From data-mining to social networks.

Aesthetic assessment of photos

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

December 2023

Objectives of this course

- 1 – We will see that aesthetic 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 for AI, 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 on-line 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**

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

Connect on
Wooclap.com
With **PYRFOD** as password

About Beauty of Photos

©Flicker


Dan Kitwood via Getty Images

©Yawar Nazir

©Pixabay

©Shah Marai

Where does the beauty of images come from?

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**
- **For this task, Beauty is competing with:**
  - **Relevance (or interestingness)**
  - **Amazingness (or surprise)**

Beauty

- Beauty is a major reason to select an image in a large set
- For this task, 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)**
  - **Relevance theory of J.L. Dessalles (2010)**

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
  - **Relevance theory of J.L. Dessalles (2010)**
    - 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)

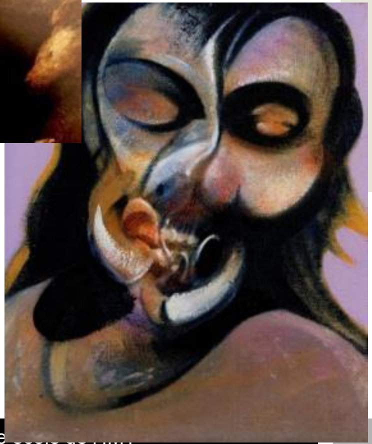
With modern definitions of art, it is possible to decide about:

- **Obviously non-aesthetical representations:**

J.S. Chardin

F. Goya

F. Bacon



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 (GAFAM, …)**
  - **Start-ups and companies**

- Beauty assessment is also enbedded in some cameras and proposed as tool-kits in some Softwares

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: Objectivist vs. Subjectivist

- **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 we have to look for rules, they are to be found in the consciousness of the observer.

Instances: **Romantism,**
Impressionism,
Surrealism
Pop Art …

Objectivist approach

Approaches based on objectivist approach:

1) Machine Learning + Classification

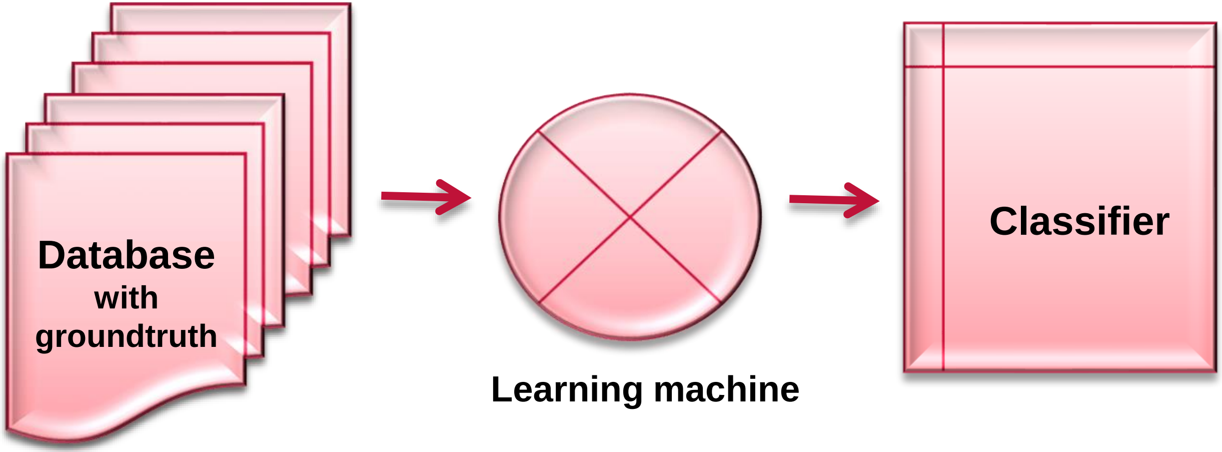
&

2) Deep Neural Networks

Automatic Aesthetic Assessment

- Machine Learning approaches: 2 steps

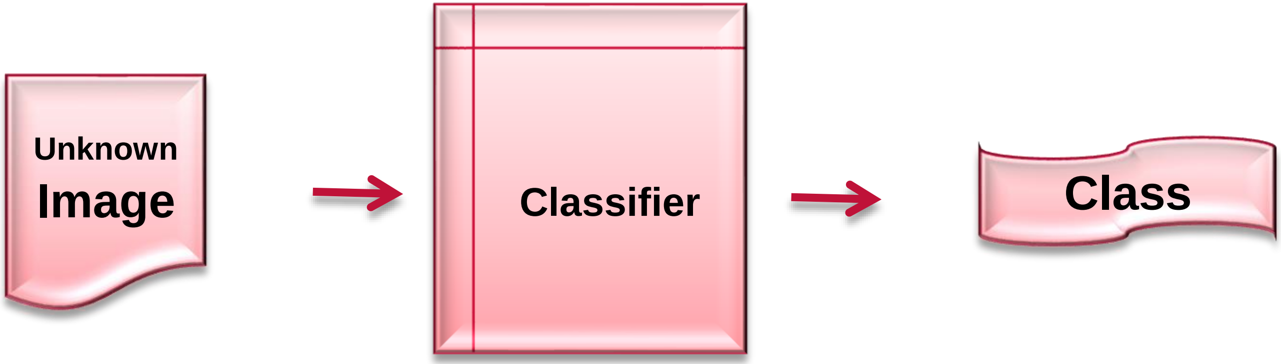
Step 1: Training stage = to build the « machine »



Automatic Aesthetic Assessment

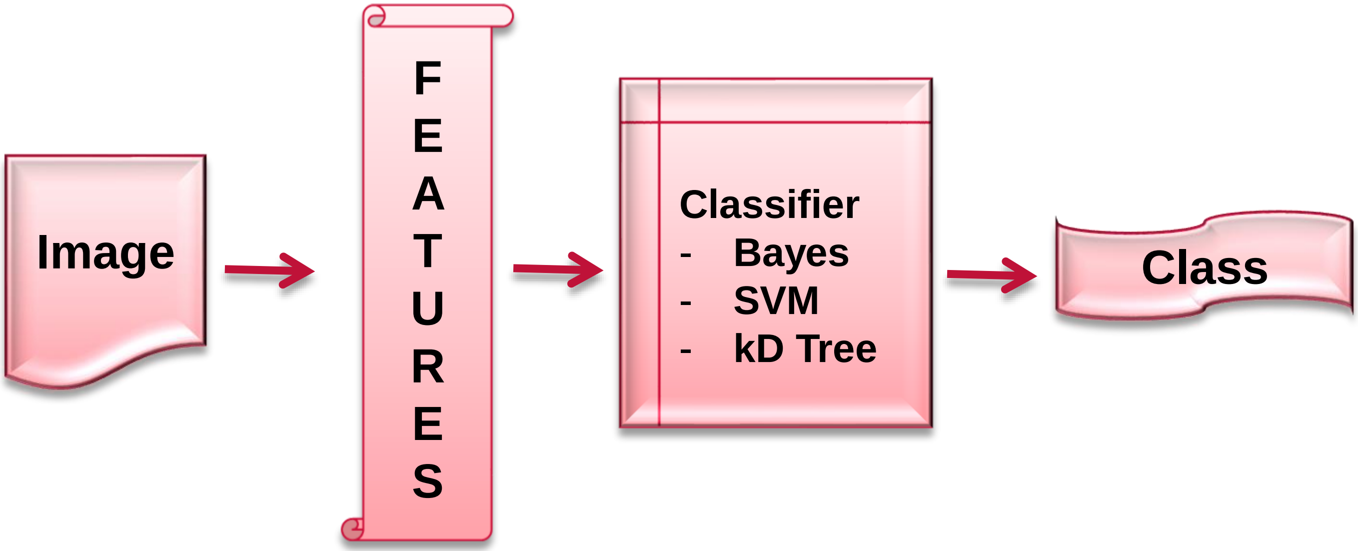
- **Machine Learning approaches: 2 steps**

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



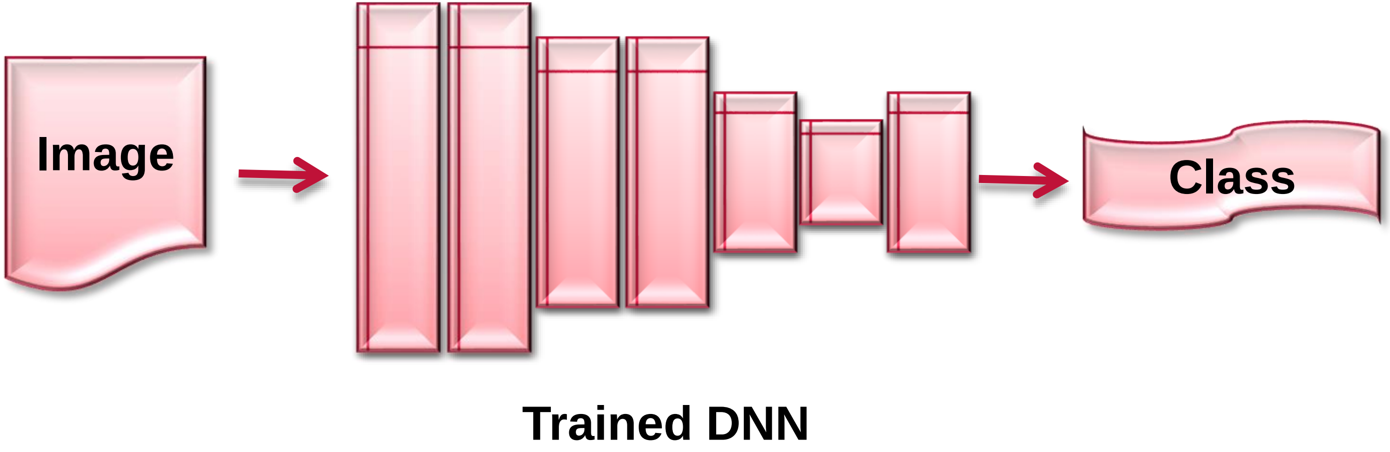
Machine Learning approaches

- **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 »

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,
  - Metadata : semantic classes (portraits, landscape …), style (B&W, abstract, night/flash, effects, …), comments (litteral appraisal)

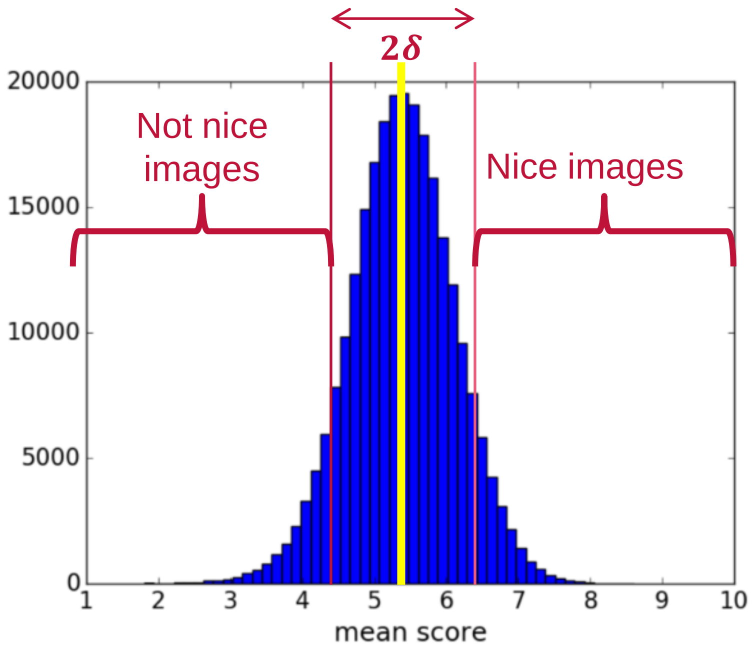
  - ➔ **AVA = Aesthetic Visual Analysis**

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 portrait. 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?



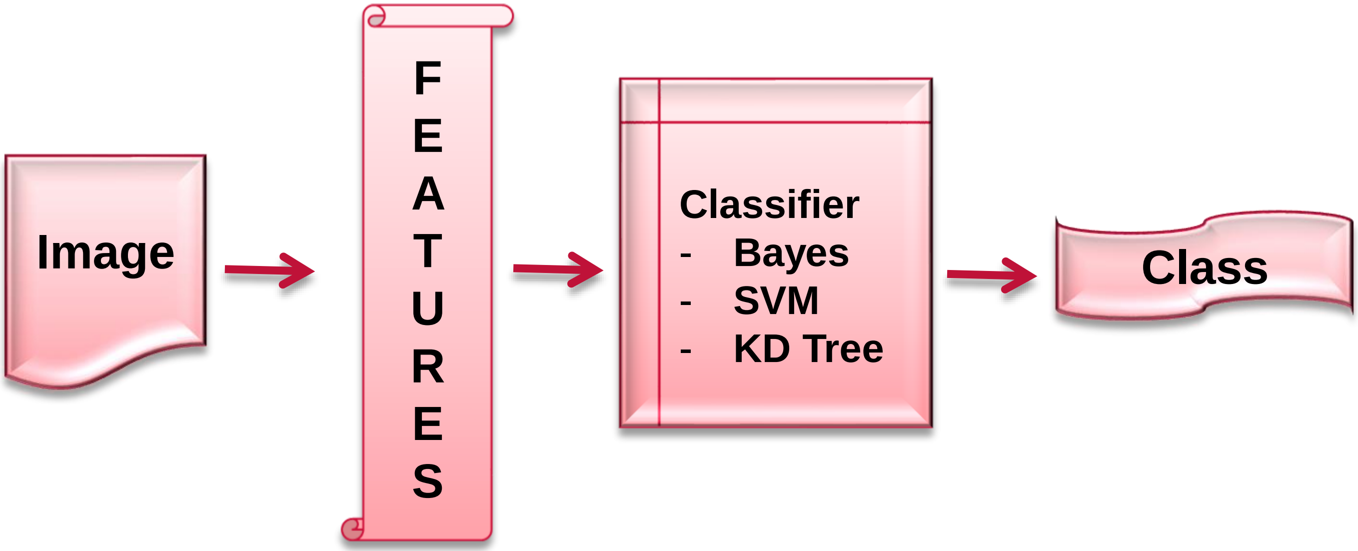
Mark distribution of the 250 000 images from AVA database

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

Mean value = 5,34
Almost Gaussian distribution

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)*

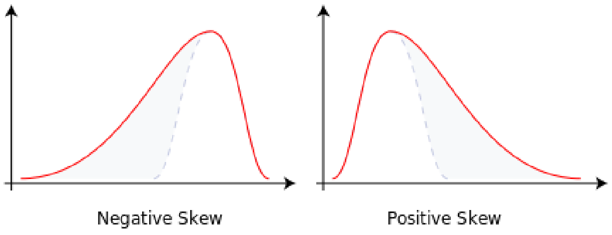
Machine learning with Photo aesthetics rules (2/4)

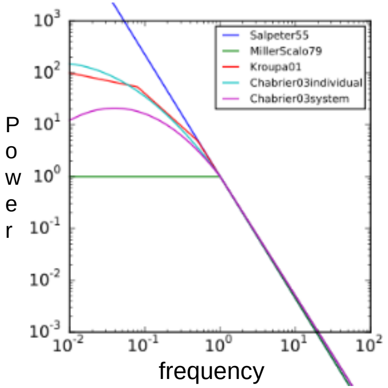
- **Composition rules**
  - **Weight of the center**
  - **Importance of order** (regular, progressive)
  - **Symetries**
  - **Alignements, diagonal, perspective …**

- **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: proba(I)**
  - **Any Mean or Variance: <I> <(i-<i>)^2>**
  - **Positively biased skew (order 3)**

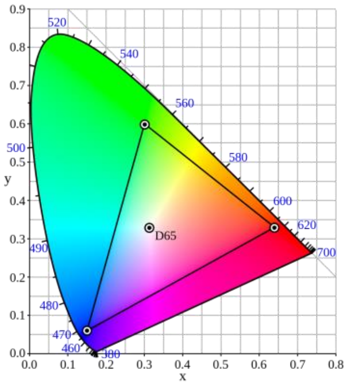
Attewell & Baddeley, 2007

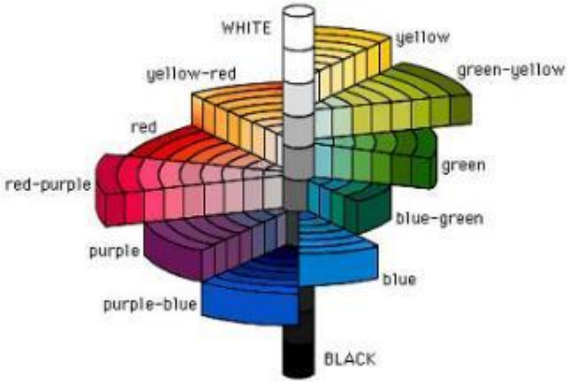
- **Power spectrum (after Fourier transform):**
  - $P_f(I) = |TF(I)|^2$
  - **Radial Symmetry**
  - $1/f^2$ **power spectrum decay**

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)*

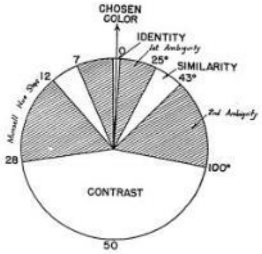


FIG. 2. Regions of similarity and contrast in a plane z=const. (constant Munsell value).

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.

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

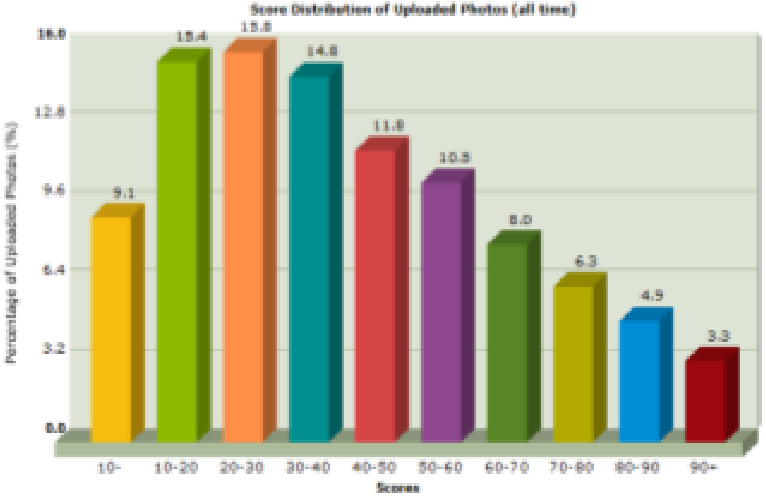
Datta, Joshi, Wang, Li - 2006-2012

Penn State University

- **Generic image features (56 features)**
- **Input: on 7 levels   Output: from 0 to 100**
- **with SVM as classifier**
- **(No longer) Available on-line: ➔ *Acquine***

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

- *Datta R., Joshi D., Li J.Wang J., Studying aesthetics in photographic images using a computational approach, Computer Vision, ECCV 2006, 3953 Lecture Notes in Computer Science, 288-301, 2006.*
- *Datta R., Li J.Wang J. Z., Algorithmic inferencing of aesthetics and emotion in natural images : An exposition, 15th IEEE International Conference on Image Processing., 105-108, 2008.*
- *Datta R.Wang J., ACQUINE : Aesthetic quality inference engine - real-time automatic rating of photo-aesthetics, ACM, , MIR'10, 2010.*



(a) distribution of all ACQUINE scores

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 »

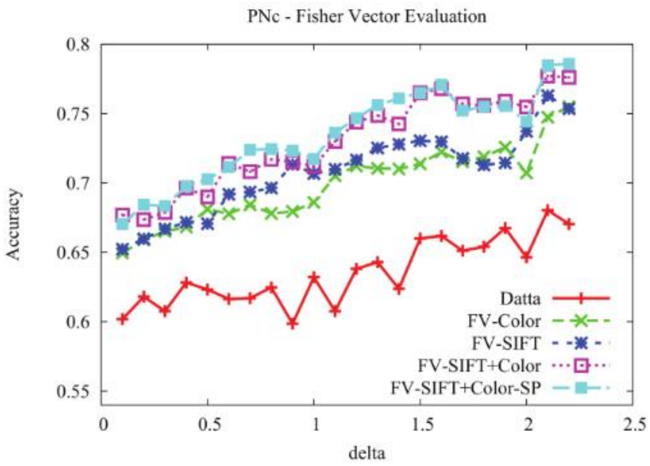
- *Lo K., Liu K., Chen C., Assessment of photo aesthetics with efficiency., IEEE International Conference on Pattern Recognition (ICPR),, 2186-2189., Nov. 2012.*

Marchesotti, Murray, Perronnin - 2011-2013

Xerox Lab, Grenoble

- **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

- *Marchesotti L., Murray N.Perronnin F., Discovering Beautiful attributes for aesthetic image analysis, International Journal of Computer Vision, 113, 246-266, 2015.*



Ke, Tang, Jing - 2006

Carnegie Mellon + Microsoft Asia

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

- *Ke Y., Tang X., Jing F., The design of high-level features for photo quality assessment, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), 1, 419-426, 2006.*

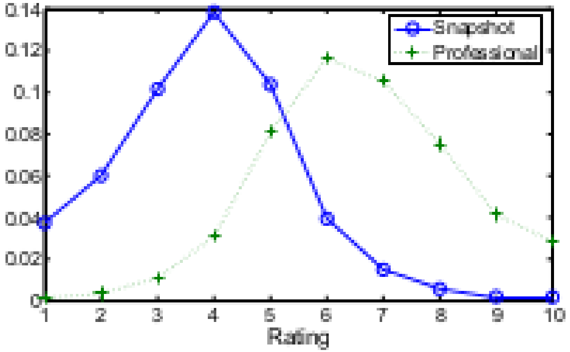
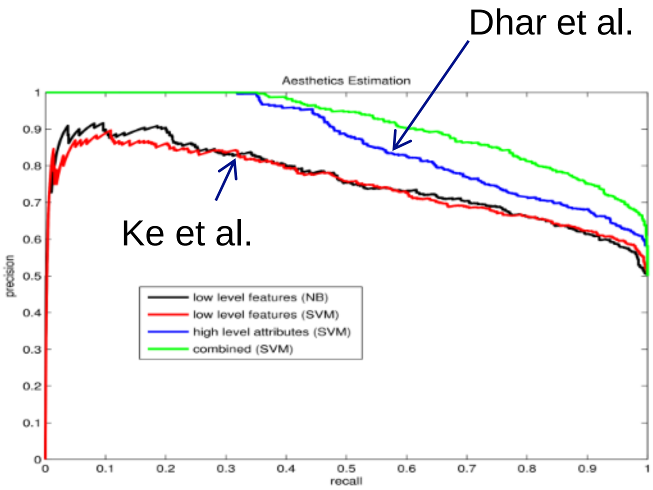


Figure 9. Distribution of people's ratings of professional photos and snapshots. There is significant overlap in the distributions, meaning there is ambiguity in the perceived quality of the photos.

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**

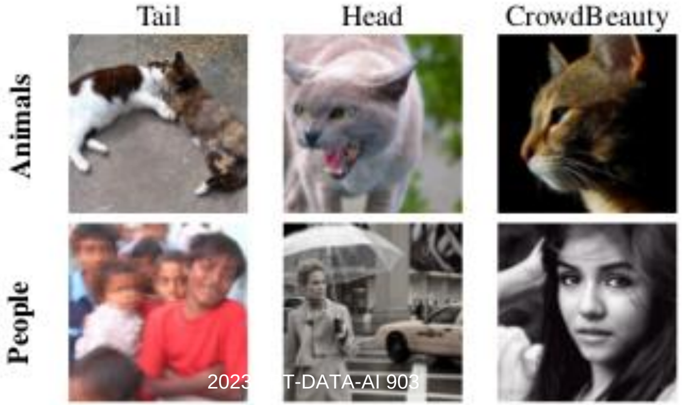


- *Dhar S., Ordonnez V., Berg T. - High level describable attributes for predicting aesthetics and interestingness, CVPR 2011, 1657-1654*

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 culturaly 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.)



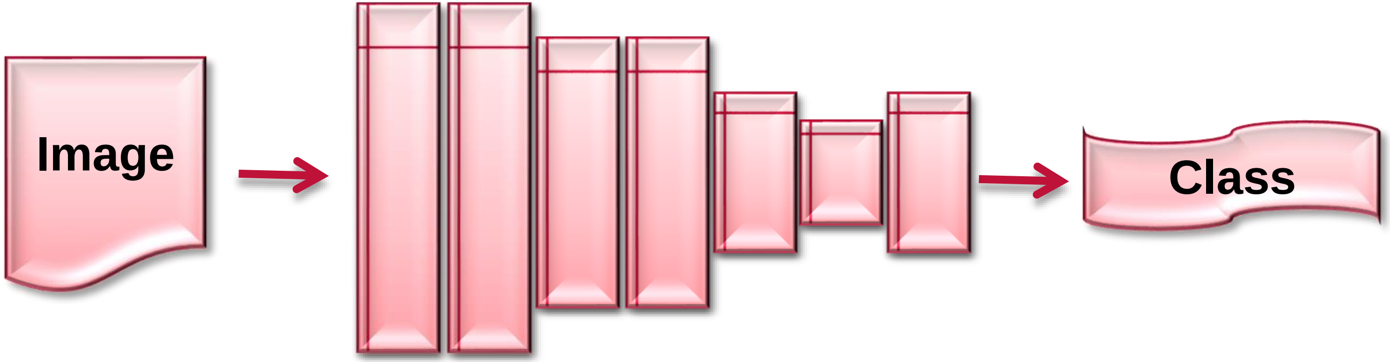
- *Schifanella R., Redi M., Aiello M., An image is worth more than thousand favorites surfacing the hidden beauty of Flicker images, ACM, arXiv preprint arXiv :1505.03358., 2015.*

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

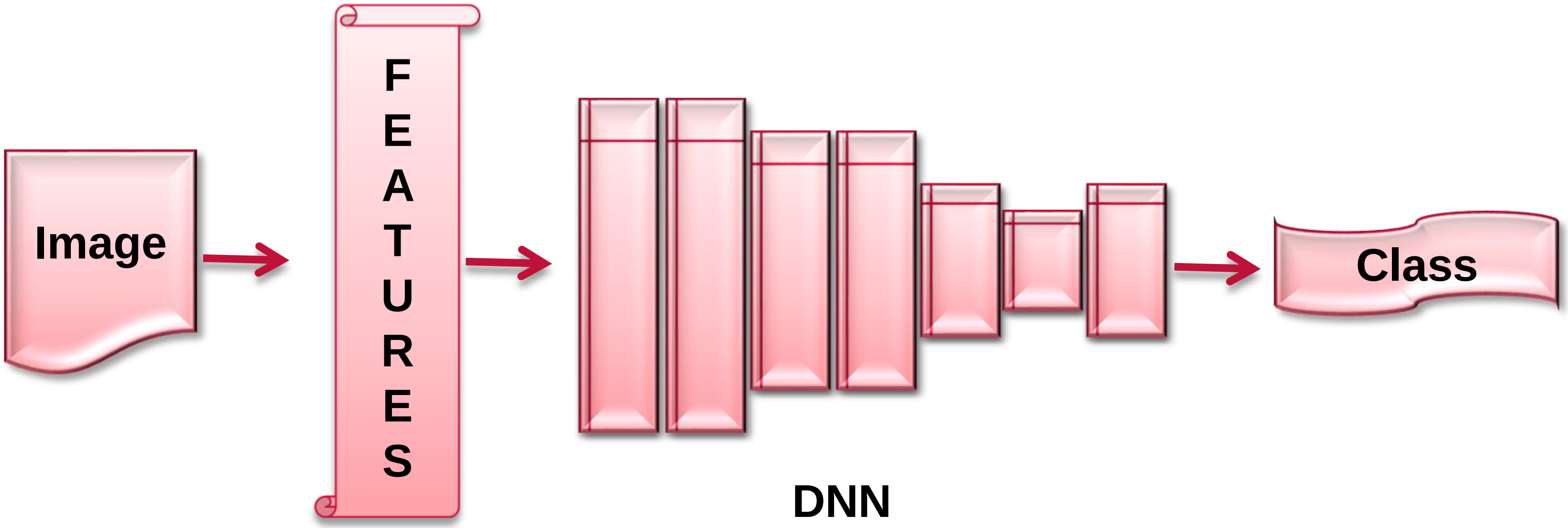
- Handcrafted Feature Selection + Classifier
- **Deep Neural Network**



DNN

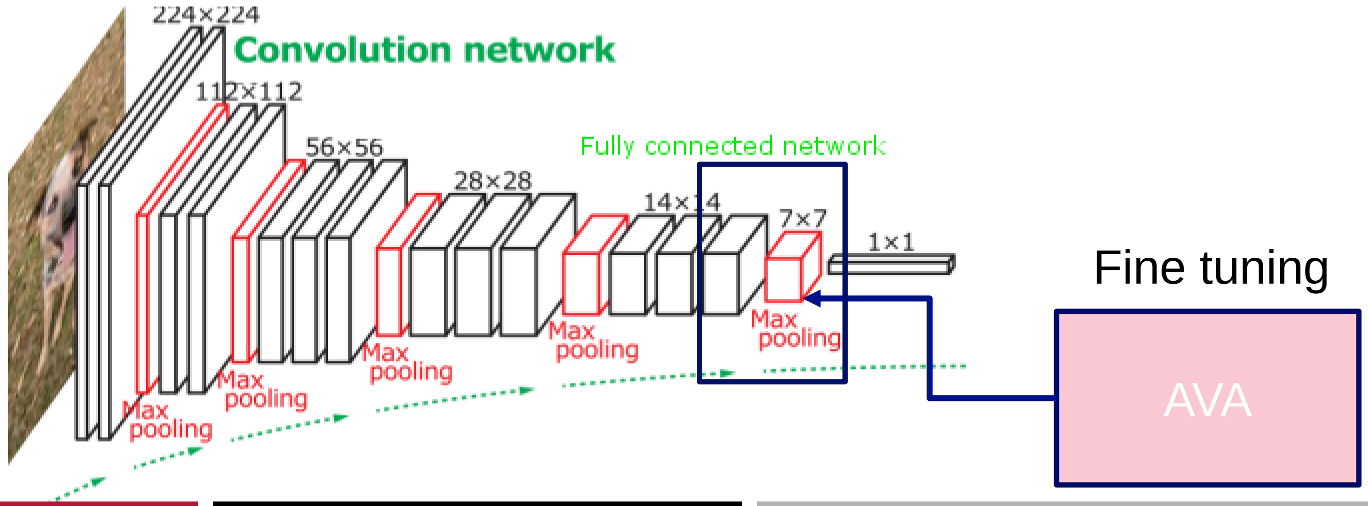
Machine Learning approaches

- 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 a conventional image database** (ImageNet, …)
- **Add a level of fine tuning with data issued from an aesthetic database**



Main publications using CNN (1/2)

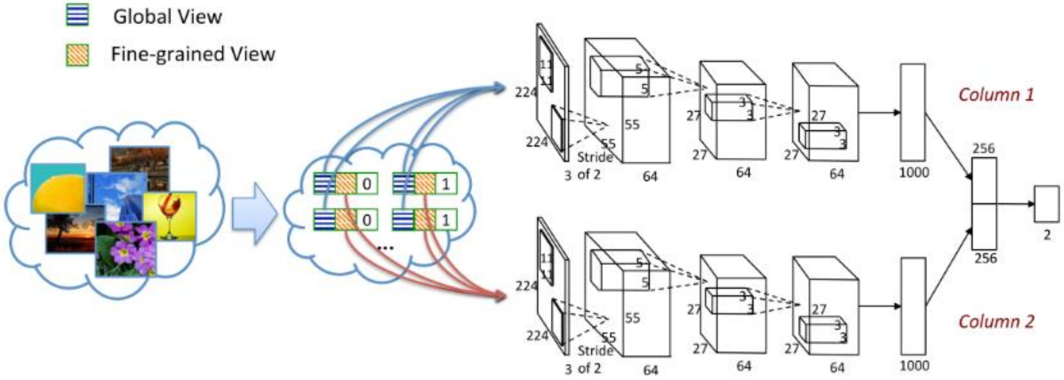
reference	Data Base	Evaluation granularity	Middleware	decision
[Lu et al. 2014] Penn State + Adobe	AVA	2 classes	2 tracks + categories	2xCNN : 50 random sub-images
[Lu et al. 2015] Penn State + Adobe	AADB + AVA	2 classes	6 tracks	1CNN + 5 random sub-images
[Kong et al. 2016] Irvine + Adobe	AADB + AVA	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

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



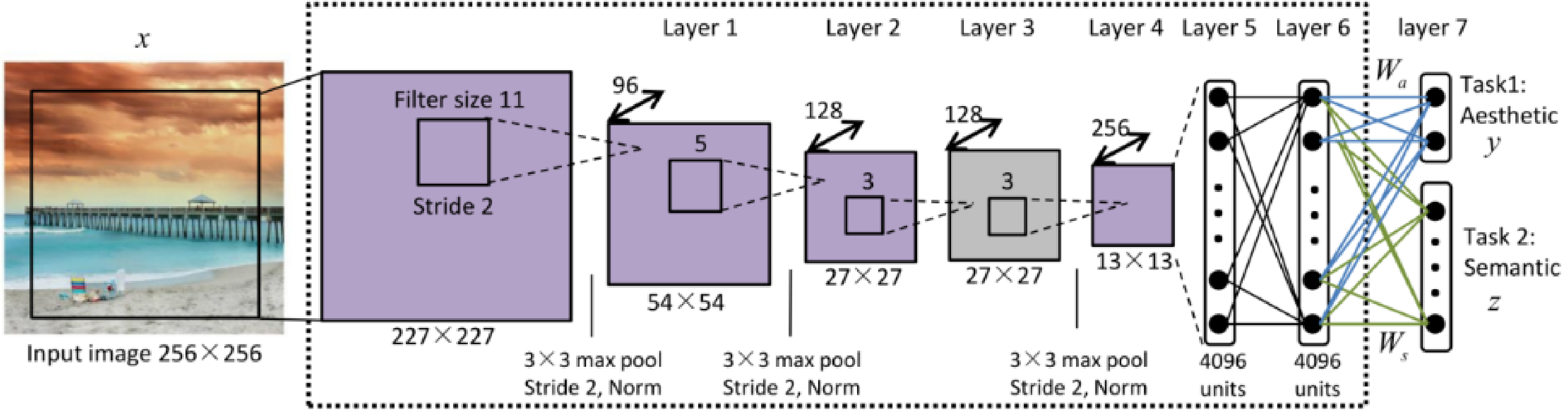
- « Style » regularisation channel
- Learning : 200 000 images from AVA

δ	[24]	SCNN	AVG_SCNN	DCNN	RDCNN
0	66.7%	71.20%	69.91%	73.25%	74.46%
1	67%	68.63%	71.26%	73.05%	73.70%

X. Lu, Z. Lin, H. Jin, J. Yang, J.Z. Wang
ACM Int. Conf on MM, Orlando 2014 - pp 457-466

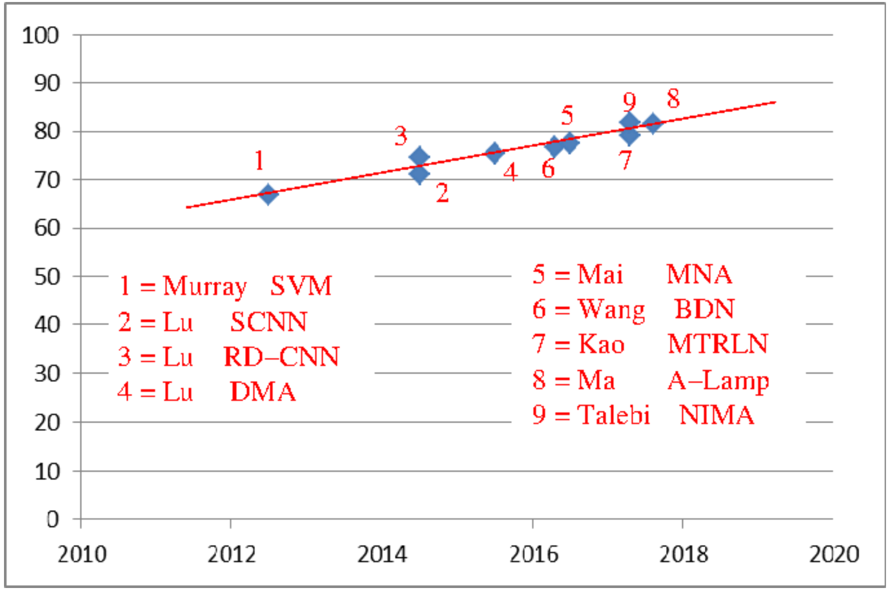
Example 2: Y. Kao et al (2017)

- **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

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 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 datase t(Behance)

II – Subjectivity from a social profile

- **Collect any information on the user, available from 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 – Subjectivity from a social profile

Alternative:

- **Make use of photos posted by the user on the web** *(Instagram, Flickr, Pinterest, …)* to categorize users *(Lovato, Perina, Sebe, … et al., 2013):*

$$x_m = \ feature\ vector\ of\ an\ image \text{ (one in M)}$$

$$y_m = \ discrete\ label\ representing\ the\ user \text{ (one in N)}$$

$$y_m = w^T\ x_m$$

$$\text{E(w)} = \sum_{m=1}^{M} (y_m - \ w^T\ x_m)^{\,2}$$

Optimization with Lasso for each user separately, on all pictures ➔ Discriminatation of each user w.r.t. the others ➔ $\boldsymbol{w}^{(\boldsymbol{n})}\ \ n = 1, \ldots N$

- *Experiment :* 40 000 images, 50 % training, 50% testing - 200 users

II – Subjectivity from a social profile

- Results *(Lovato, Perina, Sebe, … et al., 2013):*

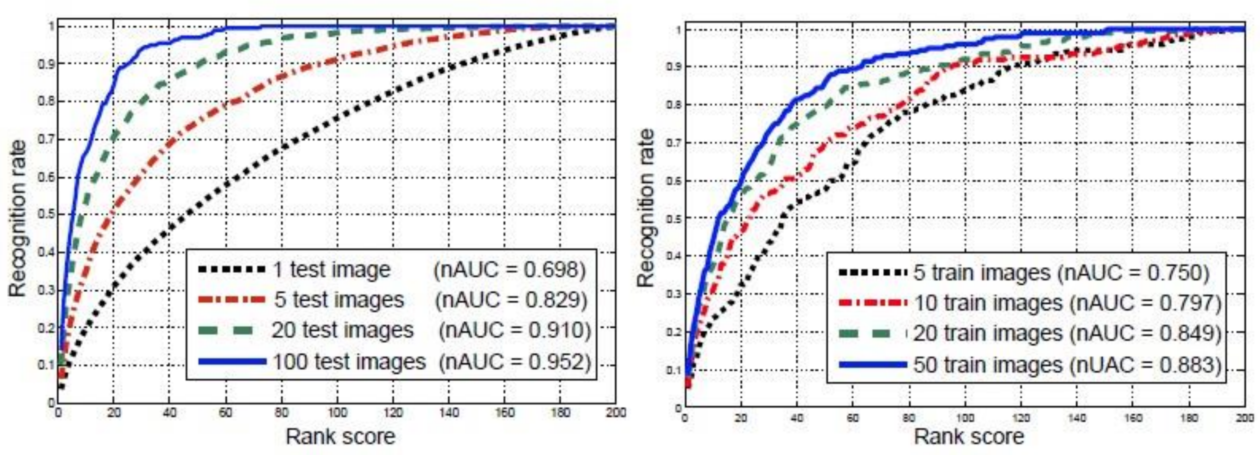
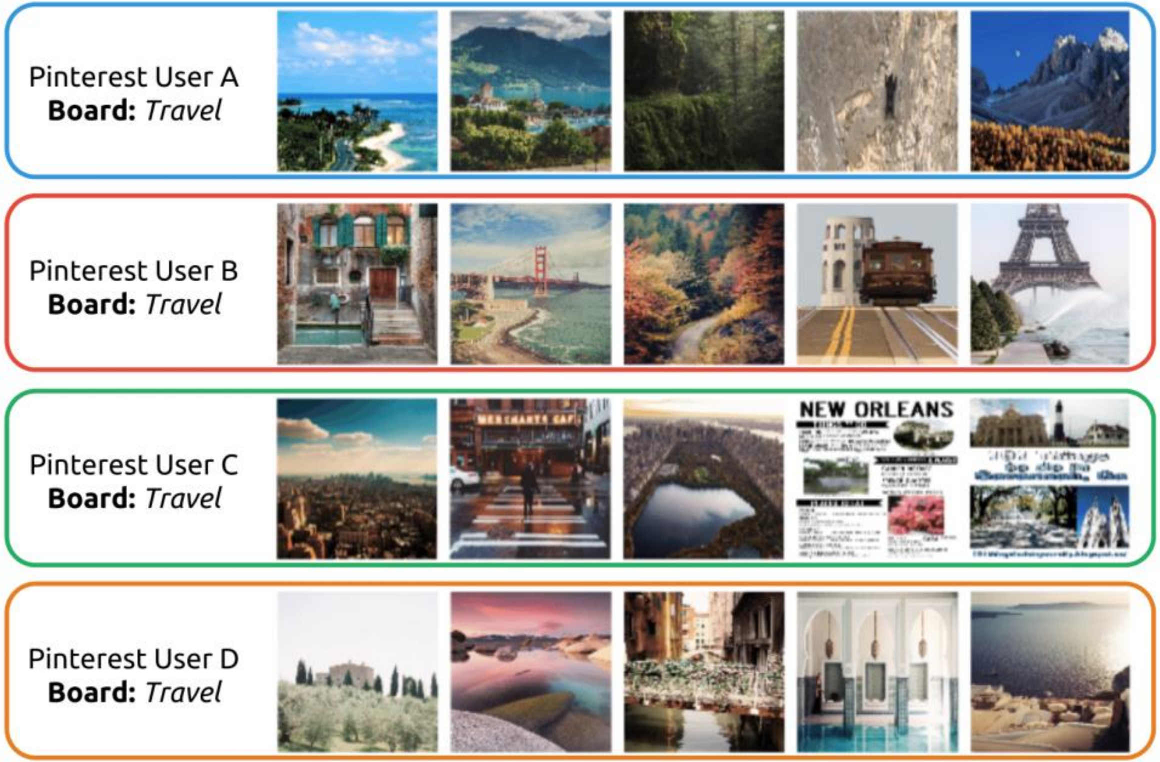


Fig. 2. CMC curves for our dataset. On the left: For each curve, we varied the number of testing images to be considered as a single "set". On the right: For each curve, we varied the number of images used to train Lasso.

II – Social classes from posted images online



Experiment based on Pinterest
- 3 790 users
- each with 200 photos
- from one of 34 categories (« boards ») = *Travel*
- for each user > 100 photos

Feature space: D=410
200 clusters

Yang, Hsieh & Estrin, 2015

II – Social classes from posted images online

- *Yang, Hsieh & Estrin, 2015*

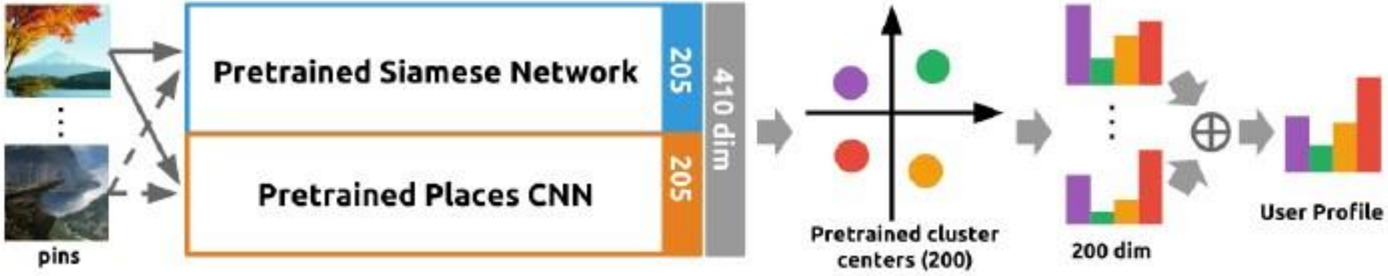


Figure 2. Algorithmic framework for user interests profiling from visual contents. *Phase 1:* Siamese Network and CNN based feature extraction; *Phase 2:* Euclidean distance based soft assignment to pre-trained visual clusters; *Phase 3:* Generate user profile by aggregating all image visual cluster features.

- 200 pre-trained clusters
- Soft distance of any image
- User profile

$$c_j^i(k) = \begin{cases} e^{-\frac{1}{2\alpha^2}\|d_j^i - r_k\|^2} & : \|d_j^i - r_k\| \le \delta \\ 0 & : \|d_j^i - r_k\| > \delta \end{cases}$$

$$\tilde{v}_i = \sum_{j=1}^{|S^i|} c_j^i; \quad v_i = \frac{1}{\|\tilde{v}_i\|_1}\tilde{v}_i$$

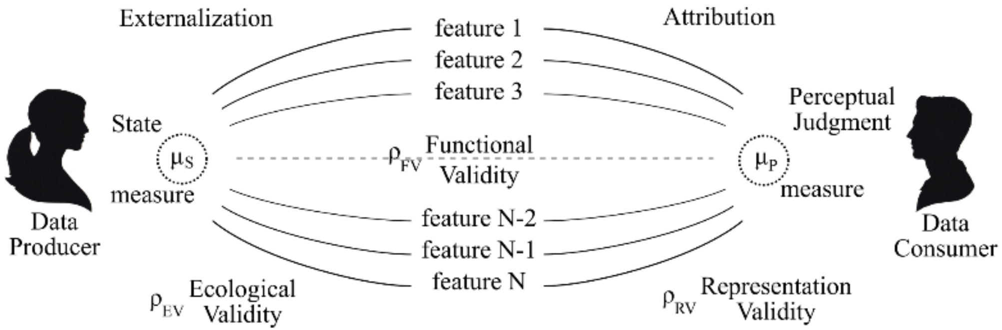
II – Social classes from posted images on line

Efficient deep distance
Metric for images

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. Brunswick, 1956)*



M. Cristani, Vinciarelli, Segalin & Perina, 2013

Measure of Big Five using Brunswick lens model

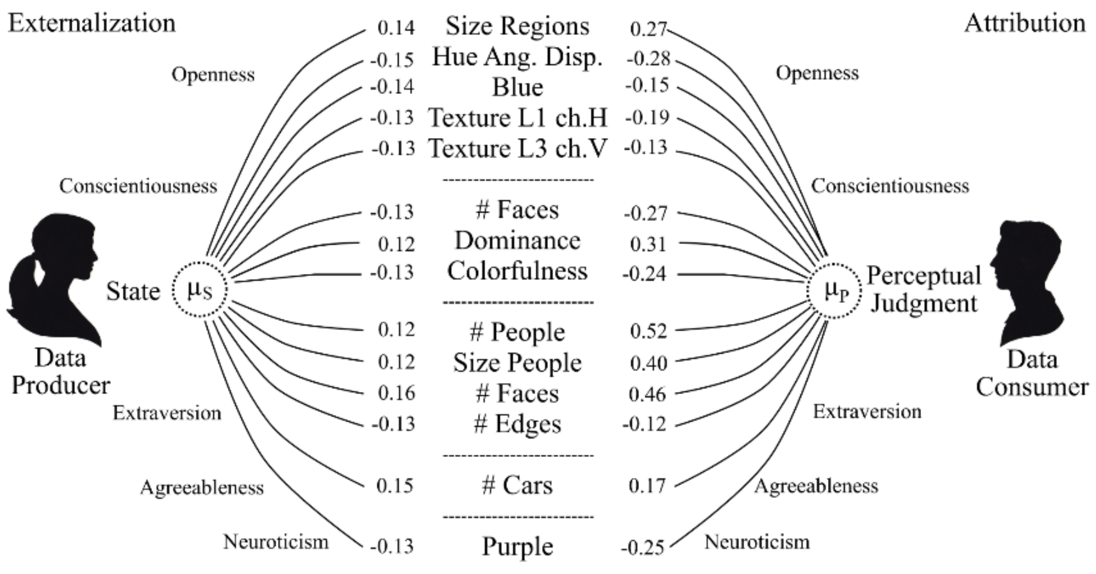


Figure 2: The picture shows the Brunswik Lens model for the PsychoFlickr dataset, where the state corresponds to the Big Five traits (as per assessed with the BFI-10). Ecological and Representation validities are measured with the Spearman Coefficient and the picture shows (for each trait) features for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

Ground truth = PsychoFlickr database :

- *60 000 images*
- *300 photographers – 200 favourite photos (made by others) each*

- *Emitter profile = Big Five from BFI_{10}*

- *Receiver profile = average of 8 assessors*

M. Cristani, Vinciarelli, Segalin & Perina, 2013

Learning Big Five

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

OCEAN
Conscientiousness = sens of responsability

C. Segalin, Chang & Cristani., 2017

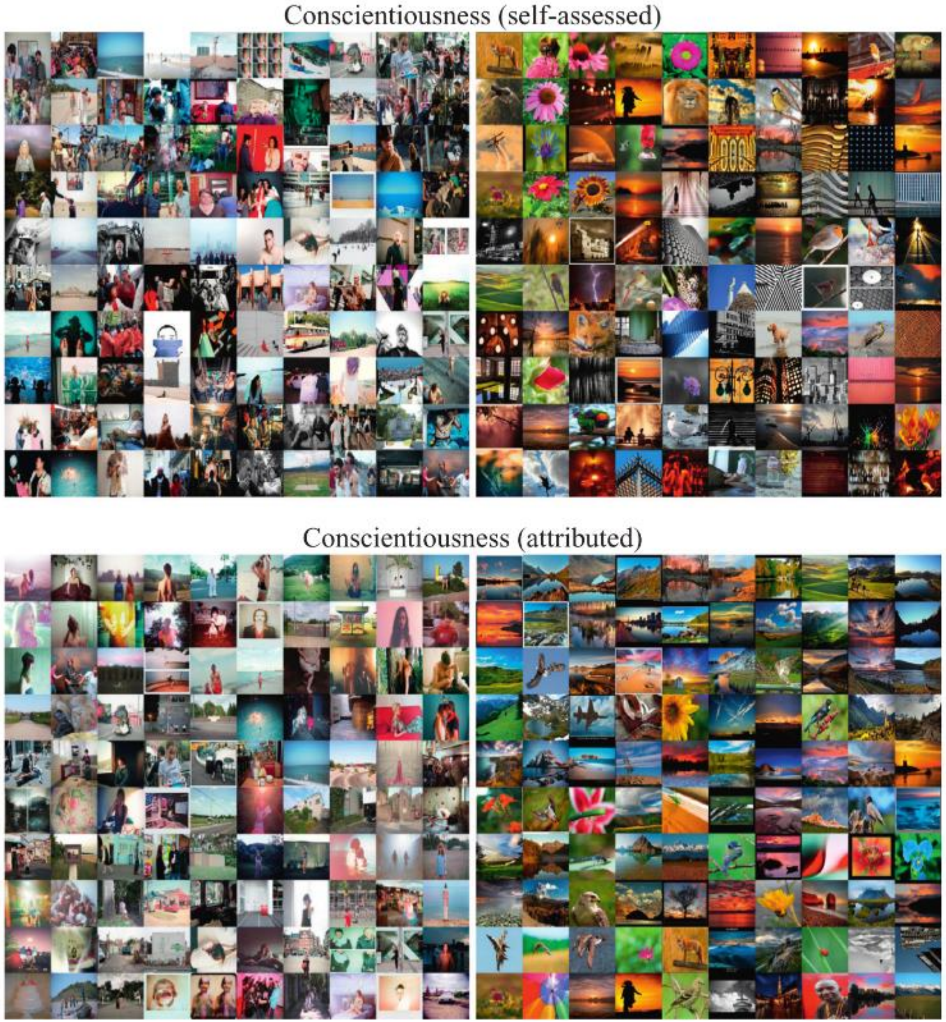
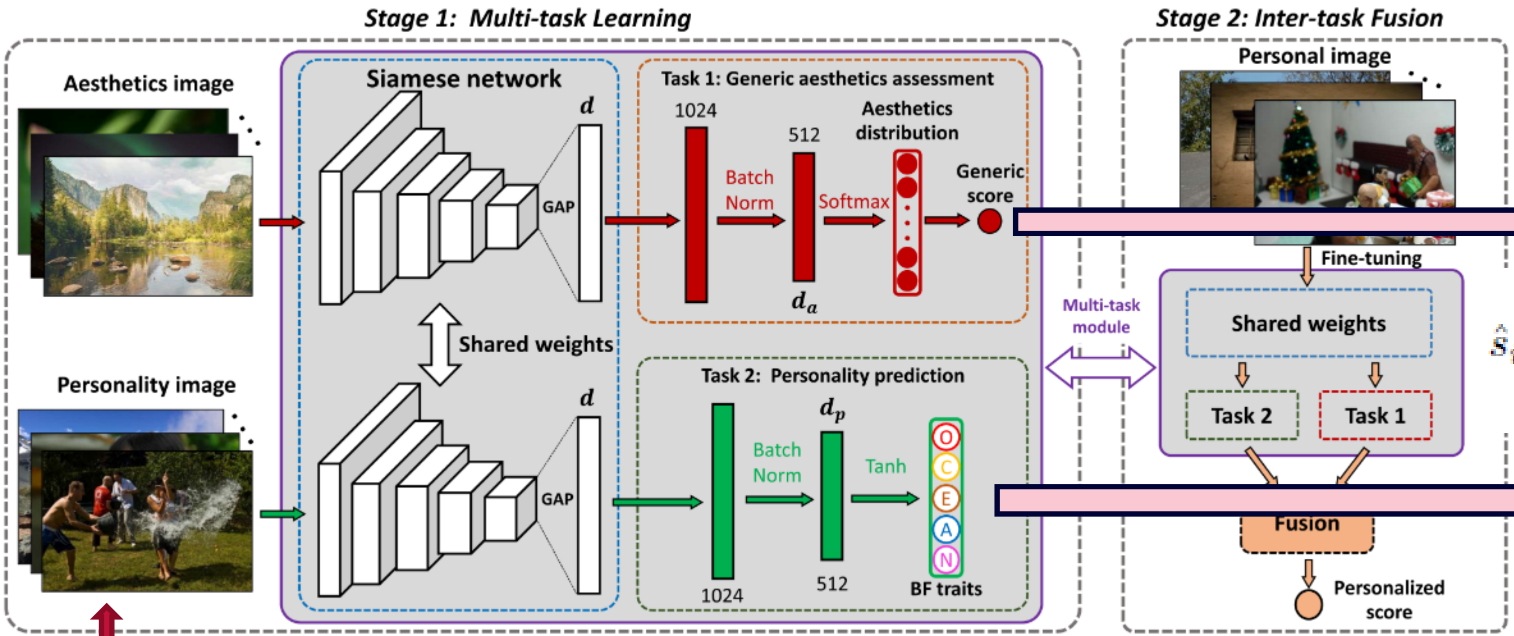


Fig. 7. Representative images in the low (left) and high (right) value classes for the Consciousness trait.

Use of Big Five for Aesthetic appraisal of photos



$$\hat{s}_i = \hat{s}_i^a + \sum_{j=1}^{5} w_j \hat{s}_{i,j}^p$$

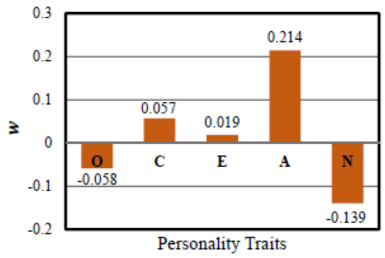
Fig. 3. The framework of the proposed personality-assisted multi-task learning for IAA.

Make use of images posted by the user (PsychoFlickr)

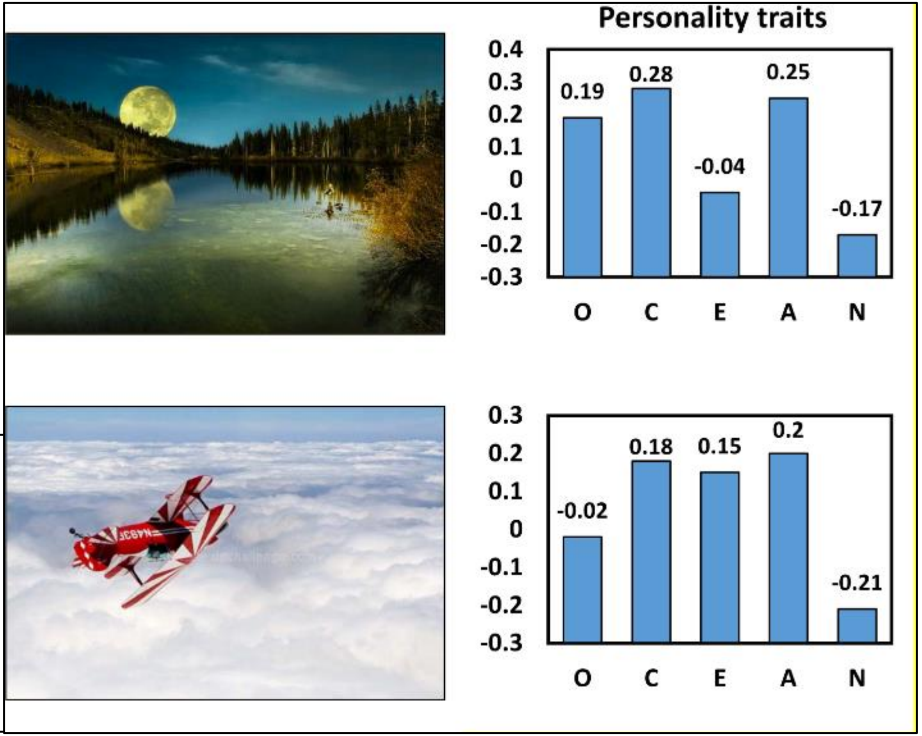
Li, Zhu & Zhao, 2020

Use of Big Five for Aesthetic appraisal of photos

Li, Zhu & Zhao, 2020



Contribution of OCEAN to Beauty



Openness	O
Conscientiousness	C
Extraversion	E
Agreableness	A
Nevrotism	N

IV – Learning aesthetical tastes with tests

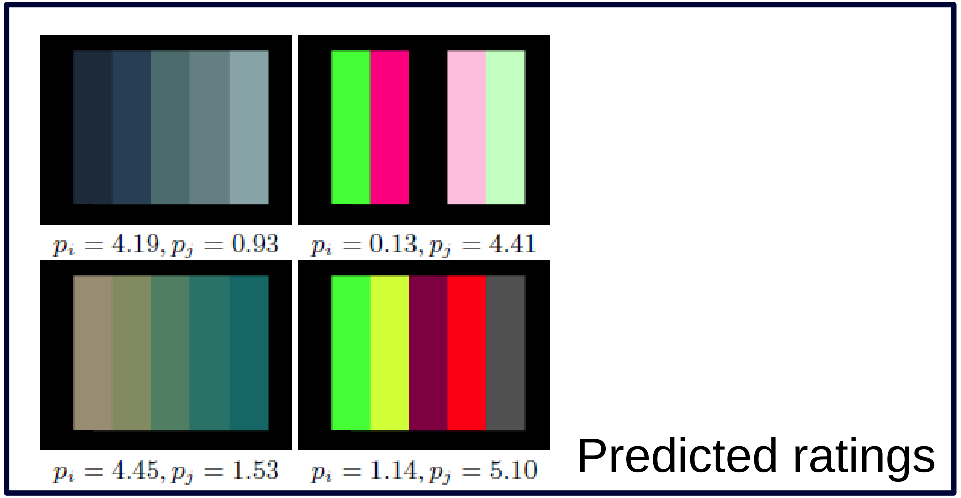
- **Submit some images to the user's judgement**
  - **How many ?** 10 ? 100 ? No clear answer
  - **Which rating** ? Ranked ? Nice/Poor ? Mark ?
  - **Fine-tune a CNN with user's preferences**

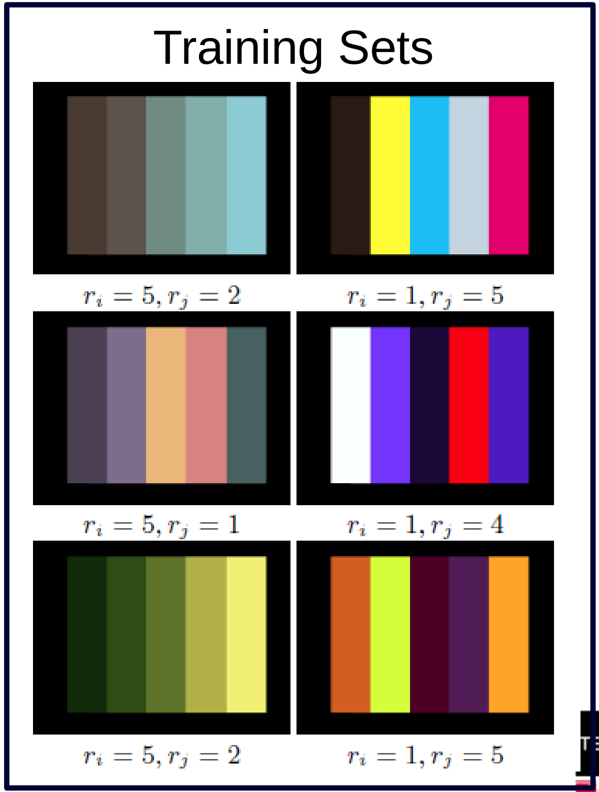
- **Either: Learn individual tastes in a separate test** *(O'Donovan et al., 2014)*
- **Or: Compare the user to one of some prototypic photographers previously used to train a CNN (for instance with** *AVA Photographer Demographic***)** *(Zhu, Li, Wu et al., 2020)*
- **Or: Compare the user's choices 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)*

Evaluate aesthetic preferences with tests

O'Donovan, Agarwala & Hertzmann, 2014

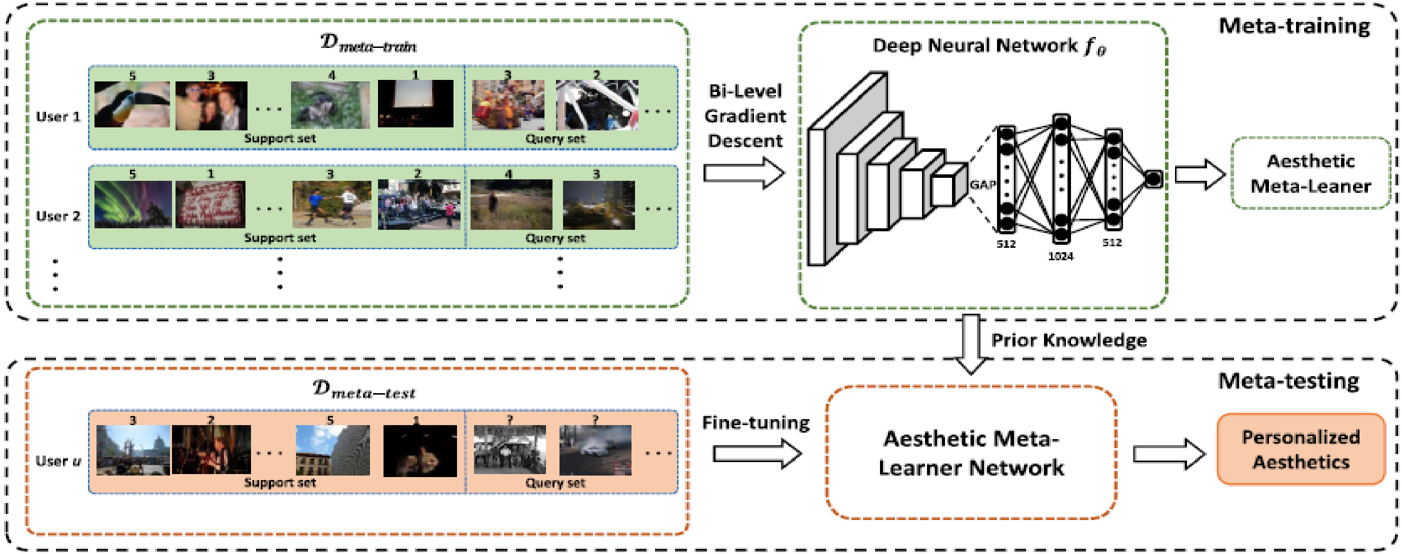
Determine the user's preferences for color palettes or color themes
- 13 343 palettes with random color distribution
- 40 evaluators giving 1 mark for each palette ➔ 528 106 ratings
- 334 features measures for each palette (in RGB, CIELab, HSV, CHSV) : color differences, saturations, mean, max, …)





Evaluate aesthetic preferences with tests

- *(Zhu, Li, Wu et al., 2022)*



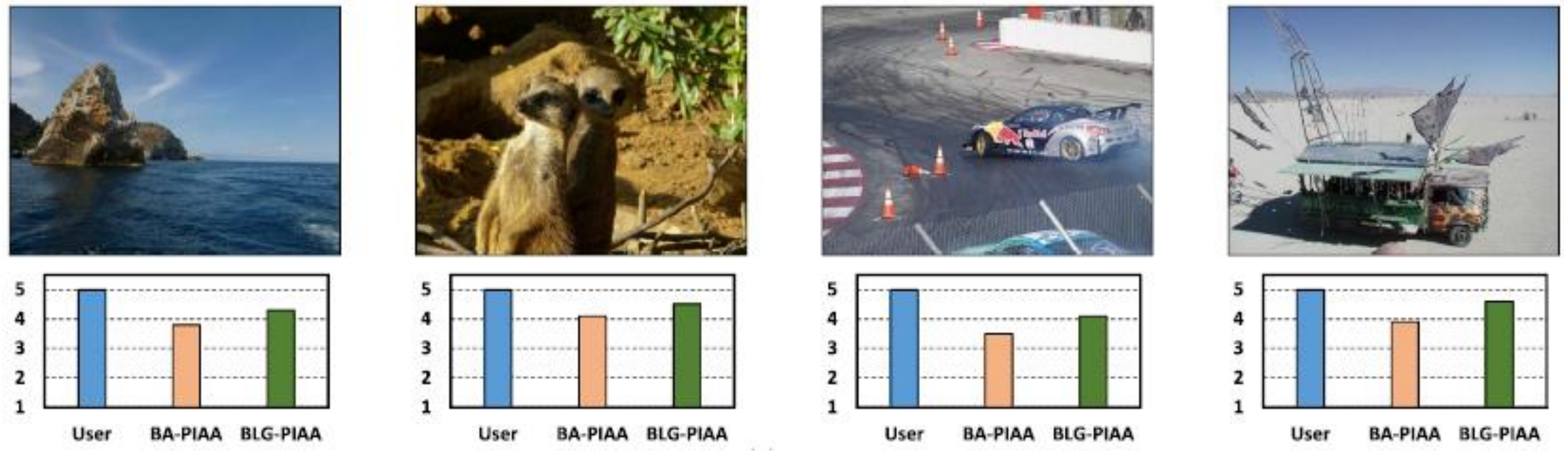
Pretrained on ImageNet, trained on Flickr-AES
173 users

100 samples/user

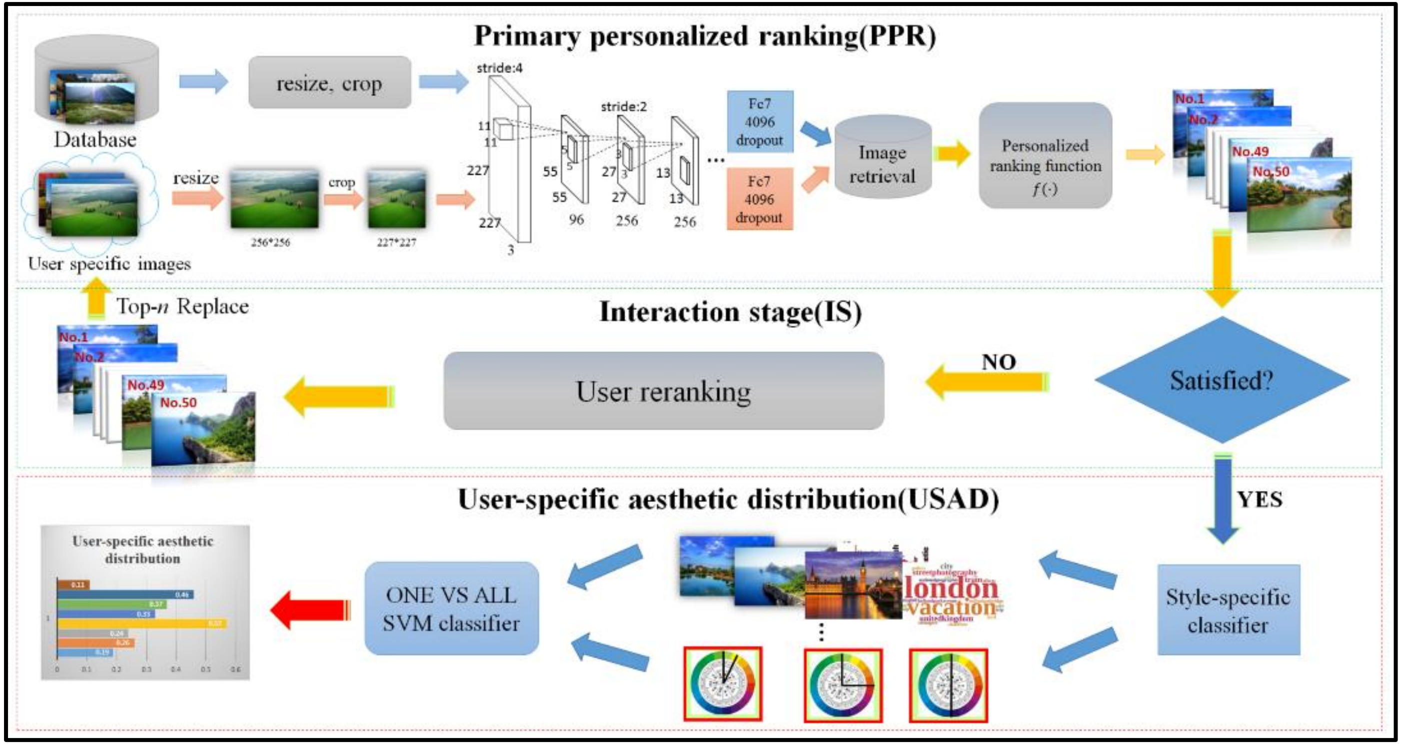
Evaluate aesthetic preferences with tests

- *(Zhu, Li, Wu et al., 2022)*

Comparaison of ratings:
 - BA-PIAA = with state of the art personnalization assessment
 - BLG-PIAA = with the proposed method



Learning user's preference through reinforcement loop



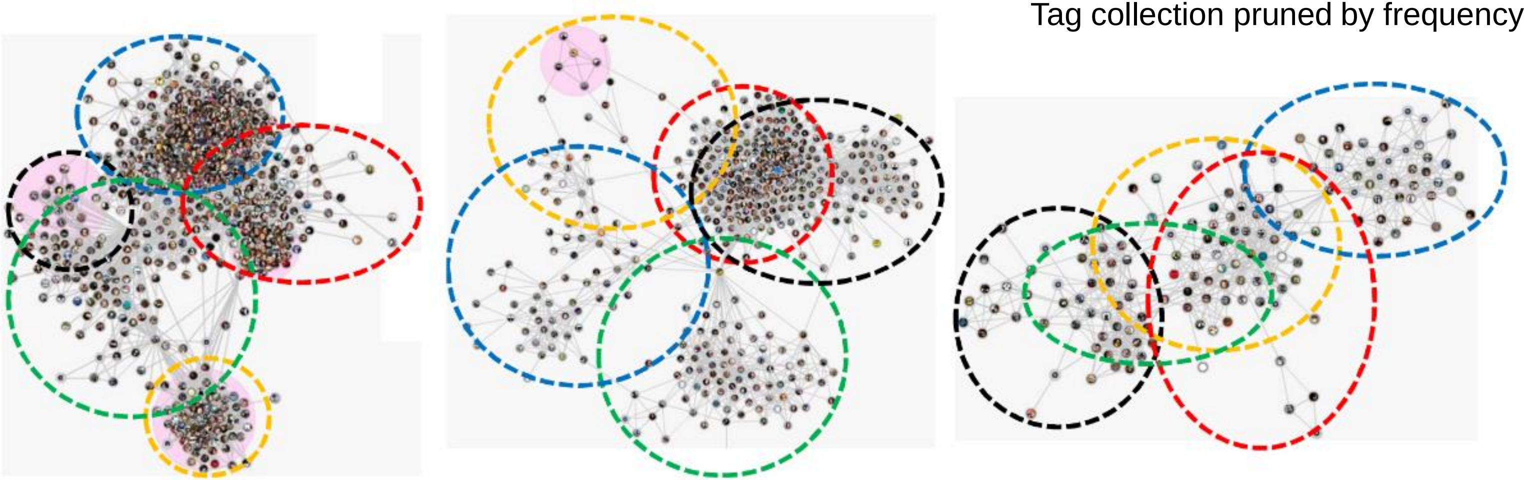
The reinforcement loop is used to detect in the learnig database those which are closed from the user's posted images.

The USAD transfers the learning to the whole database

Lv, Wang, Xu, … et al., 2018

V – Classify the user's population

Flicker Data Base
Tag collection pruned by frequency ➔ 117 tags



Using the classification obtained in the feature space: « **Designers** », « **Colors** », « **Architecture** », « **Black & White** »
- from both pictorial features and litterary features from the caption.

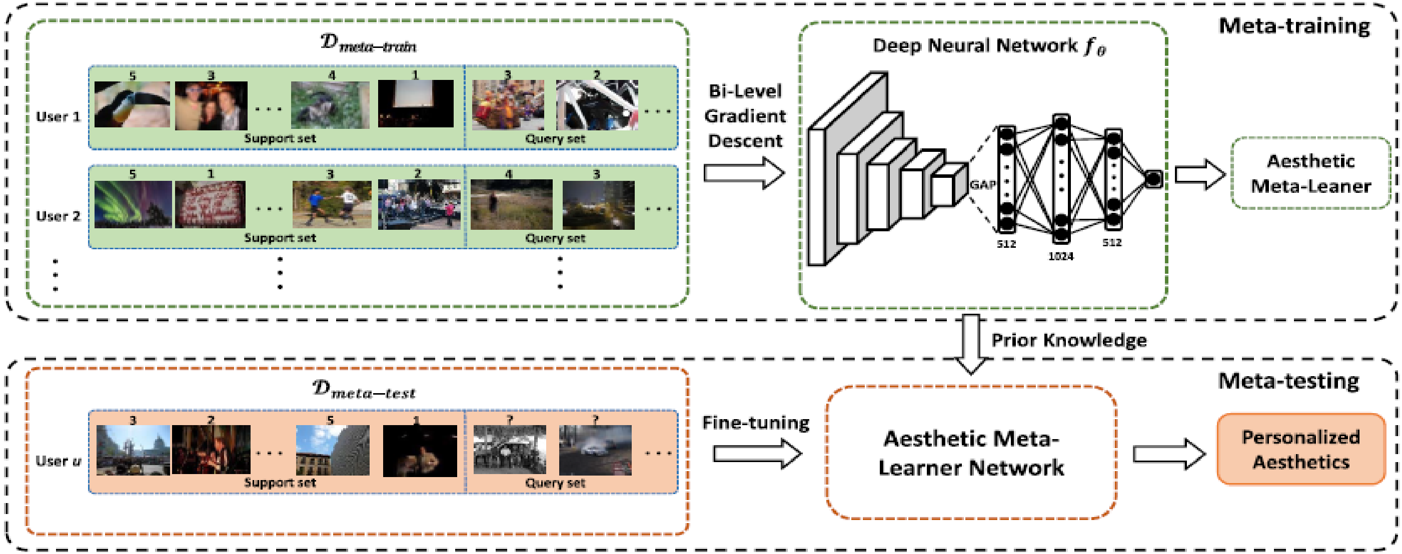
Makes use of « bag of words » :

- 1 image=1 « word »,
- 1 photograph = 1 « sentence »
- from sentences are extracted « topics » which represent « communities »

Hong, Zhang et Tao, 2016

VI Multiply expertises

- **Use experts to specialize a network** *(Zhu, Li, Wu, Zhao, et al. 2022)*
  - **How many experts ?**
  - **How chosen ?**



Critics of the recurse to AI

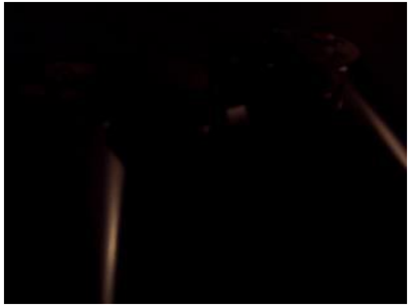
The limits of the database : AVA’s best



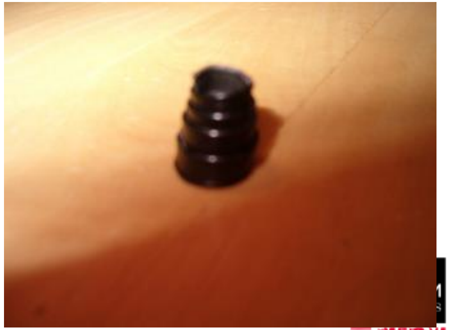




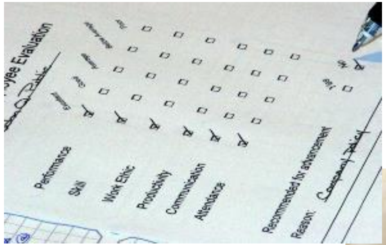
Limits of the database : AVA's worst



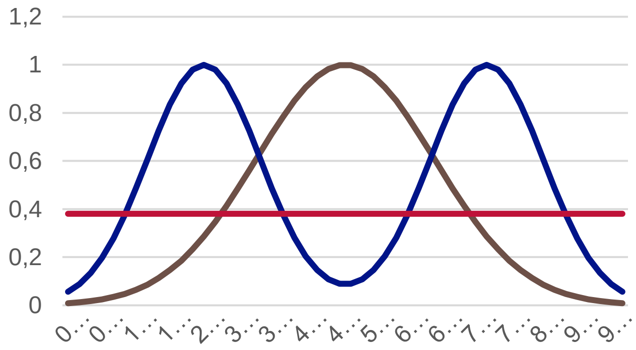



Limits of the database AVA' average B=5.43



Questionning 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

High Quality Photo or Beautiful Photo

Julia Margaret Cameron - 1866