From data-mining to social networks. Aesthetic assessment of photos. Subjectivity and AI.
Objectives of this course

1 – We will see that esthetic assessment of photos appears as an ultimate challenge for Artificial Intelligence, competing the human expertise.

2 – We will present two families of solutions based on two different conceptions of Beauty:
   • The “easy” track, based on ‘objectivism’, provides rather conclusive but limited issues.
   • The difficult one, based on ‘subjectivism’ is still in its infancy and requires huge investments to convince.

This course is the opportunity for an exemplified review of 20 years of “Machine Learning” from handcrafted feature detection+classification to end-to-end Neural Networks.

It underlines the difficulty to adapt the methods based on intensive training to individual situations.

It proposes some tracks benefitting from social networks and mail order techniques.
The course in brief

- Evaluating the beauty of a photo may be seen as a challenge for AI.
  - Solutions have been proposed using either:
    - Machine Learning and Handmade features
    - Deep Neural Networks
  - They are based on:
    - Collecting data bases of instances of photos either “nice” or “not”
    - Collecting expertise on the assessment

- These methods let no room for the user’s personal tastes
- Several attempts are made to alleviate the problem:
  - Recommendation systems
  - Assessment of the user’s psychological profile
  - Assessment of the user’s tastes
  - Proposing concurrent expertise

About Beauty of Photos
©Dan Kitwood
©Pixabay
Where comes beauty with images?
Can we measure Beauty?

- A specific attention to Photos.
- Beauty of a photo is different from Quality of a photo
- Evaluation of image beauty is important for:
  - Professional applications: Advertisement, Journals & Press, Design, Film industry, Video games, Entertainment, Archive management, etc.
  - Customer applications: Personal photo-collection management, web data mining & retrieval

- A positive answer is proposed using Artificial Intelligence, based on recent progress on:
  - Machine learning
  - On-line image data-bases
  - Social networks & Internet
Beauty

- Beauty is a major reason to select an image in a large set of images
- Beauty is competing with:
  - Relevance (or interestingness)
  - Amazingness (or surprise)
Beauty

- Beauty is a major reason to select an image in a large set
- Beauty is competing with:
  - Relevance (or interestingness)
  - Amazingness (or surprise)
Our domains of interest and Image Relevance

- Universal field = General interest
Our domains of interest and Image Relevance

- **Egocentric field**

- **Relevance received attention from:**
  - James-Lange emotion theory of arousal (1927)
  - Relevance theory of D. Sperber and D. Wilson (1986)
Interest and Relevance

- **Relevance or Interest:**
  - *Interest is an « emotion » = a complex psychophysiological experience*
  - *which results from*
    - environmental (i.e. external) stimuli and from biochemical (i.e. internal) stimuli ➔ passive
    - Or « reasoning » = a cognitive process which allows to deduce new pieces of knowledge from external stimuli ➔ active

- **Several theories of Relevance may be applied**
  - **James-Lange emotion theory of arousal (1927)**
    - The role of emotion to modify appraisal: emotion is the cause of appraisal and not the converse
  - **Relevance theory of D. Sperber and D. Wilson (1986)**
    - Emphasis of the role of cognition vs. linguistic approach of H.P. Grice (1975): relevance is depending on the receiver knowledge
    - Improves the predictive capacity of relevance theory by mean of a mathematical modeling
Our domains of interest and Image Relevance

- Personnal sphere (egocentric field): more easily modeled prediction of relevance
  - Few number of relatives, friends, etc.
  - Close spatial localization (Tannen 1984)
  - Specific and limited topics
  - Temporal coincidence with personnal white stones (availability heuristics) (Tversky & Kahneman, 1973)

  - Bayesian approaches with statistical counts, Poisson distribution,
  - Information theory, etc.
  - compositional pertinence formulation of unexpectedness by use of the W-machine (Dessalles, 2010)
Interest and surprise or amazement
Beauty, Art and Aesthetics
Beauty and Aesthetics

- **Aesthetics** is the branch of philosophy dealing with
  - the nature of art, beauty and taste,
  - the creation and appreciation of beauty.

- **Beauty** is a characteristic of an object or a person that provides a perceptual experience of **pleasure** or satisfaction.

- **Art** is a range of human activities in creating artifacts (artworks), intended to be appreciated for their **beauty** or emotional power.
Beauty and Art: the 20th century fracture

- Beauty is no longer the prime objective of Art.
- Beauty may even disappear as a subsidiary objective

- VISUAL ARTS evolved more and more towards emotional power and less towards beauty

Therefore an artwork needs no longer to be nice but:

- To be an artwork, the document should be recognized as such by:
  - The artist himself
  - Some « academic » actors: museums, galleries, critics, experts
  - General audience
    (no reference to any aesthetical predicate)

(A. Danto, 1992)
With modern definitions of art, it is possible to decide about:

- **Obviously non-aesthetical representations:**

  - F. Goya
  - F. Bacon
  - J.S. Chardin
  - E. Munch
Art is able to decide about:

- **Fakes, copies & reproductions**
  - As « beautiful » as the original but not « art-pieces »

- **Ready-made**
  - The « art-piece » is undistinguishable from the commercial product

A. Warhol  
M. Duchamp
Beauty, Art & Aesthetics

- We are definitively not concerned by Art in Photography
- But only about Beauty in Photography
Automatic Beauty Assessment

- **Beauty** is a characteristic of an object or a person that provides a perceptual experience of pleasure or satisfaction.

- An increasing offer to rank images according to their level of beauty, from:
  - Academic labs
  - Research labs (GAFA …)
  - Start-ups and companies
Beauty … a rather long story for Humanity

25 centuries of philosophical debating:
- Beauty vs Aesthetics vs Art,
- Objectivists vs Subjectivists
- Transcendental vs Naturalist

150 years of experimental & social Psychology
- Fechner, Gestalt, … ➔ Bauhaus, Design

150 years of biological studies
- Physiology of vision ➔ eye & visual paths
- Neurobiology with fNMR ➔ brain & neuro-aesthetics

180 years of Photography
- User’s manual, starting to advanced level manuals
- Universities & courses
- Galleries, museums, exhibitions and challenges,
Aesthetics: a short review of Beauty assessment (1/2)

Two opposite sides:

1 - « Objectivist » approach (Plato, Democrites, Pythagores) : beauty is part of the piece of art. Thus, Beauty is:

- Universal,
- Identically perceived by everybody
- Independent of time, space and context (Kant, Hegel)

- Beauty follows « rules » (to be discovered) linking the attributes of objects (disposition, form, color, etc.)

Instances: Hellenic and Roman Schools
Renaissance and Classical periods
Aesthetics: a short review of Beauty assessment (2/2)

2 - For « Subjectivists » (Locke, Burke, Diderot, Goethe) Beauty is only a matter of personal assessment
   - It depends on observer & context
   - It may vary in time
   - It may be faintly related to the observed object
   - If « rules » have to be found, they ly in the consciousness of the observer.

Instances: Romantism, Impressionism, Surrealism, Pop Art …
Objectivist approach
Approaches based on objectivist approach:

1) Machine Learning + Classification

&

2) Deep Neural Networks
Automatic Aesthetic Assessment

- Machine Learning approaches: 2 steps

Step 1: Training stage = to build the « machine »

Database with groundtruth → Learning machine → Classifier
Automatic Aesthetic Assessment

- Machine Learning approaches: 2 steps

Step 2: Decision stage = to use the « machine »

Diagram:
- Unknown Image
- Classifier
- Class
Machine Learning approaches

- Handcrafted Feature Selection + Classifier
- Deep Neural Network
Machine Learning approaches

- Handcrafted Feature Selection + Classifier
- Deep Neural Network
Objectivist approach

Machine Learning

1 - Feature detection & Classification (handcrafted) 2006 ➔ 2015
Decision rules issued from:
- Applied photography & Psychovision
- Image processing
- Vision & Multimedia

2 - Deep Convolutional Neural Networks (CNN) 2014 ➔ ...
Existing architectures & software exploiting: ResNet, AlexNet, VGG, GoogleNet, …
Specific adaptations to « Beauty »
Criteria to choose a database:

- Where are images taken from? Quality, topics, size, ...
- Level of expertise: professional, trained unprofessional, amateurs
- Evaluation granularity: « Nice/Not-nice », continuous grade,
- Metadata: semantic classes (portraits, landscape …), style (B&W, abstract, night/flash, effects, …), comments (literal appraisal)

AVA = Aesthetic Visual Analysis
AVA data base: Aesthetic Visual Analysis

- 250 000 photos issued from *DPChallenge* (Digital Photography Challenge)
  https://www.dpchallenge.com/
- Good quality: many almost professional images
- Are only kept the photos with at least 200 notes given by visitors
  - Notes between [0,10] ➔ global mark = mean value
- Often with literal comments
  - « I love the audacious composition, as well as the B&W conversion. Great entry. Congrats on the ribbon! »
- Labels: title of the challenge

AVA : https://computervisiononline.com/dataset/1105138637
Learning stage: Which images are beautiful?

Mark distribution of the 250 000 images from AVA database

- Continuous mark: [0,10]
- Binary mark (as a function of $\delta$)

Mean value = 5.34
Almost Gaussian distribution
Machine Learning approaches

- Handcrafted Feature Selection + Classifier
- Deep Neural Network

Image → FEATURES → Classifier (Bayes, SVM, KD Tree) → Class
Format or aspect ratio

- Commercial Formats: 3/2, 4/3, 5/3, 16/9, A, B, ...  
  Markowsky (1992) ➔ 1.83

- Rule of Third:  
  S.E Amirashi et al. (2014)

- Golden Number: \( \varphi = \frac{1+\sqrt{5}}{2} \sim 1.618 \)  
  Markowsky (1992)

- Weight of Center:  
  Arnheim (1983)
Machine learning with Photo aesthetics rules

- Composition rules
  - Weight of the center
  - Importance of order (regular, progressive)
  - Symmetries
  - Alignements, diagonal, perspective …

- Complexity vs Simplicity

- Role of attention points
  - Objet versus background
    - Number
    - Position
    - Focus
    - Contrast
Machine learning with Photo aesthetics rules (3/4)

- **Image**: $I(x,y)$
- **Grey level histogram**: $\text{proba}(I)$
  - Any Mean or Variance: $<I> = \langle i - <i> \rangle^2$
  - Positively biased skew (order 3)
  
    *Attewell & Baddeley, 2007*

- **Power spectrum (after Fourier transform)**:
  - $P_f(I) = |TF(I)|^2$
  - Radial Symmetry
  - $1/f^2$ power spectrum decay

    *Koch M, Denzler J, Redies C, (2010)*

---

\[\begin{align*}
\text{Histogram} & \quad \text{Negative Skew} & \quad \text{Positive Skew} \\
\text{Power Spectrum} & \quad 10^2 & \quad 10^3 & \quad 10^4 & \quad 10^5 & \quad 10^6 & \quad 10^7 \\
\text{Frequency} & \quad 10^2 & \quad 10^3 & \quad 10^4 & \quad 10^5 & \quad 10^6 & \quad 10^7 \\
\end{align*}\]
Machine learning with Photo aesthetics rules

- **From RVB to Lab color space**

- **Color harmony**
  - Moon & Spencer (1943)
  - Judd & Wiszecki (1967)
### An instance of vector of features (Simond et al. 2015)

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brightness AVG and STD</td>
<td>(1) Average and standard deviation of the brightness, using the V channel in the HSV space.</td>
</tr>
<tr>
<td>Color Variance</td>
<td>(1) Variance of colors in the LAB space.</td>
</tr>
<tr>
<td>Contrast</td>
<td>(1) Width of the middle 96% mass of the histogram of the V channel in the HSV space.</td>
</tr>
<tr>
<td>#Edges and #Edges L, R, T, B, C</td>
<td>(1) We split the canny map into $16 \times 16$ blocks and we compute the number of blocks containing more than 10% of edges. We also compute this number on the left, right, top, bottom and center regions of the image.</td>
</tr>
<tr>
<td>Hue Count</td>
<td>(1) Approximation of the number of unique hues [18].</td>
</tr>
<tr>
<td>Saturation AVG and STD</td>
<td>(1) Average and standard deviation of the saturation.</td>
</tr>
<tr>
<td>Sharpness</td>
<td>(1) Variance of the Laplacian. [20]</td>
</tr>
<tr>
<td>Distance to the Center</td>
<td>(2) Distance of the salient region to the center of the image.</td>
</tr>
<tr>
<td>Rule of Thirds</td>
<td>(2) Shortest distance of the salient region to a power point.</td>
</tr>
<tr>
<td>Salient Hue, Brightness and Saturation</td>
<td>(2) Average hue, brightness and saturation of the salient region.</td>
</tr>
<tr>
<td>Salient Sharpness</td>
<td>(2) Sharpness of the salient region.</td>
</tr>
<tr>
<td>Salient Size</td>
<td>(2) Size of the salient region.</td>
</tr>
<tr>
<td>Salient LOC</td>
<td>(2) We split the image into nine equal parts, and compute the proportion of the salient region in each part. LOC can then take nine values: Top-Left, Middle-Left, Bottom-Right...</td>
</tr>
<tr>
<td>Color Difference</td>
<td>(3) Difference of colors in the LAB space between the salient object and the background.</td>
</tr>
<tr>
<td>Hue, Saturation and Brightness Difference</td>
<td>(3) Difference of hue, saturation and brightness between the salient region and the background.</td>
</tr>
<tr>
<td>Sharpness Difference</td>
<td>(3) Difference of sharpness between the salient region and the background.</td>
</tr>
</tbody>
</table>
Machine learning with Image processing features

- Statistics from histogram & color distribution
- Texture statistics: wavelets, Gabor, Haralick, fractal index
- Detection of edges & contours
- Segmentation and distribution of areas

Machine learning with general vision oriented features

- Characteristic points (key points) : SIFT or SURF, Harris
Datta, Joshi, Wang, Li - 2006-2012
Penn State University

- Generic image features (56 features)
- Input: on 7 levels  
  Output: from 0 to 100
- with SVM as classifier
- Available on-line: ➔ Acquine

http://infolab.stanford.edu/~wangz/project/imsearch/Aesthetics/MIR10/

Lo, Liu, Chen - 2010-2012
Academia Sinica, Taiwan

- 24d features from image processing toolbox: color palette, layout composition, edge composition, textures, contrast, blur, dark channel ...

- Hong Kong University data base: 10,000 images separated in 7 categories: Animal, Plant, Static, Human, Night, Architecture and Landscape.

- Very small pictures (<480 pixels), very fast processing

- SVM classifier

- 2 classes: « High » or « Low »

Use of generic image descriptors for data bases: GIST + SIFT + color descriptors = 920D features in a pyramidal description

Learning with SVM: Bag-of-Visual Words (BOVs) with GMM or Fisher Vectors

Ke, Tang, Jing - 2006
Carnegie Mellon + Microsoft Asia

- Professional quality versus snapshot
- Image processing features
- Bayes classifier: 2 classes
- Images from DPChallenge


Figure 9. Distribution of people’s ratings of professional photos and snapshots. There is significant overlap in the distributions, meaning there is ambiguity in the perceived quality of the photos.
Dhar, Ordonnez, Berg – 2011
Stony Brook University

- **Similar to Ke et al. but with high level features**
  - Image content
  - Image organisation
  - lighting

- **Image features**
  - 17 SVM classifiers in parallel

- Dhar S., Ordonez V., Berg T. - High level describable attributes for predicting aesthetics and interestingness, CVPR 2011, 1657-1654
Schiffanella, Redi, Aiello – 2015
University of Torino + Yahoo

- **Low quality images** from Flickr in categories (animals, urban, nature, people)
- **Training by crowdsourcing** (non expert, heterogeneous, but culturally homogeneous, filtered and cross-validated)
- **Classification using Partial Least Square Regression (PLSR)**
- **47 features** (color (with emotions), spatial arrangement (symmetry, rule of third), textures, (Haralick, entropy, etc.))

## Feature Detection & Classification

<table>
<thead>
<tr>
<th>reference</th>
<th>Data Base</th>
<th>Evaluation result</th>
<th>features</th>
<th>classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Ke et al., 2006] Carnegie Mellon + Microsoft Asia</td>
<td>DPChallenge</td>
<td>2 classes</td>
<td>Image: Generic</td>
<td>Bayes, AdaBoost</td>
</tr>
<tr>
<td>[Datta and Wang, 2010] PennState University</td>
<td>Photo.net</td>
<td>2 classes + mark</td>
<td>Photography theory</td>
<td>SVM</td>
</tr>
<tr>
<td>[Marchesotti et al., 2011] Xerox, Grenoble</td>
<td>Photo.net + CUHK</td>
<td>mark</td>
<td>Vision: generic</td>
<td>PCA + FV + SVM</td>
</tr>
<tr>
<td>[Dhar et al., 2011] Stony Brook Uni.</td>
<td>DPChallenge</td>
<td>2 classes</td>
<td>Semantic – Hi level</td>
<td>SVM</td>
</tr>
<tr>
<td>[Lo et al., 2012] Academia Sinica, Taiwan</td>
<td>CUHK</td>
<td>2 classes</td>
<td>Photography theory</td>
<td>SVM with categories</td>
</tr>
<tr>
<td>[San Pedro et al., 2012] Telefonica, Barcelona - Maryland</td>
<td>DPChallenge</td>
<td>Continuous mark</td>
<td>Crowd-sourcing</td>
<td>SV_eps</td>
</tr>
<tr>
<td>[Lu et al., 2014a] BUPT, Beijing - Raytheon</td>
<td>AVA</td>
<td>2 classes</td>
<td>color HSV or Munsell</td>
<td>LDA + Lasso</td>
</tr>
<tr>
<td>[Schifanella et al., 2015] Uni. Torino - Yahoo</td>
<td>BEAUTY</td>
<td>5 levels</td>
<td>Image: Generic</td>
<td>PLSR</td>
</tr>
</tbody>
</table>
Machine Learning approaches

- Handcrafted Feature Selection + Classifier
- Deep Neural Network

Image → DNN → Class
Machine Learning approaches

- Handcrafted Feature Selection + Classifier
- Deep Neural Network
DNN architecture

- Use a generic DNN as optimized for pattern recognition competitions (ResNet, Photo.net, Google LeNet, VGG, Inception …)
- Train it with a conventional image database (ImageNet, …)
- Add a level of fine tuning with data issued from an aesthetic database
## Main publications using CNN (1/2)

<table>
<thead>
<tr>
<th>reference</th>
<th>Data Base</th>
<th>Evaluation granularity</th>
<th>Middleware</th>
<th>decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Lu et al. 2014]</td>
<td>AVA</td>
<td>2 classes</td>
<td>2 tracks + categories</td>
<td>$2\times$CNN : 50 random sub-images</td>
</tr>
<tr>
<td>Penn State + Adobe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Lu et al. 2015]</td>
<td>AVA</td>
<td>2 classes</td>
<td>6 tracks</td>
<td>$1\times$CNN + 5 random sub-images</td>
</tr>
<tr>
<td>Penn State + Adobe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Kong et al. 2016]</td>
<td>AVA</td>
<td>2 classes</td>
<td>reduced images + categories</td>
<td>Siamese CNN</td>
</tr>
<tr>
<td>Irvine + Adobe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Jin et al. 2016]</td>
<td>AVA</td>
<td>2 classes</td>
<td>reduced images + categories</td>
<td>CNN - ILGNet</td>
</tr>
<tr>
<td>Beijing Uni.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Mai et al., 2016]</td>
<td>AVA</td>
<td>2 classes</td>
<td>pyramid + categories</td>
<td>MNA CNN multitask</td>
</tr>
<tr>
<td>Portland Uni + Adobe</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Schwartz et al. 2016]</td>
<td>Tübingen</td>
<td>ranking</td>
<td>1 track</td>
<td>Siamese &amp; triplets</td>
</tr>
<tr>
<td>Uni. Tübingen</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Wang et al. 2016]</td>
<td>AVA</td>
<td>2 classes</td>
<td>multiples parallel paths</td>
<td>CNN = Brain Inspired</td>
</tr>
<tr>
<td>Uni Illinois</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Kao et al., 2017]</td>
<td>AVA</td>
<td>2 classes</td>
<td>aesthetics + categories</td>
<td>CNN multitask</td>
</tr>
<tr>
<td>Chinese Acad. Sci.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Main publications using CNN (2/2)

<table>
<thead>
<tr>
<th>reference</th>
<th>Data Base</th>
<th>evaluation granularity</th>
<th>Middleware</th>
<th>classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Redi et al., 2017] Bell Labs + Flickr + Yahoo</td>
<td>AVA =&gt; Redi</td>
<td>3 classes</td>
<td>web on-line training</td>
<td>CNN</td>
</tr>
<tr>
<td>[Kairanbay et al. 2017] Malaysia</td>
<td>AVA</td>
<td>2 classes</td>
<td>Global Average Pooling</td>
<td>CNN – GAP AlexNet</td>
</tr>
<tr>
<td>[Ma et al. 2017] SUNY (Buffalo) + Tianjin</td>
<td>AVA</td>
<td>2 classes</td>
<td>Aesthetical criteria</td>
<td>A-Lamp – CNN multitask</td>
</tr>
<tr>
<td>[Murray &amp; Gardo, 2017] Naver Labs Europe</td>
<td>AVA</td>
<td>Mark distribution</td>
<td>categories</td>
<td>ResNet - VGGNet</td>
</tr>
<tr>
<td>[Park et al., 2017] Postech, Corea</td>
<td>AVA + interaction</td>
<td>2 classes</td>
<td>Personnal preference learning</td>
<td>CNN + R-SVM + SVR</td>
</tr>
<tr>
<td>[Talebi &amp; Milanfar 2017] Google Mountain View</td>
<td>AVA + TID2013</td>
<td>Mark distribution</td>
<td>EMD Distance</td>
<td>VGG16-Inception-MobileNet</td>
</tr>
<tr>
<td>[Srivastava &amp; Kant, 2018] ParallelDots</td>
<td>AVA2</td>
<td>2 classes</td>
<td>LAB space</td>
<td>ILGNet</td>
</tr>
<tr>
<td>[Wang et al. 2018] Fudan (Shanghai) + Xi’an</td>
<td>AVA - Reviews</td>
<td>report</td>
<td>Vision to Language LSTM</td>
<td>CNN recurrent = NAIR</td>
</tr>
</tbody>
</table>
Example 1: RAPID system X. Lu et al. - 2014

- 2 classe Classification: « high » and « low »
- 2 track CNN : image 256x256x3 + sub-images 256x256x3 hi-resolution

- « Style » regularisation channel
- Learning : 200 000 images from AVA

<table>
<thead>
<tr>
<th>δ</th>
<th>SCNN</th>
<th>AVG_SCNN</th>
<th>DCNN</th>
<th>RDCNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>66.7%</td>
<td>71.20%</td>
<td>69.91%</td>
<td>73.25%</td>
</tr>
<tr>
<td>1</td>
<td>67%</td>
<td>68.63%</td>
<td>71.26%</td>
<td>73.05%</td>
</tr>
</tbody>
</table>

ACM Int. Conf on MM, Orlando 2014 - pp 457-466

- 5 convolutional layers + 2 fully connected layers
- 1 channel for aesthetic assessment
- 1 channel for semantic assessment

Kao, Y. He, R., Huang, K. IEEE trans on IP, 26(3): 1482-1495
Example 3: « *Brain inspired* », Wang et al. (2016)

Beauty assessment with DNN

- **Results**: good for binary classification > 80%
  - DNN > « Handcrafted+Classification »
  - Expert knowledge of no use

- **Problems remain with**:
  - Image resolution
    - Concatenate mosaïc subprocessing,
    - Parallel processing of sub-images,
    - Pyramidal pooling,
    - Random sub-image high resolution processing
  - Use of semantic metadata
  - Decision over nice/poor images
Subjectivity of aesthetic appraisal
Subjectivity

- Subjectivity is the quality or condition of an individual who possesses conscious experiences, such as perspectives, feelings, beliefs, and desires.
- Subjectivity is an explanation for that which influences, informs, and biases people's judgments about truth or reality.

“Beauty is in the eye of the beholder”, Margaret Wolfe Hungerford (1878)
Subjectivity

- It is the result of many different causes:
  - Long term observer’s personal dispositions: temper or temperament
  - Short term observer state of consciousness: mood
  - Long term cultural & social context of life (unconscious), personal experience
  - Education and training (volunteer)
  - Context of the assessment experimentation

- The role of these factors in aesthetics appraisal is ascertained but their exact implication is not known

➤ Many different ways to involve subjectivity in the task of Beauty Assessment
I - Recommendation systems

- **Popular for films, books, series ...** « you appreciated xxx, you will love yyy ... » and commercial prospection

- **Method:**
  - Collect a large series of opinions from a large number of customers
  - Collect some opinions from the user of concern
  - Deduce the user’s opinion on the object of concern from the collection of opinions using various strategies

  (Dedjoo et al. 2016, 2018, Elahi et al. 2017)

- **Not well adapted to aesthetic assessment**
  - Recommendations work well with hidden **semantic features**: actors, authors, genre,

- **Exception:** Art galleries
  - But aesthetics has only a minor role in photo selection versus artist’s name, price and previous acquisitions from the user

  (Benouaret 2017, Massina et al. 2017, 2018)
  BAM! Behance artistic media dataset (Behance)
II – Subjectivity from a social profile

- Collect any information on the user, available from online social networks
- Make use of a specific database AVA-PD (AVA Photographer Demographic)
  - Subset of AVA with the only photographers the social profiles of which are known (age, gender, profession, country of life …)
  - Proceed to an aesthetic assessment with a CNN training on AVA + a fine tuning on the only photographers of AVA-PD the profile of which are close from the user’s one.
  - Submit the unknown photo to the trained CNN

  \(\text{(Kairanbay et al., 2019)}\)

Alternative:
- Make use of photos posted by the user on the web (Instagram, Flickr, Pinterest, …) to categorize users
  \(\text{(Lovato et al., 2013, Yang et al. 2015, You et al, 2016)}\)
II – Social classes from posted images online

- Pinterest User A
  Board: Travel

- Pinterest User B
  Board: Travel

- Pinterest User C
  Board: Travel

- Pinterest User D
  Board: Travel

Yang, Hsieh & Estrin, 2015
II – Social classes from posted images on line

Yang, Hsieh & Estrin, 2015
III - Subjectivity from a psychological profile

- Components of psychological profile: the Big Five (L. Goldberg, 1990)
  - Openness (O)
  - Conscientiousness (C)
  - Extraversion (E)
  - Agreeableness (A)
  - Neuroticism (N)

- When using a media to determine a profile ➔ 2 different profiles
  - The producer psychological profile vs. the receiver image of the producer’s profile ➔ Brunswik lens (E. Brunswik, 1956)

M. Cristani et al., 2013
Measure of Big Five using Brunswik lens model

Ground truth = PsychoFlickr database:
- 60,000 images
- 300 photographers
- Emitter profile
- Receiver profile

Figure 2: The picture shows the Brunswik Lens model for the PsychoFlickr dataset, where the state corresponds to the Big Five traits (as per assessed with the BFI-10). Ecological and Representation validities are measured with the Spearman Coefficient and the picture shows (for each trait) features for which both values are statistically significant ($p < 5\%$).

M. Cristani et al., 2013
Learning Big Five

- With handcrafted features and classification: M. Cristani et al., 2013
- With CNN: C. Segalin et al., 2017

Conscientiousness = sens of responsability
C. Segalin et al., 2017
Use of Big Five for Aesthetic appraisal of photos

Fig. 3. The framework of the proposed personality-assisted multi-task learning for IAA.

Make use of images posted by the user

Li, Zhu & Zhao, 2020
Use of Big Five for Aesthetic appraisal of photos

Li, Zhu & Zhao, 2020

<table>
<thead>
<tr>
<th>Personality traits</th>
<th>O</th>
<th>C</th>
<th>E</th>
<th>A</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Openness</td>
<td>0.19</td>
<td>0.28</td>
<td>0.25</td>
<td>-0.04</td>
<td>-0.17</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extraversion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agreeableness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nevrotism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Openness O
Conscientiousness C
Extraversion E
Agreeableness A
Nevrotism N
IV – Learning aesthetical tastes with tests

Submit some images to the user’s judgement
- How many? 10? 100? No clear answer
- Which judgement? Ranked? Nice/Poor? Mark?
- Fine-tune a CNN with user’s preferences

Either: Compare the user to one of some prototype photographers previously used to train a CNN (for instance with AVA Photographer Demographic) (O’Donovan et al., 2014; Zhu et al., 2020)

Or: Compare the user’s choices to yet classified images and select from the ground truth a subset to specify a personal CNN (reinforcement technique) (Ren et al., 2017; Lv et al., 2018)
Learning user’s preference through reinforcement loop

The reinforcement loop is used to detect in the learning database those which are closed from the user’s posted images.

The USAD transfers the learning to the whole database.

Lv et al., 2018
V – Classify the user’s population


- from both pictorial features and litterary features from the caption.

Makes use of « bag of words »:
- 1 image = 1 « word »
- 1 photograph = 1 « sentence »
- from sentences are extracted « topics » which represent « communities »

Hong et al. 2016
VI Multiply expertises

- Use experts to specialize a network (Zhu et al 2020)
  - How many experts?
  - How chosen?
Critics of the recurse to AI
The limits of the database: AVA’s best
Limits of the database: AVA’s worst
Limits of the database  AVA’ average  B=5.43
Questioning the expertise

- Where are they from?
- Which training?
- How is exploited the diversity of judgements?
Conclusions
Conclusions

- Conventional AI based on:
  - Very large data base
  - Uncriticized & universal ground truth

- Efficient, easy to use and flexible
- Provides rather good results for Photo Beauty Assessment

- Adaptation for more personal needs is very difficult
  - Recommendation techniques are not adapted
  - Psychosocial profiles may be determined, but the relation with aesthetic assessment is loose
  - Individual aesthetic tastes are hard to define and to measure

- Progress are urgently needed to match automatic judgement and personal tastes