CHAPTER 11

Computer Vision for the Study of Older (and Younger) Adult Faces

Approaches, Advances, and Applications

Daniel N. Albohn, Joseph C. Brandenburg, and Reginald B. Adams, Jr.

The use of computer vision to study psychological phenomena has become increasingly popular over the last few decades, particularly for addressing facial appearance-based questions. This is perhaps unsurprising given the impressive advances that computer vision, and more specifically machine learning, have made in that time. In the span of less than twenty years, the field has gone from being able to detect faces in images within seconds (a feat heralded as breakthrough at the time) to being able to generate hyper-realistic faces of people who do not actually exist in reality (Karras et al., 2018; Viola & Jones, 2004).

As with other technological advances that have come before (e.g., fMRI), the tools that computer vision has to offer psychological and behavioral science are both exciting and cause for caution and introspection. For example, computer vision has provided a way to quantify how similar neutral faces incidentally resemble emotional expressions based solely on face features such as shape and without the confounded influence of age, race, gender, context, and so on that human observers inevitably integrate into their judgments (Adams et al., 2022; Albohn & Adams, 2020b; Zebrowitz et al., 2010). On the other hand, there have been myriad research studies and debates on the bias that machine learning models can inherit from humans, raising ethical concerns related to privacy and fairness (Coeckelbergh, 2020; Wehrli et al., 2021). In a particularly poignant example of the latter, a machine learning model that Google used to auto tag images was found to automatically categorize Black men as "gorillas" (Simonite, 2018). Critically, many of the tasks that AI has been trained to infer from faces do not align with the general public's opinion of what it *should* infer, such as personality characteristics from static facial portraits

https://doi.org/10.1017/9781009209656.012 Published online by Cambridge University Press

(Engelmann et al., 2022). However, despite exciting advances and worrisome drawbacks, what is clear is that machine learning (i.e., computer vision, artificial intelligence) is now becoming an integrated part of our society that is unlikely to leave anytime soon.

In the remainder of this chapter, we outline important computer vision techniques that can be (and are) used to inform important questions related to face perception generally, and to the perception of older adult faces specifically. We divide the use of computer vision for such research into two parts. In the first section, we discuss how computer vision has been used for problems related to classification and prediction, such as determining whether a face is present in an image or estimating the age of an individual. In the second section, we discuss how computer vision has been used to manipulate and generate images, such as aging a young adult face or creating an image of an older adult who does not actually exist in reality. Of course, the division between these two categories is only for pragmatic purposes. Oftentimes researchers use both techniques in combination to answer the questions that they pose.

At this time, it is important to note that many of the techniques and research we discuss in the remainder of this chapter can be applied to research in the domain of older adult person perception and beyond. Focusing on advances in computer vision related to the perception of older adults is interesting and important for a number of factors.

First, even though individuals can belong to many different social categories simultaneously over their lifetimes, age is unique in that it is a category that individuals 1) do not choose to be part of (unlike, e.g., Republican or Democrat) and 2) are not born into.[1] Since (old) age is a social category that no one voluntarily opts into, yet many eventually become part of, it provides an interesting social construct that can lend important insight into human behavior.

Second, age is important to examine within the lens of computer vision because it has often been overlooked when compared to other social categories such as race and gender/sex. Indeed, much of the social discourse on bias in machine learning has focused on race and gender biases, but relatively little work has discussed how, when, and where such biases can impact older adults. This issue is compounded further when one considers the research that has examined such issues from an intersectional perspective (see Hedgecoth, Chapter 5, this volume). Because of these two factors, we frame our discussion within the context of research that has examined age as a primary factor (whether young, middle, or older adults). However, we recognize that many of the techniques discussed apply more broadly.

[1] You are born into *an* age group (i.e., "infant"), but you are not born into the social category of "older adult."

https://doi.org/10.1017/9781009209656.012 Published online by Cambridge University Press

To summarize, it is clear that machine learning is now a fundamental component of society and will continue to evolve alongside its related social issues such as big data mining and privacy. While machine learning can be used across many domains that range from predicting consumer behavior to reading emotions, relatively little work has used machine learning techniques to understand these questions as they relate to older adults. Such a dearth of research is surprising given that older adults are the fastest growing age group in the world (United Nations, 2020), not to mention a social category that most individuals involuntarily "age into."

Classification and Prediction

The notion that computer algorithms are capable of computing and predicting complex real-world data has been part of scientific research and advancement for decades. The roots of machine learning date back to seminal work in the 1950s by computer scientists who developed and trained rudimentary computer algorithms to predict things such as numbers (SNARC), recognize images (perceptron), and play checkers (IBM).

However, theoretical advances quickly overtook the practical limitations of technology. By the late 1960s researchers had realized and acknowledged that advancement in the field of artificial intelligence was stymied by computer and electronic technology of the time. At the time, computers and machines were unable to handle the intense computations required for complex, real-world application (Minsky & Papert, 2017). Despite technical limitations, work on the theoretical underpinnings of machine learning and computer vision continued and paved the way for complex three-dimensional visual representation, segmentation, identification, and analysis of real-world phenomena to be used by machines, commonly referred to as computer vision.

We first discuss seminal work on face detection within computer vision. It is important to understand advances in face detection because many other approaches we discuss in the remainder of this chapter start with first determining whether a face is present in an image. After determining that a face is present, the face typically undergoes additional preprocessing such as extraction, rotation, and alignment so that the machine learning model can learn on a standardized image set.

Face Detection

Although there are many methods for automatic face detection (see, e.g., the next few paragraphs), the method proposed by Viola and Jones (2004) stands out as it was developed almost two decades ago and is still one of the most popular face detection paradigms. Viola &

Jones (2004) proposed a three-stage algorithm that 1) computes an "integral image," 2) uses AdaBoost to select a small number of important features from the integral image, and 3) uses a "cascade classifier" (quasi-decision tree) to determine if a face is present or not. This algorithm tends to be fast and accurate for frontal face images, uses relatively low computational resources, and computes very few false positives. These qualities have led to its continued contemporary use. Indeed, many of the more advanced face detection and classifications algorithms utilize the Viola-Jones algorithm as a first step in their pipelines. For example, in order to accurately classify age from the face, the first step must be to locate and extract the face from the image. Given the strengths outlined above, the Viola-Jones algorithm can outperform others, especially when combined with a dedicated graphics processing unit (GPU).

More recently, advances using deep neural networks have surpassed classic methods of face recognition. Models such as Facebook's DeepFace (Taigman et al., 2014), Google's facenet (Schroff et al., 2015), and OpenFace (Baltrušaitis et al., 2018) reach face detection accuracies of up to 99 percent on standardized image sets. One advantage of these models is their ability to recognize not only frontal face images, but also side and occluded faces. Another advantage of deep neural networks is that they do not require training on handpicked features. Instead, the network automatically learns important features from the face images, allowing for a completely bottom-up approach to feature extraction (Trigueros et al., 2018).

Though deep neural networks have several advantages over traditional methods (e.g., Viola & Jones, 2004), they do have disadvantages. Deep neural networks require training sets consisting of thousands of prototypes to reliably extract meaningful features to make accurate classifications. Because of this, traditional techniques can still outperform, and in some cases might be preferable to, deep neural networks if the correct features are utilized in the model. As such, traditional methods tend to be less complex and require less computational power to perform accurately. As stated earlier, this allows some approaches to be extremely quick and in real time (Viola & Jones, 2004).

Age Estimation

Age estimation is difficult because individuals age differently based on genetics, access to resources, lifestyle, and health (see also Schwartzman and Rule, Chapter 2, this volume). Therefore, two individuals of the same age may *appear* as though they are of different ages. This poses a significant and unique challenge for automatic age detection. For example, studies comparing identical twins have shown that factors such as smoking and body mass index change perceptions of aging (Guyuron et al., 2009; Okada et al., 2013). Similarly, exposure to certain healthcare treatments such as Botox can dramatically alter facial appearance and in some cases reduce the appearance of age-related features (e.g., wrinkles and saggy skin; see also, Han et al., Chapter 9, this volume) on the face resulting in lower age estimates and increased health and attractiveness estimated by observers (Fink & Prager, 2014).

To account for differences in aging, research has focused on estimating age using individual facial characteristics, such as utilizing an automatic wrinkle detection system (Cula et al., 2013), or facial texture (Iga et al., 2003). Similarly, there has been progress within computer vision age estimation to address specific factors that can disrupt accurate estimation, such as Botox (Sarhan et al., 2016). Despite these advances, even estimating facial wrinkles is a difficult task (Elbashir & Hoon Yap, 2020). Many algorithms are best at detecting horizontal (e.g., forehead) wrinkles and are prone to false positives (Elbashir & Hoon Yap, 2020).

While locating individual wrinkles on a face is difficult, estimating overall age is generally easier. Efforts that use deep neural networks have shown promise in estimating age and gender. Current top performing computer vision models reach accuracies of 94 percent for gender discrimination, and within a margin of 4.7 years for accurately predicting age, both of which are on par or exceed human estimates (Han et al., 2015).

Emotion Expression Prediction

Whereas face identification aims to locate and/or extract a face from an image, emotion recognition adds to this procedure the need to estimate the geometric properties of the face, standardize extracted face coordinates (e.g., to account for pitch, roll, and yaw), and finally to perform some type of emotion classification using the additional extracted features. Early approaches for expression recognition used similar methods as face detectors, including template matching of geometric coordinates (Shan et al., 2005), principal component analysis (Turk & Pentland, 1991), and optical flow for dynamic displays (Essa & Pentland, 1997). Contemporary approaches typically use deep neural networks, or a combination of deep neural networks with other approaches to preprocess the image (e.g., extract and crop the face prior to deep neural network analysis). Such deep neural network approaches for emotion estimation typically achieve accuracies close to or surpass human accuracies by taking features of the whole image into account rather than

https://doi.org/10.1017/9781009209656.012 Published online by Cambridge University Press

handpicked features (such as structure). For example, FaceReader – a commercial emotion recognition software – reports that their models can achieve an accuracy of up to 99 percent for prototypical displays of emotions and comparable to human estimates for spontaneous displays of emotion, though this number varies dramatically by a number of factors including the stimulus set used in benchmarking (Dupré et al., 2020; Krumhuber et al., 2021; Noldus Information Technology, 2022).

However, there has also been considerable effort within the research community to identify lower-order expressive units. By far the most studied expressive units are "Facial Action Units" (FACS), which were developed as a way of objectively describing what is occurring on the face (Ekman & Friesen, 1978). Each action unit roughly corresponds to the underlying muscles of the face, such that certain action units were considered "activated" when specific muscles, or groups of muscles, were also activated (e.g., action unit 12 "lip corner puller" corresponds to activation of the *zygomaticus major* muscle). Because FACS is a general descriptive measure of face behavior, parsing facial movement rather than expression allows for a more nuanced understanding of facial states. By combining action units, researchers can examine emotional states beyond basic, discrete emotions, such as complex expressions like boredom, mental states (Whitehill et al., 2014), and micro-expressions lasting less than 0.5s (Yan et al., 2013).

A lesser studied facet of utilizing machine learning is the *relative* contribution of different types of face metrics to facial expression recognition and classification. Some work already discussed attempts to use features within one category (e.g., individual shape features), or a combination of two metrics (e.g., color and shape), to gain a performance advantage for classification. However, little work has examined how each type of metric, uniquely and in combination, influences facial emotion recognition for both machines and humans. Further, decomposing each different facial metric into a separate classifier allows an aggregate algorithm to better approximate human vision, as human vision weights and integrates features. Another advantage to decomposing metrics and features for emotion classification is the ability to detect subtle expressive content (i.e., emotion cues on neutral faces).

Recently, we isolated the unique contribution of several channels of visual information when examining subtle emotional cues from neutral faces. Albohn and Adams (2021) extracted the color, structure, and texture from neutral faces and used each facial metric to predict emotional resemblance. Each channel offered unique information on its own, and when combined provided the best accuracy. Critically, all three channels aligned with gender-emotion stereotypes, suggesting that humans integrate these channels of information when forming judgments derived from neutral faces.

The structural portion of the aforementioned model has also been applied to White and Black neutral faces (Adams et al., 2022). However, in this work a counter-stereotypic pattern emerged: White neutral faces resembled anger expressions relatively more, whereas Black neutral faces resembled fear expressions relatively more – the opposite to what would be expected from judgments provided by humans based on race-emotion stereotypes. Taken together, these two studies underscore how using computer vision techniques can provide researchers with a better understanding of bottom-up phenotypic versus top-down stereotypic contributions to person perception. In the case of gender, emotion stereotypes and machine-derived face metrics both align with human judgments. On the other hand, machine-derived face metrics appear in some cases to be misaligned with race-emotion stereotypes, suggesting that facial appearance can actually attenuate the very stereotypic impressions that they otherwise serve to activate.

Research examining age-related emotion perceptions has also provided evidence for both top-down stereotypic and bottom-up phenotypic contributions. Indeed, a significant number of studies has shown that older adults are believed to express and feel more negative relative to positive emotions, presumably due to age-related facial appearance cues (e.g., wrinkles) mimicking emotion expressivity (Adams et al., 2016; Hess et al., 2012; Malatesta et al., 1987; North & Fiske, 2015). Despite this common stereotype, research that has used computer vision to examine the structural similarity of older adult neutral faces to emotion expressions has found a counter-stereotypic effect. Specifically, older adult neutral faces were structurally more similar to happy expressions compared to younger adult faces, an apparent contradiction of prevailing age-emotion stereotypes (Palumbo et al., 2017). However, more recent work employing more advanced technology has shown a stereotype-congruent effect (Albohn & Adams, 2020a). Using a deep neural network (i.e., FaceReader) to examine the valence of older adult neutral faces, it was found that they objectively resembled negative emotions more than positive emotions. The contradiction in computer vision results from these two studies suggests that there might be differences in facial appearance channels (i.e., structure, color, and texture) that either coincide or run counter to age-emotion stereotypes. One possibility is that structural face cues – which Palumbo and colleagues (2017) used – run counter to the negative age stereotype and can be overridden by other face metrics such as color and texture that are in line with common age-emotion stereotypes. Indeed, wrinkles and folds of the skin that mimic negative expressions are likely not to be

https://doi.org/10.1017/9781009209656.012 Published online by Cambridge University Press

captured by structural face metrics alone but are taken into account in other models such as FaceReader that employ pixel-wide analyses. Needless to say, this work underscores the complicated nature of quantifying age and age-related cues on the face. Future work should examine how to disentangle specific facial metrics, cues, and features and how they are related to emotion perception in older adult faces. Specifically, one might predict that older individuals with few wrinkles appear more emotionally positive than even younger individuals who have more wrinkles. A prediction that derives directly from considering computer vision algorithms.

Bias and Future Directions in Detection and Classification

Another important issue to consider in machine learning is bias present in classification algorithms. It has been well documented in the face detection literature that various algorithms have different performance accuracies depending on the gender and race of the face being evaluated (see, e.g., Buolamwini & Gebru, 2018). The ACLU provided an eye-opening example of this bias when they conducted a study showing that commercially available face-matching recognition software (Amazon's "Rekognition") falsely matched twenty-eight members of congress with actual criminal mugshots (Snow, 2018). Perhaps more concerning, however, is that of the twenty-eight false positives, eleven were people of color. Although the results of the ACLU's investigations have already ignited a host of policy changes for the use of such technology by law enforcement officials, the underlying issue remains: Bias can be trained into algorithms. This issue is of clear societal concern. Thus, it is important to examine age, race, and gender effects of any algorithm and compare them with known human responses, keeping in mind, as noted earlier, that it is often the bias in annotations by humans which are at the root of biased algorithms.

While simple emotion classification may have reached comparable levels to, and in some cases even surpassing, human accuracy, methods and research have yet to fully incorporate into machine learning models how static face features attenuate, enhance, or overall influence automatic emotion expression recognition. For example, older adult faces are often classified by younger adults as appearing more negative (Albohn & Adams, 2020a; Hess et al., 2012). Yet, this bias can disappear when older adults rate other older adult faces (Zebrowitz et al., 2013). Few machine learning models account for such changes in perceptual performance due to stimulus characteristics, and no models account for such changes due to observer characteristics.

Visual Manipulation and Generation of Faces

Aside from classification and prediction, advances in machine learning have allowed researchers to manipulate and even generate entirely new faces with specific and tightly controlled features and attributes. For example, face manipulation techniques have allowed researchers to take faces and apply attributes of interest into them such as manipulating the face's age, gender, race, or emotion expression. Here we discuss several of the most popular face morphing and manipulation techniques that have been used to study how individuals perceive faces that vary in these attributes.

JPsychoMorph is a Java application that has been used extensively in psychological and behavioral science to estimate facial averages, and create morphs and transformations of various facial attributes, such as race, gender, age, and health (Tiddeman et al., 2001). JPsychoMorph has the ability to separate facial structure, color, and texture allowing any combination of the three to be averaged together or "morphed into" other faces (see Figure 11.1). For example, in one study researchers derived color appearance from individuals high and low in fruit and vegetable consumption and then applied these color masks to composite images to mimic appearance changes in faces related to fruit and vegetable consumption (Whitehead et al., 2013). Other studies have used similar morphing or transformation approaches to examine how appearance-based cues influence perception of sleep deprivation (Holding et al., 2019), emotion-resemblance (Adams et al., 2012), and kinship (DeBruine et al., 2008) derived from the face.

Although, JPsychoMorph – and morphing in general – has been a powerful tool to manipulate faces in order to better understand perception of faces across a diverse range of attributes, it has a number of

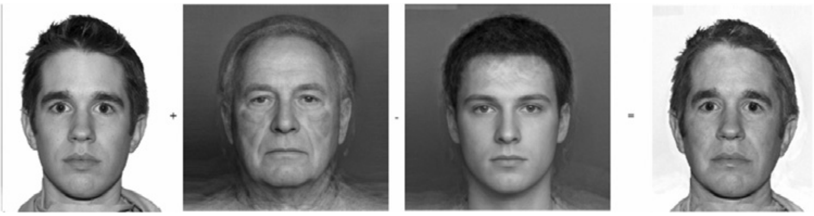

Figure 11.1. Example of transforming a young male neutral face ("A") into an older adult male neutral image ("D") using JPsychoMorph. JPsychoMorph takes the structure, color, and texture that is "uniquely older adult male" and applies it to a novel stimulus by subtracting the average young male face ("C") from the average older adult male face ("B"). Male neutral face taken with permission from the NimStim Set of Facial Expressions (Tottenham et al., 2009).

https://doi.org/10.1017/9781009209656.012 Published online by Cambridge University Press

limitations. Manipulation requires two averaged base faces to be used. If, for example, a researcher wanted to transform a face from young to old, it would require both an older adult average face prototype as well as a younger adult average prototype. This is necessary for the algorithm to know what is specific to the target transformation (in this case age) in order to only transform those attributes. Because the averaging process itself tends to leave artifacts in the images, these can sometimes be transferred to the manipulated transformation image resulting in undesirable noise. Hair and skin "smoothing" are particularly prone to such effects when averaging (see Figure 11.1).

Aside from morphing and transformation tools, there has been a concerted effort over the last two decades to create accurate and realistic synthetic avatar faces. This process allows a researcher to spontaneously generate or manipulate faces without any need for a real face image. One popular tool has been the commercial product FaceGen (Blanz & Vetter, 1999). FaceGen can manipulate avatar faces along a number of different parameters, including emotion expression, face phenotype (lip size, cheek prominence), race, gender, and age through a process much like creating a character in a video game. The strength of FaceGen is that it allows for tight experimental control when generating and manipulating faces. For example, the exact same emotion intensity, age parameters, and eye size can be applied to a whole data set of synthetic avatar faces without having to control for such variables *post hoc* (see also Metahumans, www.unrealengine.com/en-US/metahuman).

Data-driven methods in face perception have demonstrated important underlying commonalities to impression judgments. Oosterhoff and Todorov (Oosterhof & Todorov, 2008) utilized FaceGen to derive a two-dimensional social face space that consisted of orthogonal dimensions of dominance and trustworthiness. To do so, they had participants rate randomly generated FaceGen avatars on a number of social dimensions, which were then reduced to two principal components of dominance and trustworthiness that accounted for > 80 percent of the variance. Because the synthetic avatar faces are represented by a vector of numbers, a statistical model can be applied to the vector and transformed along each dimension of interest, essentially transforming any given face within the two-dimensional face space.

There have been other data-driven approaches used to reveal how individuals perceive attributes on the face (see also Sutherland and Young, Chapter 7, this volume). For example, Jack and colleagues (2016) examined facial temporal dynamics of emotional expressions using a similar approach to that of Oosterhoff and Todorov (2008) with a three-dimensional generative face grammar and combined with reverse correlation. Similarly, a multitude of studies have used psychophysical reverse correlation to examine how individuals visualize certain facial attributes in their "mind's eye" from ethnic groups to social attributes such as dominance and trustworthiness (Dotsch et al., 2008; Dotsch & Todorov, 2012), see also Gossetti and Jack, Chapter 12, this volume. In our own work, we have used reverse correlation to reveal negative aging stereotypes (Albohn & Adams, 2020a). Participants consistently created mental representations of older adults that appeared more negative and less positive than their younger adult representations, reflecting both implicit and explicit stereotypes that older adults are often perceived as more negative than younger adults (see Figure 11.2).

Although JPsychoMorph can create photorealistic manipulations, and avatar methods like FaceGen can transform any given feature in a data-driven manner, there are still some limitations that can be attenuated by recent advances in computer vision. As alluded to earlier, to get the best results from JPsychoMoprh average faces should be used during the transformation procedure, which can cause artifacts in the final transformed image (e.g., the appearance that the skin is "smoothed" as well as "frizzled" hair). Further, to achieve the best transformation and morphing results individual images used in the pipeline need to be manually processed by individually placing facial landmarks on each image. To do

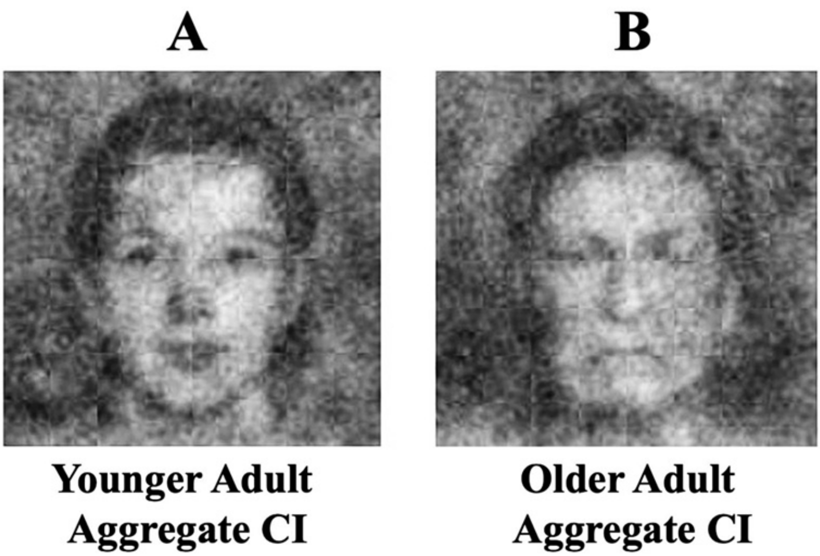

Figure 11.2. Participant mental representations for how they view younger adults ("A") and older adults ("B") in their mind's eye. Each image represents aggregated results across all participants in the study (N = 27). Figure adapted from Albohn and Adams (2020a).

https://doi.org/10.1017/9781009209656.012 Published online by Cambridge University Press

this for thousands of images is not only time-consuming, but prone to human error. Similarly, avatar methods are useful for manipulating individual facial features, but the resultant images appear unrealistic (i.e., clearly computer generated and often presented without hair). One approach that addresses these issues is StyleGAN, a generative adversarial network (GAN) model that was trained on 70,000 real face images and is capable of producing novel, high-quality, photorealistic face images (Karras et al., 2018, 2020).

Since StyleGAN's inception, a multitude of complementary models have been developed to either supplement, improve, or extend it. One model that combines several advanced features is StyleGAN-NADA (Gal et al., 2021). Under this framework, the user inputs a text description of the transformation that images should undergo. For example, the input text might be "face image" while the output text might be "Picasso painting" (see Gal et al., 2021, manuscript appendix for many more examples). The model learns "on the fly" what the text description means using a corpus of many millions image-text pairs (Radford et al., 2021).

The utility of StyleGAN-NADA is that once it learns the relationship between the input text/image and the desired outcome, it can transfer the "style" of that transformation to new images, essentially generating new images on the fly. For example, a user can ask the model to learn the transformation from "young woman" to "old woman" and then apply that learned style to a novel image to create additional stimuli (see example in Figure 11.3).

However, in order to manipulate real faces using any StyleGAN flavor of model, the real face must be "imported" into the model through a process called *inversion*, which essentially results in a reproduction of

Figure 11.3. Example of transforming an input image ("A") from "young woman" to "old woman" ("B") using StyleGAN-NADA (Gal et al., 2021) with 100 training iterations.

Figure 11.4. Example of an average neutral face (composite of four neutral faces) inverted by StyleGAN3-PTI ("A"; Roich et al., 2021) and manipulated to appear older using InterFaceGAN ("B"; Shen et al., 2020) to communicate an emotion.

the real face as if the model generated it. Inverting actual images into GAN latent space that are simultaneously realistic but still able to be manipulated is a rapidly burgeoning subfield within deep learning. One method of inversion called "Pivotal Tuning Inversion" (PTI) uses the input image to fine-tune the StyleGAN model so that a more realistic output image is obtained (Roich et al., 2021). The benefit of such a process is that while identity shifts to a more realistic-appearing generated image, the latent space that the image is projected into is still able to be manipulated (e.g., to age the face). Figure 11.4 displays the output of using PTI to create an average composite image of four neutral faces and then manipulate it to appear "older" by multiplying the averaged inverse image by a matrix that represents a specific direction in latent space that corresponds to "age." These latent directions are computed through training a separate model whereby generated images from the model (such as StyleGAN2) are categorized or rated on a specific feature such as age (Shen et al., 2020). Recent work has taken this a step further and has identified a large number of social attributes with the latent space such as "smart," "memorable," and "skinny" (Peterson et al., 2022). This approach allows for an incredible amount of control over how individual face images appear and nearly an infinite number of dimensions on which they can be manipulated.

Bias and Future Directions in Manipulation and Generation

Generative models come with advantages and disadvantages. One important advantage is that these generative models allow for large amounts of experimental control when generating faces that, in turn, allows for greater internal validity. Much like FaceGen, methods are available for applying the same exact parameters (e.g., 25 percent more trustworthy and +10 age) to any number of faces. However, the most important advantage is that these GAN images appear photorealistic and nearly indistinguishable from real face images. This is an exciting advancement that provides a high-fidelity approach with large amounts of face validity for examining questions related to face perception.

On the other hand, these models may also contain bias in the images they generate and manipulate. For example, in Figure 11.4 when the younger adult is manipulated to appear "older" the older adult image appears to have spontaneously acquired a suit rather than maintaining the plain T-shirt apparent in the younger adult image. Although this is likely an artifact of the data used to train the model, it is nonetheless a bias. In essence, the model has learned that being "older" includes wearing a suit in addition to features such as gray hair, wrinkles, and so on. Although this is a fairly innocuous form of bias, it is not hard to imagine how other biases can creep in and be extremely harmful or unwanted – particularly within a highly controlled research context.

Additionally, with the rise of hyperrealistic face images comes concerns about ethics and privacy. With such methods rapidly advancing it will not be long before fully dynamic synthetic models can be deployed easily within a research context. Such approaches coupled with machine learning that can replicate voice pitch, tone, and patterns will likely yield stimuli that are nearly indistinguishable from real video recordings. While this opens a number of exciting opportunities for research, it simultaneously can be used for nefarious purposes (e.g., Deepfakes). It is paramount that researchers utilize these emerging technologies with care, taking into consideration the ethics involved with manipulating the digital identities of real individuals. Further, researchers utilizing these types of technologies should also consider the types of biases that may be present in their models and make a concerted effort to be transparent about the limitations and boundaries that such models can achieve.

Closing Considerations

Although computer vision and machine learning has advanced considerably in recent decades, what is perhaps clear from this chapter is that relatively little work has examined age perception at the intersection of other, related facial cues including emotion, sex/gender, and race (see Hedgecoth et al., Chapter 5, this volume). Part of the reason for this dearth of research has likely been due to the limited technology available to properly conduct it. However, technology has now evolved to the point where it is both possible and accessible to researchers to start

Table 11.1 *Open source and freely available face detection, categorization, manipulation, and generation software. These resources include both highly validated and relatively new face-related technologies*

Resource	URL	Programming Language	Description
"faceplyr"	https://github.com/d-bohn/faceplyr	R and Python	Extract structure, color, and texture metrics of a face to examine resemblance to emotion expressions.
Deepface	https://github.com/serengil/deepface	Python	Facial age, gender, race, and emotion recognition.
Deepfacelab	https://github.com/iperov/DeepFaceLab	Python	Static and dynamic face swapping utility.
JPsychoMorph	https://users.aber.ac.uk/bpt/jpsychomorph/	Java with GUI	Face averaging and transformation tool.
MetaHuman	www.unrealengine.com/en-US/metahuman	C++ with Real Engine web GUI	Create and manipulate static and dynamic faces varying in appearance and expression on realistic, high-fidelity digital avatars.
OpenFace	https://github.com/TadasBaltrusaitis/OpenFace	Python and C++ with command line interface	Open face toolkit that can detect, predict emotion expression, and quantify action units on face images and videos.
Pivotal Tuning Inversion	https://github.com/danielroich/PTI	Python	Method for reproducing high quality real faces within the StyleGAN latent space.

Table 11.1 *(cont.)*

Resource	URL	Programming Language	Description
Real-time Voice Cloning	https://github.com/CorentinJ/Real-Time-Voice-Cloning	Python	Method to train and mimic voices and produce arbitrary speech.
Reverse Correlation	https://github.com/rdotsch/rcicr	R	Mental representation creation tool.
StyleGAN-nada	https://github.com/rinongal/StyleGAN-nada	Python	Style transfer GAN based on text prompt.
StyleGAN2	https://github.com/NVlabs/stylegan2-ada-pytorch	Python	GAN able to create high quality synthetic faces.
Webmorph	https://webmorph.org/	Java with web GUI	Web version of JPsychoMorph with an updated suite of tools.

https://doi.org/10.1017/9781009209656.012 Published online by Cambridge University Press

looking back and answering previously unanswerable questions (see Table 11.1 for list of free resources reviewed herein).

Arguably the two most critical advances in technology for researchers are data-driven methods to meticulously and accurately determine specific visual features involved in specific judgments along with the ability to manipulate and generate high-fidelity stimuli. Recent advances have been able to combine these two achievements to provide researchers with an incredible tool. For example, it is now relatively easy to manipulate and even generate photorealistic face images that vary both in terms of age and emotion. These techniques, coupled with emerging methods in real-time voice cloning, provide an opportunity for researchers to examine interesting questions through real-time digital interactions. It will not be long before researchers can create a synthetic version of a participant's face that is indistinguishable from a real photo and then alter his or her age, race, sex/gender, and expression in real time all while they interact with another live participant whose face has undergone similar transformations. In doing so, all parameters of a person's nonverbal behavior can be held constant, while examining how changing their identity influences impressions. In short, the addition of these computer vision techniques allows for more nuanced and difficult questions to be asked and answered at the intersection of age, emotion, and just about any other face attribute one is interested in examining.

To summarize, we began this chapter by discussing the advances that machine learning and computer vision have made in recent decades. Simultaneously, we also stressed that there is a growing concern related to the rising use of artificial intelligence throughout many aspects of research and society. While it is important to acknowledge machine learning's limitations, our goal here was to showcase the opportunities that prior and contemporary work within this burgeoning field can provide to domains such as person perception, affective science, and in the case of this book, specifically to an understanding of aging in the face and its influence on emotion and person perception. Given the incredible amount of data at our fingertips, it is clear that machine learning, computer vision, and artificial intelligence can be a useful tool to aid researchers in their attempt to understand human behavior.

References

Adams, R. B., Jr., Albohn, D. N., Hedgecoth, N., Garrido, C. O., & Adams, K. D. (2022). Angry White faces: A contradiction of racial stereotypes and emotion-resembling appearance. *Affective Science*. https://doi.org/10.1007/s42761-021-00091-5

Adams, R. B., Jr., Garrido, C. O., Albohn, D. N., Hess, U., & Kleck, R. E. (2016). What facial appearance reveals over time: When perceived expressions in neutral faces reveal stable emotion dispositions. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00986

Adams, R. B., Jr., Nelson, A. J., Soto, J. A., Hess, U., & Kleck, R. E. (2012). Emotion in the neutral face: A mechanism for impression formation? *Cognition & Emotion*, *26*(3), 431–441. https://doi.org/10.1080/02699931.2012.666502

Albohn, D. N., & Adams, R. B., Jr. (2020a). Everyday beliefs about emotion perceptually derived from neutral facial appearance. *Frontiers in Psychology*, *11*, 264. https://doi.org/10.3389/fpsyg.2020.00264

Albohn, D. N., & Adams, R. B., Jr. (2020b). Emotion residue in neutral faces: Implications for impression formation. *Social Psychological and Personality Science*, *12*(4), 479–486. https://doi.org/10.1177/1948550620923229

Albohn, D. N., & Adams, R. B., Jr., (2021). The expressive triad: Structure, color, and texture similarity of emotion expressions predict impressions of neutral faces. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.612923

Baltrušaitis, T., Zadeh, A., Lim, Y. C., & Morency, L.-P. (2018). OpenFace 2.0: Facial Behavior Analysis Toolkit. *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*, 59–66. https://doi.org/10.1109/FG.2018.00019

Blanz, V., & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques – SIGGRAPH *'99*, 187–194. https://doi.org/10.1145/311535.311556

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, *81*, 1–15.

Coeckelbergh, M. (2020). *AI ethics*. Cambridge, MA: MIT Press.

Cula, G. O., Bargo, P. R., Nkengne, A., & Kollias, N. (2013). Assessing facial wrinkles: Automatic detection and quantification. *Skin Research and Technology*, *19*(1), e243–e251. https://doi.org/10.1111/j.1600-0846.2012.00635.x

DeBruine, L. M., Jones, B. C., Little, A. C., & Perrett, D. I. (2008). Social perception of facial resemblance in humans. *Archives of Sexual Behavior*, *37*(1), 64–77. https://doi.org/10.1007/s10508-007-9266-0

Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science*, *3*(5), 562–571. https://doi.org/10.1177/1948550611430272

Dotsch, R., Wigboldus, D. H. J., Langner, O., & van Knippenberg, A. (2008). Ethnic out-group faces are biased in the prejudiced mind. *Psychological Science*, *19*(10), 978–980. https://doi.org/10.1111/j.1467-9280.2008.02186.x

Dupré, D., Krumhuber, E. G., Küster, D., & McKeown, G. J. (2020). A performance comparison of eight commercially available automatic classifiers for facial affect recognition. *PLoS ONE*, *15*(4), e0231968. https://doi.org/10.1371/journal.pone.0231968

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.

Elbashir, R. M., & Hoon Yap, M. (2020). Evaluation of automatic facial wrinkle detection algorithms. *Journal of Imaging*, *6*(4), 17. https://doi.org/10.3390/jimaging6040017

Engelmann, S., Ullstein, C., Papakyriakopoulos, O., & Grossklags, J. (2022). What People Think AI Should Infer From Faces. *2022 ACM Conference on Fairness, Accountability, and Transparency*, 128–141. https://doi.org/10.1145/3531146.3533080

Essa, I. A., & Pentland, A. P. (1997). Coding, analysis, interpretation, and recognition of facial expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*(7), 757–763. https://doi.org/10.1109/34.598232

Fink, B., & Prager, M. (2014). The effect of incobotulinumtoxin a and dermal filler treatment on perception of age, health, and attractiveness of female faces. *The Journal of Clinical and Aesthetic Dermatology*, *7*(1), 36–40.

Gal, R., Patashnik, O., Maron, H., Chechik, G., & Cohen-Or, D. (2021). StyleGAN-NADA: CLIP-guided domain adaptation of image generators. *ArXiv:2108.00946 [Cs]*. http://arxiv.org/abs/2108.00946

Guyuron, B., Rowe, D. J., Weinfeld, A. B., et al. (2009). Factors contributing to the facial aging of identical twins. *Plastic and Reconstructive Surgery*, *123* (4),1321–1331. https://doi.org/10.1097/PRS.0b013e31819c4d42

Han, H., Otto, C., Liu, X., & Jain, A. K. (2015). Demographic estimation from face images: Human vs. machine performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *37*(6), 1148–1161. https://doi.org/10.1109/TPAMI.2014.2362759

Hess, U., Adams, R. B., Jr., Simard, A., Stevenson, M. T., & Kleck, R. E. (2012). Smiling and sad wrinkles: Age-related changes in the face and the perception of emotions and intentions. *Journal of Experimental Social Psychology*, *48*(6), 1377–1380. https://doi.org/10.1016/j.jesp.2012.05.018

Holding, B. C., Sundelin, T., Cairns, P., Perrett, D. I., & Axelsson, J. (2019). The effect of sleep deprivation on objective and subjective measures of facial appearance. *Journal of Sleep Research*, *28*(6), e12860. https://doi.org/10.1111/jsr.12860

Iga, R., Izumi, K., Hayashi, H., Fukano, G., & Ohtani, T. (2003). A gender and age estimation system from face images. *SICE 2003 Annual Conference (IEEE Cat. No. 03TH8734)*, *1*, 756–761.

Jack, R. E., Sun, W., Delis, I., Garrod, O. G. B., & Schyns, P. G. (2016). Four not six: Revealing culturally common facial expressions of emotion. *Journal of Experimental Psychology: General*, *145*(6), 708–730. https://doi.org/10.1037/xge0000162

Karras, T., Laine, S., & Aila, T. (2018). A style-based generator architecture for generative adversarial networks. *ArXiv:1812.04948 [Cs, Stat]*. http://arxiv.org/abs/1812.04948

Karras, T., Laine, S., Aittala, M., et al. (2020). Analyzing and improving the image quality of StyleGAN. *ArXiv:1912.04958 [Cs, Eess, Stat]*. http://arxiv.org/abs/1912.04958

Krumhuber, E. G., Küster, D., Namba, S., Shah, D., & Calvo, M. G. (2021). Emotion recognition from posed and spontaneous dynamic expressions:

Human observers versus machine analysis. *Emotion*, *21*(2), 447–451. https://doi.org/10.1037/emo0000712

Malatesta, C. Z., Fiore, M. J., & Messina, J. J. (1987). Affect, personality, and facial expressive characteristics of older people. *Psychology and Aging*, *2*(1), 6. https://doi.org/DOI:10.1037/0882-7974.2.1.64

Minsky, M., & Papert, S. A. (2017). *Perceptrons: An introduction to computational geometry*. Cambridge, MA: MIT Press.

Noldus Information Technology. (2022). *FaceReader 9*. www.noldus.com/facereader?gclid=Cj0KCQjwj7v0BRDOARIsAGh37ipxkZpfWeDL9l_k4AEbMAWyYC596JVHVQ9KE91fhfgwbV-WAAlYGIAaAkNvEALw_wcB

North, M. S., & Fiske, S. T. (2015). Modern attitudes toward older adults in the aging world: A cross-cultural meta-analysis. *Psychological Bulletin*, *141*(5), 993–1021. https://doi.org/10.1037/a0039469

Okada, H. C., Alleyne, B., Varghai, K., Kinder, K., & Guyuron, B. (2013). Facial changes caused by smoking: A comparison between smoking and nonsmoking identical twins. *Plastic and Reconstructive Surgery*, *132*(5), 1085–1092. https://doi.org/10.1097/PRS.0b013e3182a4c20a

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Palumbo, R., Adams, R. B., Jr., Hess, U., Kleck, R. E., & Zebrowitz, L. (2017). Age and Gender Differences in Facial Attractiveness, but Not Emotion Resemblance, Contribute to Age and Gender Stereotypes. *Frontiers in Psychology*, 8. https://doi.org/10.3389/fpsyg.2017.01704

Peterson, J. C., Uddenberg, S., Griffiths, T. L., Todorov, A., & Suchow, J. W. (2022). Deep models of superficial face judgments. *Proceedings of the National Academy of Sciences*, *119*(17), e2115228119. https://doi.org/10.1073/pnas.2115228119

Radford, A., Kim, J. W., Hallacy, C., et al. (2021). Learning transferable visual models from natural language supervision. *ArXiv:2103.00020*. http://arxiv.org/abs/2103.00020

Roich, D., Mokady, R., Bermano, A. H., & Cohen-Or, D. (2021). Pivotal tuning for latent-based editing of real images. *ArXiv:2106.05744 [Cs]*. http://arxiv.org/abs/2106.05744

Sarhan, S., Hamad, S., & Elmougy, S. (2016). Human injected by Botox age estimation based on active shape models, speed up robust features, and support vector machine. *Pattern Recognition and Image Analysis*, *26*(3), 617–629. https://doi.org/10.1134/S1054661816030184

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A Unified Embedding for Face Recognition and Clustering. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 815–823. https://doi.org/10.1109/CVPR.2015.7298682

Shan, C., Gong, S., & McOwan, P. W. (2005). Robust facial expression recognition using local binary patterns. *IEEE International Conference on Image Processing 2005*, *2*, II–370.

Shen, Y., Yang, C., Tang, X., & Zhou, B. (2020). InterFaceGAN: Interpreting the disentangled face representation learned by GANs. *ArXiv:2005.09635*. http://arxiv.org/abs/2005.09635

Simonite, T. (2018). When it comes to gorillas, google photos remains blind. *Wired*. www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/

Snow, J. (2018). www.aclunc.org/blog/. ACLU. www.aclunc.org/blog/amazon-s-face-recognition-falsely-matched-28-membe

Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). DeepFace: Closing the Gap to Human-Level Performance in Face Verification. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 1701–1708. https://doi.org/10.1109/CVPR.2014.220

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, *21*(4),42–50. https://doi.org/10.1109/38.946630

Trigueros, D. S., Meng, L., & Hartnett, M. (2018). Face recognition: From traditional to deep learning methods. *ArXiv Preprint ArXiv:1811.00116*.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*(1), 71–86.

United Nations. (2020). *World population ageing, 2019 highlights*.

Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision*, *57*(2), 137–154.

Wehrli, S., Hertweck, C., Amirian, M., Glüge, S., & Stadelmann, T. (2021). Bias, awareness, and ignorance in deep-learning-based face recognition. *AI and Ethics*, *2*, 509–522. https://doi.org/10.1007/s43681-021-00108-6

Whitehead, R. D., Ozakinci, G., & Perrett, D. I. (2013). A randomized controlled trial of an appearance-based dietary intervention. *Health Psychology*, *33*(1), 99. https://doi.org/10.1037/a0032322

Whitehill, J., Serpell, Z., Lin, Y.-C., Foster, A., & Movellan, J. R. (2014). The faces of engagement: Automatic recognition of student engagementfrom facial expressions. *IEEE Transactions on Affective Computing*, *5*(1), 86–98. https://doi.org/10.1109/TAFFC.2014.2316163

Yan, W.-J., Wu, Q., Liang, J., Chen, Y.-H., & Fu, X. (2013). How fast are the leaked facial expressions: The duration of micro-expressions. *Journal of Nonverbal Behavior*, *37*(4), 217–230. https://doi.org/10.1007/s10919-013-0159-8

Zebrowitz, L. A., Franklin, R. G., Hillman, S., & Boc, H. (2013). Older and younger adults' first impressions from faces: Similar in agreement but different in positivity. *Psychology and Aging*, *28*(1), 202–212. https://doi.org/10.1037/a0030927

Zebrowitz, L. A., Kikuchi, M., & Fellous, J.-M. (2010). Facial resemblance to emotions: Group differences, impression effects, and race stereotypes. *Journal of Personality and Social Psychology*, *98*(2), 175–189. https://doi.org/10.1037/a0017990