From data-mining to social networks.

Subjectivity and AI

Cours TPT-DATAAI-903
M2 AI (Artificial Intelligence) - Université Paris Saclay
Master CPS (Cyber Physical Systems) - IPP
Master DATA AI - IPP

Henri Maître - Télécom Paris
henri.maitre@telecom-paris.fr

October 2021

Objectives of this course

- 1 - To present some limits of Artificial Intelligence (AI) based on deep neural networks (DNN),
- 2 - To emphasize the role of subjectivity in the consumer’s demand,
- To present some tracks to conciliate 1) and 2) based on tools developed for social networks managing.

- The course takes aesthetic assessment of photos as experimental field.

- It lights up the place of individual criteria in human perception of images

--/---

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**

These slides are extended in the document “assessment-of-beauty.pdf**”, translation of “**Juger du Beau avec subjectivité : le défi de l’esthétique computationnelle”, H. Maître, ISTE OpenScience, 2020. All references in the slides may be found there

About Beauty of Photos

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
  - Galeries, 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 not be related at all to the observed object
  - If « rules » have to be found, they ly in the consciousness of the observer.

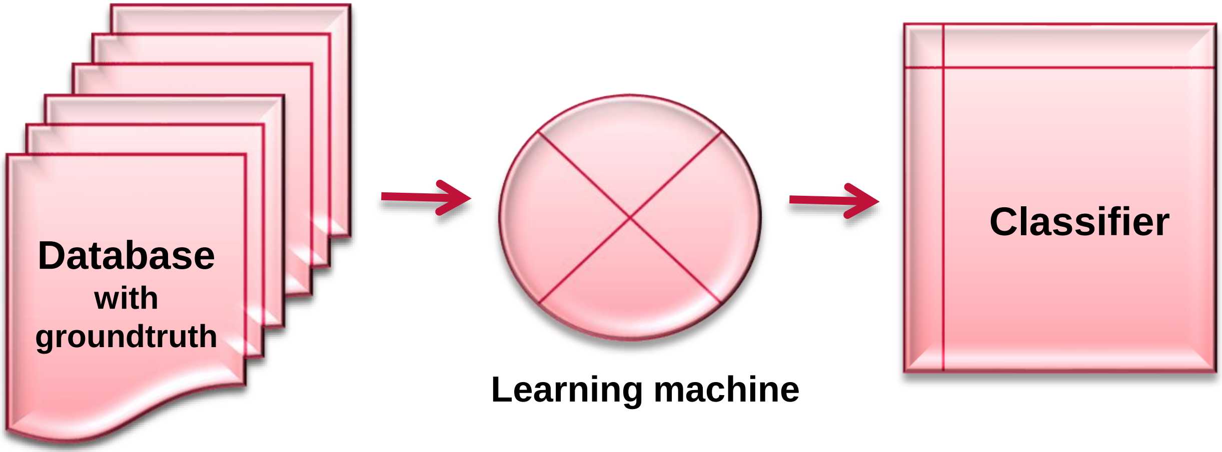
Instances: Romantism,
Impressionism,
Surrealism
Pop Art …

Approaches based on objectivist approach: Machine Learning & Deep Neural Networks

Automatic Aesthetic Assessment

- **Machine Learning approaches: 2 steps**

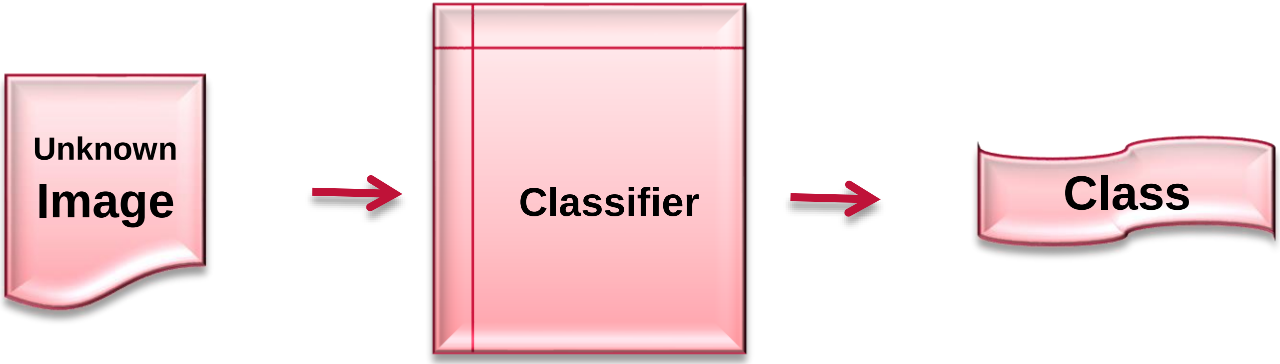
Step 1: Training stage



Automatic Aesthetic Assessment

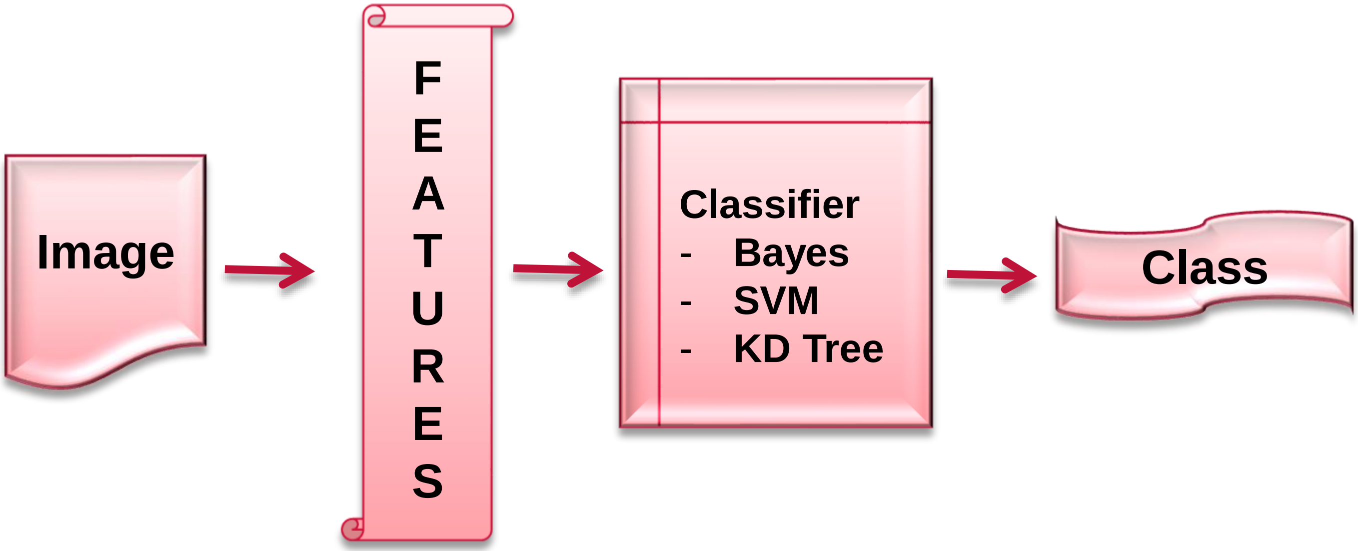
- **Machine Learning approaches: 2 steps**

Step 2: Decision stage



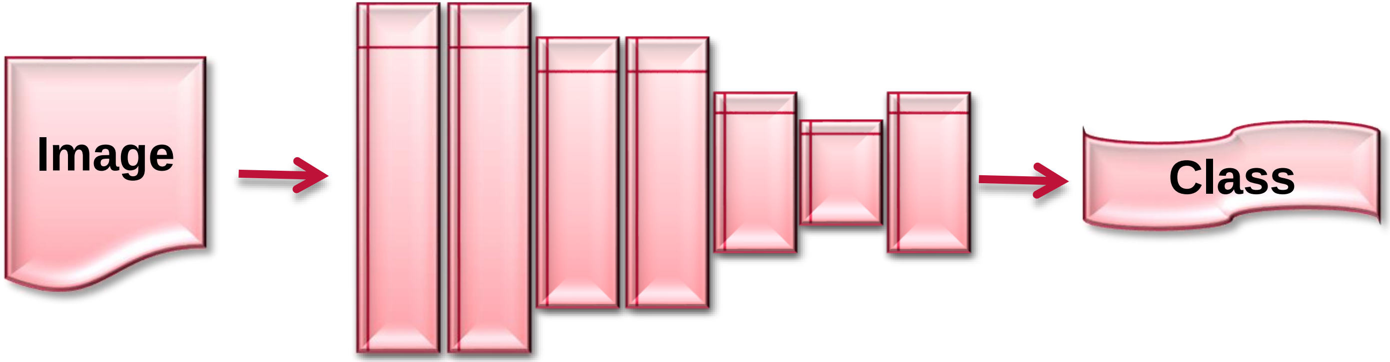
Machine Learning approaches

- **Handcrafted Feature Selection + Classifier**
- Deep Neural Network



Machine Learning approaches

- Handcrafted Feature Selection + Classifier
- **Deep Neural Network**



Trained DNN

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 »

Data Bases for Machine Learning & CNN

criteria	BEAUTY	AADB	Redi's base	AVA	AVA-2 CUHK-DB	AVA-PD	Uni Tübingen	Psycho Flickr	Flickr AES
size ($\times$1000)	15	10	100	250	50	120	380	60	40
aesthetical quality	weak	weak	weak	high	high	high	weak	high	weak
aesthetic mark	3	5	4	10	10 / 2	10	10	5	5
semantic classes	4	no	yes	44	44	44	no	yes	yes
style labels	no	11	no	14	14	14	no	no	no
annotation	no	yes	yes	yes	yes	yes	no	no	yes
origine	Flickr	Flickr	web	DP Chal.	AVA	AVA	Flickr	Flickr	Flickr

- 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,
  - Used Metadata : semantic classes, style, comments
  - ➔ **AVA = Aesthetic Visual Analysis**

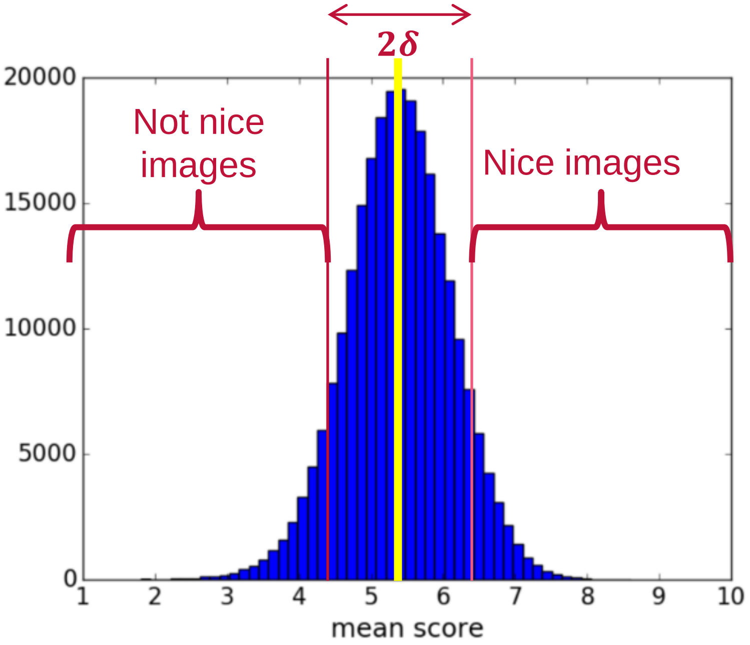
AVA data base: Aesthetic Visual Analysis

- 250 000 photos issued from *DPChallenge* (Digital Photography Challenge)

  https://www.dpchallenge.com/
- Good quality: almost professional
- 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: tittle of the challenge
  - « textures & materials », « self-portrait », « spring season », « upward & angled », …

AVA : https://computervisiononline.com/dataset/1105138637

Learning stage: Which images are beautiful?

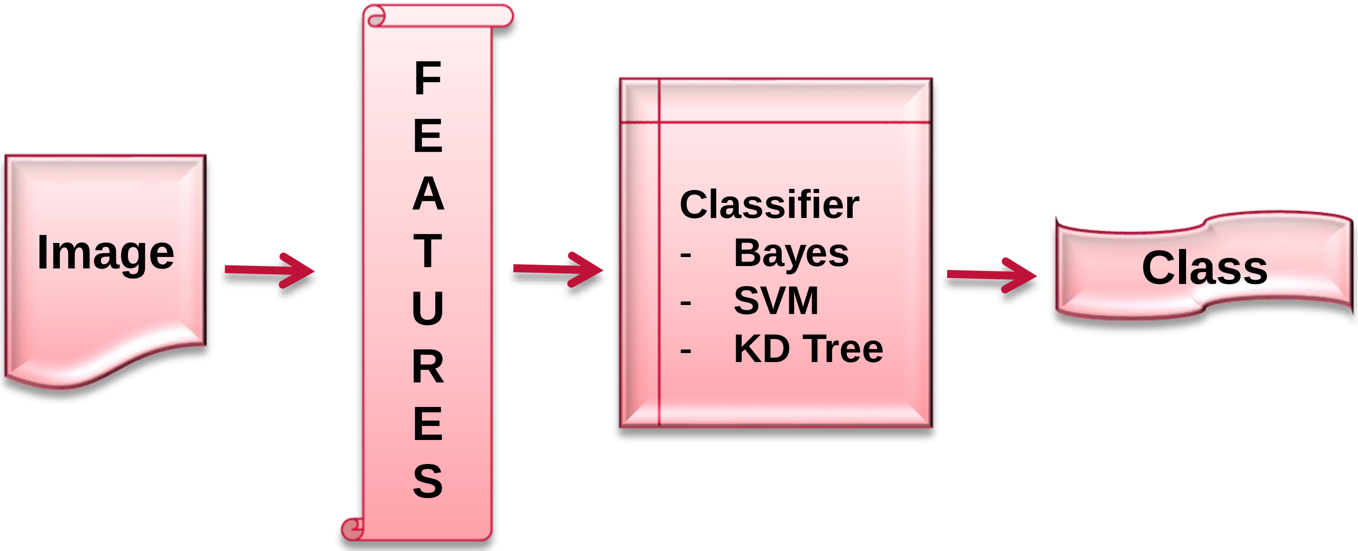


Mark distribution of the 250 000 images from AVA database

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

Machine Learning approaches

- **Handcrafted Feature Selection + Classifier**
- Deep Neural Network



Machine learning with Photo aesthetics rules (1/4)

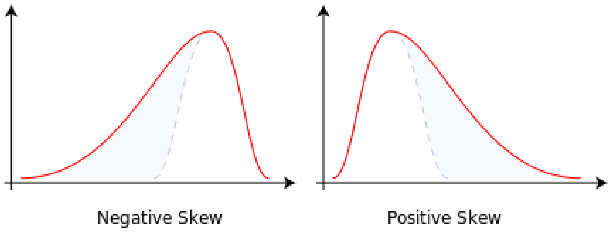
- **Format or aspect ratio**
  - **Commercial Formats : 3/2, 4/3, 5/3, 16/9, A, B, …**
    Markowsky (1992) ➔ **1.83**
  - **Rule of Third:** *S.E Amirashi et al. (2014)*
  - **Golden Number:** $\varphi = \frac{1+\sqrt{5}}{2}$ ~ 1.618 *Markowsky (1992)*
  - **Weight of Center**: *Arnheim (1983)*

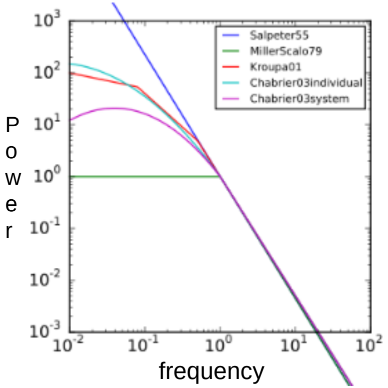
Machine learning with Photo aesthetics rules (2/4)

- **Composition rules**
  - **Weight of the center**
  - **Importance of orders** (regular, progressive)
  - **Symetries**
  - **Alignements, diagonal, perspective**

- **Role of attention points**
  - **Objet versus background**
    - Position
    - Focus
    - Contrast

Machine learning with Photo aesthetics rules (3/4)





- Image : $I(x,y)$
- Grey level histogram: proba(I)
  - Any Mean or Variance: $<I>$ $<(i-<i>)^2>$
  - Positively biased skew (order 3)

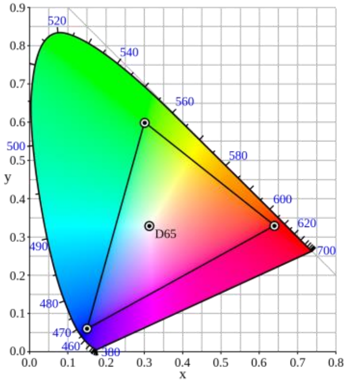
Attewell & Baddeley, 2007

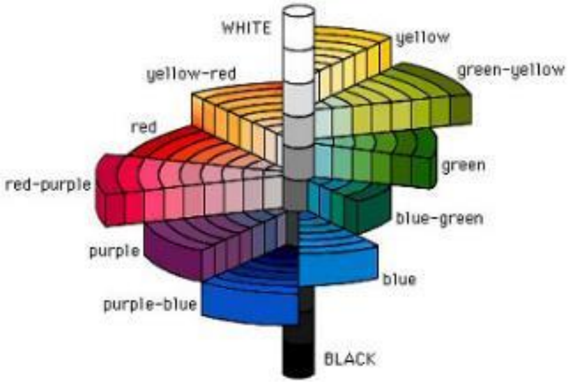
- Power spectrum (after Fourier transform):
  - $P_f(I) = |TF(I)|^2$
  - Radial Symmetry
  - $1/f^2$ power spectrum decay

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

Machine learning with Photo aesthetics rules (4/4)

- **From RVB to Lab color space**





- **Color harmony**
  - *Moon & Spencer (1943)*
  - *Judd & Wiszecki (1967)*
  - *Matsuda (1995)*

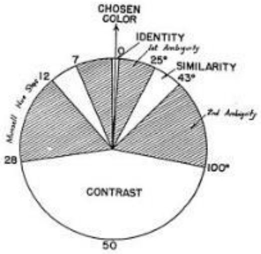


FIG. 2. Regions of similarity and contrast in a plane z=const. (constant Munsell value).

An instance of vector of features (*Simond et al. 2015*)

Name	Description
Brightness AVG and STD	(1) Average and standard deviation of the brightness, using the V channel in the HSV space.
Color Variance	(1) Variance of colors in the LAB space.
Contrast	(1) Width of the middle 96% mass of the histogram of the V channel in the HSV space.
#Edges and #Edges L, R, T, B, C	(1) We split the canny map into 16×16 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.
Hue Count	(1) Approximation of the number of unique hues [18].
Saturation AVG and STD	(1) Average and standard deviation of the saturation.
Sharpness	(1) Variance of the Laplacian. [20]
Distance to the Center	(2) Distance of the salient region to the center of the image.
Rule of Thirds	(2) Shortest distance of the salient region to a power point.
Salient Hue, Brightness and Saturation	(2) Average hue, brightness and saturation of the salient region.
Salient Sharpness	(2) Sharpness of the salient region.
Salient Size	(2) Size of the salient region.
Salient LOC	(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...
Color Difference	(3) Difference of colors in the LAB space between the salient object and the background.
Hue, Saturation and Brightness Difference	(3) Difference of hue, saturation and brightness between the salient region and the background.
Sharpness Difference	(3) Difference of sharpness between the salient region and the background.

An instance of vector of features (*Simond et al. 2015*)

Luminosité	moyenne et variance du canal V d'une représentation HSV
Variance de la couleur	Variance des couleurs dans l'espace Lab
Contraste	Largeur de l'intervalle central à 96\ % du canal V de HSV
nombre de Contours	L'image est divisée en blocs 16x16. Nombre de blocs contenant plus de 10 \% de contours
nombre de Contours L R T C B	Nombres de contours dans les blocs à gauche, à droite, en haut au centre et en bas
nombre de Teintes	Nombre de teintes uniques
Saturation	Moyenne et variance de la saturation
Netteté	Variance du laplacien
Distance au centre	Distance du centre de la zone S au centre de l'image
Règle du tiers	Plus courte distance du centre de la zone S à un axe de tiers
Couleurs de la zone S	Teinte, luminosité et saturation moyennes de la zone S
Netteté de la zone S	Variance du laplacien de la zone S
Taille de la zone S	Nombre de pixels de la zone S
Localisation de la zone S	L'image étant divisée en 9, proportions de la zone S dans chaque zone
Différences de couleur	Différence des couleurs dans l'espace Lab entre la zone S et le fond
Différences HSV	Différences teinte, luminosité et saturation entre la zone S et le fond
focalisation sur l'objet	Différence de netteté entre la zone S et le fond

Machine learning with Image processing features

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

Machine learning with general vision oriented features

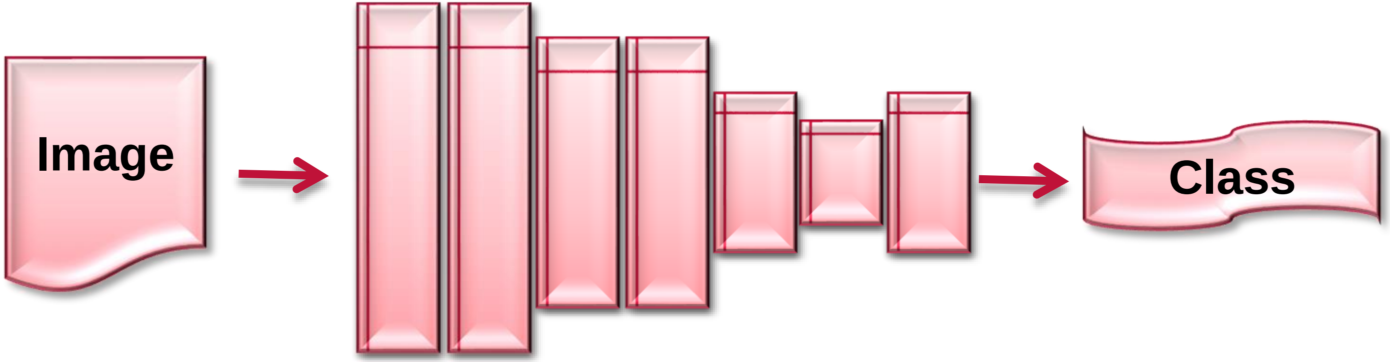
- Characteristic points (points of interest) : SIFT or SURF

Feature Detection & Classification

reference	Data Base	Evaluation result	features	classifier
[Ke et al., 2006] Carnegie Mellon + Microsoft Asia	DPChallenge	2 classes	Image : Generic	Bayes, AdaBoost
[Luo and Tang, 2008] Chinese Uni. Hong-Kong	DPChallenge	2 classes	Photography theory	Bayes,SVM, AdaBoost
[Datta and Wang, 2010] PennState University	Photo.net	2 classes + mark	Photography theory	SVM
[Marchesotti et al., 2011] Xerox, Grenoble	Photo.net + CUHK	mark	Vision : generiic	PCA + FV + SVM
[Dhar et al., 2011] Stony Brook Uni.	DPChallenge	2 classes	Semantic – Hi level	SVM
[Lo et al., 2012] Academia Sinica, Taiwan	CUHK	2 classes	Photography theory	SVM with categories
[San Pedro et al., 2012] Telefonica, Barcelona - Maryland	DPChallenge	Continuous mark	Crowd-sourcing	SV_eps
[Lu et al., 2014a] BUPT, Beijing - Raytheon	AVA	2 classes	color HSV or Munsell	LDA + Lasso
[Schifanella et al., 2015] Uni. Torino - Yahoo	BEAUTY	5 levels	Image : Generic	PLSR

Machine Learning approaches

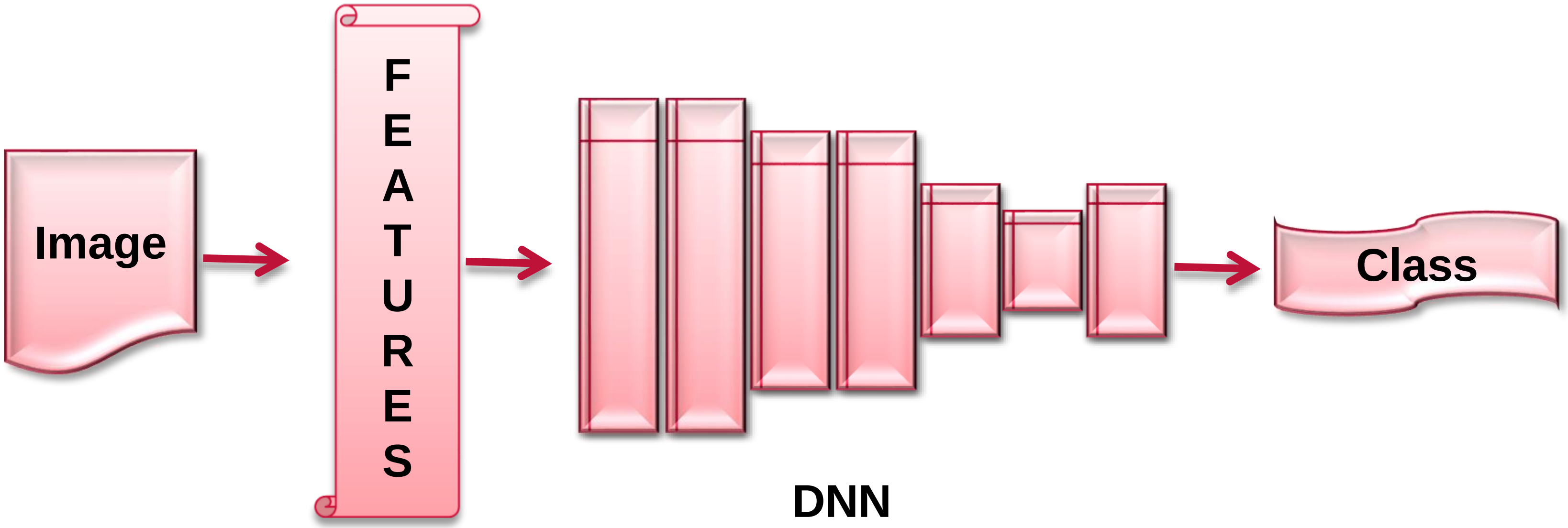
- Handcrafted Feature Selection + Classifier
- **Deep Neural Network**



DNN

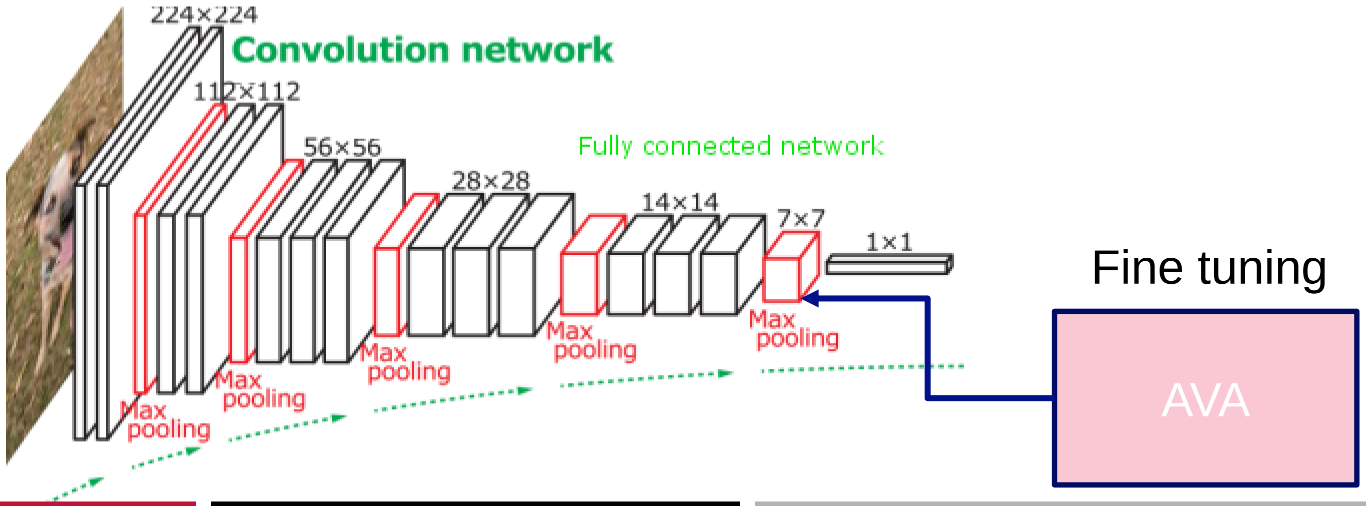
Machine Learning approaches

- 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 it with ordinary databases?
- Add a last level of fine tuning with data issued from an aesthetic database



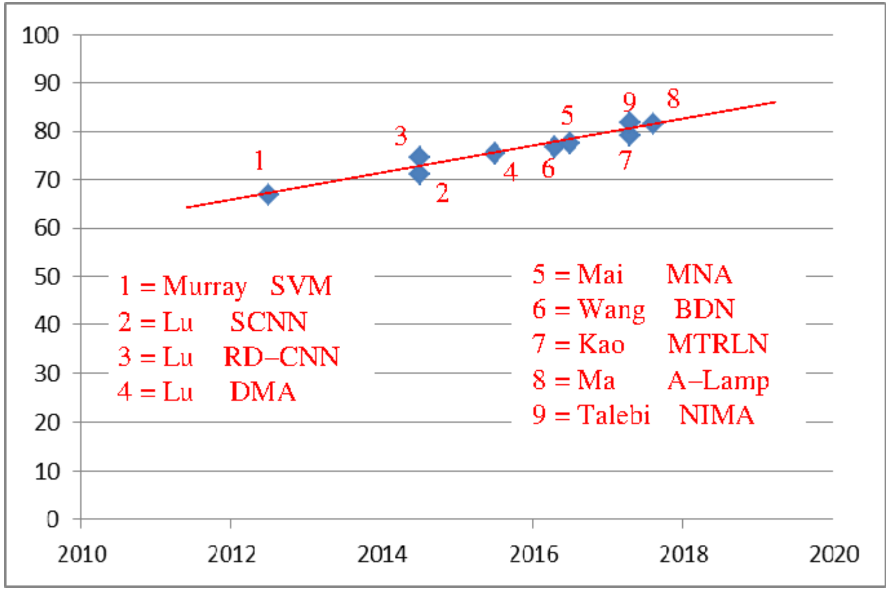
Main publications using CNN (1/2)

reference	Data Base	Evaluation granularity	Middleware	decision
[Lu et al. 2014] Penn State + Adobe	AVA	2 classes	2 tracks + categories	2xCNN : 50 random sub-images
[Lu et al. 2015] Penn State + Adobe	AADB + AVA	2 classes	6 tracks	1CNN + 5 random sub-images
[Kong et al. 2016] Irvine + Adobe	AADB + AVA	2 classes	reduced images + categories	Siamese CNN
[Jin et al. 2016] Beijing Uni.	AVA	2 classes	reduced images + categories	CNN - ILGNet
[Mai et al., 2016] Portland Uni + Adobe	AVA	2 classes	pyramid + categories	MNA CNN multitask
[Schwartz et al. 2016] Uni. Tübingen	Tübingen	ranking	1 track	Siamese & triplets
[Wang et al. 2016] Uni Illinois	AVA	2 classes	multiples parallel paths	CNN = Brain Inspired
[Kao et al., 2017] Chinese Acad. Sci.	AVA	2 classes	aesthetics + categories	CNN multitask

Main publications using CNN (2/2)

reference	Data Base	evaluation granularity	Middleware	classifier
[Redi et al., 2017] Bell Labs + Flickr + Yahoo	AVA ==> Redi	3 classes	web on-line training	CNN
[Kairanbay et al. 2017] Malaisia	AVA	2 classes	Global Average Pooling	CNN – GAP AlexNet
[Ma et al. 2017] SUNY (Buffalo) + Tianjin	AVA	2 classes	Aesthetical criteria	A-Lamp – CNN multitask
[Murray & Gardo, 2017] Naver Labs Europe	AVA	Mark distribution	categories	ResNet - VGGNet
[Park et al., 2017] Postech, Corea	AVA + interaction	2 classes	Personnal preference learning	CNN + R-SVM + SVR
[Talebi & Milanfa,r 2017] Google Mountain View	AVA + TID2013	Mark distribution	EMD Distance	VGG16-Inception-MobileNet
[Srivastava & Kant, 2018] ParallelDots	AVA2	2 classes	LAB space	ILGNet
[Wang et al. 2018] Fudan (Shanghaï) + Xi'an	AVA - Reviews	report	Vision to Language LSTM	CNN recurent = NAIR

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 is not enough known today

I - Recommendation systems

- **Popular for films, books, series …** *« you appreciated xxx, you will love yyy … »*
- **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**

- **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 rôle in photo selection versus artist's name, price and previous acquisitions from the user

II – Subjectivity from a social profile

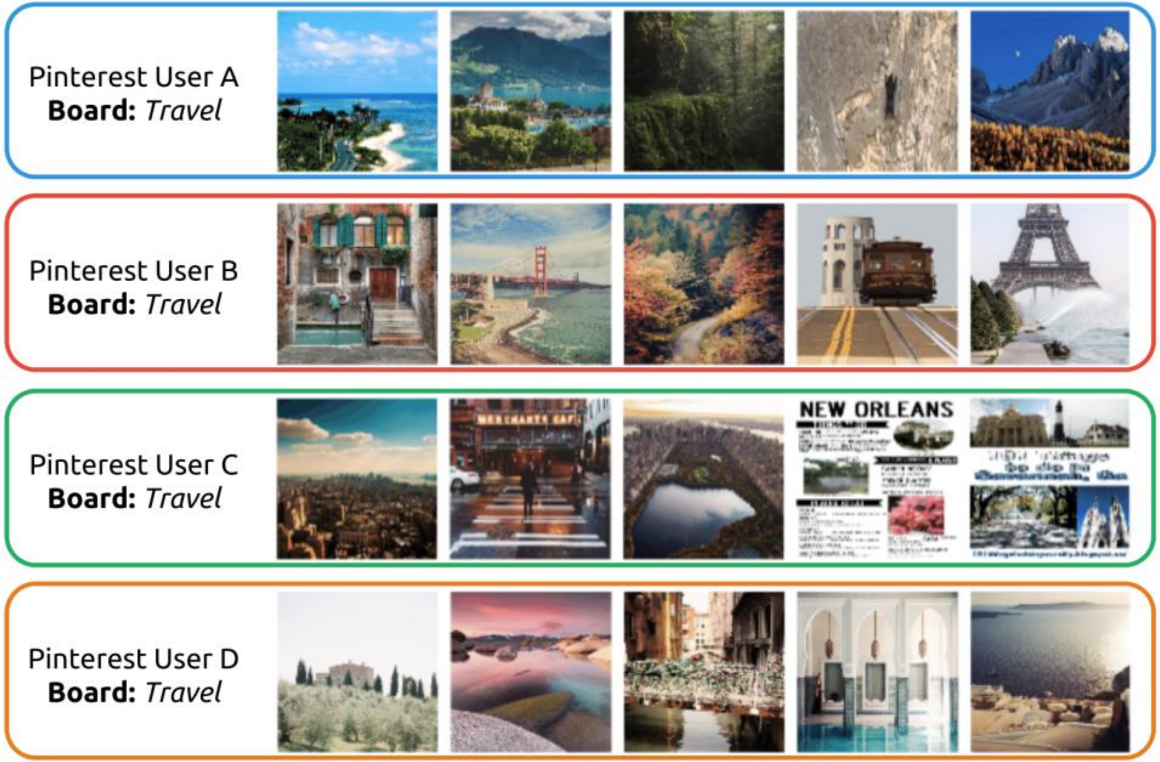
- **Collect any information on the user available fro, on line social networks**
- **Make use of a specific database AVA-PD** *(AVA Photographer Demographic)*
  - Subset of AVA with the only photographers the social profiles of which are known (age, gender, profession, country of life …)
  - Proceed to an aesthetical 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

(Kairanbay et al., 2019)

Alternative:

- **Make use of photos posted by the user on the web** *(Instagram, Flickr, Pinterest, …)* to categorize users *(Lovato et al., 2013, Yang et al. 2015, You et al, 2016)*

II – Social classes from posted images online



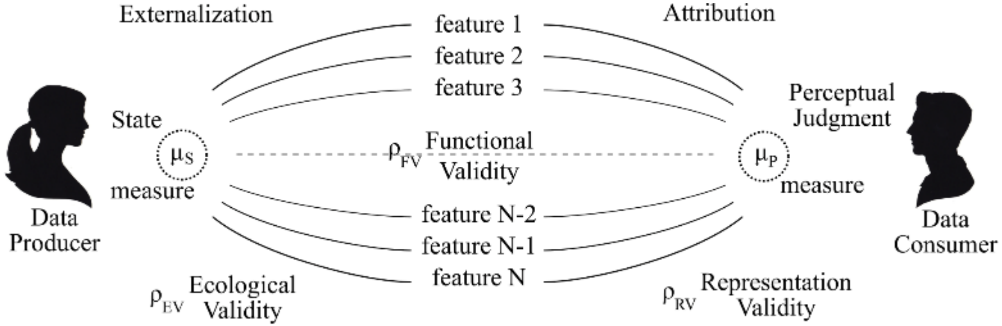
Yang, Hsieh & Estrin, 2015

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**
  - Agreableness **A**
  - Nevrotism **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

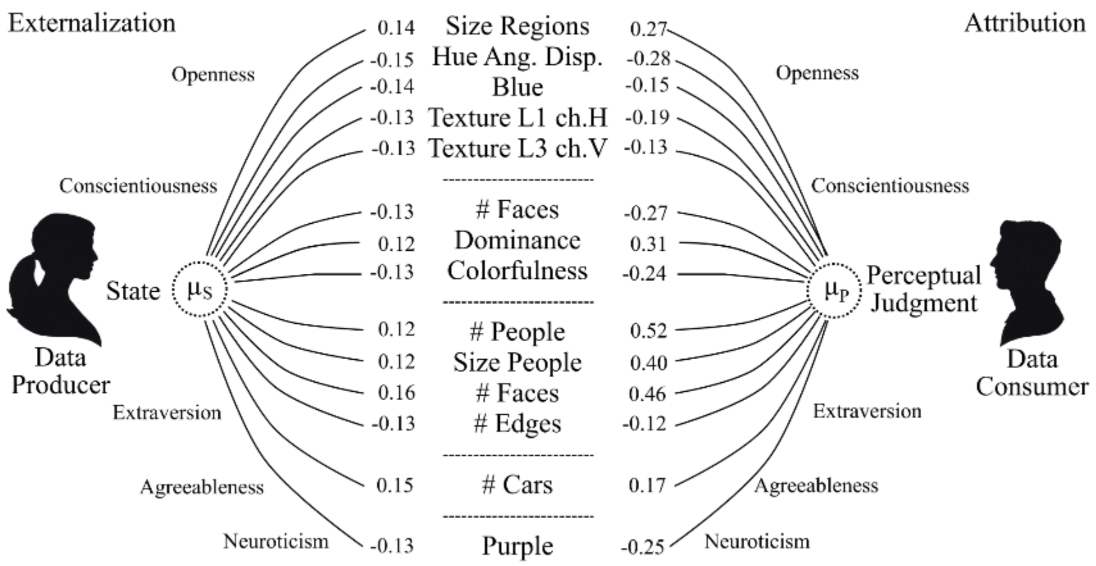


Figure 2: The picture shows the Brunswik Lens model for the PsychoFlickr dataset, where the state corresponds to the Big Five traits (as per assessed with the BFI-10). Ecological and Representation validities are measured with the Spearman Coefficient and the picture shows (for each trait) features for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

Ground truth = PsychoFlickr database :

- 60 000 images
- 300 photographers
- Emitter profile
- Receiver profile

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

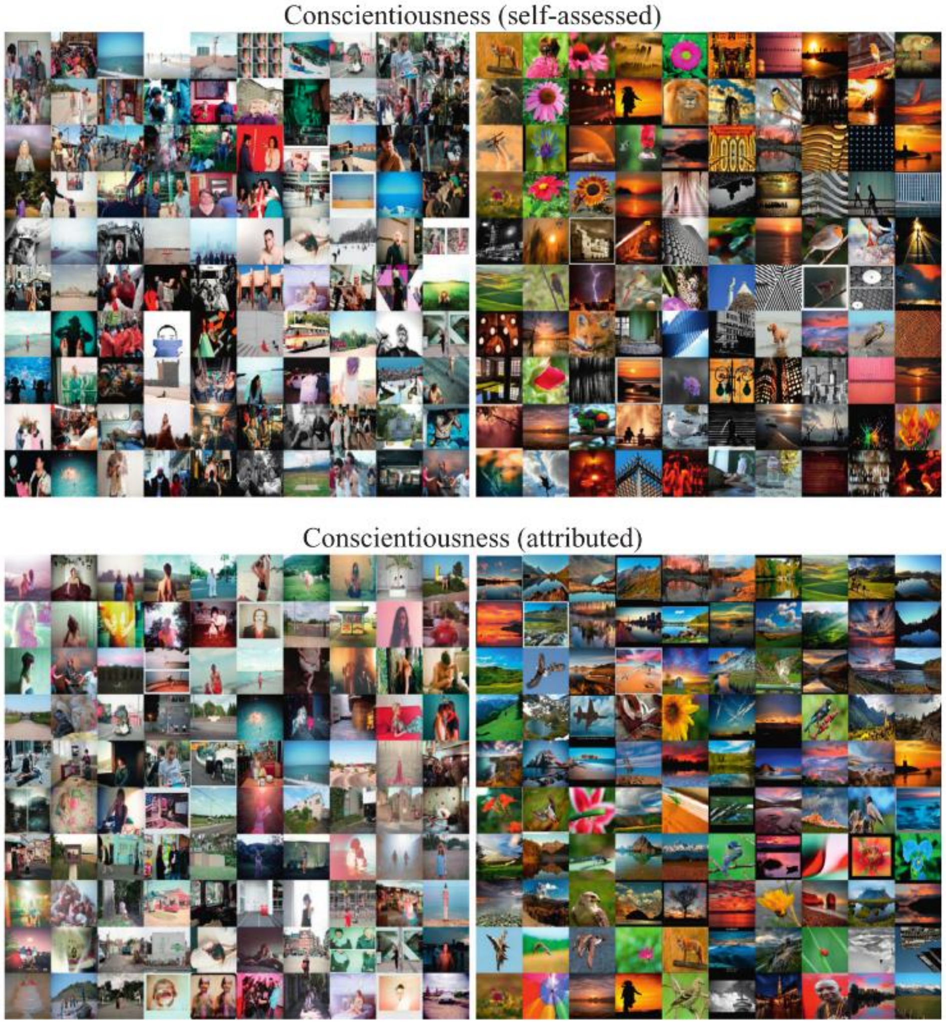


Fig. 7. Representative images in the low (left) and high (right) value classes for the Consciousness trait.

Use of Big Five for Aesthetic appraisal of photos

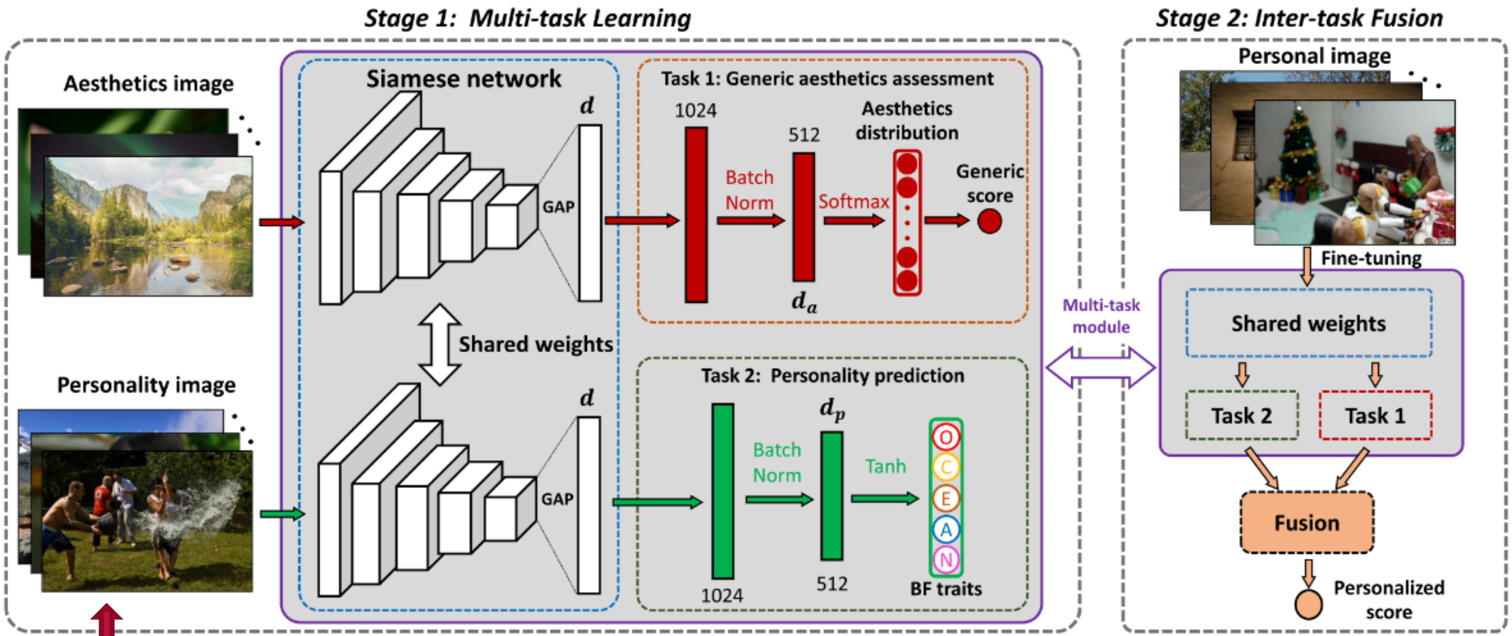


Fig. 3. The framework of the proposed personality-assisted multi-task learning for IAA.

Make use of images posted by the user

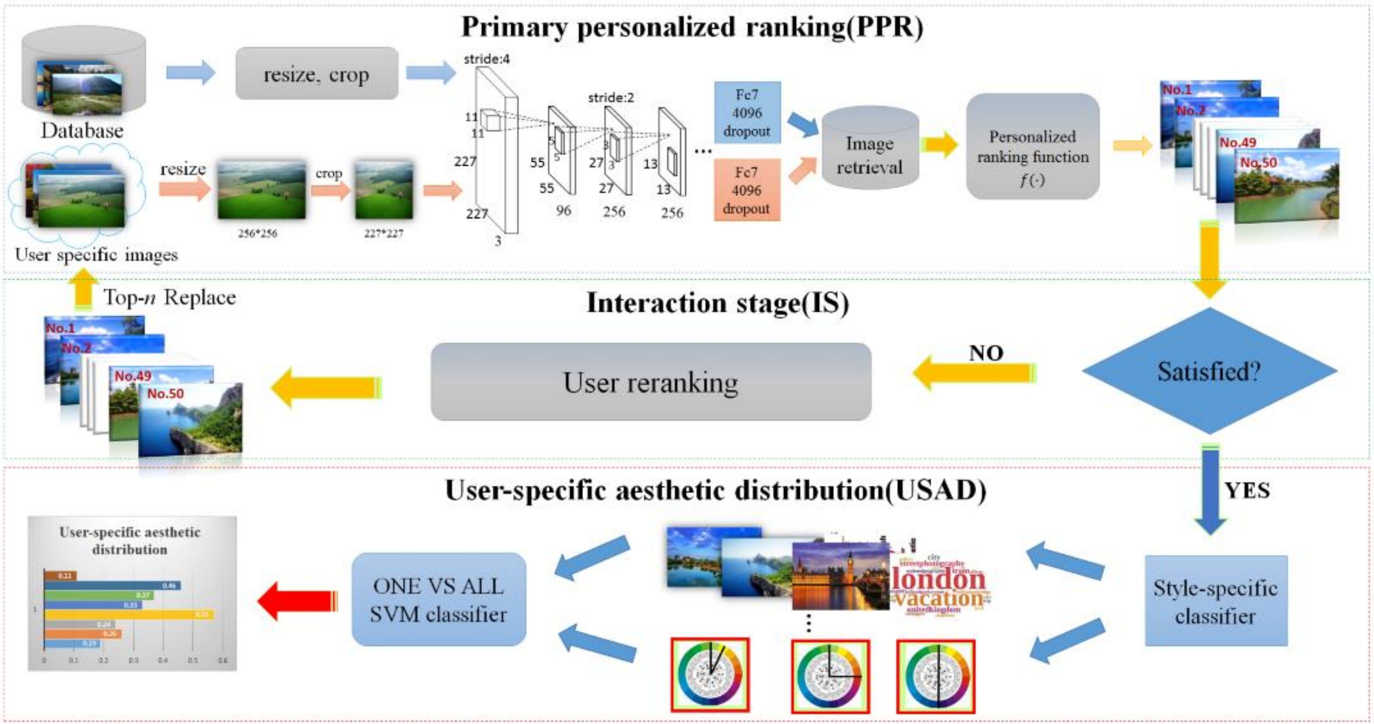
Li, Zhu & Zhao, 2020

IV – Learning aesthetical tastes with tests

- **Submit some images to the user's judgement**
  - **How many ? 10 ? 100 ? No clear answer**
  - **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** (reinforment technique) *(Ren et al., 2017; Lv et al., 2018)*

Learning user’s preference through reinforcement loop



Lv et al., 2018

Conclusions

Conclusions

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

Is efficient, easy to use and flexible

It 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 aesthetic assessment is loose**
  - **Aesthetic tastes are hard to define and to measure**

- **Progress are needed**