AI Aesthetics

Lev Manovich

Abstract:

AI plays a crucial role in the global cultural ecosystem. It recommends what we should see, listen to, read, and buy. It determines how many people will see our shared content. It helps us make aesthetic decisions when we create media. In professional cultural production, AI has already been adapted to produce movie trailers, music albums, fashion items, product and web designs, architecture, etc. In this short book, Lev Manovich offers a systematic framework to help us think about cultural uses of AI today and in the future. He challenges existing ideas and gives us new concepts for understanding media, design, and aesthetics in the AI era.

> "[The Analytical Engine] might act upon other things besides number...supposing, for instance, that the fundamental relations of pitched sounds in the science of harmony and of musical composition were susceptible of such expression and adaptations, the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent."
> *- Ada Lovelace, 1842*

> Итак, кто же я такой? С известными оговорками, я и есть то, что люди прошлого называли «искусственным интеллектом». (Виктор Пелевин, *iPhuck 10*, 2017.)
> "So who am I exactly? With known caveats, I am what people of the past have called "artificial intelligence."" (Victor Pelevin, *iPhuck 10*, 2017.)

Writing in 1842, Ada Lovelace imagines that in the future, Babbage's Analytical Engine (general purpose programmable computer) will be able to create complex music. In the 2017 novel by famous Russian writer Victor Pelevin, set in the late 21th century, the narrator is an algorithm solving crimes and writing novels about them. Today we exist somewhere between

these two visions of *cultural AI* (Artificial Intelligence). Algorithms are frequently used to write music, but they don't really "understand" the human world and human meanings. Whether the latter will ever happen is unclear.

The original vision of AI was about automation of cognition. Today, AI also plays a crucial role in culture, increasingly influencing our choices, behaviours, and imaginations. For example, it is used to recommend photos, videos, music, and other media. AI is also used to suggest people we should follow on social networks, to automatically beautify selfies and edit user photos to fit the norms of "good" photography, and to generate and control characters in computer games.

While algorithms have been employed in artistic creation by artists since the 1960s, today industrial scale "cultural AI" is built into devices and services used by billions of people. Instead of being an instrument of a single artistic imagination, AI has become a mechanism for influencing the imaginations of billions. Gathered and aggregated data about the cultural behaviours of multitudes is used to model our "aesthetic self," predicting our future aesthetic decisions and likes – and potentially guiding us towards choices preferred by the majority.

The integration of AI into the everyday cultural lives of billions of people raises important questions about the future of culture, aesthetics, and taste. In this short book, I discuss some of these questions.

I. AI and Production of Culture

The Meaning of "Artificial Intelligence"

In the original vision of AI in the 1950s-60s, the goal was to teach a computer to perform a range of cognitive tasks. In this vision, a computer would simulate many operations of a single human mind. They included playing chess, solving mathematical problems, understanding written and spoken language, and recognizing the content of images. Sixty years later, AI became a key instrument of modern economies, deployed to make them more efficient, secure, and predictable by automatically analyzing medical images, making decisions on consumer loans, filtering job applications, detecting fraud, and so on. AI is also seen as an enhancer of our everyday lives, saving us time and effort. A good example of this is the use of voice interface instead of typing.

But what exactly is "Artificial Intelligence" today? Besides original tasks that defined AI such as playing chess, recognizing objects in a photo, or translating between languages, computers today perform endless "intelligent" operations. For example, your phone keyboard gradually adapts to your typing style. Your phone may also monitor your usage of apps and adjust their work in the background to save battery. Your map app automatically calculates

the fastest route, taking into account traffic conditions. There are thousands of intelligent, but not very glamorous, operations at work in phones, computers, web servers, and other parts of the IT universe.

Therefore, in one sense, AI is now everywhere. While some AI roles attract our attention – such as Google's Smart Reply function that suggests automated email replies (used for 10% of all answers in Google's Inbox app in 2017) – many others operate in the gray everyday of the digital society.

Why are some intelligent tasks that computers can accomplish seen as "real" AI, and others are not? Observers and historians of the AI field talk about "AI effect." It means that "when we know how a machine does something 'intelligent,' it ceases to be regarded as intelligent" (*Promise of AI Not So Bright*, 2006).

That is, after the AI field solves a problem and the solution is implemented in the industry, it is no longer seen as part of the field. Paradoxically, we tend to only see the challenging and not-yet-solved problems as belonging to Artificial Intelligence – and this creates the impression that AI research has not been successful throughout its long history.

The dramatic increase in computer capacities, the ubiquity of digital devices and networks, and the challenges and opportunities brought by the "big data" trend of the 2000s have also affected AI. We moved from automation of a single mind to a kind of "super-cognition." Think, for example, of search engines such as Baidu, Yandex, Bing, and Google that continuously scan the web and index billions of websites and blogs. When you enter a search query, a search engine instantly returns relevant results drawn from such an index. No single human could ever perform such a feat. The scale of digital culture demands intelligence that is similar to a human qualitatively but operates on a quantitatively different scale.

Aesthetic AI

As I already pointed out, the original vision of AI was about automation of cognition. Despite the difference in scale, super-cognition still follows this paradigm. So, when people talk about the great successes of AI in recent years, the examples used are the same tasks defined at the field's start many decades earlier: natural speech understanding, automated translation, and recognition of objects in photos. But what is perhaps less obvious is that AI now plays an equally important role in our cultural lives and behaviours, increasingly automating the processes of aesthetic creation and aesthetic choices.

Consider, for example, these selected examples of AI adoption in just one single cultural field – digital photography. I divided them into two categories: assistance in selecting appropriate images from large (often massive) collections, and assistance in the creation/editing of new content. (Note that AI can assist the human selection process or be employed for completely automatic selection.)

Selecting from existing content:

1. Image sharing services and marketplaces use AI to predict the content of images and assign keywords to them (Grigonis, 2016).
2. Instagram's Explore screen recommends images and videos to each user based on a combination of many factors (not only based on what the user liked in the past).
3. Yelp automatically selects the best photos to illustrate the numerous businesses listed.
4. The Roll app from EyeEm automatically rates the aesthetic quality of user photos, while EyeEm image marketplace assigns such scores to submitted photos.
5. Huawei ran a photo contest where submitted photos were judged by AI: "Trained using 4,000,000 images taken by professional photographers and picture editors, the AI will then give each photo a personalized AI score based on parameters such as focus, jitter, deflection, color, and composition" (Hillen, 2018).

Creation / Editing of new content

1. Photo apps can automatically modify captured photos according to the norms of "good photography."
2. Other apps "beautify" selfies and portraits. Tencent's (leading Chinese IT company) long list of capabilities in this area demonstrates the range of automated adjustments possible: "dermabrasion, skin whitening, eye enlargement, face thinning, removing acne, adding eyelids, changing skin color, recognizing face color and applying foundation, applying lip gloss, eyebrow shaping, and applying other makeup" (Tencent YoTu Lab, 2018).
3. Photoshop 9.1 uses AI in its "subject select" function to automatically select objects from a background.
4. Phone cameras analyze the 3D layout of a scene and blur backgrounds in portraits and selfies (Associated Press, 2016).
5. The Huawei Mate 10 phone camera (released 10/2017) uses AI to analyze what it sees. It then classifies it into one of several scene types and selects the appropriate parameters for capturing a given scene – even before you decide to take a photo.

Other applications of AI in photography are experimental at the time of writing this book or are only emerging and have not yet been implemented in products. For example, EyeEm engineers described an experiment where their system learned the styles of different photo curators from only 20 sample photos from each, and then selected similar images from EyeEm's large collection, so the curators could make further selections (Shaji & Yildirim, 2017). Google designed a system that mimics the skills of a professional photographer, such as selecting photos suitable for editing, cropping, and applying filters. Progress in commercial implementation of new ways that AI can understand aspects of photographs can be tracked by periodically visiting the website of Clarifai company, one of the leaders in this field (Clarifai, 2018).

Outside of photography, cultural uses of AI include music recommendations in Spotify, iTunes, and other music services, apps that automatically edit a user's raw video to create

short films in a range of styles, and the creation of new fashion items and styles (Shah, 2017). As the adoption of AI in culture continues to grow, the concept of AI is also changing, and it is challenging to come up with a definite taxonomy to describe it. Building on the two categories I used to organize examples of applications of AI in photography, here is one possible taxonomy of types of cultural AI I see today:

1) *Selecting* content from larger collections: search, discovery, curation, recommendations and filtering. (Asrar, 2016).

2) *Targeting* content (e.g., one-to-one marketing, behavioral targeting, and market segmentation).

3) *Assistance* in creation/editing of new content. (If we are to think of AI as intelligent in the biological sense, we can call this "participation" in content creation.)

4) *Fully autonomous* creation (e.g., AI composing music tracks in a particular style, writing business and sports news articles (Kafka, 2016), creating visualizations from given datasets, designing websites, generating email responses, etc.)

AI and Aesthetic Diversity

One important trend we can see in the examples above is a movement towards gradual automation (semi or full) of aesthetic decisions – recommendation engines suggesting what we should watch, listen to, read, write, or wear; devices and services that automatically adjust the aesthetics of captured media to fit certain criteria; software that rates the aesthetic quality of our photos, etc. This development raises big questions about the future of culture. Does such automation lead to a *decrease in aesthetic diversity over time*? Is this inevitable, or are there other forces that may counteract this, increasing diversity?

To illustrate what this means for image culture, the question may be rephrased, for example, as following: Do automated enhancements and edits that mobile cameras and photo sharing services apply to user photos make them less diverse aesthetically? Will further AI integration in user photography devices and image sharing sites lead to standardization of "photo imagination?" Do search and recommendation engines or functions such as Instagram's Explore tend to show the same images (or many variations of images with certain content, or perhaps only images with a certain "professional" aesthetic) to lots of people, thus diminishing the diversity of what we see?

But AI, algorithms, and user interfaces of digital services, apps, and products may also be increasing aesthetic diversity. For example, digital cameras and photo apps have many functions for customization. On my camera (Fuji E-3), I can choose the shutter speed, aperture, ISO, desired levels of highlight and shadow tone, color density, sharpening, grain, dynamic range, noise reduction, and film simulation filters (Fujifilm Corporation, 2017). Free photo editing apps such as Snapseed, which are used by many people to prepare their photos for Instagram, also offer a large number of editing tools comparable to professional

desktop software such as Photoshop and Lightroom. Over time, phone cameras and photo editing apps have been adding more and more controls, and a lot of them are now free. Therefore, while the gradual AI integration into phone cameras and sharing sites may contribute to a decrease in aesthetic diversity, the simultaneous addition of more and more controls to cameras and photo apps may have the opposite effect.

Or consider recommendation engines. They can be programmed to recommend items that are already the most popular among other users, thus decreasing your chances of seeing more varied items. Alternatively, they can be programmed to expose users to more diverse items including ones that they most likely would not find on their own. A person may use many engines for different media daily, and they may all be programmed differently, so it would be unwise to assume that all engines together push users in any one particular direction. And since any industry engine uses many different inputs to come up with its recommendations, the items it recommends to a user may both expose her to what is already popular, and to what she would not find on her own. In one 2010 quantitative study, researchers set out to "investigate the impact of YouTube's recommendation system on view diversity to understand whether the recommendation system helps users to discover videos of interest, but not necessarily popular [ones], or is more likely to recommend popular videos only." Their conclusion was "that the current recommendation system helps to increase the diversity of video views in aggregation" (Zhou, Gao, & Khemmarat, 2010) – but this can be different in different times, as YouTube changes its algorithms.

Of course, a number of other trends also influence aesthetic diversity in contemporary culture besides computational technologies. The rise of the world wide web and social networks, growth of international travel, globalization of consumer economies and advertising, zero cost telecommunication, growth of foreign student enrollment, growth of remote work, and the rise of Japan, followed by Korea and then China, as exporters of cultural products and images are just some examples among many other developments all playing a role. On the one hand, they are making the world into a single global village – or if you like, a single cultural marketplace, where certain images, ideas, values, narratives, products, and styles are marketed to everybody and available everywhere, and this may decrease diversity. On the other hand, the same trends may also be increasing diversity because local cultural DNAs become available globally.

Given that many developments are influencing global aesthetic diversity, the role played by cultural AI is probably not the most significant yet – but it is likely to grow in the future for at least two reasons. First, billions of people who still don't have access to internet and smartphones will get this access and start using the same AI-driven recommendation engines, automated aesthetic editing of captured media, selfie-beautifying apps, and so on. Second, the automation of aesthetic decisions we have seen so far is still at an early stage, with many more things to come. For example, right now people take photos themselves, cameras apply some aesthetic adjustments at the time of capture, and then a person may use editing software to do further adjustments. But it is easy to imagine a future scenario where cameras will themselves choose what and when to capture to give us the most satisfying photos that fit certain concepts and aesthetic ideals. In fact, the Google Clips video camera released in January 2018 is already doing this. The camera is fully AI-based. It uses

computer vision to recognize people and pets and certain emotional expressions, and was trained by professional photographers to make "good" videos with proper composition, interesting actions, etc. (Lovejoy, 2018).

Measuring Diversity

So how can we know if aesthetic diversity in contemporary culture – or even only in one area such as photography – is growing or decreasing? Maybe we can use AI itself to start answering such questions more precisely, as opposed to guessing or just following our intuitions that are often wrong?

Since the middle of the 2000s, hundreds of thousands of computer and social scientists have been quantitatively analysing massive samples of contemporary digital culture, including billions of posts and user interactions on Facebook, Twitter, Instagram, Flickr, Pinterest, and other social networks, as well as on recommendation sites such as Yelp, creatives networks such as Behance, etc. (I will describe relevant examples of such research below). They developed many quantitative measures that describe – or attempt to – some aspects of culture such as structures of sharing in social networks or uniqueness and originality of user-created images. If we think within this paradigm, we can also propose *measures of aesthetic diversity*, and apply them to some cultural areas and types of media. Since we can often access user content shared online in the past (for example, content shared on Flickr since 2004, or Instagram since 2010), we can also calculate how diversity in some cultural areas is changing over time.

This would require developing formal *measures of aesthetic diversity* for different media and cultural fields, from fashion and interior design to cinema and music. And this would be very useful by itself because it would allow us to look at contemporary culture in new ways. Although such formal definitions will never fully account for our aesthetic experiences in many cases, they can still help us by providing new concepts for thinking about global digital culture.

Note that we would need to differentiate between different types of diversity. One is *diversity of content*, i.e. the items being created in a particular cultural area. For example, in the case of photography, this diversity includes types of subjects, techniques, and styles in photographs. Another is *diversity in users' choices*, similar to how it was analysed in the YouTube study previously cited. For example, contemporary fashion designers around the world may be creating items that can have very diverse styles, silhouettes, forms, volumes, materials, textures, and colours, but the diversity of items being purchased and worn by people worldwide may be much smaller. Or, it can be much bigger since today many people mix different items in their outfits, creating composite looks that designers and retailers do not offer. Other types of diversity can also be defined as appropriate.

The idea of measuring aesthetic diversity in global contemporary culture allows us to make other interesting distinctions. One is between *global and local diversity*. If we measure a sufficient number of items globally, culture in many local places may look very homogeneous

in comparison to the full range of choices available worldwide. But, if we zoom into such places, what looked like small hills will now look like mountains, so to speak – i.e., we will realize that these places are quite diverse if viewed on their own.

Related to this distinction is another one – between *objective and subjective measurements*. (We can also call this analyst vs. users' perspectives.) So far, we assumed that our measurements are taken from an abstract, all-seeing point of view. All items or user choices are placed on a single scale and viewed and measured from the outside. Such perspective is standard in biology when we are measuring Earth's biodiversity – for example, counting the number of distinct species in a given habitat, or on Earth as a whole. But in the case of culture, we may alternatively consider how diversity and differences between items inside a habitat is perceived by members of this habitat itself. In this perspective, local traditions and conventions determine whether some items or choices are perceived as radical or not – and not their objective characteristics alone. To continue with fashion examples, in many Western cities people commonly wear many colours, and this is not seen as radical. But in a city like Seoul where the palettes of greys, whites, and blacks dominate how people dress, the appearance of outfits with very saturated colours, and especially more than one such colour together, will be noticed and perceived as being outside the norm.

Limits to Automation

Will AI replace professional cultural creators – media, industrial, and fashion designers, photographers and cinematographers, architects, urban planners, and so on? Will countries and cities worldwide compete as to who can more quickly and better automate their creative industries? Will countries and cities (or separate companies) that figure out how best to combine AI and human skills and talents get ahead of the others?

Today AI gives us the option to automate our aesthetic choices (via recommendation engines), assists in certain areas of *aesthetic production* such as consumer photography, and automates other cultural experiences (for example, automatically selecting ads we see online). But in the future, it will play a larger part in professional cultural production. Its use of helping to design fashion items, logos, music, TV commercials, and works in other areas of culture is already growing. But currently, human experts usually make the final decisions or do actual production based on ideas and media generated by AI.

The well-known example of *Game of Thrones* (American fantasy television drama series that premiered in 2011) is case in point. The computer suggested plot ideas, but the actual writing and the show's development was done by humans. We can only talk about fully *AI-driven culture* where AI will be allowed to create the finished design and media from beginning to end. In this future, humans will not be deciding if these products should be shown to audiences; they will just trust that AI systems know best – the way AI is already fully trusted to choose when and where to show particular ads, as well as who should see them.

We are not there yet. For example, in 2016 IBM Watson created the first "AI-made movie trailer" for the feature film *Morgan* (Mix, 2016). However, AI only chose various shots from the completed movie that it "thought" were suitable to include in the trailer, and a human editor did the final selection and editing. In another example, to create a system that would automatically suggest suitable answers to the emails users receive, Google workers first created a dataset of all such answers manually. AI chooses what answers to suggest in each possible case, but it does not generate them. (The head of Google's AI in New York explained that even one bad mistake in such a scenario could generate bad press for the company, so Google could not risk having AI come up with suggested answer sentences and phrases on its own.)

It is logical to think that any area of cultural production which either follows explicit rules or has systematic patterns can be in principle automated. Thus, many commercial cultural areas such as TV dramas, romance novels, professional photography, music video, news stories, website and graphic design, and residential architecture are suitable for automation. For example, we can teach computers to write TV drama scripts, do food photography, or compose news stories in many genres (so far, AI systems are only used to automatically compose sports and business stories). So rather than asking if any such area will be automated one day or not, we need to assume that it will happen and only ask "when."

This sounds logical, but the reality is not so simple. Starting in the 1960s, artists, composers, and architects used algorithms to generate images, animations, music, and 3D designs ("Computer Art," n.d.). Some of these works have entered the cultural canons. They display wonderful aesthetic inventiveness and refinement. However, in most cases they are *abstract compositions* with interesting and complex patterns, but *without direct references to the human world*. Think of such classics as abstract geometric images by Manfred Mohr (1969-1973) (Mohr, n.d.), John Whitney's computer animation *Arabesque* (1975), or Iannis Xenakis's musical compositions *Atrées* and *Morsima-Amorsima* (1962) (Maurer, 1999). There is no figuration in these algorithmically generated works, no characters like in novels, and no shots of the real world edited together into narratives like in feature films.

Now compare these abstract algorithmic classics with current attempts to automatically synthesize works that are about human beings, their worlds, their interests, emotions and meanings. For example, today Google Photos and Facebook offer users automatically created slideshows and videos edited from their photos. The results are sometimes entertaining, and sometimes useful, but they can't be yet compared to professionally created media. The same applies to images generated by Google engineers using DeepDream neural net (2015-) and later by others who used the same technology (DeepDream, n.d.). These AI creations in my view are more successful than the automatically generated slideshows of user photos, but this is not because DeepDream is a better AI. The reason is that 20th century visual art styles tolerate more randomness and less precision than, for example, a photo narrative about a trip that has distinct conventions and restrictions on what and can be included and when. Thus, in the case of DeepDream, AI can create artistically plausible images which do refer to the human world because we consider it "modern art" and expect big variability. But in the case of automatically edited slideshows, we immediately know that the computer does not really understand what it is selecting and editing together.

AI and Genre Conventions

Creating aesthetically satisfying and semantically plausible media artefacts about human beings and their world may only become possible after sufficient progress in "artificial general intelligence" (AGI, also referred to as "strong AI") is made. In other words, a computer would need to have approximately the same knowledge of the world as an adult human.

However, this is not only a matter of progress in AI research. Whether an algorithmic creation looks plausible or not also depends on genre conventions. In some cases, even very simple AI can produce satisfying results.

In 2002-2005, I collaborated with Andreas Kratky on *Soft Cinema* – a semi-automated system for the making of procedural narrative films (Manovich & Kratky, 2005; Manovich, 2005). Using software we developed, we produced three such films. Each film used a database of a few hundred short video shots we recorded. The choice of shots for the films and their editing in time was done by software using parameters we selected.

The project was shown as an installation in 45 exhibitions. During the duration of a singular exhibition, which sometimes would last a few months, the software was continually editing a film in real-time by selecting short video clips from a database. The complete narrative for each film lasted between 7 and 13 minutes, depending on the film. After one version played, a computer immediately started editing and showing the next version.

Rather than trying to simulate conventions of mainstream narrative cinema or documentary filmmaking, we instead took as our inspiration experimental 20th century films. Specifically, we used the principle that narration and visuals do not always have to correspond.

As a voice-over narrated a story, a computer program was selecting short video clips from a database and editing them together using metadata and rules we established. The program was also generating screen layouts which used a Mondrian-like grid to display anywhere from one to six videos in the same frame. The sequences of clips did not directly illustrate the narrative. However, since all clips in each film database shared certain semantic and visual references, the overall result of this semi-automated process looked meaningful. The mind of a viewer created connections between the narrative content and visuals of the clips playing on the screen. Thus, the more "loose" and associative conventions of "avant-garde" or "experimental" cinema turned out be much easier to simulate than a conventional narrative film.

The latter requires much tighter coordination between all shots. The viewer would immediately notice all mistakes AI made, while in our *Soft Cinema* mistakes were not possible in principle, since the selection and screen placement of clips did not directly illustrate the narrative.

However, it is likely that we will also gradually learn to use AI to generate works in cultural genres which have a lot of rules and constraints. For example, synthesized 3D human characters and conversational agents are gradually becoming more realistic. This is a very challenging area because human verbal communication is certainly a very strict "genre" – if a speaker makes a gesture or facial expression that does not correspond to what he or she is saying, we will notice this immediately.

The use of algorithms in design and architecture (often called "parametric design") is another mature area. Perhaps one day we can delegate to computers the creation of all details of an entire city, from planning to architecture, landscaping, traffic management, and all infrastructure. But when this happens, will we be able to prevent such an AI-metropolis from imposing on, regulating, and rationalizing our existence in the name of progress and human happiness – like the urban visions of the 17th – 19th centuries, European utopias of ideal human communities, Le Corbusier's 1920s radical urban rationalization proposals, and cybernetic and mechanical visions in Godard's *Alphaville* and Tati's *Playtime*?

If all creative and knowledge work will become the domain of AI, what will be left for humans? What will be the purpose of our existence? Watching endless films created by AI, listening to AI-generated music, and being driven in driverless cars around AI-generated cities?

Many modern thinkers and artists have envisioned a future where humans, liberated by machines from mechanical and boring work, will be engaging only in play and art (e.g., Constant's *New Babylon*). But if automation of cultural production by AI continues, eventually it will be these AI playing and making art – not us.

AI as a Culture Theorist

Automation of cultural production may use AI based on a system of explicit rules, or it may use a different approach called "supervised machine learning." Advances in "deep learning" (particular methods for supervised machine learning) in the 2010s have made the latter approach very popular today. In one common cultural application of supervised learning, a computer is first fed many examples of works of particular genres, styles, and media, and it gradually learns the patterns common to each genre or style. After that, a computer is given new works it has not seen before and uses the patterns it learned to classify those works. For example, in the case of creating the movie trailer for *Morgan*, a computer was given 100 horror movies. In these movies, "each scene was tagged with an emotion from a broad bank of 24 different emotions and labels from across 22,000 scene categories, such as eerie, frightening, and loving" (Smith, 2016). The computer learned "the types of scenes that categorically fit into the structure of a suspense/horror movie trailer." So, after it was given *Morgan*, it selected the scenes that it believed to be the best ones for the trailer.

We can say that AI, implemented in this way, acts as a cinema (or art, video game, fashion, etc.) theorist and historian. These researchers also study many works created in particular

places and historical periods to find *common patterns*. Their findings become part of the *history and theory* of this area.

However, this is a crucial difference between an "AI culture theorist" and a human theorist/ historian. The latter comes up with explicit principles that describe how a cultural area function. For example, the standard textbook used in universities in numerous film studies courses – *The Classical Hollywood Cinema* – provides answers to questions such as "How does the typical Hollywood film use techniques and storytelling forms of the film medium?" (Bordwell, Staiger, & Thompson, 2010) But at present, the frequent result of training AI ((i.e., use of deep neural networks) on many cultural examples is a *black box*. Given new examples, it can classify them correctly – for example, it can decide whether a particular film belongs to "classical Hollywood cinema" or not. But we often don't know how a neural network came up with this decision. Similarly, a neural net can be trained to distinguish between works of different artists, fashion designers, or film directors. And it can also generate new objects in the same style. But often we don't know what exactly the computer has learned. (However, many computer scientists are working to develop methods to make black boxes created by neural networks more transparent and for it to be possible to "audit" them.)

This is one of the key issues surrounding cultural uses of AI. Are the results of machine learning *interpretable*, or are they only a black box which is efficient in production but useless for human understanding of the domain? Will the expanding use of machine learning to create new cultural objects make explicit the patterns in many existing cultural fields that we may not be aware of? And if it does, will it be in a form that will be understandable for people without degrees in computer science? Will the companies deploying machine learning to generate movies, ads, designs, images, music, urban designs, etc. expose what their systems have learned?

II. AI and Analysis of Culture

Quantitative Analysis of Large Cultural Data

Most of my examples of practical cultural uses of AI listed above involve commercial devices, services, or research done in companies. But they are not the only players in cultural AI. Many academic researchers have been using various computer methods to quantitatively study culture using large datasets of cultural objects, or records of users' interactions and behaviours. This research is published in academic journals or presented at conferences, so it is accessible, and code and datasets used in studies are also made available. (Note that computer scientists working in companies also periodically publish papers describing in detail the techniques used in these companies, such as for example the description of Pinterest's visual recommendation engine [Jing, et al., 2015]).

Much of such research in computer science uses large samples of user content shared on social networks and data about people's behaviours on these networks, such as numbers of views, likes, and shares. Researchers have published hundreds of thousands of papers that computationally analyse the characteristics of image, video, and text posts, as well as user behaviour on most popular social networks and media sharing services such as Weibo, Facebook, Instagram, Flickr, YouTube, Pinterest, and Tumblr. In one research area called "computational aesthetics," scientists create mathematical models that predict which images and videos will be popular, and how this popularity is affected by their content and other characteristics such as "memorability," "interestingness," "beauty," or "creativity" (Redi, O'Hare, Schifanella, Trevisiol, & Jaimes, 2014; Schifanella, Redi, & Aiello, 2015). (The researchers also proposed metrics to measure these characteristics).

Let's look at computer science publications that quantitatively analyse Instagram as an example. To locate only quantitative papers, I usually add the word "dataset" to the name of the social network I want to find research about, and then search for this combination of words on Google Scholar. The search for "Instagram dataset" conducted on July 15, 2017 returned 9,210 journal articles and conference papers. Here are some of these papers from 2014-2017 that will give you an idea about topics pursued by researchers.

One publication analysed the most popular Instagram subjects and types of users in terms what kinds of subjects frequently appear in photos in their feeds (Hu, Manikonda, & Kambhampati, 2014). Another paper used a sample of 4.1 million Instagram photos to quantify the effect of using filters on numbers of views and comments (Bakhshi, Shamma, Kennedy, & Gilbert, 2015). In another paper, a group of researchers analyzed temporal and demographic trends in Instagram selfies by using 5.5 million photos with faces they collected from the network. They have also tested three alternative hypotheses about the reasons behind posting selfies in each of 117 countries in their dataset (Souza, et al., 2015). Yet another paper investigated clothing and fashion styles in 44 world cities using 100 million Instagram photos (Matzen, Bala, & Snavely, 2017).

Other publications use computational methods to analyse patterns in historical culture using the available digitized archives. Interesting examples of such research are *Toward Automated Discovery of Artistic Influence* (Saleh, Abe, Arora, & Elgammal, 2014), *Measuring the Evolution of Contemporary Western Popular Music* (Serrà, Corral, Boguñá, & Haro, 2012), *Shot durations, shot classes, and the increased pace of popular movies* (Cutting & Candan, 2015), and *OmniArt: Multi-task Deep Learning for Artistic Data Analysis* (Strezoski & Worring, 2017).

The first paper presents a mathematical model for the automated discovery of influence between artists. The model was tested using 1,710 images of paintings by 66 well-known artists. While some of the discovered influences are the same ones often described by art historians, the model also suggests other visual influences between artists that had not been discussed previously. The second paper investigates changes in popular music using a dataset of 464,411 songs produced between 1955 and 2010. The dataset included "a variety of popular genres, including rock, pop, hip hop, metal, or electronic." The authors concluded

that over time, there was "the restriction of pitch transitions" and "the homogenization of the timbral palette" – in other words, some of the musical variability has decreased. The third paper analyses gradual changes in average shot duration across 9,400 English-language narrative films created between 1912 and 2013. The fourth paper uses machine learning to predict artist name, period, and material of artworks from a dataset of 432,217 historical art images.

The Limits of Statistical Reason

These and many other studies that analyse large cultural samples and collections using AI and statistical methods present valuable and original insights. They would be impossible to arrive at without such datasets and methods, using only "armchair" academic theorizing or small group ethnographic observations. But because most of these studies use large datasets and a statistical approach, they also have some common limitations.

The discipline of statistics has a few areas, and one of them (historically the earliest) is *descriptive statistics*. Here "statistics" refers to a *summary of a collection of information*. For example, we can measure the height of a large number of people in a particular demographic group in a given city or a country and calculate their average height. If we measure one million people, we can then replace one million different numbers with a single number representing this average.

Any summary of the information by default is going to omit some details because it is smaller than the original information. Therefore, if we use a statistical approach to summarize a collection of cultural artefacts, or information about cultural behaviours, our findings will not apply to every item in the collection or a record of behaviour. Similarly, mathematical models that describe patterns and relationships in cultural data and attempt to predict future data will only be correct some of the time. Such summaries and models describe only some patterns in a cultural dataset, omitting many others.

For example, the authors of the paper about the effects of Instagram image filters analysed data on millions of photos and reported that "filtered photos are 21% more likely to be viewed and 45% more likely to be commented [on]" (Bakhshi, Shamma, Kennedy, & Gilbert, 2015). But certainly, this large sample also contains many filtered photos that were not viewed more frequently or did not receive more comments than unfiltered photos. This is the nature of statistical summaries and models – they describe general trends but usually can't predict every single data point. For example, when measuring the height of a population, some people will be significantly taller or shorter than the average. Similarly, in the case of filtered Instagram photos, some filtered images are likely to receive significantly more views than the 21% average; others will receive the same number of views as unfiltered photos, and some will receive less views.

Of course, most qualitative cultural theories and histories are also summaries of information. They don't describe every single artefact within a given genre, medium, or period, but present only certain trends. Specifically, they may propose that certain artistic techniques,

themes, and conventions are common to a particular set of feature films, TV drama series, literary narratives, or works in other media, genres, and periods. Both the writers and the readers understand that these techniques, themes, and conventions are present in some of these artefacts, but not in all of them.

So, in this way, qualitative and quantitative studies of culture are similar – they account for only some of the information present in a cultural dataset. The artefacts that are not covered may contain other techniques, themes, structures, or conventions that qualitative theory or quantitative models do not describe. And even the artefacts that are covered are likely to have other patterns a particular theory or model does not describe.

The same goes for the accounts of user experience. Qualitative theorists often talk about the effects of certain artefacts on the readers, viewers, or listeners by extrapolating from their own singular experience, and this is a very big generalization. Quantitative studies of social media instead rely on the explicit signs and traces of the experience of a large number of users such as likes and shares on social media. While larger samples allow them to draw more accurate conclusions, such signs and traces can't cover the full range of aesthetic experiences, and this is a very big limitation of this approach.

Here it is relevant to recall the famous essay *S/Z* by French literary critic and semiotician Roland Barthes which analyses Balzac's short story (Barthes, 2002 [1973]). The essay does not claim to describe all semiotic codes in Balzac's short story. It also does not claim that they form a system, contrary to an earlier structuralist semiotics view. Instead, Barthes in his analysis only scanned Balzac's story, showing examples of many kinds of codes, without assuming that a comprehensive description of all of them even in such a short story is possible or desirable. This explicit recognition of limits of any cultural reading remains very relevant today, as the use of computational and statistical methods for analysis of cultural data is becoming more popular.

Against Reduction

For scientists, having a statistical model which can predict only some future data is not ideal, because science assumes that such a model should accurately capture characteristics of the phenomena being modelled. A model that is accurate 90% of the time is better than one with only 60% accuracy. Having a model that predicts only some of the data means that we don't fully understand the process that generates it, or that we did not capture all the relevant characteristics. (For example, in object and scene recognition research, computer vision scientists compete every year to see which model can most accurately recognize objects and types of scenes in a large number of photos [Russakovsky, et al., 2015]).

But we can also use AI and other computational methods to study culture without such assumptions, and this is part of the alternative research paradigm I refer to as Cultural Analytics (2007-) (Manovich, 2009). In this paradigm, we do not want to "explain" most or even some of the cultural data using simple mathematical models and treat the rest as "error" or "noise" because our models cannot account for it. We also do not want to assume

that certain patterns such as "hero's journey," "golden ratio," or "binary oppositions" exist in a lot of cultural artefacts. In other words, the ultimate goal of Cultural Analytics is *to avoid reductive summarization* typical of both traditional qualitative cultural theory and history and recent quantitative computational research.

To achieve this, we have been creating very high-resolution visualizations that show all artefacts in a visual dataset together, rather than quantifying only some dimensions of these artifacts and then building statistical models. In the visualizations, the artefacts are sorted in different ways according to properties such as colour, texture, composition, or content. Our software measures such properties using techniques from computer vision, and then organizes all artefacts according to these measurements. Therefore, some reduction is also inevitable – for example, if we organize 50,000 images shared on Instagram in Tokyo or Bangkok according to the time of day and average color saturation, the patterns in time and saturation will stand out, and other patterns will be harder to see. However, rather than producing a single visualization of each dataset, we create multiple visualizations where images are sorted according to many different properties, so each can reveal a different pattern.

"To avoid reductive summarization" – does this mean that we are only interested in differences between artefacts and that we want to avoid any kind of reduction at all cost? To postulate existence of cultural patterns is to accept at least some reduction in analysing data. Without this, we cannot compare anything, unless we are dealing with extreme minimalism or seriality, where the artist makes everything else equal and only varies values on one or two dimensions, like in1960s sculptures by Sol LeWitt.

Therefore, some reduction is actually desirable. And to claim that we see any pattern at all is to practice such reduction. We just need to be explicit about what our patterns reveal and what is left invisible. Therefore, if we are to define Cultural Analytics as *analysis and visualization of cultural patterns on different scales*, we will need to immediately qualify this statement. While we want to discover repeating patterns in cultural artefacts and behaviours, we should always remember that they only account for *some* aspects of these artefacts and behaviors.

Unless it is a 100% human copy of another cultural artefact or produced mechanically or algorithmically to be identical with other copies, every expression and interaction is unique at least in some ways. In some cases, this uniqueness is not important for our analysis, and in other cases it is. For example, in our project *Selfiecity*, we extracted automatically many characteristics from thousands of Instagram self-portraits shared during the same week in six global cities (Manovich, et al., 2014-2015). Our interactive software allows you to sort these selfies according to parameters such as angle and tilt of a head, degree of smile, age, gender, and others. For instance, you can select all female selfies from São Paolo that tilt their head more than 10 degrees, or all selfies from Bangkok with strong smiles (strength of a smile was measured on a 0-100 scale). In this way, you can discover many trends in how people pose for their selfies and compare them between cities.

In our everyday perceptual experiences, we also constantly recognize common patterns. This is what allows us to make judgements that something is typical, unusual, or unique. Such recognition is also a form of reduction – we feel that something either fits into familiar categories we experienced previously or that it does not fit into any of them. But this is not all we do. I think that the reason we do not get tired looking at endless faces and bodies when we browse Instagram is that each of them is unique, and we notice this. *What fascinates us is not repeating patterns but unique details and their combinations*. No two human faces are exactly the same, and we enjoy noticing these differences.

The ultimate (perhaps utopian) vision of Cultural Analytics is to map in detail and understand the full *diversity* of contemporary professional and user-generated cultural artefacts created globally – i.e. to focus on what is different between numerous artefacts and not only on what they share (i.e. common patterns). In the 19th and 20th centuries, the lack of appropriate technologies to store, organize, and compare large cultural datasets was contributing to the popularity of reductive cultural theories – for example, seeing every human culture as going through the same three stages (early development, reaching perfection, followed by decay). Such categorical schemes reduce the variability and messiness of actual cultural histories. Technological limitations in data management may have also contributed to the 19th century obsession with cultural classification – another way to reduce the endless diversity of human cultural expressions to a manageable system.

Seeing Differences

Today we can use a single computer to capture, compare, quantify, and visualize thousands of differences between tens of millions of objects. We do not have any more excuses to only focus on what cultural artefacts or behaviours share (as members of the proposed cultural category or a cultural period) – which is what we do when we categorize them or perceive them as instances of general types. Although we may have to start with extracting patterns just to draw our initial maps of contemporary cultural production and dynamics given their scales and dynamics, eventually they may recede in the background or even completely dissolve, as we focus only on the differences between individual objects. This is to me the ultimate promise behind using AI methods and cultural datasets to observe and describe cultures in the 21st century.

We established our Cultural Analytics Lab in 2007 and spent the next 10 years assembling, analysing, and visualizing dozens of datasets of artefacts in many genres of visual media, from pages of manga books, modernist paintings, and avant-garde films to images shared on Instagram, Twitter, and Flickr. As I learned, the Cultural Analytics ideals presented above are easy to state but challenging to execute well in practice. Human brains and natural languages are categorizing machines. Our perception constantly processes sensory information and categorizes it. A pattern we observe is like constructing a category: a recognition that some things or some aspects of these things have something in common. So, can we learn *to think about cultural processes without categories*? How can we hope to eventually dissolve the patterns that AI-based analysis will reveal in numerous artefacts to

see only their differences? How can we refuse the very idea of summarization ingrained in us by a hundred years of statistics?

How do we move away from the assumption of the humanities (which until now "owned" thinking and writing about culture) that their goal is discovery and interpretation of general cultural *types*, be they "modernism," "narrative structures," "images of working class," "selfies," or "amateur digital photographers?" How do we instead learn to see cultures in more detail, without immediately looking for, and noticing, only types, structures or patterns?

First, we need sufficiently large cultural data samples. Next, we can extract sufficiently large numbers of features that capture characteristics of the artefacts, their reception and use by audiences, and their circulation. (We also need to think more about how to represent cultural *processes*, interactions, and dynamics – especially since today we use interactive digital cultural media as opposed to historical static artefacts. Once we have such datasets, we can explore them using a variety of *unsupervised machine learning* methods such as visualization of multi-dimensional data, clustering, and others.

Supervised machine learning is commonly used in the culture industry to classify artefacts, people, and behaviours. Such classifications often reinforce already existing and taken-for-granted classifications and ways of seeing the world. For example, some photo sharing sites and apps automatically detect and tag familiar classes of objects and scenes in user photos ("cities," "faces," "food," "landscapes," "sea," etc.).

Unsupervised machine learning methods allow us to discover new categories for which we don't have names and to see connections we were not previously aware of. Thus, rather than *reducing* cultural data to familiar categories, unsupervised machine learning can expose limitations of such categories and suggest new ways of seeing culture.

In this short book I outlined some of the issues raised by the ongoing integration of AI into many parts of culture, as well as the use of AI to analyse it quantitatively. I presented examples of AI use in digital devices and services and one possible taxonomy of these uses, as well as examples of computational analysis of digital culture. The key theme that guided this discussion was the question of cultural variability. Does AI integration in cultural production and reception lead to a decrease in aesthetic variability? Or does it, on the contrary, increase it? And what are the different ways in which aesthetic variability can be defined and measured? I see the last question productive in itself, since it may open new ways of thinking about aesthetics in the digital era.

Our cultural period is characterized by an unprecedented scale of production and circulation, and also by global integration in cultural production, reception, and reuse. Today people around the world create, share, and interact with billions of new digital artefacts every day. We need new methods for seeing culture at its new scale, velocity, and connectivity that can combine both qualitative and quantitative approaches and that can reveal full variability of this new ecosystem without reducing it to a small number of categories. AI plays a crucial role in this new global cultural ecosystem, suggesting to people whom to follow and what to see, helping them edit media they create, making aesthetic decisions for them, determining

how many people will see their content, deciding which ads will be shown to them, etc. Understanding the basic principles of AI techniques used today in culture is therefore important if you want to be culturally literate today, and essential if you are a creator yourself.

References

Al-Halah, Z., Stiefelhagen, R., & Grauman, K. (2017, August 3). *Fashion Forward: Forecasting Visual Style in Fashion*. Retrieved August 6, 2018 from arXiv.org: https://arxiv.org/abs/1705.06394v2

Asrar, S. (2016, March 28). *A quick look at recommendation engines and how the New York Times makes recommendations*. Retrieved August 6, 2018 from Knight Lab: https://knightlab.northwestern.edu/2016/03/28/a-quick-look-at-recommendation-engines-and-how-the-new-york-times-makes-recommendations/

Associated Press. (2016, October 24). *Apple's latest iOS brings blurring 'portrait mode' to the iPhone 7 Plus*. Retrieved August 6, 2018 from The Guardian: https://www.theguardian.com/technology/2016/oct/24/apple-ios-10-1-iphone-7-plus-portrait-mode

Bakhshi, S., Shamma, D. A., Kennedy, L., & Gilbert, E. (2015). Why We Filter Our Photos and How It Impacts Engagement. AAAI. Retrieved August 6, 2018 from http://comp.social.gatech.edu/papers/icwsm15.why.bakhshi.pdf

Barthes, R. (2002 [1973]). *S/Z.* (R. Miller, Trans.) London: Blackwell. Retrieved August 6, 2018 from https://monoskop.org/images/d/d6/Barthes_Roland_S-Z_2002.pdf

Bordwell, D., Staiger, J., & Thompson, K. (2010, September). *The Classical Hollywood Cinema Twenty-Five Years Along*. Retrieved August 6, 2018 from http://www.davidbordwell.net/essays/classical.php

Clarifai. (2018). *Models: Explore all our ready-to-use image recognition models to suit your specific needs*. Retrieved August 6, 2018 from Clarifai: https://clarifai.com/models/

Computer Art. (n.d.). Retrieved August 6, 2018, from Wikipedia: https://en.wikipedia.org/wiki/Computer_art

Cutting, J. E., & Candan, A. (2015). Shot durations, shot classes, and the increased pace of popular movies. *Projections: The journal for movies and mind, 9*(2), 40-62. doi:http://people.psych.cornell.edu/~jec7/pubs/cutting&candanProj15.pdf

DeepDream. (n.d.). Retrieved August 6, 2018, from Wikipedia: https://en.wikipedia.org/wiki/DeepDream

Fujifilm Corporation. (2017). *Fujifilm X-E3 Owner's Manual*. Retrieved August 6, 2018 from
    http://fujifilm-dsc.com/en/manual/x-e3/

Grigonis, H. (2016, August 26). Searching for a photo? Updated EyeEm app's AI uses
    keywords to find it. *Digital Trends*. Retrieved August 6, 2018 from
    https://www.digitaltrends.com/photography/eyeem-adds-auto-keywords/

Hillen, B. (2018, July 19). Huawei launches photography contest with an AI judge. *DP
    Review*. Retrieved August 6, 2018 from
    https://www.dpreview.com/news/4273899128/huawei-launches-spark-a-renaissance-
    photography-contest-with-an-ai-judge

Hu, Y., Manikonda, L., & Kambhampati, S. (2014). What We Instagram: A First Analysis of
    Instagram Photo Content and User Types. *Proceedings of the Eighth International
    AAAI Conference on Web and Social Media* (pp. 595-598). AAAI. Retrieved August
    6, 2018 from
    https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8118/8087

Jing, Y., Liu, D., Kislyuk, D., Zhai, A., Xu, J., Donahue, J., & Tavel, S. (2015). *Visual Search
    at Pinterest.* doi:10.1145/2783258.2788621

Kafka, P. (2016, August 5). *The Washington Post will use robots to write stories about the
    Rio Olympics*. Retrieved August 6, 2018 from Recode:
    https://www.recode.net/2016/8/5/12383340/washington-post-olympics-software

Khosla, A., Bernstein, M., Berg, A.C., and Fei-Fei, L. (2015, January 30). ImageNet Large
    Scale Visual Recognition Challenge. Retrieved August 6, 2018 from arXiv.org:
    https://arxiv.org/abs/1409.0575v3

Китов, А. (1956). Глава VI. Неарифметические применения Электронных Цифровых
    Машин. In *Электронные цифровые машины.* Retrieved August 6, 2018 from
    http://www.computer-museum.ru/books/kitov_ecm.htm

Lovejoy, J. (2018, January 25). *The UX of AI: Using Google Clips to understand how a
    human-centered design process elevates artificial intelligence*. Retrieved August 6,
    2018 from Google Design: https://design.google/library/ux-ai/

Manovich, L. (2005). *Soft Cinema*. Retrieved August 6, 2018 from Ubu Web:
    http://www.ubu.com/film/manovich_soft.html

Manovich, L. (2009). Cultural Analytics: Visualizing Cultural Patterns in the Era of "More
    Media". *Domus*. Retrieved August 6, 2018 from
    http://manovich.net/index.php/projects/cultural-analytics-visualizing-cultural-patterns

Manovich, L. (2017). *Instagram and Contemporary Image.* Retrieved August 6, 2018 from http://manovich.net/index.php/projects/instagram-and-contemporary-image

Manovich, L., & Kratky, A. (2005). *Soft Cinema. Navigating the Database.* Cambridge, MA: The MIT Press. Retrieved August 6, 2018 from https://mitpress.mit.edu/books/soft-cinema

Manovich, L., Stefaner, M., Yazdani, M., Baur, D., Goddemeyer, D., Tifentale, A., Hochman, N., and Chow, J. (2014-2015). *Selfiecity.* Retrieved August 6, 2018 from http://selfiecity.net/

Matzen, K., Bala, K., & Snavely, N. (2017, June 6). StreetStyle: Exploring world-wide clothing styles from millions of photos. Retrieved August 6, 2018 from arXiv.org: https://arxiv.org/abs/1706.01869

Maurer, J. A. (1999, March). *A Brief History of Algorithmic Composition*. Retrieved August 6, 2018 from https://ccrma.stanford.edu/~blackrse/algorithm.html

Mix. (2016, September 1). *Watch: IBM Watson creates the first AI-made movie trailer – and it's really eerie*. Retrieved August 6, 2018 from The Next Web: https://thenextweb.com/insider/2016/09/01/watch-ibm-watson-creates-first-entirely-ai-made-movie-trailer-really-eerie/

Mohr, M. (n.d.). *Art Works*. Retrieved August 6, 2018 from http://www.emohr.com/ww4_out.html

Promise of AI Not So Bright. (2006, April 13). *The Washington Times*. Retrieved August 6, 2018 from https://www.washingtontimes.com/news/2006/apr/13/20060413-105217-7645r/

Redi, M., O'Hare, N., Schifanella, R., Trevisiol, M., & Jaimes, A. (2014). 6 seconds of sound and vision: Creativity in micro-videos. Retrieved August 6, 2018 from arXiv.org: https://arxiv.org/abs/1411.4080

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A., Fei-Fei, L. (2015, January 30). ImageNet Large Scale Visual Recognition Challenge. Retrieved August 6, 2018 from arXiv.org: https://arxiv.org/abs/1409.0575v3

Saleh, B., Abe, K., Arora, R. S., & Elgammal, A. (2014, August 14). Toward Automated Discovery of Artistic Influence. Retrieved August 6, 2018 from arXiv.org: https://arxiv.org/abs/1408.3218v1

Schifanella, R., Redi, M., & Aiello, L. (2015, May 15). An Image is Worth More than a Thousand Favorites: Surfacing the Hidden Beauty of Flickr Pictures. Retrieved August 6, 2018 from arXiv.org: https://arxiv.org/abs/1505.03358

Serrà, J., Corral, Á., Boguñá, M., & Haro, M. A. (2012, July 26). Measuring the Evolution of Contemporary Western Popular Music. *Scientific Reports*. Retrieved August 6, 2018 from https://www.nature.com/articles/srep00521

Shah, S. (2017, August 24). *Amazon's AI could create the next must-have fashion brand*. Retrieved August 6, 2018 from Endgadget: https://www.engadget.com/2017/08/24/amazon-ai-fashion-brands/

Shaji, A., & Yildirim, G. (2017, March 29). *Personalized Aesthetics: Recording the Visual Mind using Machine Learning*. Retrieved August 6, 2018 from NVIDIA Developer Blog: https://devblogs.nvidia.com/parallelforall/personalized-aesthetics-machine-learning/

Smith, J. R. (2016, August 31). *IBM Research Takes Watson to Hollywood with the First "Cognitive Movie Trailer"*. Retrieved August 6, 2018 from IBM: https://www.ibm.com/blogs/think/2016/08/cognitive-movie-trailer/

Souza, F., Las Casas, D. d., Flores, V., Youn, S., Cha, M., Quercia, D., & Almeida, V. (2015). Dawn of the Selfie Era: The Whos, Wheres, and Hows of Selfies on Instagram. *Proceedings of the 2015 ACM on Conference on Online Social Networks (COSN '15)* (pp. 221-231). New York: ACM. Retrieved August 6, 2018 from http://dx.doi.org/10.1145/2817946.2817948

Strezoski, G., & Worring, M. (2017). OmniArt: Multi-task Deep Learning for Artistic Data Analysis. Retrieved August 6, 2018 from arXiv.org: https://arxiv.org/abs/1708.00684

Tencent YoTu Lab. (2018). *Automatical Facial Beauty*. Retrieved August 6, 2018 from https://bestimage.qq.com/facebeautify.html

Zhou, R., Gao, L., & Khemmarat, S. (2010). The impact of YouTube recommendation system on video views. *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement 2010.* Melbourne, Australia - November 1-3, 2010. Retrieved August 6, 2018 from https://www.researchgate.net/publication/220269659_The_impact_of_YouTube_recommendation_system_on_video_views