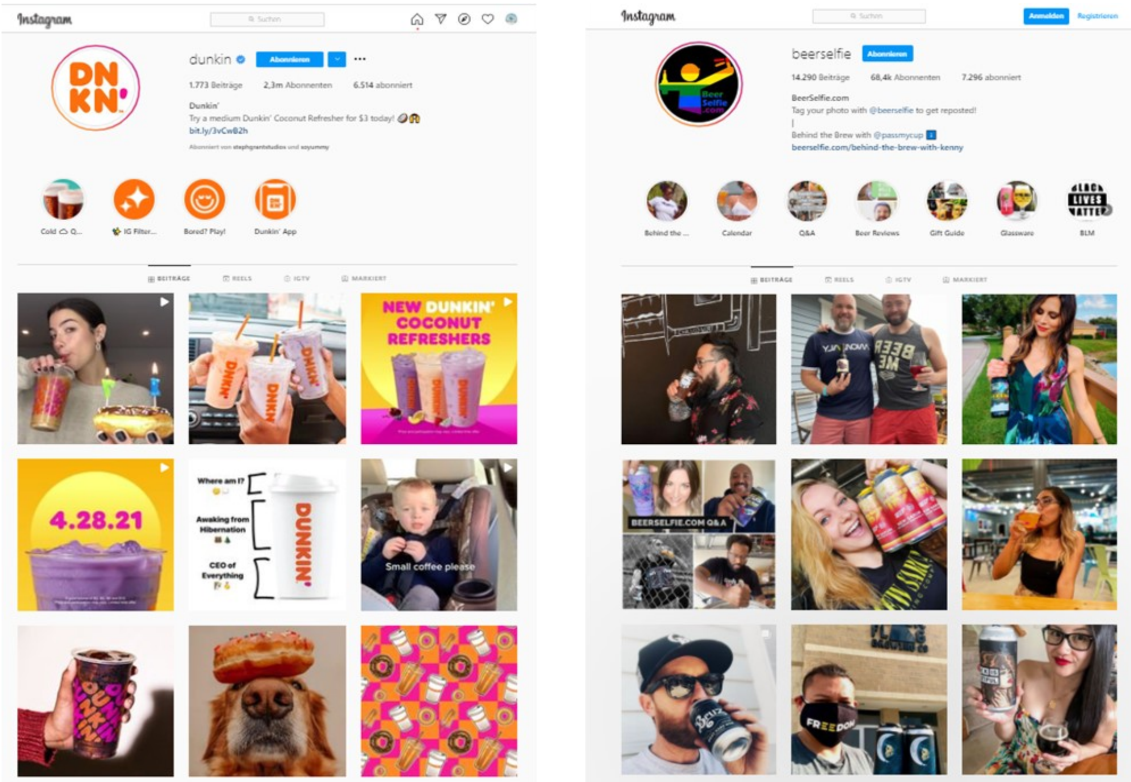


Figure 3. Example Screenshots of instagram brand profile @dunkin (2021) and community profile @beerselfie (2021)

The chosen profiles work with user generated content. This means the brand/community reposts an image taken on their own brand or community profile. There, the count of the likes starts at zero again. All posts on the brand/community profile are exposed to the followers of this profile, which makes the like count of the posts comparable (except for time posted, see chapter 3.3.3). See Figure 4 for an example.

Figure 4

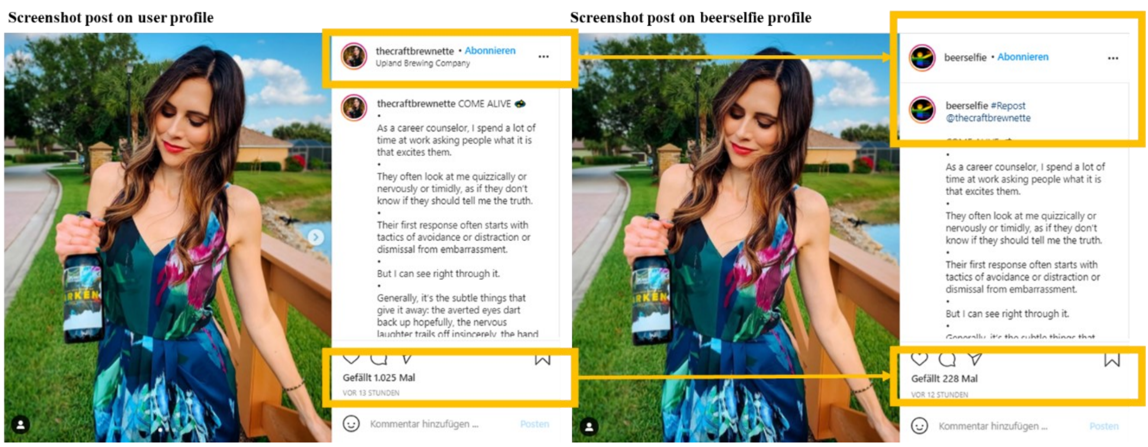


Figure 4. Repost example - Post from «thecraftbrewnette» (2021) that is reposted on the profile « beerselfie » (2021)

Not all content that is found on the selected profiles are pictures with a branded product and a person as defined in chapter 2.4. This is a common problem. Hartmann, Heimann, Schamp and Netzer (2020) conducted a study about brand selfies. They analysed a quarter million brand images and categorized them using convolutional neural networks. They identified that only 9% of them were consumer brand selfies. The problem is adressed partly by only using images with exactly one face on it (see chapter 3.3). With the choice of the brand and community profiles it is made sure that the content points towards a brand.

The focus of this thesis should be on user generated content. Patterns on the post to differenciate between user generated content and content the brand produced could not be detected. Brands or communities were chosen who repeatedly repost user generated content. Still, some brand generated content will be included.

3.2.2. Decription of data

The underying data consists of an image and metadata about the image. See Figure 5 and Figure 6 for an extract of the first five pictures and metadata entries of the pilot dataset.

Figure 5

id	permalink	media_url	caption	media_type	like_count	timestamp
17843820266338937	https://www.instagram.com/p/CEhWDAcHskl/	https://scontent-hkt1-1.cdninstagram.com/v/t51...	#Repost brewnette_barbie\n · · · \nTwo weeks off o...	IMAGE	100	2020-08-30T16:58:38+0000
17854301786181053	https://www.instagram.com/p/CEhLljGncDl/	https://scontent-hkt1-1.cdninstagram.com/v/t51...	#Repost beer.brat\n · · · \nWHAT IS A PAW PAW? wro...	IMAGE	177	2020-08-30T15:27:14+0000
17846054312335344	https://www.instagram.com/p/CEhLZ7DHq9B/	https://scontent-hkt1-1.cdninstagram.com/v/t51...	#Repost queer_for_beer\n · · · \nBoys on a bridge,...	IMAGE	155	2020-08-30T15:25:39+0000
17874530710890503	https://www.instagram.com/p/CEgBVNvneoA/	https://scontent-hkt1-1.cdninstagram.com/v/t51...	#Repost ilikeitsour\n · · · \nI love when you read...	IMAGE	231	2020-08-30T04:38:23+0000
17894485000602340	https://www.instagram.com/p/CEgA_68H2Ft/	https://scontent-hkt1-1.cdninstagram.com/v/t51...	#Repost lindsaybrewtour\n · · · \nI lift her up, s...	IMAGE	301	2020-08-30T04:35:28+0000

Figure 5. Extract of the metadata of the first 5 pictures in the dataset

The following datafields of metadata are collected. The images are downloaded and saved with the id as file name.

- id: unique identifier
- permalink: link to the post on instagram
- media_url: link to the picture on instagram
- caption: the caption written by the user
- media_type: only type «image» is used, carousel (more than one picture) and videos are not used
- like_count: number of likes to post recieved
- timestamp: time of post

Figure 6

Picture 1 ID: 17843820266338937 | Picture 2 ID: 17854301786181053 | Picture 3 ID: 17846054312335344 | Picture 4 ID: 17874530710890503 | Picture 5 ID: 17894485000602340

Figure 6. First 5 pictures in the dataset corresponding to the metadata above

3.3. Data pre-processing

3.3.1. Image preprocessing

For the face analysis, only one person can be on the image and the face of the person needs to be visible.

The python library « opencv » offers an implemented face detection algorithm using a haar cascades classifier. The implementation in this thesis uses the frontal face eye haar cascade classifier and the eye haar cascade classifier. With eye classifier, the faces are aligned so that the eyes of the person are on a horizontal line based on Serengil (2020). The angle of the eyes to a horizontal line is calculated (*Angle A* in Figure 7) and then the image was rotated by this angle.

Figure 7

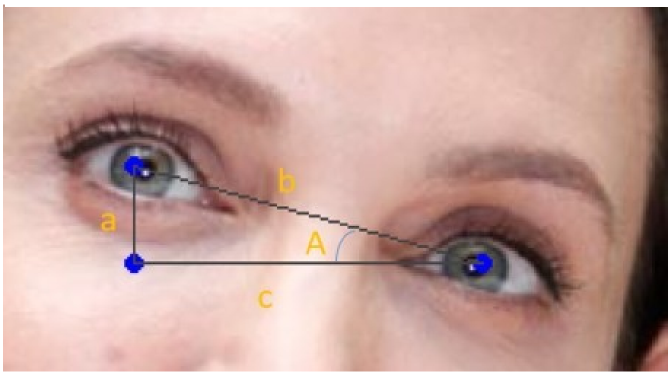


Figure 7. Example of face alignment using eye angles (Serengil, 2020)

The alignment step is necessary to perform the image decomposition / eigenface analysis and also delivers better results in the face detection that follows as the next step.

After the alignment, the whole faces are detected using the frontal face classifier following Menon (2019). The result of the classifier are coordinates of every face on the picture that then allows to cut out the face from the image. For this thesis the criteria is that there is only one face on the picture. The implemented function automatically selects the images where exactly one face is detected and disregards the rest. Priadina and Habibi (2019) use the same approach to filter selfies from Instagram.

An excerpt of the test run (without alignment) with the scraped images is shown in Figure 8. In the green frames are the faces detected. Pictures 1 and 2 are suitable. Picture 3 is not selected because there are 2 faces and picture 4 is not selected because the face is not visible well and not suitable for analysis. Pictures with starbucks cups as shown in picture 5 are problematic as the algorithm often detects the face on the starbucks logo as face. To avoid such cases, the sensitivity of the face detector is set high. Another common problem is shown in picture 6. More than one person is on the picture but only one face is detected. This means the picture would be selected for analysis but shouldn't be. By setting the sensitivity of the face detector high, some of the cases could be eliminated but it is possible that some remained in the dataset.

Figure 8

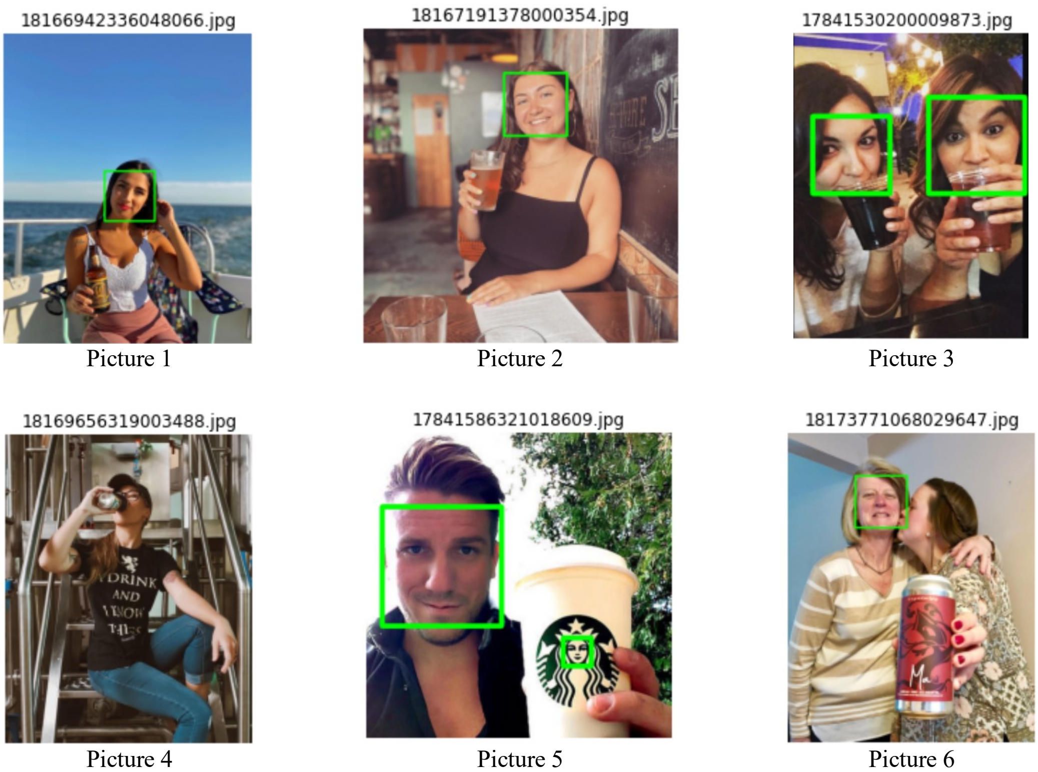


Figure 8. Sample images from the scraped data using Open CV face detection

After the alignment, face detection and selection of images with only one face, the faces are cropped and saved for further use. The result of the image preprocessing is 2300 aligned and cropped faces in resolution 224x224. An excerpt is shown in Figure 9.

Figure 9

Figure 9. Excerpt of preproprocessed selfie images

3.3.2. Engagement calculation

The industry standard to calculate engagement is the total interactions an image revieves divided by the followers of the profile.

$$engagement_{image} = \frac{total\ interactions_{image}}{followers_{profile}}$$

In this thesis, the following formula is applied[2]:

$$engagement_{image} = \frac{likes_{image}}{\text{mean } likes_{profile}}$$

The total interactions in this case are only the likes, because comments are not scraped and analyzed. Figure 10 and Table 2 show that followers and mean likes of all posts correlate. Followers or mean likes per profile could be used to calculate engagement. Average likes are chosen because differences in user activity between the profiles can be eliminated.

Figure 10

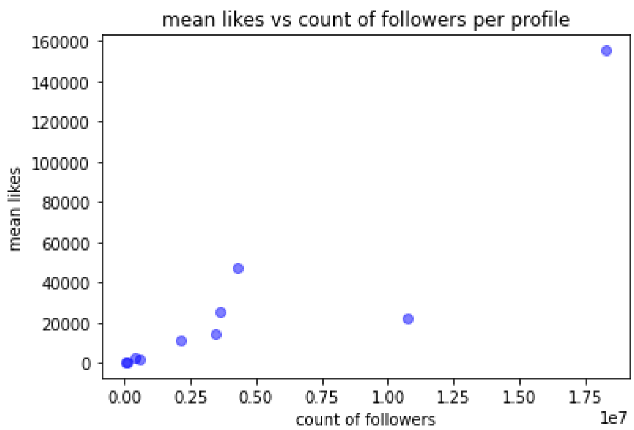


Figure 10. Scatterplot of mean likes vs count of followers per profile

Table 2
Table of of mean likes vs count of followers per profile

profile	mean likes of all posts	count of followers
profile-beerselfie	352	51’427
profile-izandco	567	80’236
profile-birchbox	1’671	578’158
profile-summerfridays	2’546	418’084
profile-dunkin	11’547	2’142’158
profile-lululemon	14’141	3’447’860
profile-asos	22’198	10’748’803
profile-americaneagle	25’722	3’626’955
profile-lushcosmetics	47’724	4’318’187
profile-starbucks	155’452	18’258’250

[2] In the dataset, the engagement variable is called «relative likes»

3.3.3. Engagement time normalization

The engagement calculation from chapter 3.3.2 assumes that the number of followers of a profile stay the same throughout time. This is mostly not the case. When the profile is created it has zero followers, so the posted images at the start of a profile lifecycle get less exposure. The longer the profile exists, the more follower it typically gains. This gets an image more exposure and therefore the chance of getting more likes is higher. However, images that are posted recently had less exposure time and are still in their process of getting the likes, so they will have less likes. This phenomenon should be accounted for. Three steps are needed to smoothen the effect of the time posted on the engagement.

Normalize time

The time of posting is described relative to the time of the first profile post and the time of the last profile post:

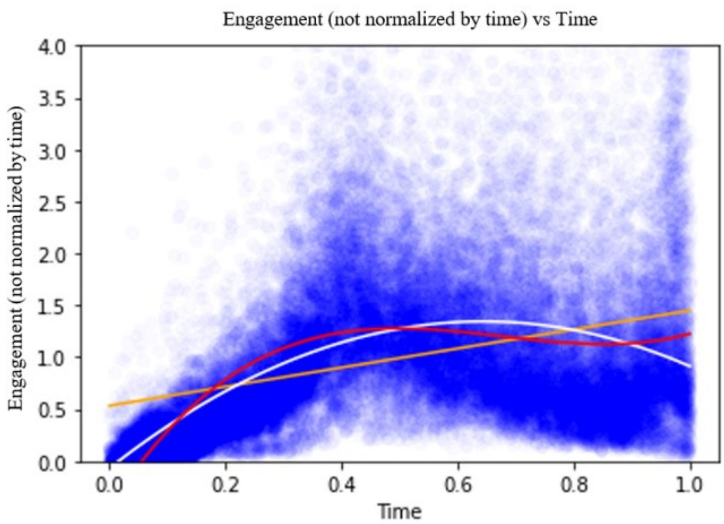
$$time\_normalized_{image} = \frac{\max time_{profile} - \min time_{profile}}{time_{image} - \min time_{profile}}$$

This means time is expressed as a number between 0 and 1, showing how close it is to the first post on the profile.

Create time model for engagement

The engagement is plotted versus the normalized time. The described pattern of starting with low engagement, then in the middle the highest and versus the end lower again can be observed. The pattern is modelled using first, second and third degree polynomials. Although the third degree polynomial performs slightly better judged by R^2, it could be due to overfitting. The second degree polynomial is chosen to calculate an engagement rate that is normalized by time[3].

Figure 11



Degree of Polynomial	Correlation (R^2)
1	0.10
2	0.19
3	0.21

Figure 11. Modelling of engagement (not normalized by time) vs time (normalized) using different polynomials

[3] In the dataset, the variable is called «relative likes time norm»

Normalize engagement by time model

After the normalisation, the pattern observed above is almost removed. The engagement is now less relient on the time a post was made.

Figure 12

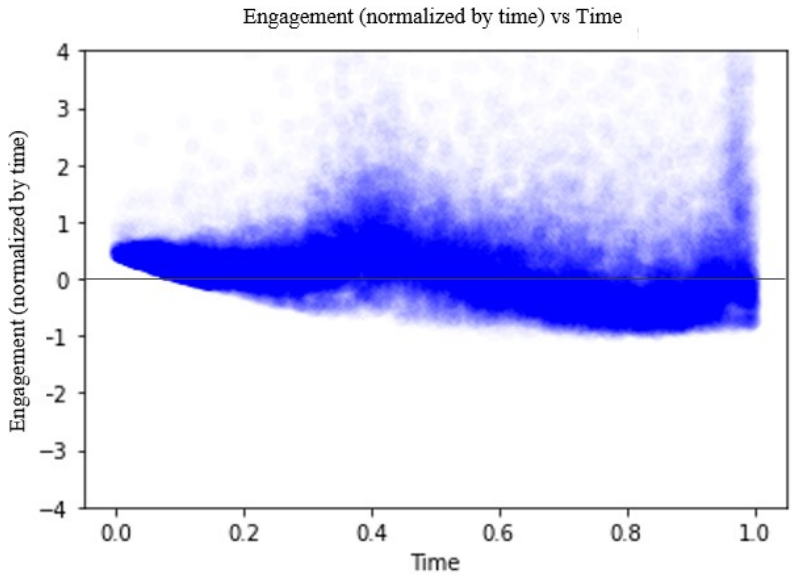


Figure 12. Modelling of engagement (normalized by time) vs time (normalized)

The models in chapter 3.7 were tested using the unnormalized engagement and the time-normalized engagement. They showed similiar results, but the models using time-normalized engagement as dependent variable performed better (measured by correlation).

3.4. Analysis of facial characteristics with convolutional neural networks

3.4.1. CNN for face analysis

Convolutional neural networks have become a standard in computer vision (see chapter 2.8). They are applied for face analysis. The models typically provide probabilities that a face belongs to a certain group (categorical), for example male/female, or estimate a number (continuous), for example age.

A few examples of existing applications are : hownormalami.eu (gender, age, attractiveness, BMI, life expectancy), clarifai.com (gender, age, multicultural appearance), hotness.ai (attractiveness), Google Vision AI (emotions). Figure 13 shows an example of an analysis on carifai.com.

Figure 13

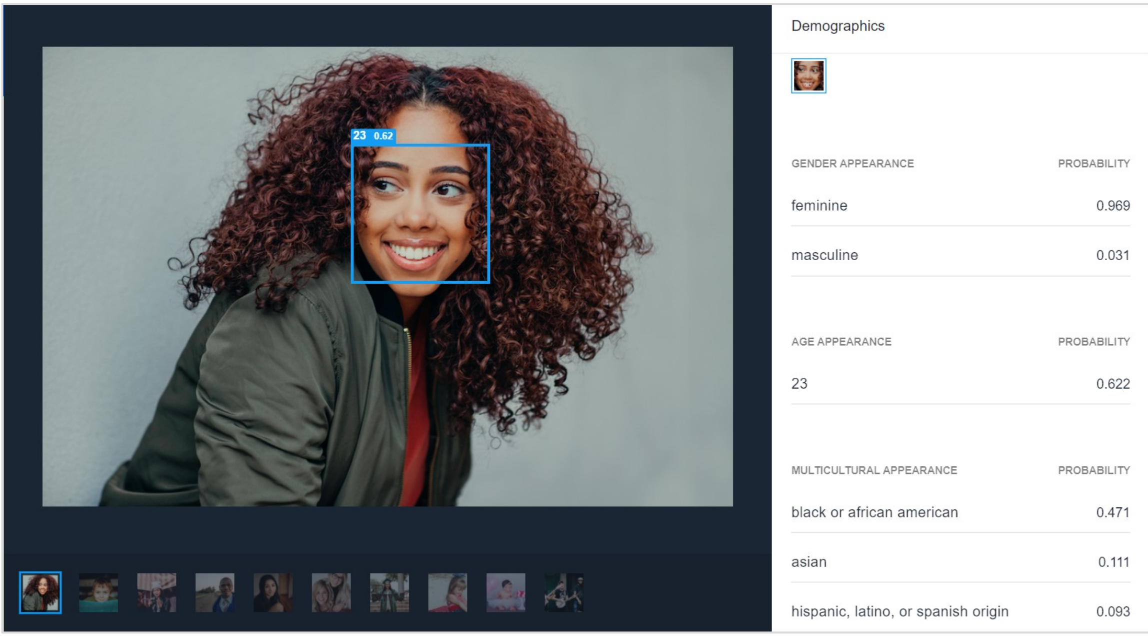


Figure 13. Example of face analysis on clarifai.com **(**Clarifai, 2020)

3.4.2. Application of CNN for face analysis for the selfie use case

Although models already exist, they are usually not freely available. Also, it is normally not disclosed what training data has been used and how accurate the models work. In this thesis, own models are developed using different labeled datasets. Five models are trained to identify a total of five characteristics of the faces: gender, age, ethnicity, emotions, attractiveness. The selection is based on previous findings about face perception and attractiveness (see chapter 2.6, 2.7, 2.8).

The python code was derived from the Jupyter Notebooks provided in the lecture *Deep Learning in Vision* by Prof. Dr. Mirko Birbaumer (2020). Pretrained CNN models are available and implemented in the framework Keras, which are used for the analysis. Different pretrained models were tested. VGG Face was used for 4 characteristics. VGG-Face is a VGG-16 architecture with pretrained weights for face recognition presented by Parkhi et al (2015). The weights were downloaded from Serengil (2018). VGG Face was already applied successfully for different face detection problems, also for gender recognition (Foggia et al., 2019). It was chosen because it performed better than the regular Keras implementation of VGG-16 model for the dataset.

In Figure 14 the VGG Face architecture is shown. At the end of the model, the layers are substituted by two dense layers and one output layer (see Figure 14 in orange). The output layer is matched to the desired result, in the image this is a sigmoid activation function to give out binary gender. One layer of the predefined weights was included in the training, the weights of the layers before that remained fixed (see Figure 14 in yellow). To model emotions, a special model («EMO») is used, see chapter 3.4.7.

Figure 14

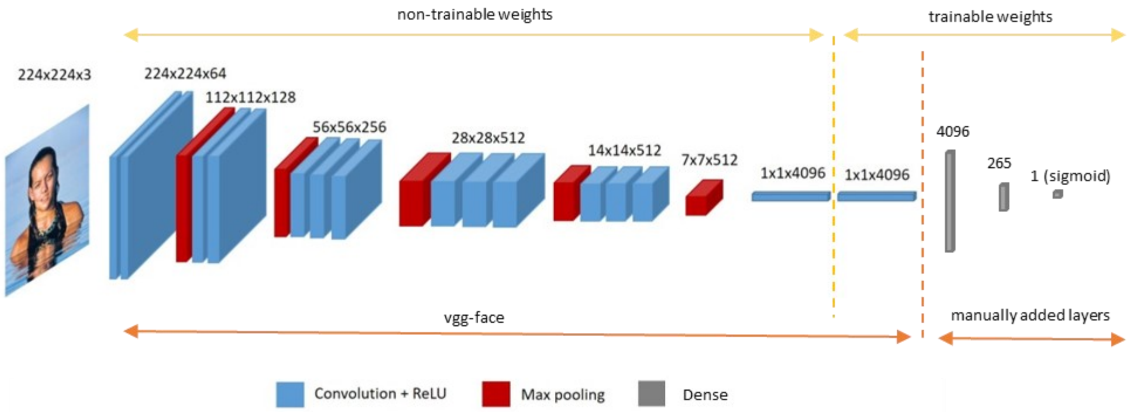


Figure 14. Model VGG-Face (Serengil, 2018), adapted by author

The training data is divided in training, validation and test datasets. The validation datasets are used to tune the hyperparameters (such as learning rate). The test datasets are used in the end to test the model's performance. The actual scraped instagram data can not be tested systematically as there are no labels available.

All python scripts are available in the appendix. All training for the five models follow the same process:

1. import Images
2. import Metadata
3. prepare labels
4. upsample if dataset is not balanced
5. split data into train, validation and test sets
6. prepare keras generators
7. prepare model
8. set metrics and hyperparameters
9. train model
10. evaluate model performance
11. save model

In Table 3 an overview of the model training is provided, the the next chapters the details of the model training is explained.

Table 3
Overview CNN models

feature	data_used	data_size	values	loss_value (validation)	Accuracy / R^2 (valid.)	Accuracy / R^2 (test)
gender	imdb wiki	62’171	2 genders	CCE[4] 0.42	Accuracy 0.836	Accuracy 0.842
age	imdb wiki	62’171	1-89 years	RMSE[5] 9.41	R^2 0.167	R^2 0.185
attractive-ness	SCUT-FBP5500	5’496	1-5 rating	RMSE 0.25	R^2 0.699	R^2 0.649
ethnicity	fairface	108’501	7 ethnicities	CCE 1.01	Accuracy 0.635	Accuracy 0.624
emotion	FER-2013	35’887	7 emotions	CCE 1.62	Accuracy 0.513	Accuracy 0.514

3.4.3. Model « Gender »

Data source : IMDB WIKI

The IMDB WIKI database contains over 500’000 faces labeled with age and gender. They are images of the 100’000 most popular actors in 2015 from the IMDb website and from Wikipedia. Rothe, Timofte and van Gool (2015) created the database and won the ChaLearn LAP 2015 challenge on apparent age estimation with it. The dataset can be downloaded from a website documenting the project (*IMDB-WIKI - 500k+ Face Images with Age and Gender Labels*, n.d.).

The images are preprocessed in the same way as the selfie data : face alignment, face detection, cropping of the face. 62’171 images remained after the preprocessing and were used for training.

Model : VGG-Face

The convolutional neural network model used is VGG-Face. At the end of the model, the layers are substituted by 2 dense layers and one output layer to generate a binary result (gender 0/1).

Training and performance

The model is trained for 50 epochs using a batch size of 64 (using all images per epoch) and adam optimizer. It reaches a 0.84 accuracy on the validation dataset and also 0.84 on the test dataset.

Figure 15

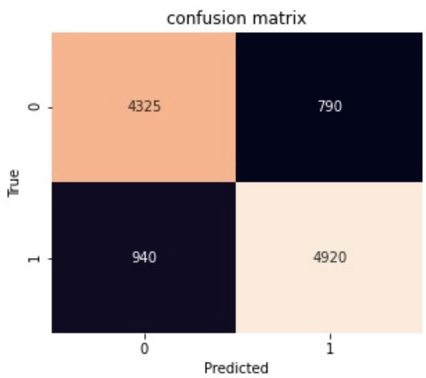


Figure 15. Confusion Matrix prediction for test dataset « gender » with accuracy 0.84

[4] CCE = categorical crossentropy
[5] RMSE = root mean squared error

3.4.4. Model « Age »

Data source : IMDB WIKI

The same data as for model *Gender* (see chapter 3.4.3) is used. The ages are in a range from 1 to 89 years. To train the model, the data is upsampled for ages that have a lower number of pictures to get a balanced dataset.

Model : VGG-Face

Age is trained as a continuous variable. VGG-Face is used as model. Instead of a sigmoid activation funtion like in model *Gender*, the output layer is not given any activation function to get the linear output.

Training and performance

The model is trained for 19 epochs using a batch size of 64 (using all images per epoch) and adam optimizer. It reaches a 9.41 root mean squared error on the validation dataset and 9.67 on the test dataset.
The R^2 on the test data is 0.185.

Figure 16

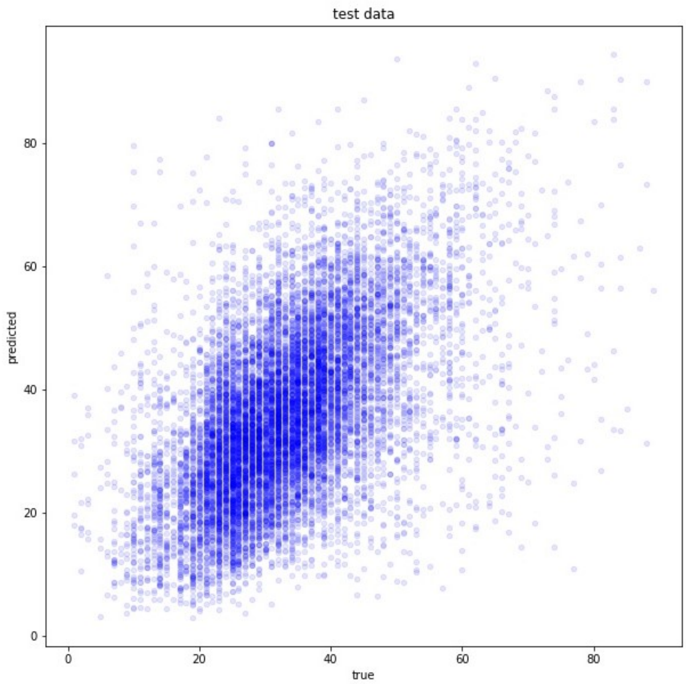


Figure 16. Scatterplot of true vs predicted age 1-89 on test dataset

3.4.5. Model « Attractiveness»

Data source : Beauty Score with SCUT-FBP5500 dataset

Liang et al. (2018) created the SCUT-FBP5500 Database with 5500 faces from the categorized into male/female and Asian/Caucasian. The are labeled with face landmarks and beauty scores from 1-5. The pictures were rated by a total of 60 persons, the score is an average of this rating. There are only caucasian and asian faces in the dataset, which might result in a bias later on when applying the mode. The dataset is available on Github (SCUT, 2017/2021).

Figure 17

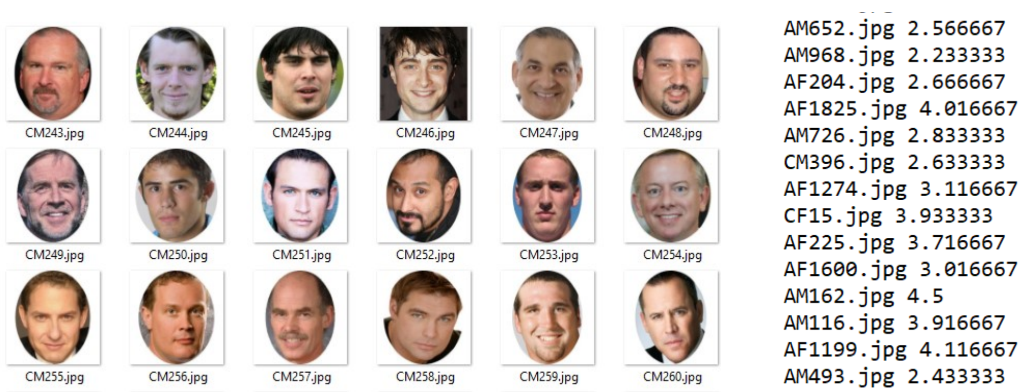


Figure 17. Extract of the 5500 downloaded pictures with faces and the text file with the corresponding rating

Model : VGG-Face

The same model as for model *Gender* (see chapter 3.4.3) is used, except that the output layer was not given any activation function to get the linear output.

Training and performance

The model is trained for 10 epochs using a batch size of 64 (using all images per epoch) and SGD optimizer with learning rate 1e-3 and decay 1e-5. It reaches a 0.25 root mean squared error on the validation dataset and 0.27 on the test dataset (attractiveness scale 1-5). The R^2 on the test dataset is 0.649.

Figure 18

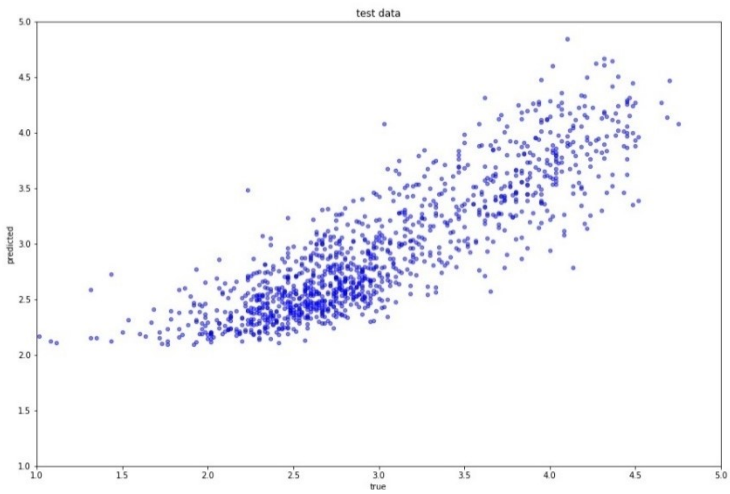


Figure 18. Scatterplot of true vs predicted attractiveness scores 1-5 on test dataset

3.4.6. Model « Ethnicity»

Data source: FAIR FACE

Karkkainen & Joo (2021) designed an face image database of 108'501 images 7 ethnicity groups (White, Black, Indian, East Asian, Southeast Asian, Middle Eastern, and Latino). The groups all have the same size. The dataset was originally used to detect racial bias. The dataset is available on Github (Joo, 2019/2021).

Model: VGG-Face

The same model as for model «Gender» (see chapter 3.4.3) is used, except that the output layer were 7 nodes with sigmoid activation functions representing the 7 ethnicity classes.

Training and performance

The model is trained for 11 epochs using a batch size of 64 (using all images per epoch) and adam optimizer. It reaches an accuracy of 0.62 on the validation dataset and also 0.62 on the test dataset (7 categories of ethnicities).

The algorithms perform well on *Black* (Accuracy 0.83), *East Asian* (Accuracy 0.76) and *White* (Accuracy 0.74). The least accurate prediction was for *Middle Eastern* (0.32), classifying a lot of pictures as *Hispanic* or *White*.

Figure 19

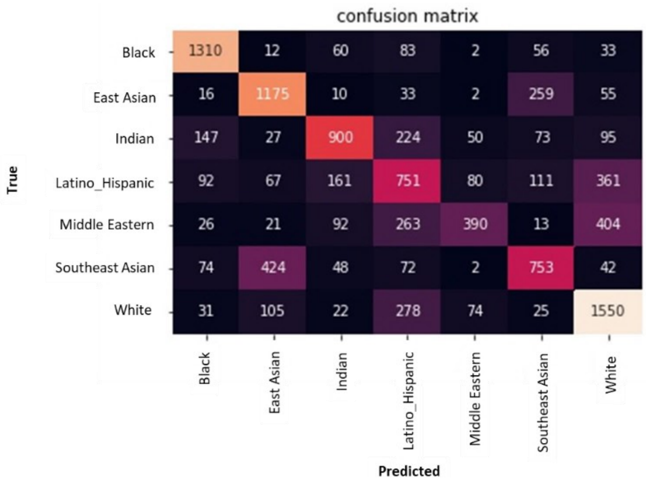


Figure 19. Confusion matrix «Ethnicity» on test dataset

3.4.7. Model « Emotion»

Data source: FER-2013 Emotion Recognition Dataset
FER-2013 was first used by Goodfellow et al. (2013, p. 3). According to them, « FER-2013 was created by Pierre Luc Carrier and Aaron Courville. It is part of a larger ongoing project. The dataset was created using the Google image search API to search for images of faces that match a set of 184 emotion-related keywords». After the API search, human labelers encoded the data manually. The dataset was enlarged since 2013, the version used consists of 35887 images in resolution 48x48 grayscale. It was downloaded from a database on Kaggle (Sambare, 2020).

Model: EMO
Because the images in the dataset are only 48x48x1, the VGG-Face model with pretrained weights can not be applied. A simpler network consisting of two convolutional layers, a dense layer and an output layer for the seven categories is used without pretrained weights.

Training and performance
The model is trained for 20 epochs using a batch size of 64 (using all images per epoch) and adam optimizer. It reaches an accuracy of 0.51 on the validation dataset and also 0.51 on the test dataset (7 categories of emotions). In the confusion matrix it can be interpreted that some emotions are difficult to classify, especially *angry* and *sad*.

Figure 20

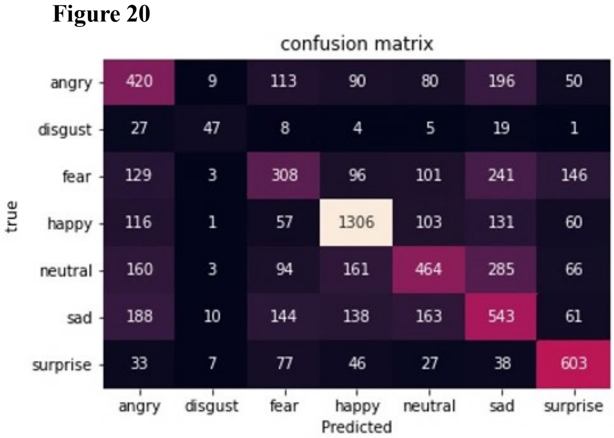


Figure 20. Confusion matrix « Emotion » on test dataset

3.4.8. Application of models to selfie dataset

The five models to determine gender, age, attractiveness, ethnicity and emotion are applied to the selfie dataset. In that way, the selfie images are labeled with the five characteristics of a person. In Figure 21 are two examples from the selfie dataset, labeled using the five trained CNN models.

Figure 21

Figure 21. Example of selfie images labeled using the five trained CNN models

3.5. Analysis of face proportions

This analysis aims to measure the proportions of a face. Studies show that attractiveness of a face can partly be assessed by certain ratios in face proportions (Hong et al., 2017).

3.5.1. Facial landmarks

Facial landmarks provide coordinates to a face and are used widely in face research. There are different implementations of facial landmarks, some giving out over 100 coordinates. In this thesis the 68 landmark predictor from the python libary *dlib* is used. An example of the application to a selfie image is shown in Figure **22**. Those 68 coordinates are the basis to calculate face proportions in the next step.

Figure 22

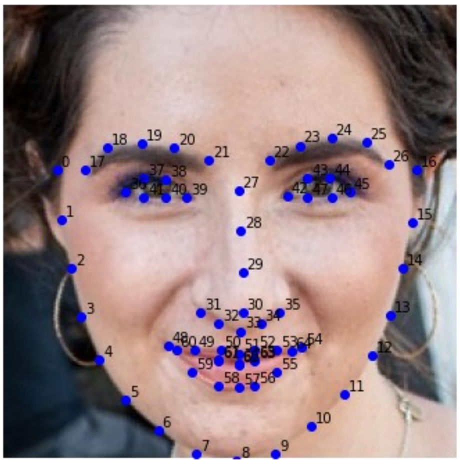
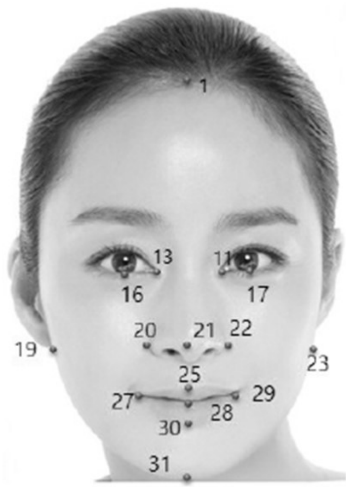


Figure 22. Dlib 68 Facial landmarks applied to example from selfie dataset

3.5.2. Face ratios

Hong et al. (2017, p. 7) provide a metastudy with different approaches to define face ratios. The approach with the highest correlation to human labeled attractiveness are the *golden ratios*. Therefore the golden ratios are calculated with the facial landmarks of the selfie dataset. The golden ratios are shown in Figure 23.

Figure 23



RFS-2: Golden Ratio	
Description (Distance/Distance)	Ratio vector
under eyes/interocular	dist(16,17)/dist(13,11)
under eyes/nose width	dist(16,17)/dist(20,22)
mouth width/interocular	dist(27,29)/dist(13,11)
upper lip-aw/interocular	dist(25,31)/dist(13,11)
upper lip-jaw/nose width	dist(25,31)/dist(20,22)
interocular/lip height	dist(13,11)/dist(25,30)
nose width/interocular	dist(20,22)/dist(13,11)
nose width/upper lip height	dist(20,22)/dist(25,28)
interocular/nose-mouth	dist(13,11)/dist(21,28)
lip height/nose-mouth	dist(25,30)/dist(21,28)
face height/face width	dist(1,31)/dist(19,23)
nose-jaw/mouth-jaw	dist(21,31)/ dist(28,31)
nose width/nose-mouth	dist(20,22)/dist(21,28)
mouth width/nose width	dist(27,29)/dist(20,22)

Figure 23. Golden ratios based on facial landmarks (Hong et al., 2017, p. 7)

The numbers of the facial landmarks used in this thesis are different from the ones used by Hong et al. The following table outlines how distances and ratios are calculated with the available facial landmarks and names them.

Table 4
Distances and ratios applied to selfie dataset with dlib 68 Facial landmarks

distances	ratios
d_under_eyes = dist(46,41)	r1 = d_under_eyes/d_interocular
d_interocular = dist(42,39)	r2 = d_under_eyes/d_nose_width
d_nose_width = dist(31,35)	r3 = d_mouth_width/d_interocular
d_mouth_width = dist(64,60)	r4 = d_upper_lip_jaw/d_interocular
d_upper_lip_jaw = dist(51,8)	r5 = d_upper_lip_jaw/d_nose_width
d_lip_height = dist(51,57)	r6 = d_interocular/d_lip_height
d_nose_width = dist(35,31)	r7 = d_nose_width/d_interocular
d_upper_lip_height = dist(51,66)	r8 = d_nose_width/d_upper_lip_height
d_nose_mouth = dist(33,66)	r9 = d_interocular/d_nose_mouth
d_face_height = dist_face_height(8)	r10 = d_lip_height/d_nose_mouth
d_face_width = dist(15,1)	r11 = d_face_height/d_face_width
d_nose_jaw = dist(33,8)	r12 = d_nose_jaw/d_mouth_jaw
d_mouth_jaw = dist(66,8)	r13 = d_nose_width/d_nose_mouth
	r14 = d_mouth_width/d_nose_width

The distance «d_face_height» could not be calculated using the simple coordinates because dlib does not provide a coordinate at the upper limit of the face. Instead, the coordinates of the upper image border (vertical) and in between the eyes (point 8, horizontal) were used to define that point.

3.6. Analysis with image decomposition (« Eigenfaces »)

This approach aims to decompose images into components that highlight different aspects of an image. Where the face proportion analysis focused on locating features on the face and the CNN analysis aimed to attribute certain labels to an image, this is the most holistic approach of the three. Some of the decompositions may concentrate of face attributes, others for example on contrasts and brightness of an image. As a result of this analysis, each image has 30 weights (*loadings*) that express how much of each component is present in the image.

The *Eigenface* approach originally proposed by Turk and Pentland (1991) is used. The application to the brand selfie topic is inspired by a study of Xiao and Ding (2014). They study if a person's face has an impact on the success of a print advertisement. The method of Eigenfaces applies principal component analysis on the image pixel values. Because the resulting eigenvectors can be visualized as a face, they are called *Eigenfaces*.

The python code used is adapted from scipy-lectures.org (*3.6.10.14. The Eigenfaces Example: Chaining PCA and SVMs — Scipy Lecture Notes*, n.d.). Each image is converted to grayscale and transformed into a 50'176x1 vector with the pixel values. The dataset is decomposed into 30 eigenvectors using PCA. The 30 most prominent eigenvectors make up for 78% of all the variance. Each image of the dataset is projected on the 30 eigenvectors to get a loading vector of 30x1. With this loading vector and the eigenvectors the faces could be reconstructed. In this study the loading vector is used to characterize each face in 30 numbers. Those 30 loadings are used in the next step of the analysis. In the following figure, the 30 Eigenfaces are shown.

Figure 24

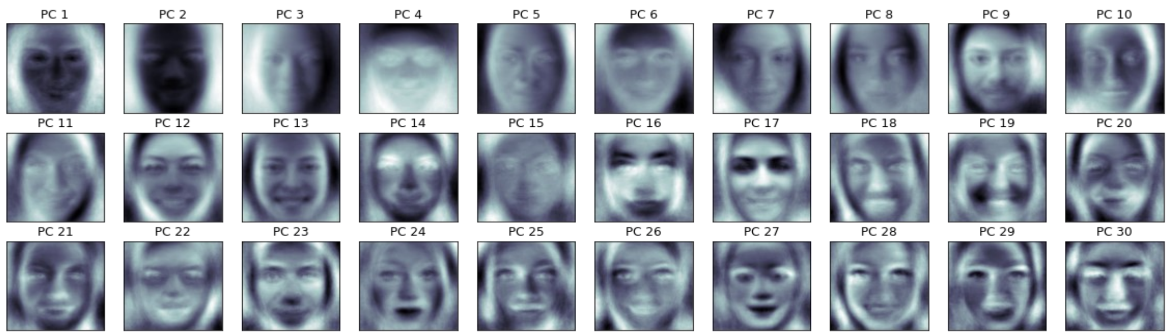


Figure 24. 30 Principal Components of Selfie Dataset

The Eigenfaces can be interpreted as being representations of image components. Each image has a loading (weight) for each image component. In the following Table is one example image that shows a high positive loading of the component PC 3 (positive number close to maximum) and one example image that shows a high negative loading of the component (negative number close to minimum). Judging from the eigenface itself and the examples, the Eigenface probably represents images where the lighting is coming from the left side.

Table 5
Example for the interpretation of Eigenface PC 3

Eigenface PC 3	Example image with component loading high	Example image with component loading low
		

The other Eigenfaces that have significant effects on the engagement in the following analysis are interpreted in chapter 3.7.2.

3.7. Analysis of the relation of faces and engagement

In this analysis, the variables generated in the last analysis steps (facial characteristics, face proportions, image decomposition) are tested for their influence on the engagement rate. An overview of all variables is provided in the appendix (Appendix A: Overview of variables).

3.7.1. Relation of facial characteristics and engagement

A multiple linear regression is calculated to predict *engagement* based on gender, emotion, ethnicity, age and attractiveness. In model 1 (Table 6) all variables generated by the CNN models are included.

The detected emotions do not have a significant impact on engagement (p-value < 0.05). Gender, age and attractiveness are significant in model 1. However when they are modeled without emotion and ethnicity, age proves not to be significant anymore. Therefore Model 2 is created with only gender and attractiveness. Some of the ethnicities are significant, others are not. The relationship is analysed further in Table 10.

Table 6
Model 1 – all characteristics as independent variables

Model 1: characteristics as independent variables

variable	coef	SE	p
Intercept	-1.233	0.173	0.000
Gender: Female*	0.000		
Gender: Male	-0.192	0.043	0.000
Emotion: angry*	0.000		
Emotion: fear	-0.413	0.241	0.086
Emotion: happy	0.041	0.077	0.595
Emotion: neutral	-0.001	0.079	0.993
Emotion: sad	-0.004	0.068	0.958
Emotion: surprised	-0.253	0.334	0.449
Ethnicity: Black*	0.000		
Ethnicity: East Asian	0.039	0.092	0.672
Ethnicity: Indian	-0.045	0.139	0.745
Ethnicity: Latino_Hispanic	0.238	0.081	0.004
Ethnicity: Middle Eastern	0.080	0.119	0.499
Ethnicity: Southeast Asian	0.232	0.120	0.054
Ethnicity: White	0.149	0.073	0.040
Age	0.004	0.002	0.039
Attractiveness	0.316	0.048	0.000
R2 of model	0.084		
Adjusted R2 of model	0.076		

* = those categorical values are used as base, they are added to the chart to provide a complete picture

Table 7
Model 2 – attractiveness and gender as independent variables with interaction

Model 2: characteristics as independent variables (significant variables) with interaction			
variable	**coef**	**SE**	**p**
Intercept	-1.175	0.153	0.000
Gender: Male	0.397	0.309	0.200
Attractiveness	0.377	0.048	0.000
Gender: Male interaction with Attractiveness	-0.199	0.066	0.066
R2 of model	0.072		
Adjusted R2 of model	0.070		

In a model with the independent variables *gender* and *attractiveness*, those variables interact (model 2). The variables are are modelled separately in model 3 and 4. The linear regression with only gender as the independent variable in model 3 shows that the attractiveness has a positive effect on the engagement rate. Here, every increase of 1 in the attractiveness rating (scale 1-5) has a positive effect of 0.411 on engagement.

Table 8
Model 3 – attractiveness as independent variable

Model 3: attractiveness as independent variable			
variable	**coef**	**SE**	**p**
Intercept	-1.327	0.119	0.000
Attractiveness	0.411	0.039	0.000
R2 of model	0.061		
Adjusted R2 of model	0.061		

The linear regression with only *gender* as the independent variable in model 4 shows that the gender male has a negative effect on the engagement rate.

Table 9
Model 4 – gender as independent variable

Model 4: gender as independent variable			
variable	**coef**	**SE**	**p**
Intercept	0.024	0.220	0.271
Gender: Male	-0.312	0.039	0.000
R2 of model	0.036		
Adjusted R2 of model	0.036		

In model 5, the linear regression of ethnicities as independent variables show that the detected ethnicity of a face does seem to have an impact on the engagement. In descending order, *Latino_Hispanic*, *White* and *Southeast Asian, Middle Eastern* and *East Asian* have a more positive impact on engagement compared to *Black*. The result for *Indian* is not significant (p>0.05).

Table 10
Model 5 – ethnicities as independent variables

Model 5: ethnicities as independent variables

variable	coef	SE	p
Intercept	-0.415	0.063	0.000
Ethnicity: Black	0.000		
Ethnicity: East Asian	0.196	0.092	0.034
Ethnicity: Indian	0.139	0.141	0.322
Ethnicity: Latino_Hispanic	0.432	0.080	0.000
Ethnicity: Middle Eastern	0.317	0.116	0.006
Ethnicity: Southeast Asian	0.340	0.123	0.006
Ethnicity: White	0.390	0.123	0.006
R2 of model	0.025		
Adjusted R2 of model	0.022		

3.7.2. Relation of attractiveness with eigenfaces and face proportions

A multiple linear regression with attractiveness as dependent variable showed correlations (R^2) of 0.31 with the eigenfaces and 0.30 with the face proportions (ratios). To explore this relation further, a lasso regression is performed for feature engineering.

The first step is to normalize the eigenface loadings and ratios so they all have a standard deviation of 1 and a mean of 0.

Eigenfaces vs attractiveness

By using grid search, an alpha of 0.02 is defined. In the following chart, all significant eigenfaces are shown (named PC+ number 1-30, PC = principal component). For the eigenfaces with Top 5 positive and Top 5 negative effects an interpretation attempt by the author is shown in Table 11 along with example images having a high and low weight of that eigenface.

Figure 25

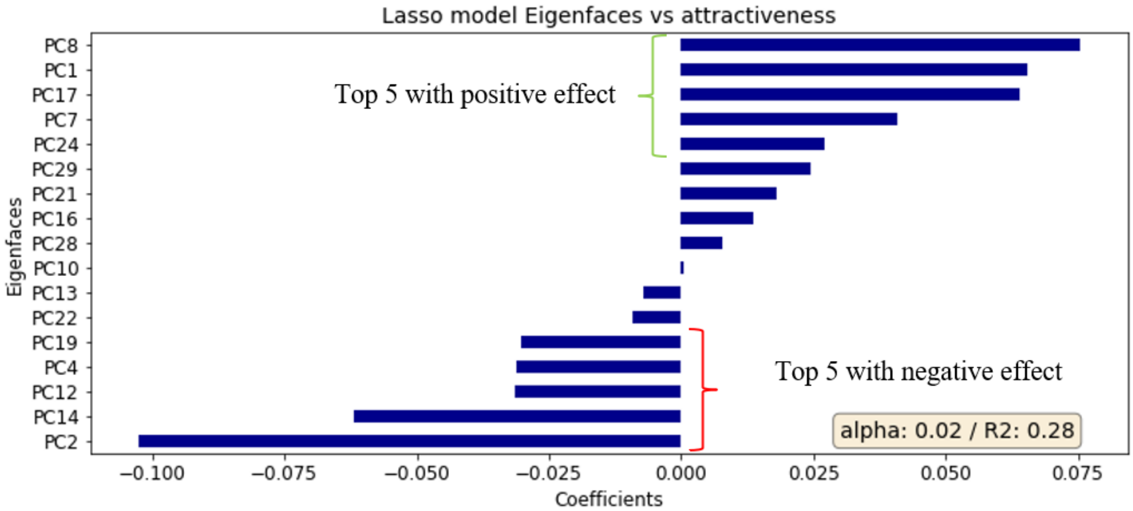
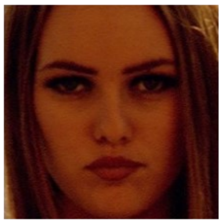
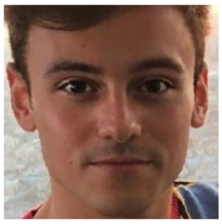
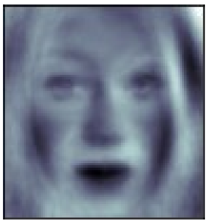
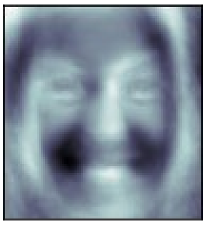
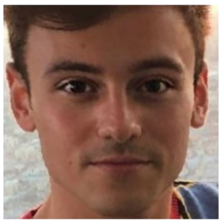


Figure 25. Significant Eigenfaces to predict attractiveness.

Table 11
Top 10 Eigenfaces with impact on attractiveness – Interpretation attempts

Component (Eigenface) Name	Component (Eigenface) Image	Component weight high	Component weight low	Interpretation
Positive effects				
PC 8				indirect soft light, object and background balanced, femininity
PC 1				Clear distinction face - background
PC 17				Accentuated eyes, makeup
PC 7				Indirect soft light, shadow on the sides, slim face
PC 24				Accentuated lips and eyes, makeup
Negative effects				
PC 19				Moustache or lower light on mouth area

Component (Eigenface) Name	Component (Eigenface) Image	Component weight high	Component weight low	Interpretation
PC 4				Light from below, hat
PC 12				Photographed too close, fisheye perspective
PC 14				Bright eye area, lack of accentuated eyes
PC 2				Light from behind the face, dark face

Face proportions vs attractiveness

The analysis for eigenfaces is repeated for face proportions. The result is shown in Figure 26. With an alpha of 0.001, 11 significant variables remained. An interpretation attempt is provided in Table 12.

Figure 26

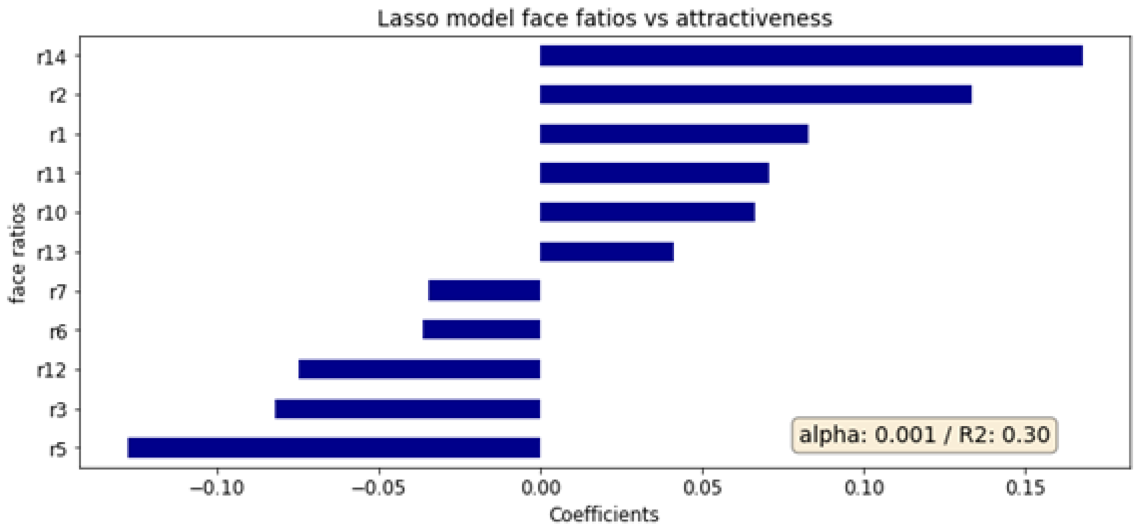


Figure 26. Significant face proportions to predict attractiveness

Table 12
Top 11 Face ratios with impact on attractiveness – Interpretation attempts

Ratio Name	Ratio definition	Interpretation
Positive effects		
r14	d_mouth_width/d_nose_width	wider mouth / narrower nose
r2	d_under_eyes/d_nose_width	eyes far apart (middle) / narrower nose
r1	d_under_eyes/d_interocular	big eyes
r11	d_face_height/d_face_width	face form (more oval)
r10	d_lip_height/d_nose_mouth	no interpretation
r13	d_nose_width/d_nose_mouth	no interpretation
Negative effects		
r5	d_upper_lip_jaw/d_nose_width	bigger jaw / narrower nose
r3	d_mouth_width/d_interocular	wider mouth / eyes far apart (point at nose)
r12	d_nose_jaw/d_mouth_jaw	mouth low between nose and jaw
r6	d_interocular/d_lip_height	no interpretation
r7	d_nose_width/d_interocular	no interpretation

Combination of eigenfaces and face proportions vs attractiveness

A combination of all the eigenfaces and face proportions reaches an R^2 of 0.42 with attractiveness. With an alpha of 0.05, the R^2 is 0.25 and 11 significant variables remain.

To put the interpretation of this model in words, one could say: an attractive image is an image that is photograped well lit[6] with a person having accentuated big eyes and a narrow nose.

Table 13
Eigenfaces and face proportions with impact on attractiveness – Interpretation attempts

Name	Ratio definition / Eigenface image	Interpretation
Positive effects		
r2	d_under_eyes/d_nose_width	eyes far apart (middle) / narrower nose
r1	d_under_eyes/d_interocular	big eyes
PC1		Clear distinction face - background
PC8		Indirect soft light, shadow on the sides, slim face
PC17		Accentuated eyes, makeup
r9	d_interocular/d_nose_mouth	no interpretation
r10	d_lip_height/d_nose_mouth	no interpretation
r14	d_mouth_width/d_nose_width	wider mouth / narrower nose

[6] The findings concerning light correspond with what is known in photography to be «well lit» (Dunsford, 2017)

Name	Ratio definition / Eigenface image	Interpretation
Negative effects		
PC2		Light from behind the face, dark face
PC14		Bright eye area, lack of accentuated eyes
r4	r4 = d_upper_lip_jaw/d_interocular	Longer face

3.7.3. Relation of image decompositions and face proportions with engagement

A multiple linear regression is calculated to predict engagement based on the 30 loadings of Eigenfaces and the 14 face ratios. With this direct approach, feature engineering of the variables is difficult. Only few show a significant p-values[7] and the correlation is low. When removing some of the variables, other variables showed significant p-values than before. Therefore, those models can not be viewed as statistically sound and are not shown here.

Using the insight from the previous chapter, the variables that had a significant impact on attractiveness are also tested against engagement. Only the two most significant eigenfaces from the previous analysis showed to have an impact, although the R^2 is low with 0.015 (model 6).

Table 14
Model 6 – eigenfaces as independent variables

Model 6: Eigenfaces as independent variables (significant variables)

variable	coef	SE	p
Intercept	-0.081	0.018	0.000
PC8	0.070	0.019	0.000
PC2	-0.063	0.018	0.001
R2 of model	0.015		
Adjusted R2 of model	0.014		

[7] 5 of the Eigenfaces reached significant results (PC2, PC8, PC15, PC12, PC15). The same was done separately with the face ratios where only 4 were significant (r3, r10, r11, r14).

Three of the face proportion ratios show a significant impact on engagement as shown in model 7.

Table 15
Model 7 – face proportions as independent variables

Model 7: Face proportions as independent variables (significant variables)			
variable	**coef**	**SE**	**p**
Intercept	-2.450	0.391	0.000
r14	0.240	0.075	0.001
r2	0.243	0.051	0.000
r1	0.659	0.183	0.000
R2 of model	0.030		
Adjusted R2 of model	0.029		

A combination of the ratios and eigenfaces is possible, resulting in model 8 in Table 16. The variables used in this model and their interpretation attempt are shown in Table 17.
.

Table 16
Model 8 – eigenfaces and face proportions as independent variables

Model 8: Eigenfaces and Face proportions as independent variables (significant variables)			
variable	**coef**	**SE**	**p**
Intercept	-2.190	0.399	0.000
r14	0.198	0.076	0.009
r2	0.204	0.053	0.000
r1	0.614	0.183	0.001
PC8	0.047	0.019	0.012
PC2	-0.037	0.019	0.052
R2 of model	0.036		
Adjusted R2 of model	0.033		

Table 17
Eigenfaces and face proportions with impact on engagement – Interpretation attempts

Ratio Name	Ratio definition	Interpretation
Positive effects		
r14	d_mouth_width/d_nose_width	wider mouth / narrower nose
r2	d_under_eyes/d_nose_width	eyes far apart (middle) / narrower nose
r1	d_under_eyes/d_interocular	big eyes
PC8		Indirect soft light, shadow on the sides, slim face
Negative effects		
PC2		Light from behind the face, dark face

Because of the interactions between the variables, especially with the variable attractiveness, it is not possible to create a combined model of face characteristics and eigenfaces/face proportion.

In order to assess if a result was significant, the p-values are used. The correlations with engagement are all under 0.1, which can be regarded as low. As this is topic studying peoples behaviour and there are many factors that might influence engagement for a post other than a specific feature of a face, no high correlations could be expected and the results are seen as valid as long as the p-value is in the significant range.

4. Results and discussion

4.1. Results

In summary, significant results were found when modelling attractiveness, gender, ethnicity, face proportions and image decompositions ("Eigenfaces") against engagement. Attractiveness had the biggest impact. Emotion and age did not have a significant effect.

Hypothesis H1: The emotion of a face on a brand selfie influences the engagement rate

Emotion can not be proven to have a significant effect on the engagement rate.

Hypothesis H2: The age of a face on a brand selfie influences the engagement rate

Age can not be proven to have a significant effect on the engagement rate.

Hypothesis H3: The ethnicity of a face on a brand selfie influences the engagement rate

The ethnicity labels do have an impact on the engagement rate. The two with the most positive effect on engagement are *White* and *Latino_Hispanic* compared to *Black*. The impact is not very strong with an R^2 of 0.025.

Hypothesis H4: The gender of a face on a brand selfie influences the engagement rate

Gender does show significant impacts on the engagement rate in models where no interaction with attractiveness is taken into account. In a model with only gender as the independent variable (Table 9), gender is significant. Male faces, categorized by the trained model, recieve less engagement than female faces.

Further, the models show that there is an interaction between gender and attractiveness, meaning the trained CNN models label female faces as more attractive than male faces. This indicates gender bias.

Hypothesis H5: The attractiveness of a face on a brand selfie influences the engagement rate

Attractiveness, rated by the trained CNN model, does have the clearest effect on the engagement rate of the brand selfie. If a face is rated as more attractive, the engagement rate is higher. This is shown in Model 4 (Table 8), where attractiveness as the independent variable in the ordinary least squared regression has a R^2 of 0.061. Every increase of 1 in the attractiveness rating (scale 1-5) has a positive effect of 0.411 on engagement (range ~8).

Hypothesis H6: The face proportions of a face on a brand selfie influence the engagement rate

11 out of the 14 face proportions correlate with attractiveness with an R^2 of 0.30. 3 of the face proportions also have a significant effect on engagement rate. The three ratios having an impact on engagement rate indicate that a narrow nose and big eyes are favorable.

It needs to be noted however that only 3 of the 14 proportions have a significant effect and that this might also be due to interactions with attractiveness.

Hypothesis H7 Images can be broken down into components which have an influence on the engagement rate

17 out of the 30 principal components correlate with attractiveness with an R^2 of 0.28. 2 of the components also have a significant effect on engagement rate. The two components having an impact on engagement rate indicate that indirect soft light, shadow on the sides and a slim face are favorable and light from behind the face or a dark face are not favorable. It needs to be noted that this is an interpretation by the author.

4.2. Discussion, limitations and recommendations for further work

Attractiveness impacts engagement

This thesis shows that the attractiveness of a face has an impact on the engagement, meaning a brand selfie with an attractive person recieves more likes. This corresponds with findings that people who are more attractive generally get preferential treatment (see chapter 2.7). It is also in line with the practises of the advertisement industry to generally employ models that are attractive.

In this thesis, the assessment of attractiveness is based on a machine learned model that tries to learn the complex concept of attractiveness. It does find patterns that ultimately have an relationship with the engagement rate. How those patterns really express attractiveness and what the criteria are to rate the attractiveness are not clear with this approach.

The analysis of the face proportions and image decompositions give some insight in how the attractiveness model works. It identifies that proportions indicating a narrow nose and big eyes relate to higher attractiveness. The image decompositions indicate that certain features like accentuated eyes are favorable for attractiveness while facial hair is not. This gives some indication on the beauty standards that the model uses.

Separating photography technique from physionometry of the face

It is challenging to extract only the physionometry of a face out of a picture and to disregard angles in which a picture was taken or the lighting of the image. In this thesis, only the frontal face was aligned so the eyes were horizontal. Other angles or aspects of photography were not changed or analysed. The findings in the eigenface analysis suggest that not only facial features are important but also the lighting and angles in which a person was photographed. This could be analysed further.

Gender bias and racial bias

Faces that were attributed the gender *male* by the trained CNN model recieved less engagement. According to an article in HopperHQ (‘Who Runs The World?’, 2017) and Fowler (2017) it is known that males recieve less likes and followings on social media.
In this thesis, ethnicity also had an small impact on engagement. Although the impact is not very strong, this indicates a racial bias that would need to be analysed further.

No relation between emotion/age and engagement

Emotion and age, labelled by the trained CNN models, did not have significant impacts on the engagement. When training the CNN model for age, the R^2 of predicted vs actual age was only 0.185. It is possible that this model did not perform well enough when labelling the selfie data that it would have made an impact on engagement or it is simply not relevant.

Emotions also did not have a significant relation with engagement. It is possible that emotions are too complex to model with an OLS regression or, as in age, that they are just not relevant.

Methods for data scraping and preprocessing

Choosing the profiles manually to scrape posts made it possible to compare the posts because one could assume that a post on a profile has similar conditions (exposure, type of followers) than the next post on the profile. The downside to this approach is that the choice of pictures is not random and from 10 profiles only.

The rigourous selection of the posts to analyse was necessary to have a solid dataset for the analysis. Only posts were selected with one frontal, well visible face. From the ~40'000 images that were originally scraped, only ~2'000 were used for the final analysis (chapter 3.7). Although many images were disregarded during the preprocessing, the dataset was still sufficiently large to perform the regression analysis.

Methods for the face analysis

A broad analysis could be performed by applying three different face analysis approaches. Those approaches were all applied for face analysis before and are known in their fields of study.

Using separate training datasets for the CNN models made it possible to select datasets that were designed specifically for that purpose. What a model learns and how it labels new pictures in the end is dependent on the training dataset. The dataset used to train the attractiveness model only includes asian and white people which indicates that it will probably learn beauty standards that are typical for those ethnicities. Other datasets did not have any obious shortcomings and none were detected in the process of this thesis.

The face proportions and the image decompositions/eigenfaces left room for interpretation. An interpretation attempt is given in this thesis. It is important to be aware that this is a subjective interpretation based on observation.

Methods for the relation analysis between face and engagement

Analysing visual content poses the challenge that oftentimes this content is not labelled. By training the convolutional neural networks to label the images, this could be achieved. However it is important to note that there was no way to control if the model labelled the selfie dataset correctly other than using models that perform well on labelled testing data and visual inspection of some samples of the selfie data. Those labels (age, gender, attractiveness, ethnicity, emotions) are then used in the next step to analyse the relationship with engagement. All results involving those variables need to take this into account.

The linear regression methods (OLS and Lasso) were chosen because they are easily explainable. Other methods using more spohisticated and/or non-linear models might add to the insights of this thesis.

Forming groups

The thesis concentrated on consumer products in the product categories food, beverages, clothing, cosmetics and jewelery. More groups were not formed. For example distinctions that are common in marketing like target groups could be made in a next analysis.

An analysis based on combined labels might give interesting insights as well. For example one might form groups like males of a certain age group vs females of a certain age group.

Engagement is not equal to purchase propensity

The ultimate goal of a brand is usually not to get engagement from users but for people to buy their products and to promote their products further. It can not simply be said that a higher engagement for a post that includes a brand leads to a higher purchase propensity for the brand as this topic is more complex. It would be interesting what role the faces play when it comes to the actual purchase decisions.

4.3. Conclusion

This thesis aimed to analyse if facial characteristics in brand selfies influence the engagement rate for the selfie on social media. Based on three different, state-of-the-art face analysis techniques that provided data on the face and regression models that put the face data into relation with the engagement, it can be concluded that facial characteristics do have an impact on the engagement for a brand selfie post.

The findings support that the preferential treatment of more attractive faces also applies to social media engagement. Gender and ethnicity also have a relation with engagement, although weaker than attractiveness. The lighting and angle of the photo might also be relevant.

To get a deeper understanding of the topic, future studies could analyse the different characteristics in more depth and also make combinations. Photo techniques of a brand selfie could also be a topic of research. Using more complex and non-linear model to analyse the relations between faces and engagement might give further insights. They could also look at different segments of selfie posters, selfie viewers or product segements or broaden the research question from the impact on engagement to the impact on buying propensity.

Bibilography

3.6.10.14. The eigenfaces example: Chaining PCA and SVMs—Scipy lecture notes. (n.d.). Retrieved 23 April 2021, from https://scipy-lectures.org/packages/scikit-learn/auto_examples/plot_eigenfaces.html

Ashleigh | The Craft Brewnette (@thecraftbrewnette). (2021, May 1). https://www.instagram.com/p/COTsWlmFZI6/

Bakhshi, S., Shamma, D. A., & Gilbert, E. (2014). Faces engage us: Photos with faces attract more likes and comments on Instagram. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 965–974. https://doi.org/10.1145/2556288.2557403

BeerSelfie.com (@beerselfie). (2021, May 1). https://www.instagram.com/p/COT0y_ILYUQ/

Berger, J., & Milkman, K. L. (2012). What Makes Online Content Viral? *Journal of Marketing Research*, *49*(2), 192–205. https://doi.org/10.1509/jmr.10.0353

Bharti, M. (2020). *When too much "me" is bad for "us": The detrimental effect of selfie on self-brand connection*. https://dr.ntu.edu.sg//handle/10356/142791

Birbaumer, M. (2020). *Deep Learning in Vision—Autumn Semester 2020*. https://elearning.hslu.ch/

Brooks, A. (2018, August 7). Why Influencers Are the Missing Link in Brand Loyalty. *PerformanceIN*. https://performancein.com/news/2018/08/07/why-influencers-are-missing-link-brand-loyalty/

Caliandro, A., & Graham, J. (2020). Studying Instagram Beyond Selfies. *Social Media + Society*, *6*(2), 2056305120924779. https://doi.org/10.1177/2056305120924779

Clarifai. (2020). *Clarifai's Demographics Model | Analyze Detected Faces in Images*. https://www.clarifai.com/models/demographics-image-recognition

Dunkin' (@dunkin). (2021, May 3). https://www.instagram.com/dunkin/

Dunsford, R. (2017, September 15). *8 Portrait Photography Tips Every Photographer Should Know*. Photography Pro. https://photographypro.com/portrait-photography/

Eisenthal, Y., Dror, G., & Ruppin, E. (2006). Facial Attractiveness: Beauty and the Machine. *Neural Computation*, *18*(1), 119–142. https://doi.org/10.1162/089976606774841602

Etcoff, N. (2011). *Survival of the Prettiest: The Science of Beauty*. Knopf Doubleday Publishing Group.

Foggia, P., Greco, A., Percannella, G., Vento, M., & Vigilante, V. (2019). A system for gender recognition on mobile robots. *Proceedings of the 2nd International Conference on Applications of Intelligent Systems*, 9. https://doi.org/10.1145/3309772.3309781

Fowler, D. (2017, March 15). *Women Are More Popular On Instagram Than Men According To New Study*. Grazia. https://graziadaily.co.uk/life/real-life/instagram-study-gender-likes/

Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D.-H., Zhou, Y., Ramaiah, C., Feng, F., Li, R., Wang, X., Athanasakis, D., Shawe-Taylor, J., Milakov, M., Park, J., … Bengio, Y. (2013). Challenges in Representation Learning: A report on three machine learning contests. *ArXiv:1307.0414 [Cs, Stat]*. http://arxiv.org/abs/1307.0414

Haenlein, M., Anadol, E., Farnsworth, T., Hugo, H., Hunichen, J., & Welte, D. (2020). Navigating the New Era of Influencer Marketing: How to be Successful on Instagram, TikTok, & Co. *California Management Review*, *63*(1), 5–25. https://doi.org/10.1177/0008125620958166

Halberstadt, J., Pecher, D., Zeelenberg, R., Ip Wai, L., & Winkielman, P. (2013). Two Faces of Attractiveness: Making Beauty in Averageness Appear and Reverse. *Psychological Science*, *24*(11), 2343–2346. https://doi.org/10.1177/0956797613491969

Hartmann, J., Heitmann, M., Schamp, C., & Netzer, O. (2021). The Power of Brand Selfies in Consumer-Generated Brand Imagery. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3354415

Hofstetter, R., Kunath, G., & John, L. K. (2020). From Sweetheart to Scapegoat: Brand Selfie-Taking Shapes Consumer Behavior. *Harvard Business School Working Paper, No. 20-085, February 2020.* https://www.hbs.edu/faculty/Pages/item.aspx?num=57648

Hong, Y., Nam, G., Choi, H., Cho, J., & Kim, I.-J. (2017). A Novel Framework for Assessing Facial Attractiveness Based on Facial Proportions. *Symmetry*, *9*, 294. https://doi.org/10.3390/sym9120294

How The Tiktok Algorithm Rates Your Face | Beauty Culture Ep. 6. (2020, June 26). https://www.youtube.com/watch?v=PqTXpJafq2M&ab_channel=QOVESStudio&app=desktop

Husain, S., Ghufran, A., & Chaubey, D. S. (2016). *Relevance of Social Media in Marketing and Advertising*. https://search.proquest.com/docview/1906047118?pq-origsite=gscholar&fromopenview=true

Huynh, L. (2021, March 8). Like Follower Ratio: How Many Likes to Get From 100 to 1-Million. *AskNugget*. https://www.asknugget.com/like-follower-ratio/

IMDB-WIKI - 500k+ face images with age and gender labels. (n.d.). Retrieved 31 October 2020, from https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/

Joo, J. (2021). *Joojs/fairface*. https://github.com/joojs/fairface (Original work published 2019)

Karkkainen, K., & Joo, J. (2021). *FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age for Bias Measurement and Mitigation*. 11.

Kavanagh, C. (2020). *An exploration of how Instagram Influencers build a reputable self-brand both with their followers and company brands.* [Masters, Dublin, National College of Ireland]. http://norma.ncirl.ie/4608/

Langlois, J. H., & Roggman, L. A. (1990). Attractive Faces Are Only Average. *Psychological Science*, *1*(2), 115–121. https://doi.org/10.1111/j.1467-9280.1990.tb00079.x

Li, Y., & Xie, Y. (2019). Is a Picture Worth a Thousand Words? An Empirical Study of Image Content and Social Media Engagement: *Journal of Marketing Research*. https://doi.org/10.1177/0022243719881113

Liang, L., Lin, L., Jin, L., Xie, D., & Li, M. (2018). SCUT-FBP5500: A Diverse Benchmark Dataset for Multi-Paradigm Facial Beauty Prediction. *ArXiv:1801.06345 [Cs]*. http://arxiv.org/abs/1801.06345

Liu, S., Fan, Y.-Y., Samal, A., & Guo, Z. (2016). Advances in computational facial attractiveness methods. *Multimedia Tools and Applications*, *75*(23), 16633–16663. https://doi.org/10.1007/s11042-016-3830-3

Lou, C., Tan, S.-S., & Chen, X. (2019). Investigating Consumer Engagement with Influencer- vs. Brand-Promoted Ads: The Roles of Source and Disclosure. *Journal of Interactive Advertising*, *19*(3), 169–186. https://doi.org/10.1080/15252019.2019.1667928

Menon, A. (2019, April 23). *Face Detection in 2 Minutes using OpenCV & Python*. Medium. https://towardsdatascience.com/face-detection-in-2-minutes-using-opencv-python-90f89d7c0f81

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review*, *98*(2), 164–181. https://doi.org/10.1037/0033-295X.98.2.164

Oxford English Dictionary—Selfie. (n.d.). In Oxford English Dictionary (Ed.), *OED Online*. Oxford University Press. Retrieved 7 May 2021, from https://www.oed.com/view/Entry/390063

Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). Deep Face Recognition. *Procedings of the British Machine Vision Conference 2015*, 41.1-41.12. https://doi.org/10.5244/C.29.41

Priadana, A., & Habibi, M. (2019). Face Detection using Haar Cascades to Filter Selfie Face Image on Instagram. *2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT)*, 6–9. https://doi.org/10.1109/ICAIIT.2019.8834526

Rothe, R., Timofte, R., & Gool, L. V. (2015). DEX: Deep EXpectation of Apparent Age from a Single Image. *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, 252–257. https://doi.org/10.1109/ICCVW.2015.41

Sambare, M. (2020, July 19). *FER-2013*. https://kaggle.com/msambare/fer2013

SCUT, D. L. and V. C. L. (2021). *HCIILAB/SCUT-FBP5500-Database-Release* [Python]. https://github.com/HCIILAB/SCUT-FBP5500-Database-Release (Original work published 2017)

Serengil, S. (2018, August 6). Deep Face Recognition with VGG-Face in Keras | sefiks.com. *Sefik Ilkin Serengil*. https://sefiks.com/2018/08/06/deep-face-recognition-with-keras/

Serengil, S. (2020, February 23). Face Alignment for Face Recognition in Python within OpenCV. *Sefik Ilkin Serengil*. https://sefiks.com/2020/02/23/face-alignment-for-face-recognition-in-python-within-opencv/

Sung, Y., Kim, E., & Choi, S. M. (2018). #Me and brands: Understanding brand-selfie posters on social media. *International Journal of Advertising*, *37*(1), 14–28. https://doi.org/10.1080/02650487.2017.1368859

Turcanu, D. (2018, November 22). The Instagram Graph API for Hashtag Search | What Changes in Walls.io. *Walls.Io Social Wall - Blog*. https://blog.walls.io/product/instagram-graph-api-for-hashtag-search-walls-io/

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*(1), 71–86. https://doi.org/10.1162/jocn.1991.3.1.71

Who Runs The World? Female Instagram users get five times more likes on average than men. (2017, March 13). *Hopper HQ Instagram Scheduler*. https://www.hopperhq.com/blog/runs-world-female-instagram-users-get-five-times-likes-average-men/

Xiao, L., & Ding, M. (2014). Just the Faces: Exploring the Effects of Facial Features in Print Advertising. *Marketing Science*, *33*(3), 338–352. https://doi.org/10.1287/mksc.2013.0837

Images for Figure 1 are retrieved on 15.10.2020 from :

asos (2020). [Guy with animal print jacket]. Instagram. https://www.instagram.com/p/B548Te0Desc/

catseyeanitsirc (2018). [Woman with Coca Cola]. Instagram. https://www.instagram.com/p/Bd6rsJnHvrI/

izandco (2020). [Hand with ring]. Instagram. https://www.instagram.com/p/l5H9pIO1N_/

Neuesichten.com. Wie machst du das perfekte Selfie. https://neuessichten.com/das-ist-noch-blindtext/

Stefancik, S. (2018). Smiling woman wearing white and black pinstriped collared top. Unsplash. https://unsplash.com/photos/QXevDflbl8A

Tandon, R. (2016). Coral labeled bottle. Unsplash. https://unsplash.com/photos/vbx7D2rDgUI

Declaration of originality

I hereby confirm that I have written this Thesis independently and without the help of any third party, have provided all the sources and cited all the literature I used and respect the copyright regulations of Lucerne University of Applied Sciences and Arts.

Date and Signature

C. Matt

Carole Mattmann
Lucerne University of Applied Sciences and Arts
Master in Applied Information and Data Science (MScIDS)
Lucerne, 29th May 2021