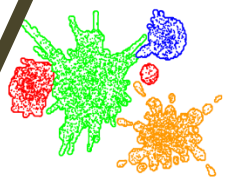
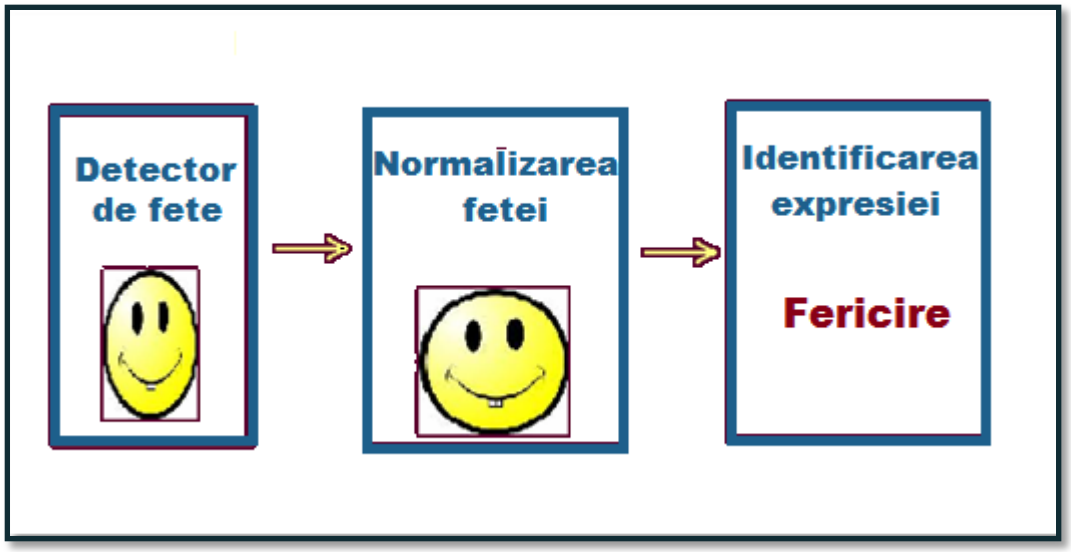
Recunoasterea expresiilor faciale

Interfatare vizuala om-masina



Recunoasterea expresiei

Problema principala: identificarea expresiei unei persoane prin analiza trasaturilor faciale

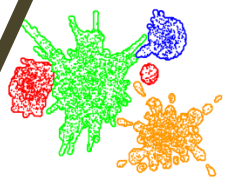


Probleme secundare:

- Identificarea *intensitatii* emotiei generatoare
- Detectia *veridicitatii* (expresia este cauzata de emotie genuina sau simulata)

Aplicatii

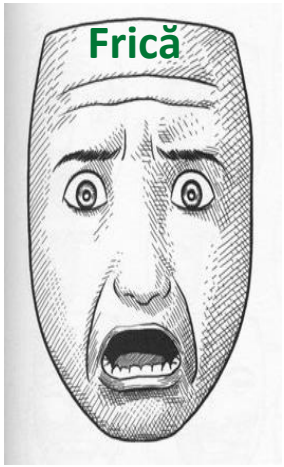
- Monitorizarea comunicatiei - utilizatorul este atentionat in momentul cand vrea sa dea un raspuns sub influenta unei emotii puternice
- Securitate – expresia este utilizata pentru a evidentia daca cineva are ceva de ascuns
- Medicina – studii psihologice:
  - Antrenarea persoanelor cu probleme de recunoastere a emotiilor (autism)
  - Estimarea intensitatii durerii
  - Detectia depresiei
- Marketing
- Jocuri

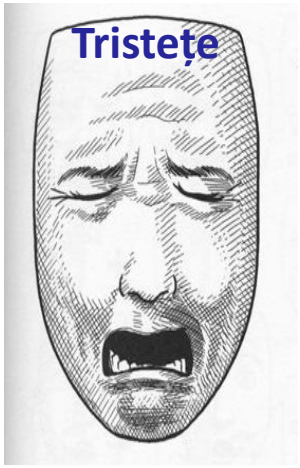


Expresii fundamentale

Ekman and Friesen







Ușor de recunoscut

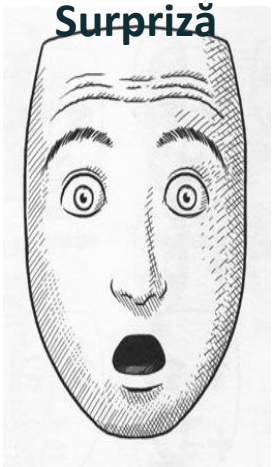
Greu de etichetat

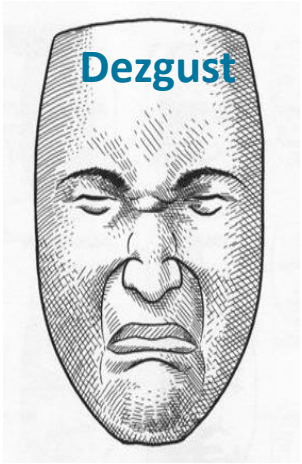
- Emoții – veridicitate
- Control voluntar - simulare

Neutru

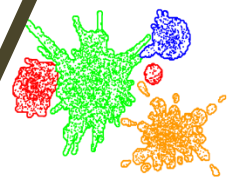
Dispreț







https://www.jeremyshuback.com/blog/the-six-fundamental-facial-expressions



Confuzie intre expresiile de baza

Participant chosen label	*Predicted label*															
	Anger		*Contempt*		*Disgust*		*Fear*		*Happiness*		*Sadness*		*Surprise*		*Other*	
Decoders	W	NW	W	NW	W	NW	W	NW	W	NW	W	NW	W	NW	W	NW
Anger	57	44	7	8	11	16	9	11	0	1	4	5	1	2	4	6
Contempt	12	11	39	30	17	17	2	2	0	1	9	4	1	1	9	6
Disgust	8	9	20	7	57	42	6	8	0	1	4	2	2	2	5	4
Fear	4	4	1	2	2	5	50	31	0	1	2	5	8	7	3	5
Happiness	0	1	1	4	0	1	1	2	84	80	0	2	17	16	9	9
Sadness	3	4	5	5	4	5	1	3	0	1	59	54	1	2	8	10
Surprise	8	9	4	6	1	3	28	35	5	3	2	3	64	54	5	5
Other	9	19	24	40	9	13	5	13	10	13	20	25	8	16	57	57

Predicted and participant-chosen emotion labels for facial expressions (portrayed in still photographs), as reported in two crosscultural studies of emotion expression (Elfenbein et al., 2007—Canadian and Gabonese samples; Yik & Russell, 1999—Canadian and Chinese samples). Figures given are the percentage of trials in which participants chose the specified label for each target. Columns are the emotions that the targets were instructed to express; rows are the labels participants were given, from which they were to choose *one*, according to the emotion they believed the targets to be experiencing. For each target emotion, the nonshaded column on the left (W) displays the rate at which decoders from a Western culture chose each label, and the shaded column on the right (NW) displays the rate at which decoders from a non-Western culture chose each label (all encoders were Caucasian, and Canadian or American). The heavily shaded areas correspond to correct answers. Total percentages may deviate from 100%, as reported rates were rounded.

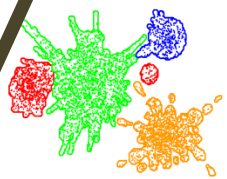
Expresii derivate

1. Grade de intensitate pentru expresiile de baza

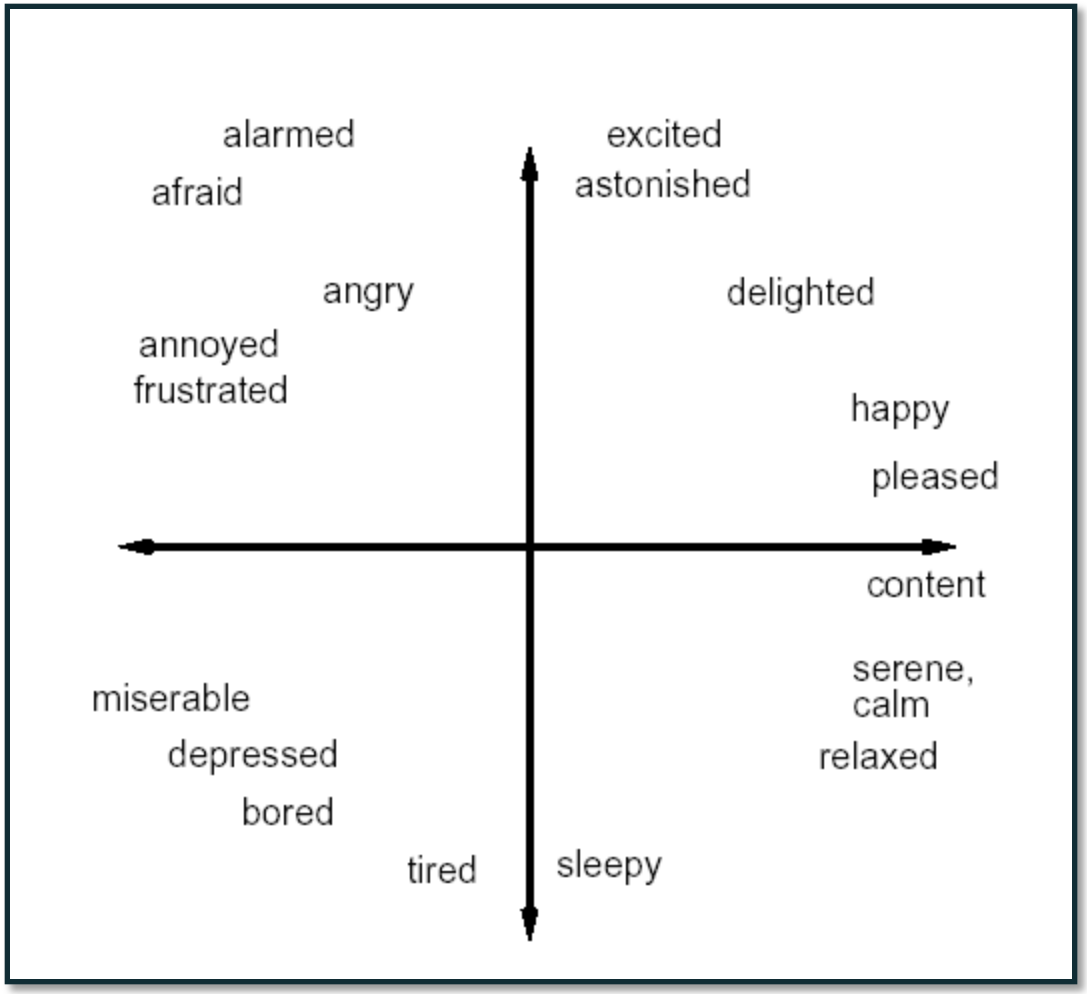
Tristete	Furie	Fericire	Teama	Dezgust	Surpriza
Dezamăgire Tristeţe Deprimare	Frustrare Enervare Furie Furie nebuna	Placere Fericire Extaz	Spaima Frica Sperietura	Dezgust Sila	Surpriză Uimire Uluire

2. Expresii obtinute prin amestecul expresiilor de baza
   ex: **Surpriza + Fericire = Supriza placuta**

[Zucker2007]

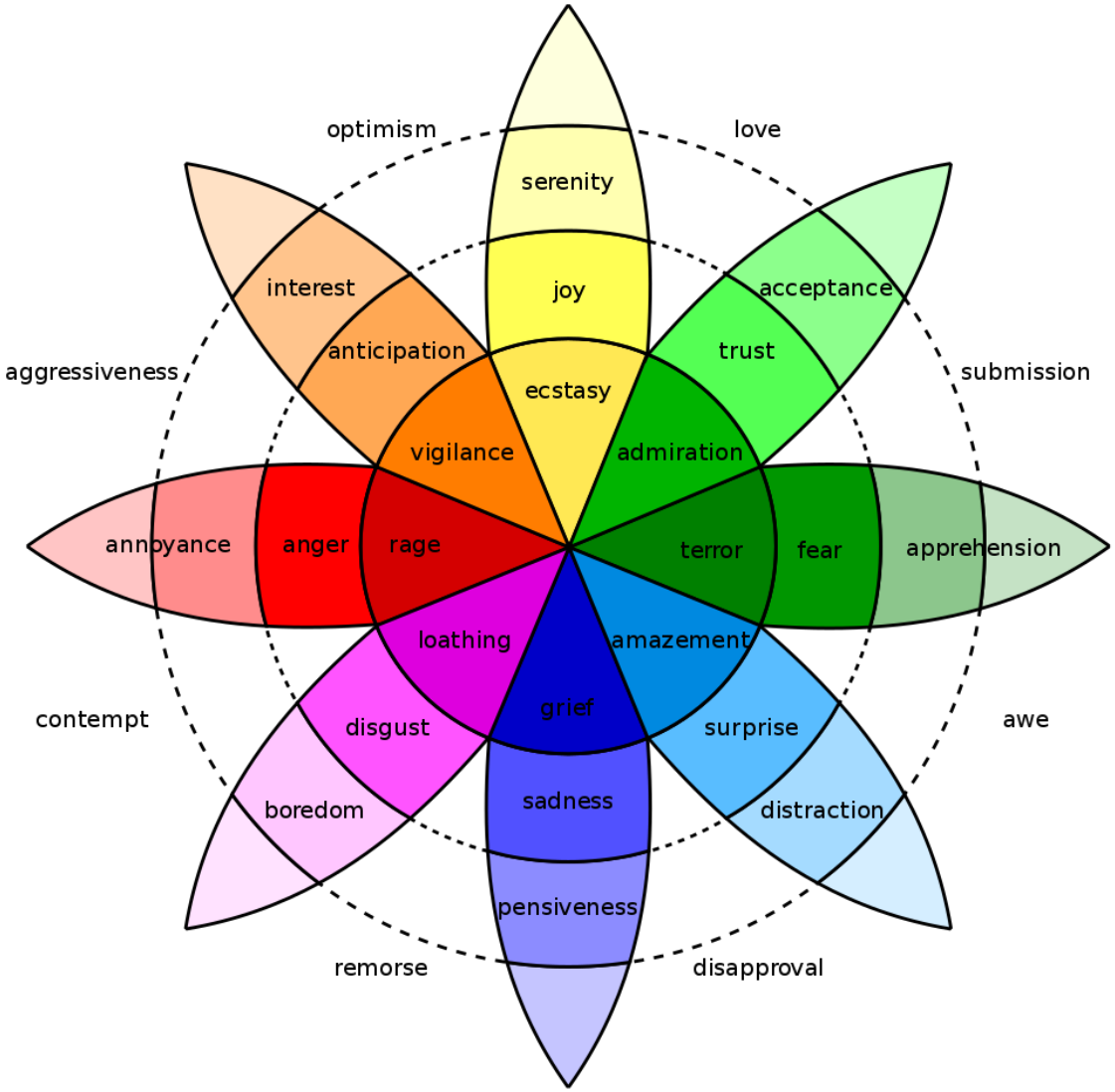


Axele Emotionale

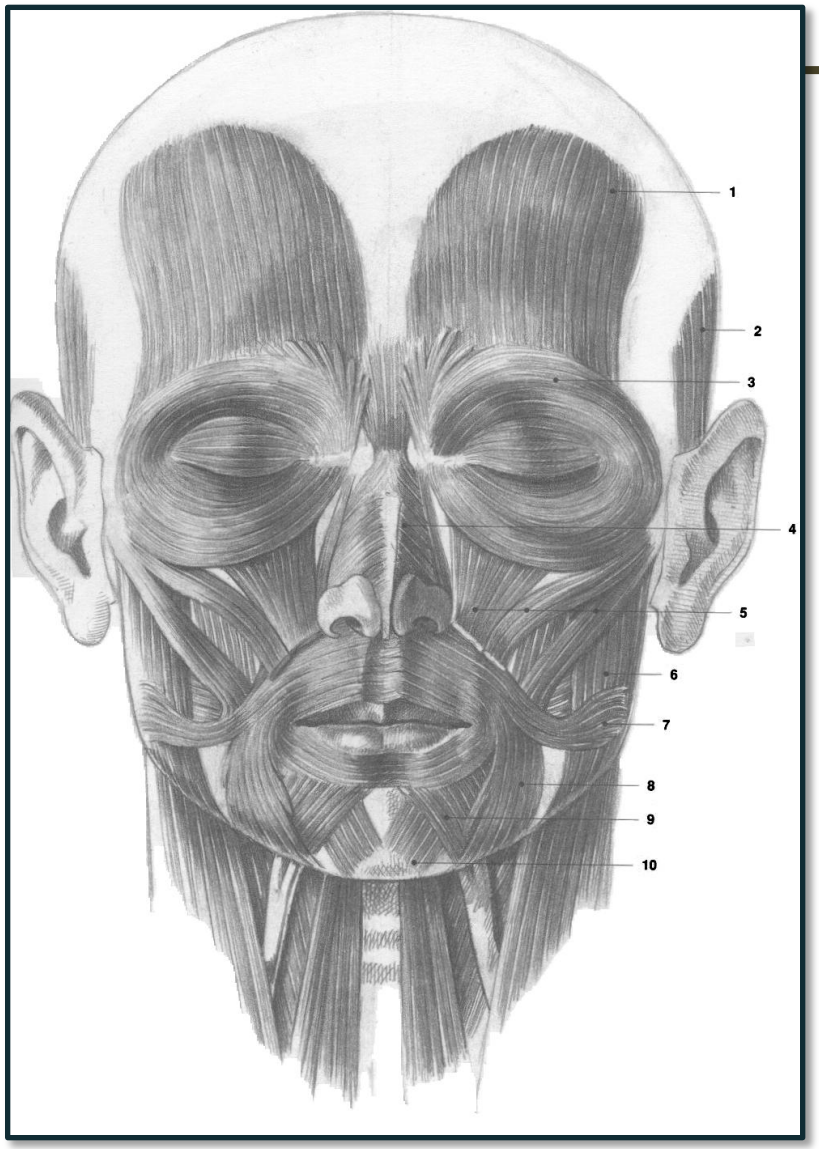


- Starile emotionale pot fi reprezentate pe un sistem cu doua coordonate: valenta si excitare

Axele Emotionale



Muschii fetei

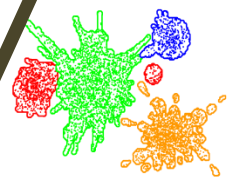


- Miscarile unui anumit muschi al fetei determina aparitia unei anumite *unitati de actiune (Action Unit – AU)*.
- O sau mai multe *unitati de actiune* participa la formarea unei expresii.

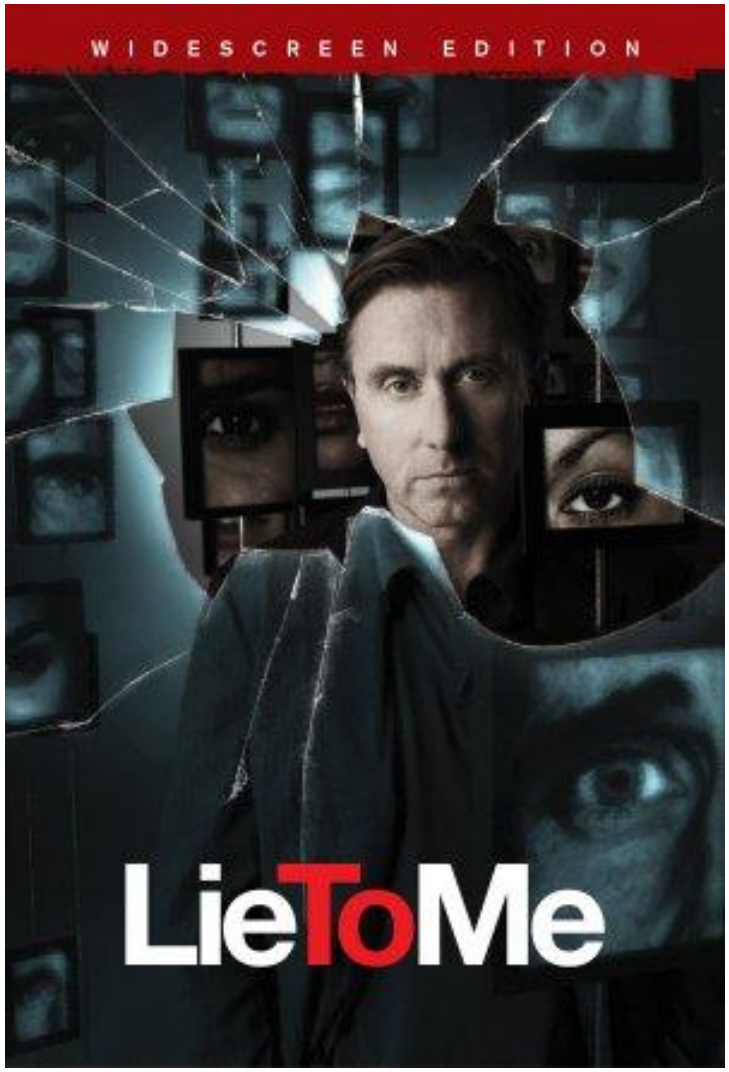
Facial Actions Coding System

Facial Action Coding System (FACS)

- Propus de catre Ekman si Friesen - 1976
- Descrie un set de 46 ‘Unitati de Actiune’ (Action Units - AU) ce corespund miscarilor elementare (independente) ale muschilor fetei
  - Unele corespund unui singur muschi
  - Majoritatea implica un set de muschi sau o miscare complexa
- Se refera la **miscare** (nu pozitii statice).



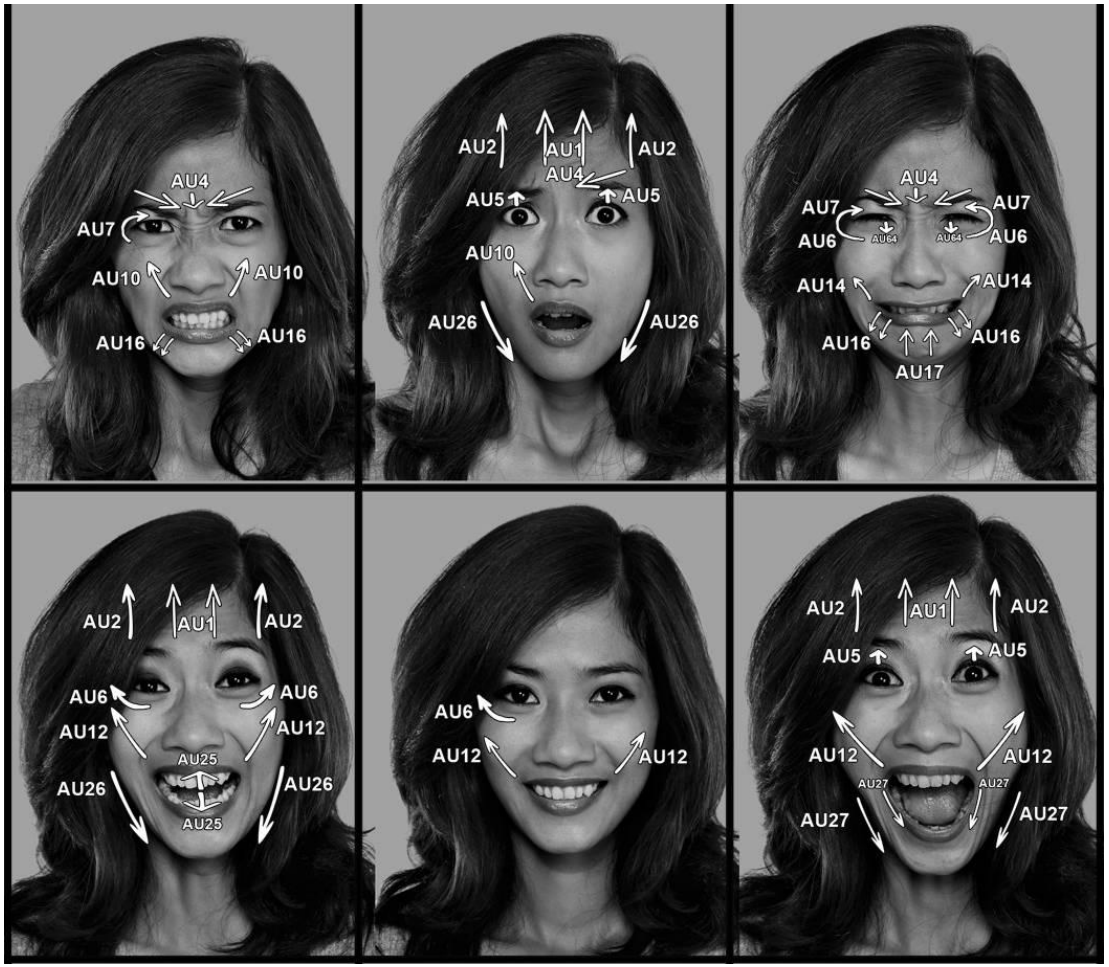
Paul Ekman


WIDESCREEN EDITION
LieToMe


F.A.C.E. Training
BY DR. PAUL EKMAN
Facial Expression. Awareness.
Compassion. Emotions.



FACS – exemple AU



AU1. Interiorul sprancenei se ridica

AU2. Exteriorul sprancenei se ridica

…

AU17. Barbia se ridica

…

AU19. Limba scoasa

AU20. Buza se intinde

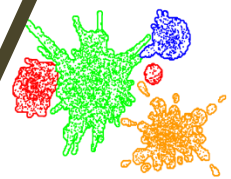
…

AU29. Mandibula iesita in afara

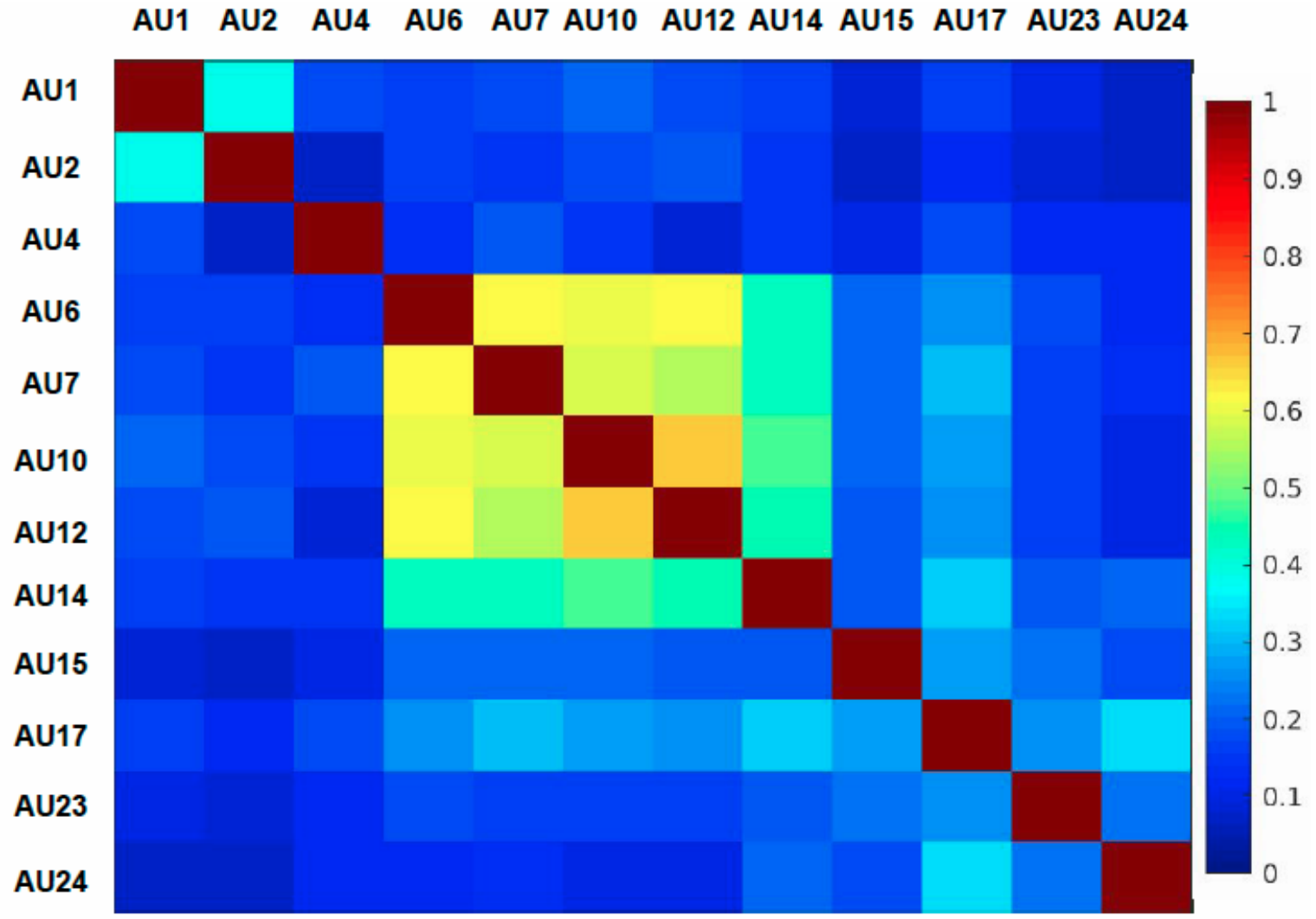
AU30. Mandibula in lateral

AU31. Falca inclestata

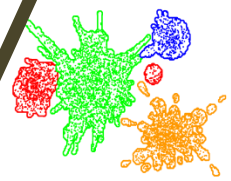
…



AU



[Onal Ertugrul2019]



AU in expresii asociate emotiilor

- Expresiile faciale sunt constituite din unitatile de actiune de baza
- Expresiile asociate emotiilor de baza au la baza un set de AU-uri universal valabil
- Ex: **Fericirea**:

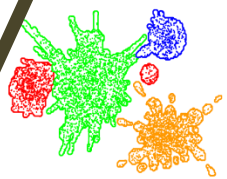
  AU1. Interiorul sprancenei se ridica (Frontalis, Pars Medialis)

  AU6. Pometele se ridica (Orbicularis Oculi, Pars Orbitalis)

  AU12. Coltul buzei se ridica (Zygomatic Major)

  AU14. Riduri la coltul gurii (Buccinator)

- Teorie simplista si controversata



AU in expresii asociate emotiilor

TABLE I
FACIAL EXPRESSION SYNTHESIS RULES AS PROPOSED IN [3]

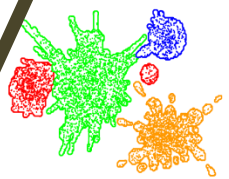
Expression	FAUs coded description [3]
Anger	$4+7+(((23 \text{ or } 24) \text{ with or not } 17) \text{or } (16+(25 \text{ or } 26)) \text{ or } (10+16+(25 \text{ or } 26)))$with or not 2
Disgust	((10 with or not 17) or (9 with or not 17)) + (25 or 26)
Fear	$(1+4)+(5+7)+20+(25 \text{ or } 26)$
Happiness	$6+12+16+(25 \text{ or } 26)$
Sadness	$1+4+(6 \text{ or } 7)+15+17+(25 \text{ or } 26)$
Surprise	$(1+2)+(5 \text{ without } 7)+26$

[Pantic2000]

Exemplu de combinatie de AU asociate emotiilor.

$$
\begin{aligned}
AU_6 + AU_{12} &= \text{Happiness} \\
AU_1 + AU_4 + AU_{15} &= \text{Sadness} \\
AU_1 + AU_2 + AU_{5B} + AU_{26} &= \text{Surprise} \\
AU_1 + AU_2 + AU_4 + AU_5 + AU_7 + AU_{20} + AU_{26} &= \text{Fear} \\
AU_4 + AU_5 + AU_7 + AU_{23} &= \text{Anger} \\
AU_9 + AU_{15} + AU_{16} &= \text{Disgust} \\
AU_{R12A} + AU_{R14A} &= \text{Contempt}
\end{aligned}
$$

[Ekman2005]



AU in expresii asociate emotiilor

Action units	*Hot Anger*	*Cold Anger*	*Panic Fear*	*Anxiety*	*Despair*	*Sadness*	*Elated Joy*	*Happiness*	*Disgust*
Inner brow raiser	0.31	0.31	0.94	0.63	0.88	0.56	0.44	0.31	0.25
Outer brow raiser	0.44	0.38	0.69	0.38	0.44	0.31	0.44	0.38	–
Brow lowerer	0.25	0.31	0.69	0.56	0.94	0.63	–	–	0.56
Upper lid raiser	0.31	–	0.50	0.19	0.25	–	0.19	–	–
Cheek raiser	0.13	–	–	–	0.44	–	0.63	0.81	0.38
Lid tightener	–	0.19	–	–	–	–	–	–	–
Nose wrinkler	0.13	–	–	–	–	–	–	–	0.13
Upper lip raiser	0.13	0.19	0.19	–	0.19	–	–	–	0.81
Lip corner puller	–	–	–	–	–	–	0.94	0.88	–
Lip corner depressor	–	–	–	–	–	0.13	–	–	0.19
Chin raiser	–	–	–	–	0.19	–	–	0.13	0.19
Lip stretcher	0.44	0.25	0.25	0.19	0.19	–	–	–	0.25
Lip tightener	0.19	0.19	–	–	–	–	–	–	–
Lips part	0.13	0.31	0.13	0.19	0.25	0.50	–	0.31	0.44
Jaw drop	0.38	0.56	0.56	0.63	0.44	0.19	0.81	0.56	0.38

Numerical data indicate the proportion of cases in which actors displayed activation for each action unit or emotion combination. A dash indicates that the respective AU was not used by any actor. Shading indicates theoretically predicted patterns of activation: light shading = increased activation; dark shading = greatly increased activation (as compared with baseline).

Trasaturi relevante expresiilor

- Analiza faciala bazata pe trasaturi trebuie sa tina cont de:
  - **Localizare** – identificarea trasaturilor de interes
  - **Intensitate** – se determina:
    - Deformarea geometrica a trasaturilor
    - Cantitatea de riduri provocata
  - **Dinamica** - in general detectia pe imagini statice este dificila
-

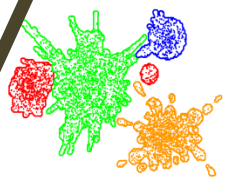


Dinamica trasaturilor

- Trasaturile fetei sunt:
  - Trecatoare :
    - Descriu emotiile
    - Localizate in jurul ochilor, sprancenelor, gurii, pometilor
  - Permanente
    - De ex: parul, cicatrici, textura pielii
- Deformarea trasaturilor trecatoare releva emotiile

Emotiile si expresiile sunt **dinamice**

Baseline – neutru ar ajuta



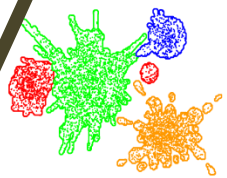
Generare imagini

Achiziții în laborator:

- Expresii simulate
  - Actori – scumpi
  - Persoane obișnuite – simulare mai puțin calitativă
- Fundal cu variție limitată
- Posturi limitate

Imagini *in the wild*

- Varietate mare
- Trebuie adnotate
  - Specialiști – rari deci scumpi
  - Persoane obișnuite ???



Adnotarea imaginilor cu expresii

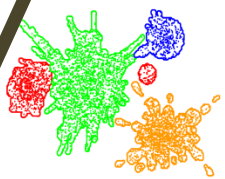
Expresii și AU

- 89.2% - 6 expresii - 23 studenti la Psihologie [S2007]
- 100 ore antrenament pt 70% in recunoaștere de AU
- Oamenii obișnuiți sunt sub 70% pe FACS
- 4-5 experți pentru 90-95%
- 40 -50 oameni antrenati pt 95%

Clase generale:

- CIFAR-10 94% - Andrey Karpathy
- 8 clase Catech 101 – 99% [D2017]

[Susskind2007]
[Dodge2017]



Baze de date publice

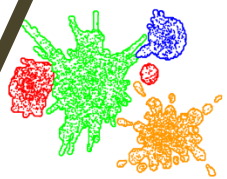
Expresii

- **FER2013/Fer+**
  - 40k imag, gri
  - 10 adnotatori - crowd
- **RafDB**
  - 12k img color
- **AffectNet**
  - Specialiști

AU (Adnotare cu Experți)**:**

- **CK +**
- 100, Simulate, exagerate
- **DISFA**
- Veridice, Laborator, intesitati
- **EmotioNet**
- Adnotate 50k
- **Bu4de**
- **BP4D**

MegaFace - 1 M neadnotate



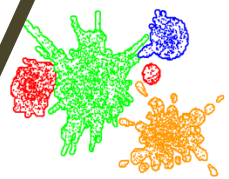
Date de test

		RGB										
		CK+	MPIE	JAFFE	MMI	RU_FACS	SEMAINE	CASME	DISFA	AFEW	SFEW	AMFED
Content	Intention(Posed/Spontaneous)	*P*	*P*	*P*	*P*	*S*	*S*	*S*	*S*	*S*	*S*	*S*
	Label(Primary/AU/DA)	*P/AU*	*P*	*P*	*AU + T*	*P/AU*	*P/AU/DA*[1]	*P/AU*	*AU + I*	*P*/[2]	*P*	*P/AU/Smile*
	Temporality(Static/Dynamic)	*D*	*S*	*S*	*D*	*D*	*D*	*D*	*D*	*D*	*S*	*D*
Capture	Environment(Lab/Non-lab)	*L*	*L*	*L*	*L*	*L*	*L*	*L*	*L*	*N*	*N*	*N*
	Multiple Perspective	○	●	○	●	●	○	○	○	●	●	●
	Multiple Illumination	○	●	○	●	○	○	○	○	●	●	●
	Occlusions	○	●	○	○	●	○	○	○	●	●	●
Subjects	# of subjects	201	337	10	75	100	150	35	27	220	68	5268
	Ethnic Diverse	●	●	○	●	○	○	○	●	●	●	●
	Gender(Male/Female(%))	31/69	70/30	100/0	50/50	-	62/38	37/63	44/56	-	-	58/42
	Age	18-50	$\mu = 27.9$	-	19-62	18-30	22-60	$\mu = 22$	18-50	1-70	-	-

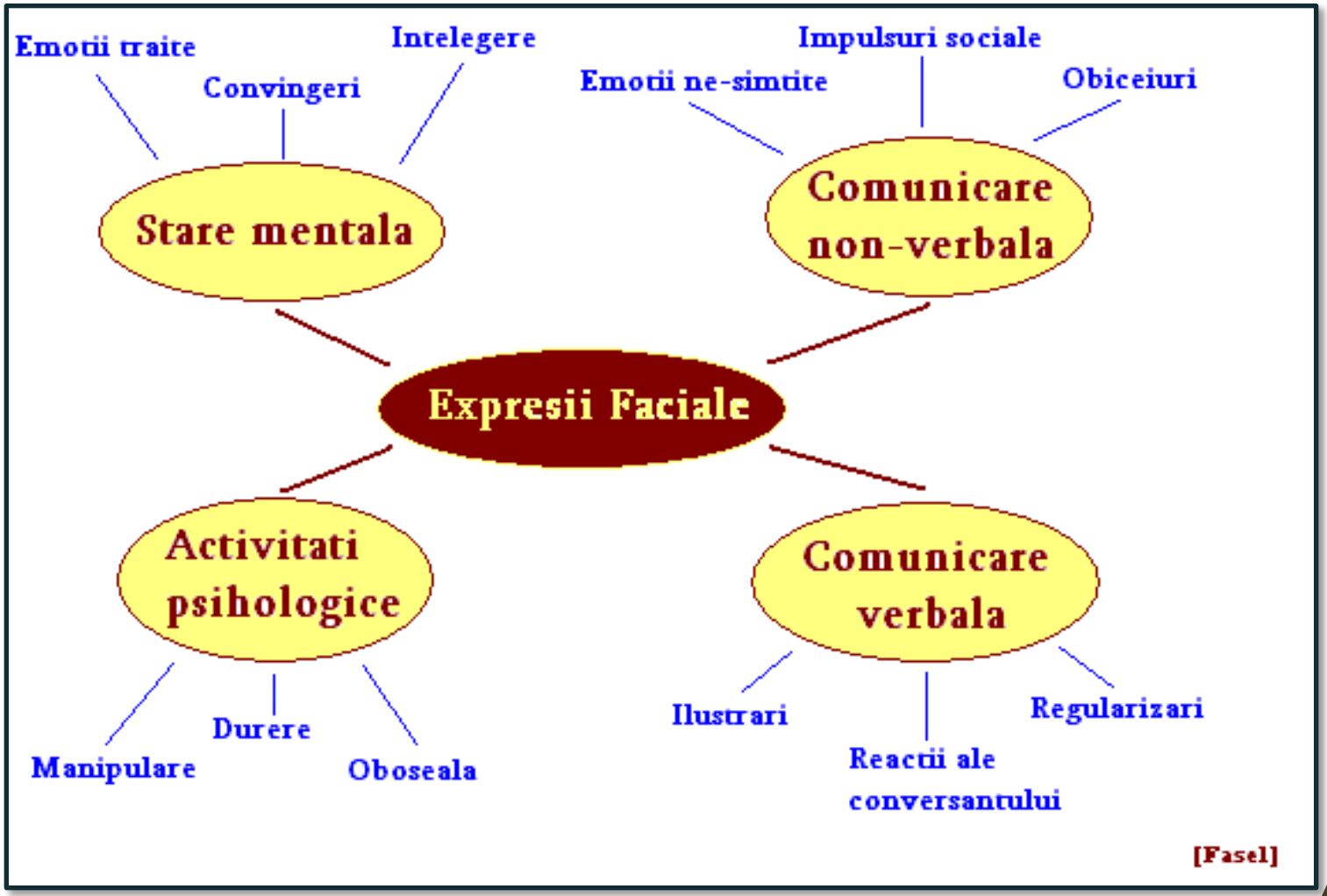
= Yes, ○ = No, - = Not enough information. DA: Dimensional Affect, I = Intensity labelling, T = Temporal segments. [1] Other labels include Laughs, Nods, Epistemic states(e.g. Certain, Agreeing, Interested etc.) etc. Refer to original paper for details [205]. [2] Pose, Age, Gender. Refer to original paper for details [203].

		3D				RGB+Thermal			
		BU-3DFE	BU-4DFE	Bosphorus	BP4D	IRIS	NIST	NVIE	KTFE
Content	Intention(Posed/Spontaneous)	*P*	*P*	*P*	*S*	*P*	*P*	*S/P*	*S/P*
	Label(Primary/AU)	*P + I*	*P*	*P/AU*	*AU*	*P*	*P*	*P*	*P*
	Temporality(Static/Dynamic)	*S*	*D*	*S*	*D*	*S*	*S*	*D*	*D*
Capture	Environment(Lab/Non-lab)	*L*	*L*	*L*	*L*	*L*	*L*	*L*	*L*
	Multiple Perspective	●	●	-	●	●	●	●	●
	Multiple Illumination	○	○	○	○	●	●	●	●
	Occlusions	●	○	●	○	●	●	●	●
Subjects	# of subjects	100	101	105	41	30	90	215	26
	Ethnic Diverse	●	●	○	●	●	-	○	○
	Gender(Male/Female(%))	56/44	57/43	43/57	56/44	-	-	27/73	38/62
	Age	18-70	18-45	25-35	18-29	-	-	17-31	12-32

= Yes, ○ = No, - = Not enough information, I = Intensity labelling.



Producerea expresiilor



Inducerea emotiilor

Tristete:

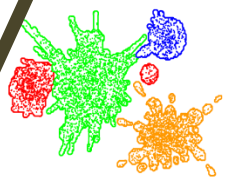
- Poveste care evocă simpatia
- Muzică de fundal trista;
- Jucărie cu aspect sumbru;
- Iluminare albastra

Furie:

- Poveste care subiecţii sunt inselati spusa pe un ton sarcastic;
- Jucărie cu aspect neplăcut;
- Iluminare rosie

Surpriză:

- Creşterea bruscă a volumului de muzica de fundal,
- Un sunet intermitent de sonerie şi de sticlă sparta



Estimarea intensitatii durerii

No Pain			
Pain			

The Prkachin – Solomon Pain Score (PSPI)

Pain = AU4 + max (AU6,AU7) + max (AU9,AU10) + AU43

AU4- Brow Lowerer



AU6 - Cheek Raiser



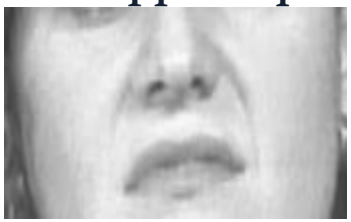
AU7 - Lid Tightener



AU9 - Nose Wrinkler

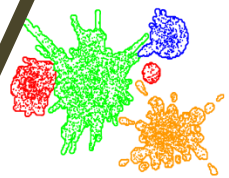


AU10 - Upper Lip Raiser



AU43- Eyes Closed





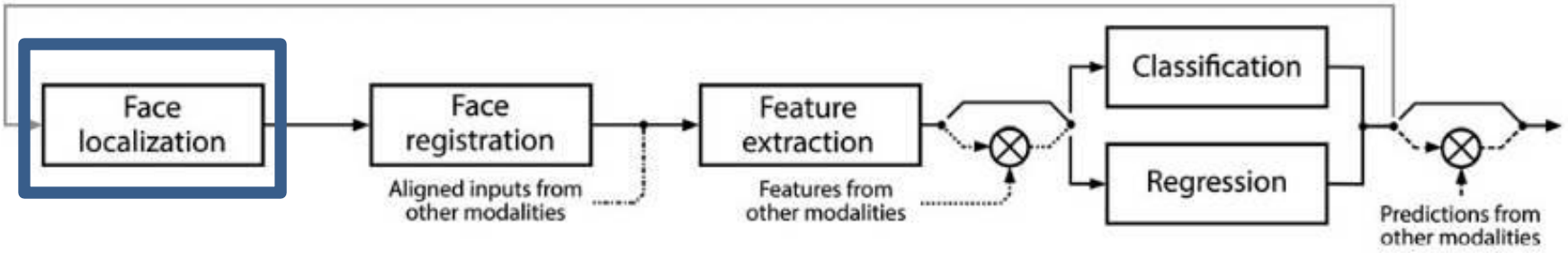
K. Prkachin, P. P. Solomon, The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain, Pain 139 (2008) 267–274.
http://www.cs.cmu.edu/~face/facs.htm

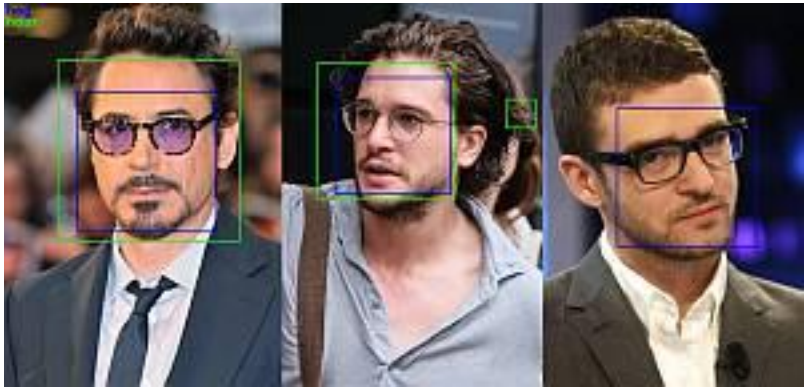
Probleme conexe:Microexpresii

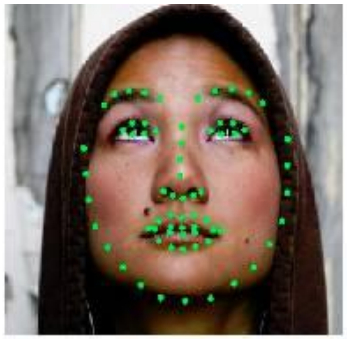


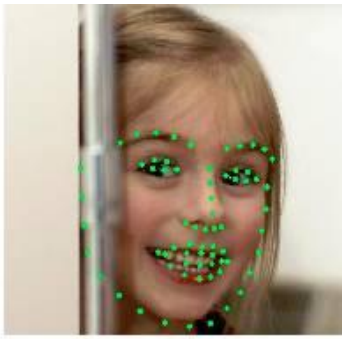
- Foarte scurte: 1/25 secunde
- Mereu veridice
- Sunt considerate “leakage” cand se incearca ascunderea unei emotii

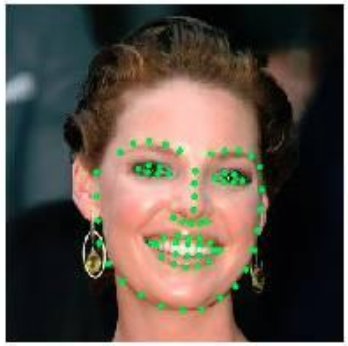
Pipeline standard





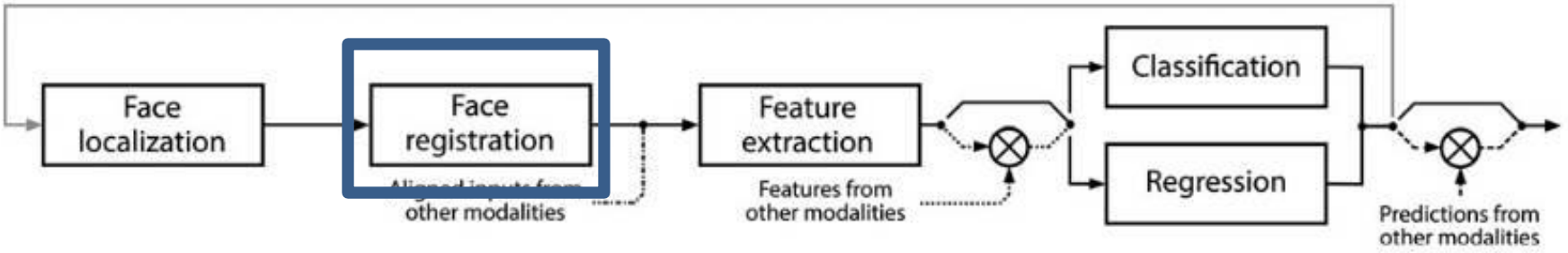


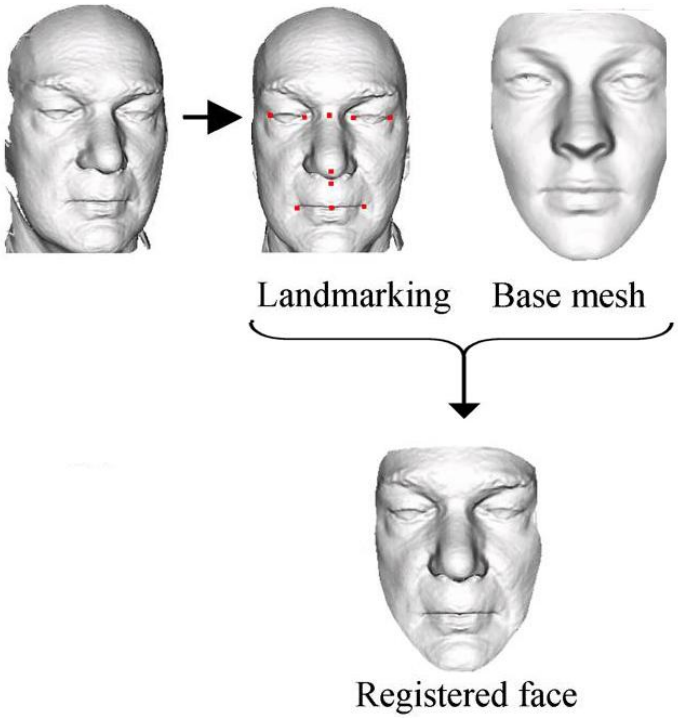


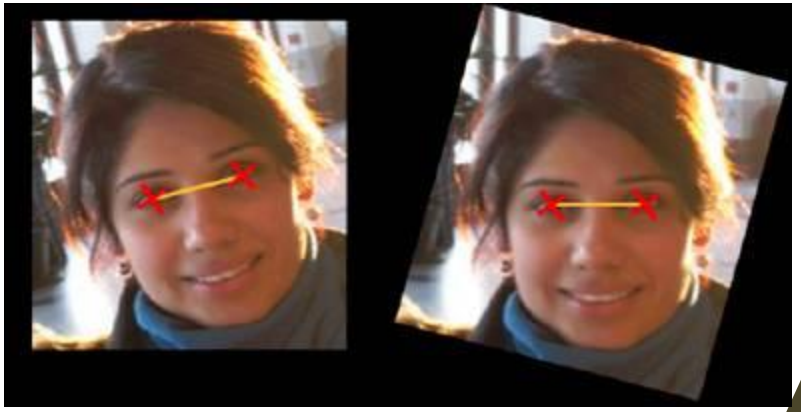




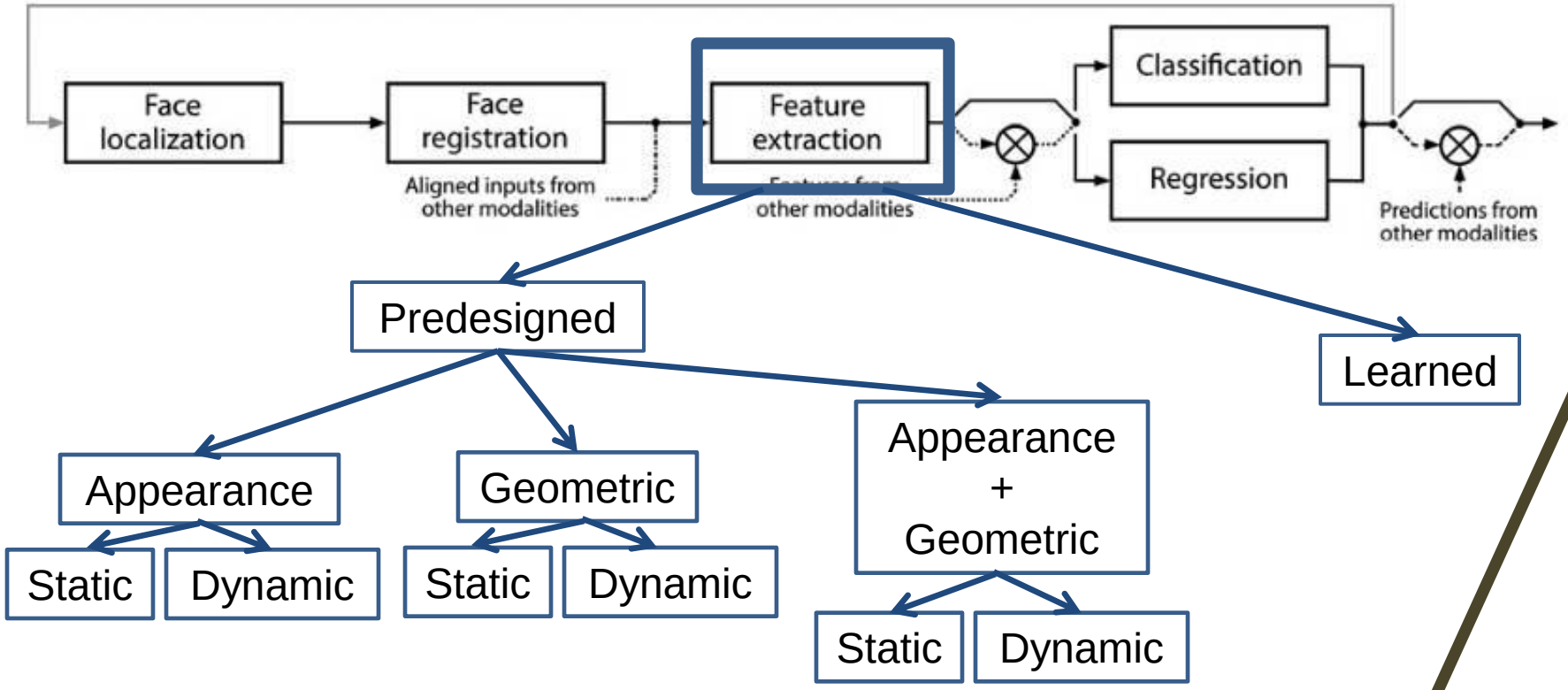
Pipeline standard





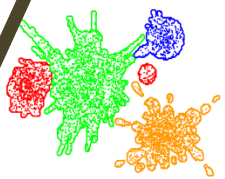


Pipeline standard



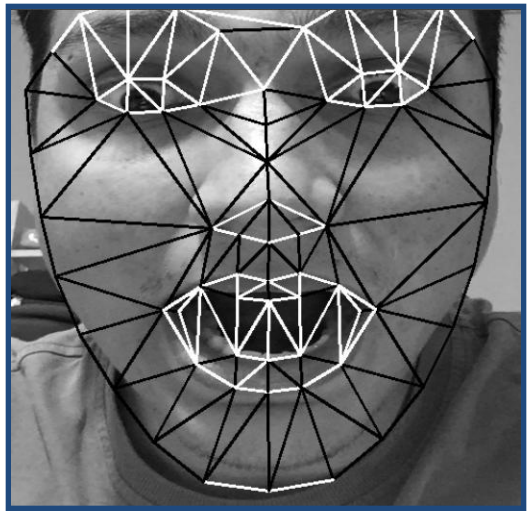
Trasaturi

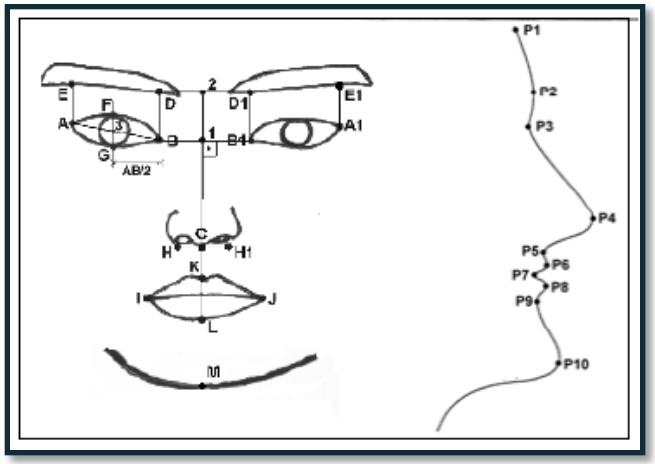
- Predesigend
  - Alese manual prin experimente
  - Nu necesita f. multe exemple
  - Munca manuala pentru alegerea parametrilor
- Learned
  - Alese automat prin invatare
  - Necesita f. multe exemple pt a nu se ajunge la overfitting
  - Performanta mai buna
  - Alegerea arhitecturii poate fi problematica

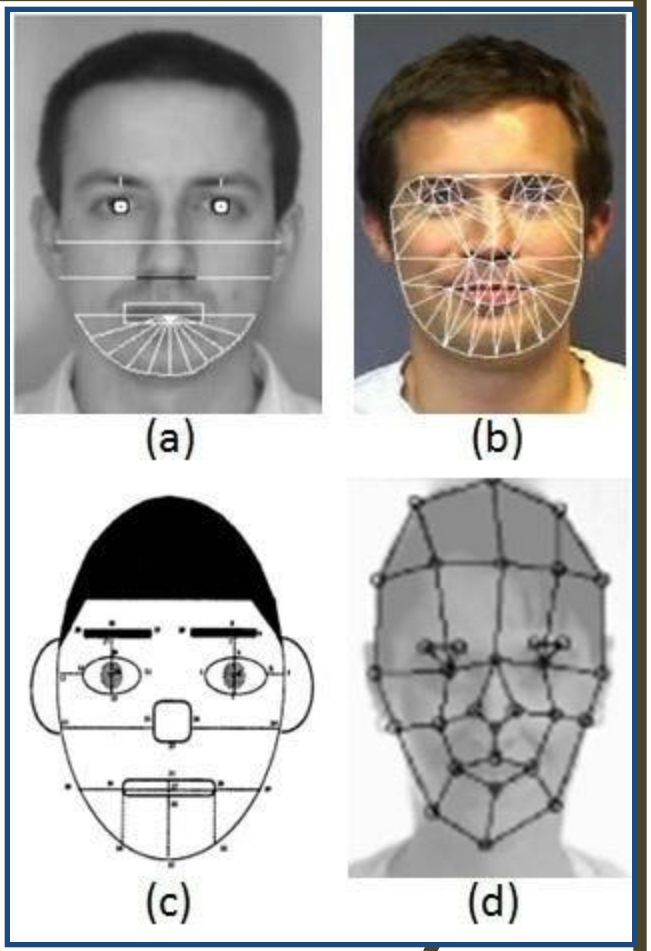


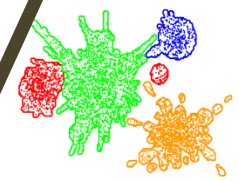
Trasaturi geometrice

- Distante intre perechi de puncte potrivit alese
- Unghiuri intre triplete de puncte





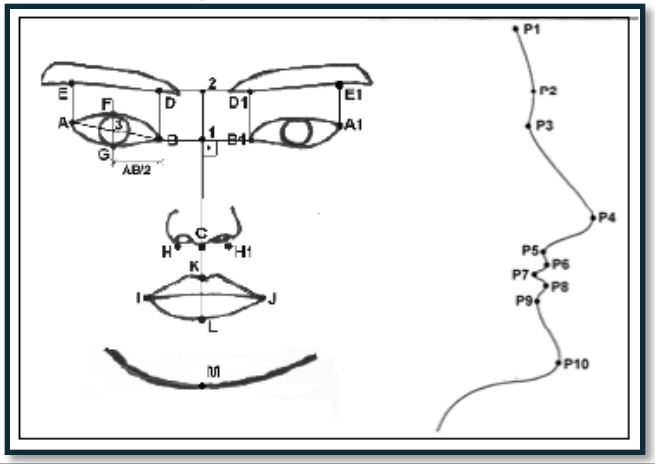




Exemplu: Trasaturi geometrice

[Pantic2000]

- Masuratori geometrice din puncte de control de pe conturul sprancenelor, ochilor, narilor, gurii si barbiei



AU	FACS description	Mapped on model
1	Raised inner brows	increased f1 & f2
2	Raised outer brow	increased f1 or f2
4	Lowered brows or Frowned brows	non-increased f20, (decreased f1 & f2) or increased f19
5	Raised upper lid	increased f5 or f6
6	Raised cheek	increased f21 or f22
7	Raised lower lid	non-increased f20, non-increased f21, non-increased f22, f9 > 0, f10 > 0, f5 > 0, f6 > 0, decreased f7 or f8

Feature	Feature description
f1	Angle ∠BAD
f2	Angle ∠B1A1D1
f3	Distance AE
f4	Distance A1E1
f5	Distance 3F, 3 is the centre of AB
f6	Distance 4F1, 4 is the centre of A1B1
f7	Distance 3G
f8	Distance 4G1
f9	Distance FG
f10	Distance F1G1
f11	Distance CK, C is 0.5HH1 (f0)
f12	Distance IB
f13	Distance JB1
f14	Distance CI
f15	Distance CJ
f16	Distance IJ
f17	Distance KL
f18	Distance CM
f19	Image intensity in circle (r(0.5BB1), C(2)) above line (D, D1)
f20	Image intensity in circle (r(0.5BB1), C(2)) below line (D, D1)
f21	Image intensity in circle (r(0.5AB), C(A)) left from line (A, E)
f22	Image intensity in circle (r(0.5A1B1), C(A1)) right from line (A1, E1)
f23	Image intensity in the left half of the circle (r(0.5BB1), C(I))
f24	Image intensity in the right half of the circle (r(0.5BB1), C(J))
f25	Brightness distribution along the line (K, L)

Exemplu: Trasaturi geometrice

[Kotsia2007]

- Foloseste analiza temporala – urmareste modul in care se misca punctele pe fata umana

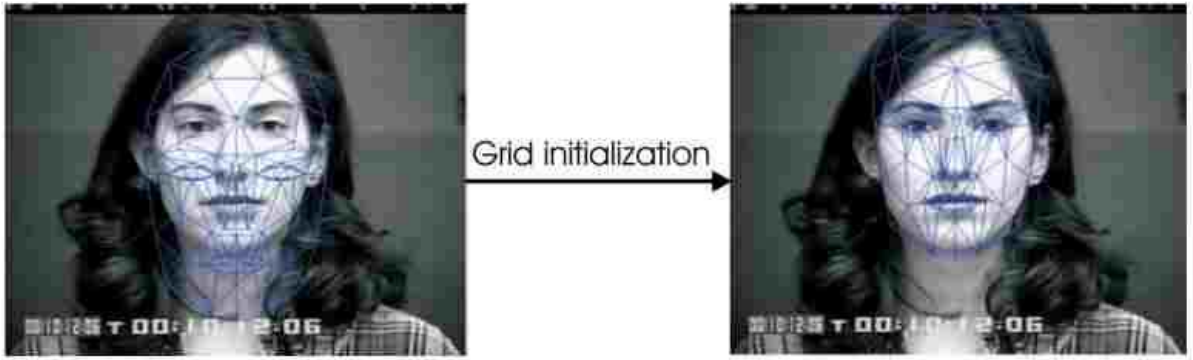
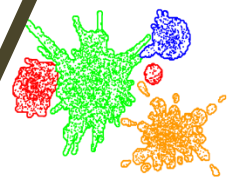
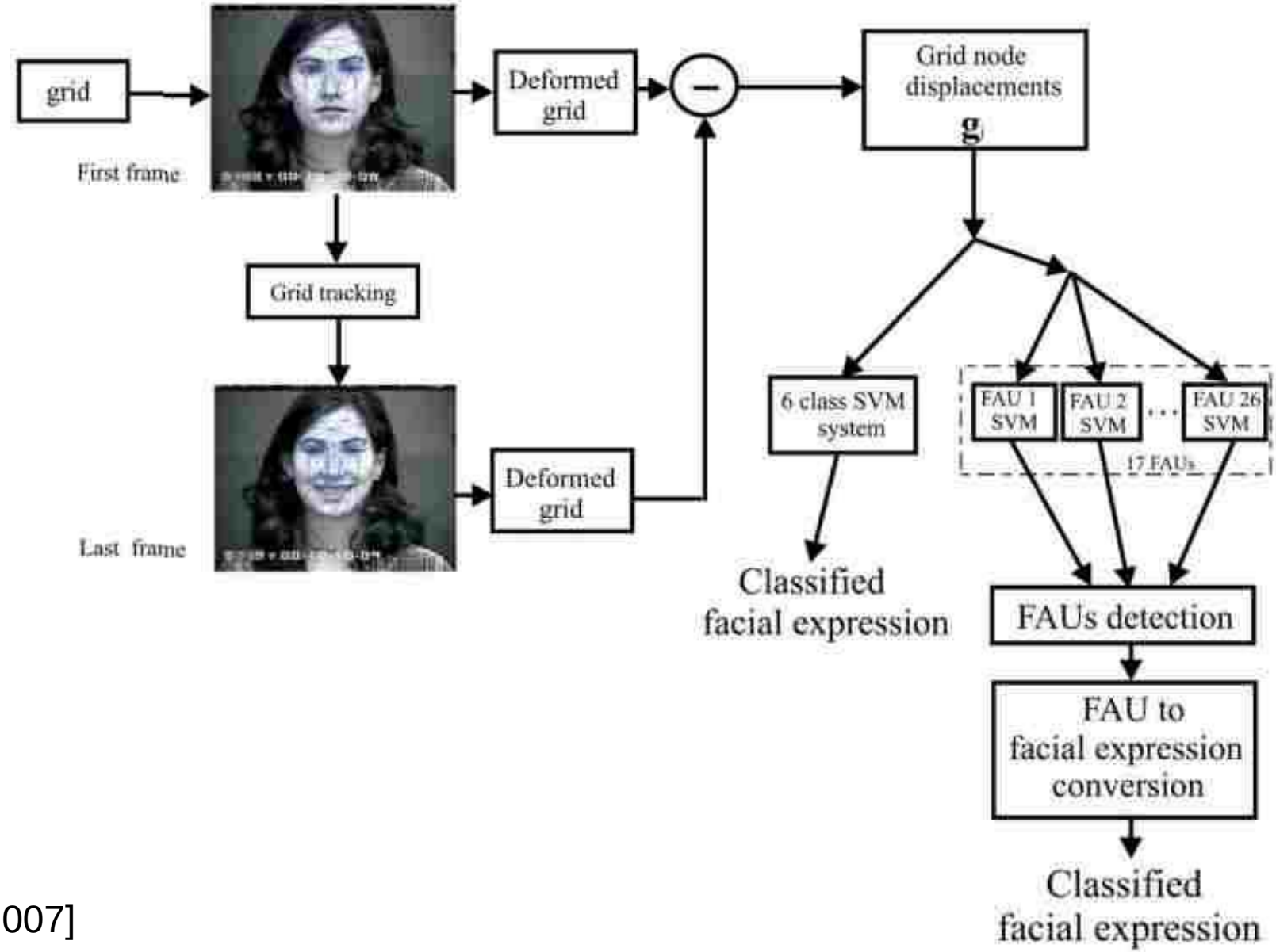


Fig. 2. Result of initialization procedure when seven Candide nodes are placed by the user on a facial image.

- Grid initializat de utilizator
- Punctele de interes sunt urmarite cu un tracker de tip Kanade–Lucas–Tomasi
- Deformarile sunt clasificate cu SVM (2-class modificat sau multiclass)
- Trasaturi – distante intre puncte (208 dimensiuni)





[Kotsia2007]

Fig. 1. System architecture for facial expression recognition in facial videos.

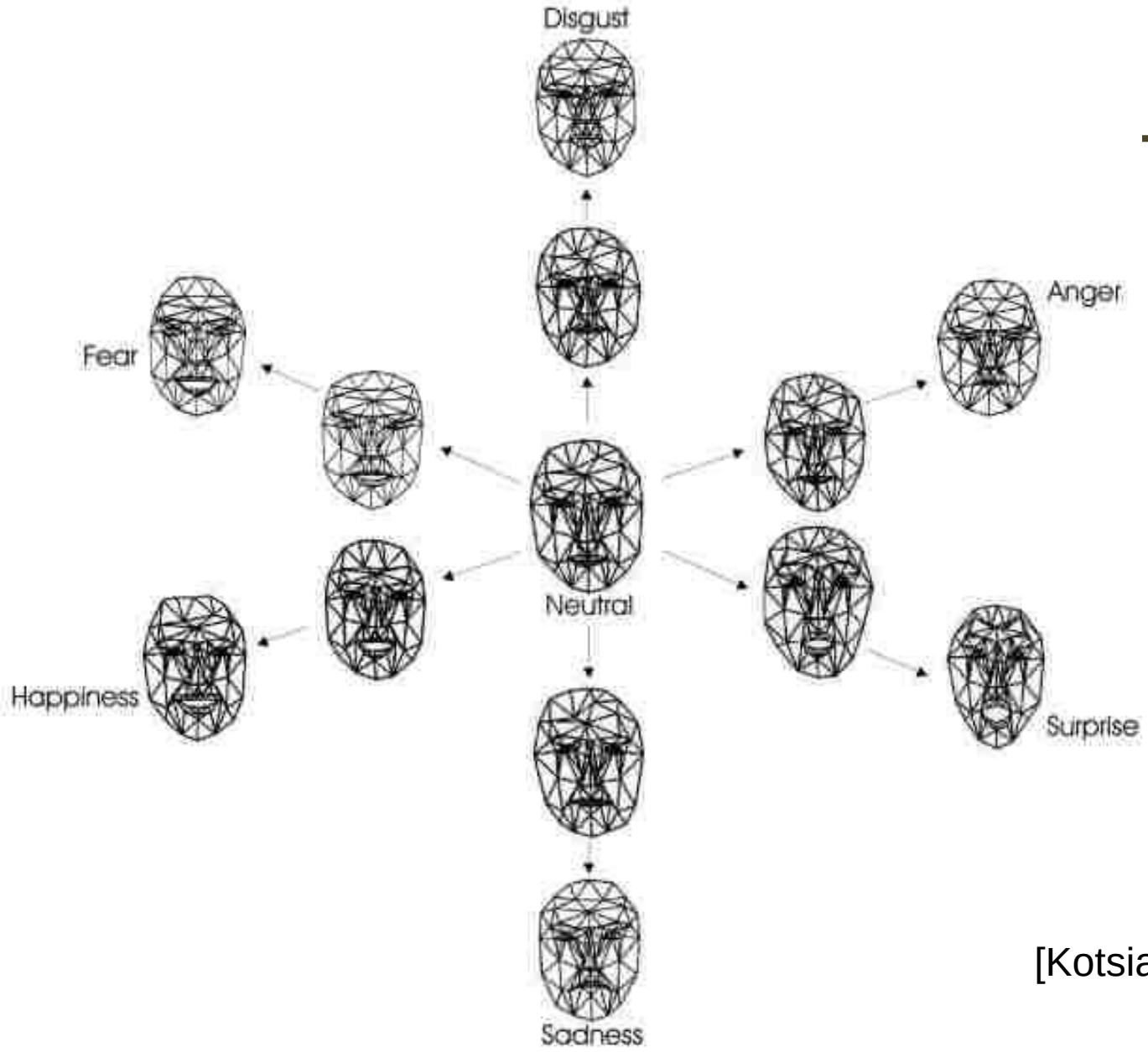
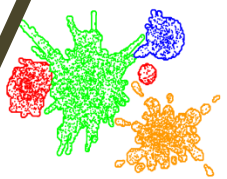
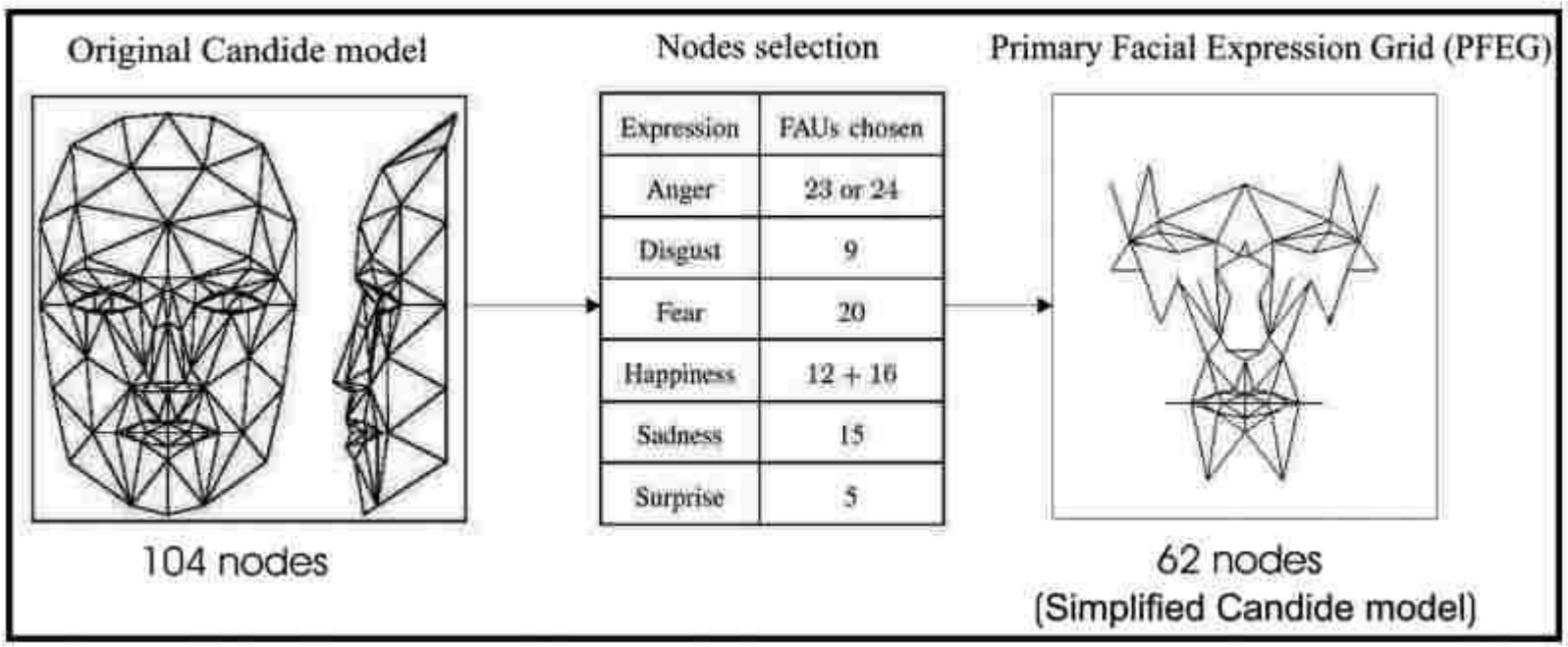


Fig. 3. Example of the deformed Candide grids for each one of the six facial expressions.

[Kotsia2007]

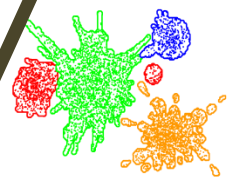


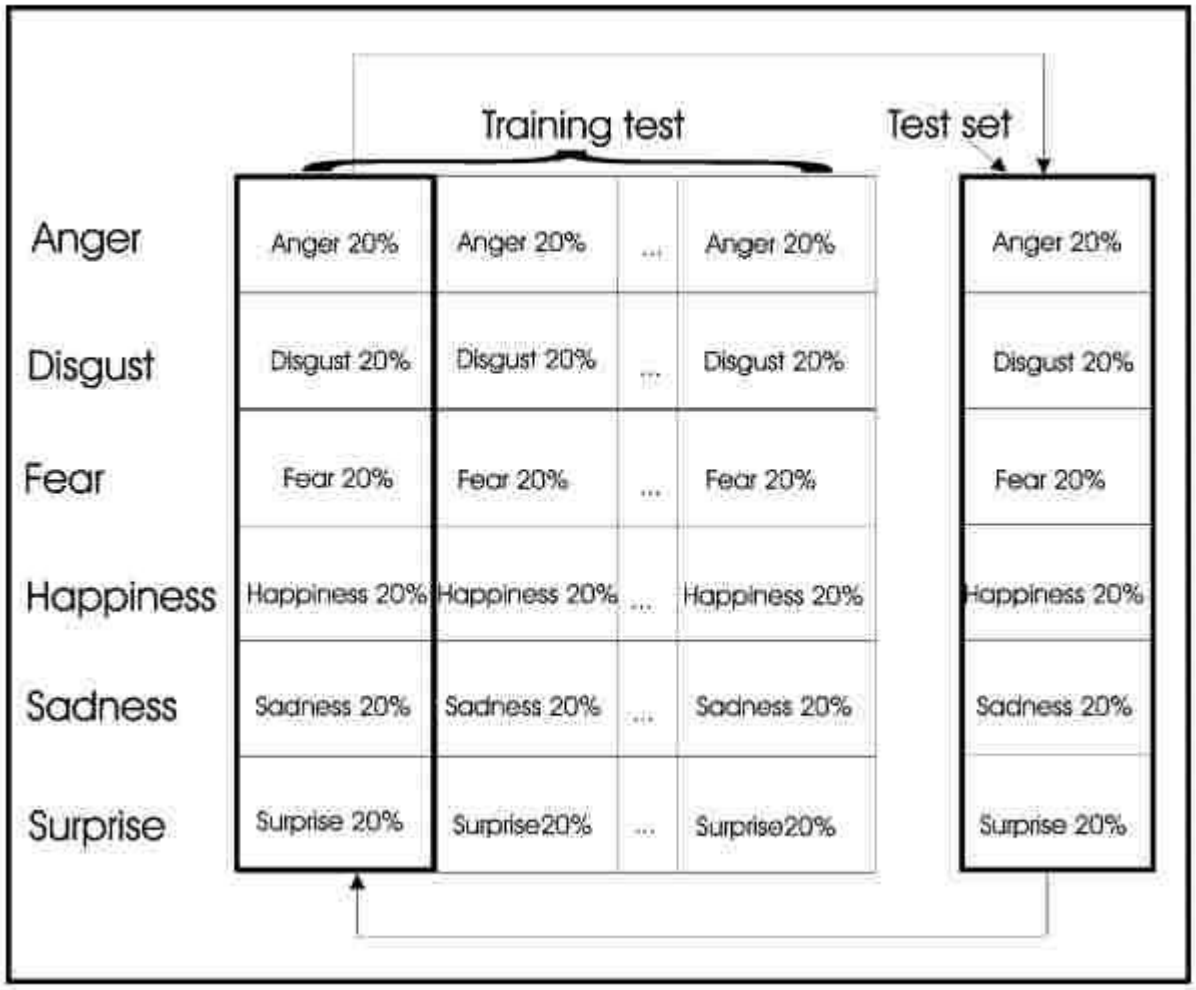


Expression	FAUs chosen
Anger	23 or 24
Disgust	9
Fear	20
Happiness	12 + 16
Sadness	15
Surprise	5

Fig. 7. PFEG according to FACS, used for the experiments.

[Kotsia2007]





- Testare pe CK
- ! Aceeasi persoana poate fi si la antrenare si la testare

[Kotsia2007]

[Kotsia2007]

TABLE II
CONFUSION MATRICES FOR FAU DETECTION-BASED FACIAL EXPRESSION RECOGNITION WHEN USING THE CANDIDE AND PFEG GRID AND THE PROPOSED SET OF FAU RULES

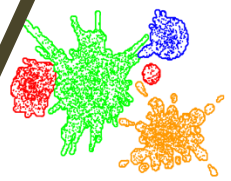
$lab_{ac} \backslash lab_{cl}$	an	di	fe	ha	sa	su
an	80%	9.6%	0%	0.3%	5.5%	2.2%
di	6.7%	77%	0.3%	0.8%	1.5%	1.5%
fe	0%	3%	98.8%	3.2%	0%	1.9%
ha	0%	0%	0.9%	95.4%	0%	0%
sa	13.3%	8.9%	0%	0.3%	93%	0%
su	0%	1.5%	0%	0%	0%	94.4%

$lab_{ac} \backslash lab_{cl}$	an	di	fe	ha	sa	su
an	91.3%	9.5%	0%	0.3%	4.4%	1.9%
di	2.7%	80%	0.3%	0.5%	0.7%	1.1%
fe	0%	3%	99.1%	0.8%	0%	1.5%
ha	0%	0%	0.6%	98.1%	0%	0%
sa	6%	6%	0%	0.3%	94.9%	0%
su	0%	1.5%	0%	0%	0%	95.5%

TABLE III
CONFUSION MATRICES FOR FACIAL EXPRESSION RECOGNITION USING MULTICLASS SVMS WHEN APPLYING THE MODIFIED SVMS TO THE CANDIDE AND PFEG GRID

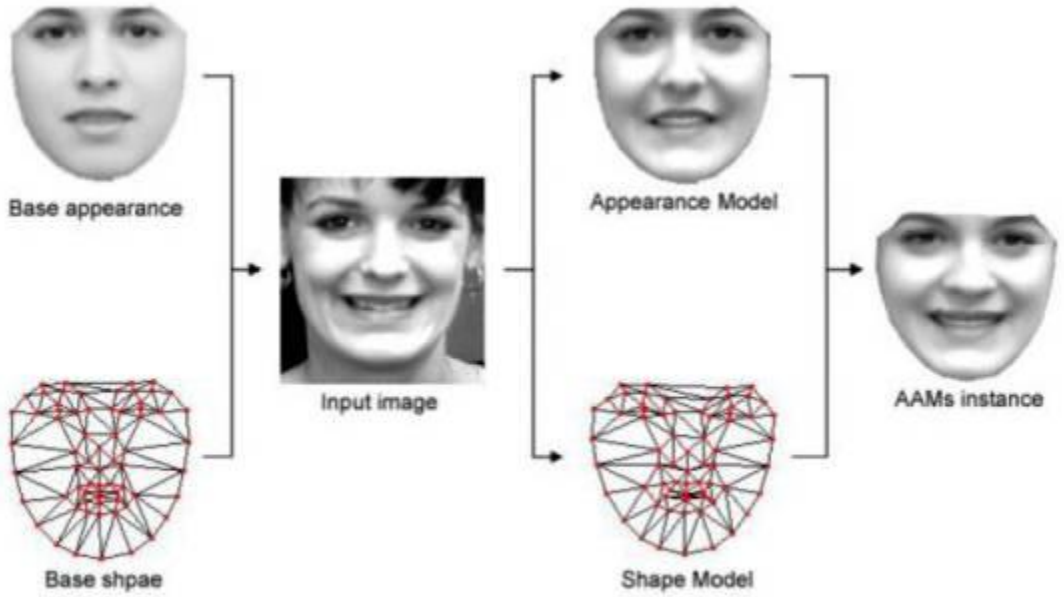
$lab_{ac} \backslash lab_{cl}$	an	di	fe	ha	sa	su
an	81.3%	0%	0%	0%	0%	0%
di	4.7%	100%	0%	0%	0%	0%
fe	0%	0%	100%	0%	0%	0%
ha	0%	0%	0%	100%	0%	0%
sa	14%	0%	0%	0%	100%	0%
su	0%	0%	0%	0%	0%	100%

$lab_{ac} \backslash lab_{cl}$	an	di	fe	ha	sa	su
an	96.7%	0%	0%	0%	0%	0%
di	0%	100%	0%	0%	0%	0%
fe	0%	0%	100%	0%	0%	0%
ha	0%	0%	0%	100%	0%	0%
sa	3.3%	0%	0%	0%	100%	0%
su	0%	0%	0%	0%	0%	100%



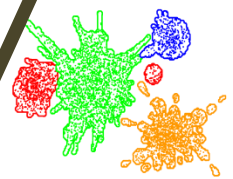
Trasaturi de aparenta

- Modele deformabile ale trasaturilor – **ASM, AAM**



'igure 2: Generation of an AAM instance for an input image

- Proiectii integrale, Haar, LBP, VLBP-TOP, LBP-TOP, HOG, Gabor, etc



Exemplu: LBP

[Shan2009]

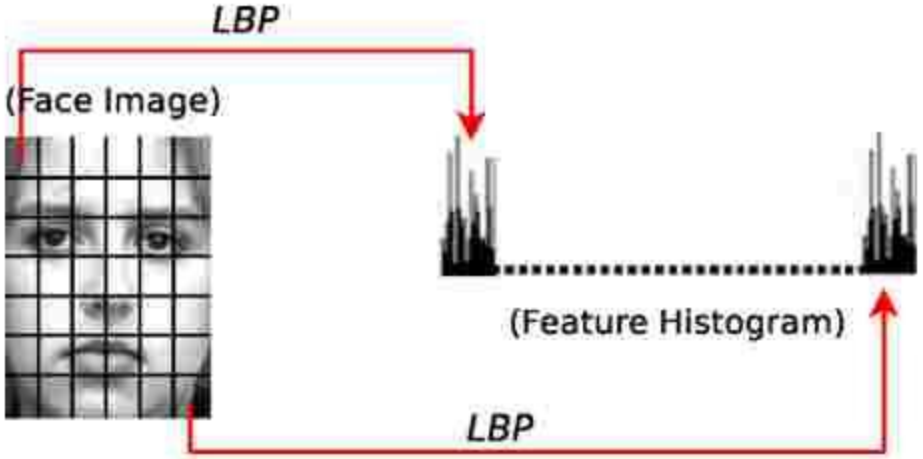


Fig. 6. A face image is divided into small regions from which LBP histograms are extracted and concatenated into a single, spatially enhanced feature histogram.

- Fata e impartita in regiuni
- Pe fiecare regiune se calculaza histograma LBP
- Trasaturi = Histogramele concatenate si ponderate

- Se clasifica cu:
  - -NN fata de trasatura medie a clasei
  - -SVM

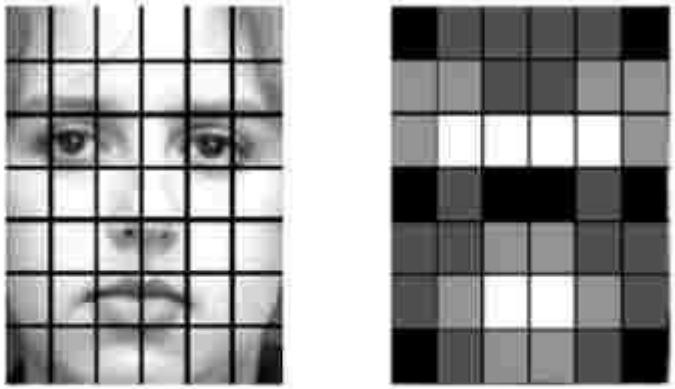


Fig. 7. (Left) A face image divided into 6×7 sub-region. (Right) The weights set for weighted dissimilarity measure. Black squares indicate weight 0.0, dark gray 1.0, light gray 2.0 and white 4.0.

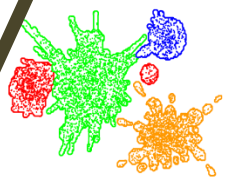


Table 2
Confusion matrix of 7-class facial expression recognition using template matching with LBP features

	Anger (%)	Disgust (%)	Fear (%)	Joy (%)	Sadness (%)	Surprise (%)	Neutral (%)
Anger	**58.7**	5.5	0	0	26.7	0	9.1
Disgust	3.3	**85.0**	2.5	0	2.5	0	6.7
Fear	1.0	0	**61.7**	24.0	10.3	0	3.0
Joy	0	0	6.0	**90.4**	0	0	3.6
Sadness	4.9	0	0	0	**72.4**	1.7	21.0
Surprise	0	0	1.3	0	2.7	**92.4**	3.6
Neutral	2.0	0.8	0.4	0.8	25.7	0	**70.3**

Testare pe CK

! Person-independent

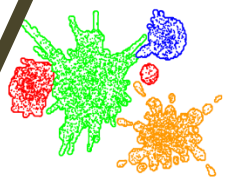
Table 3
Recognition performance of LBP-based SVM with different kernels

	6-Class recognition (%)	7-Class recognition (%)
SVM (linear)	91.5 ± 3.1	88.1 ± 3.8
SVM (polynomial)	91.5 ± 3.1	88.1 ± 3.8
SVM (RBF)	92.6 ± 2.9	88.9 ± 3.5

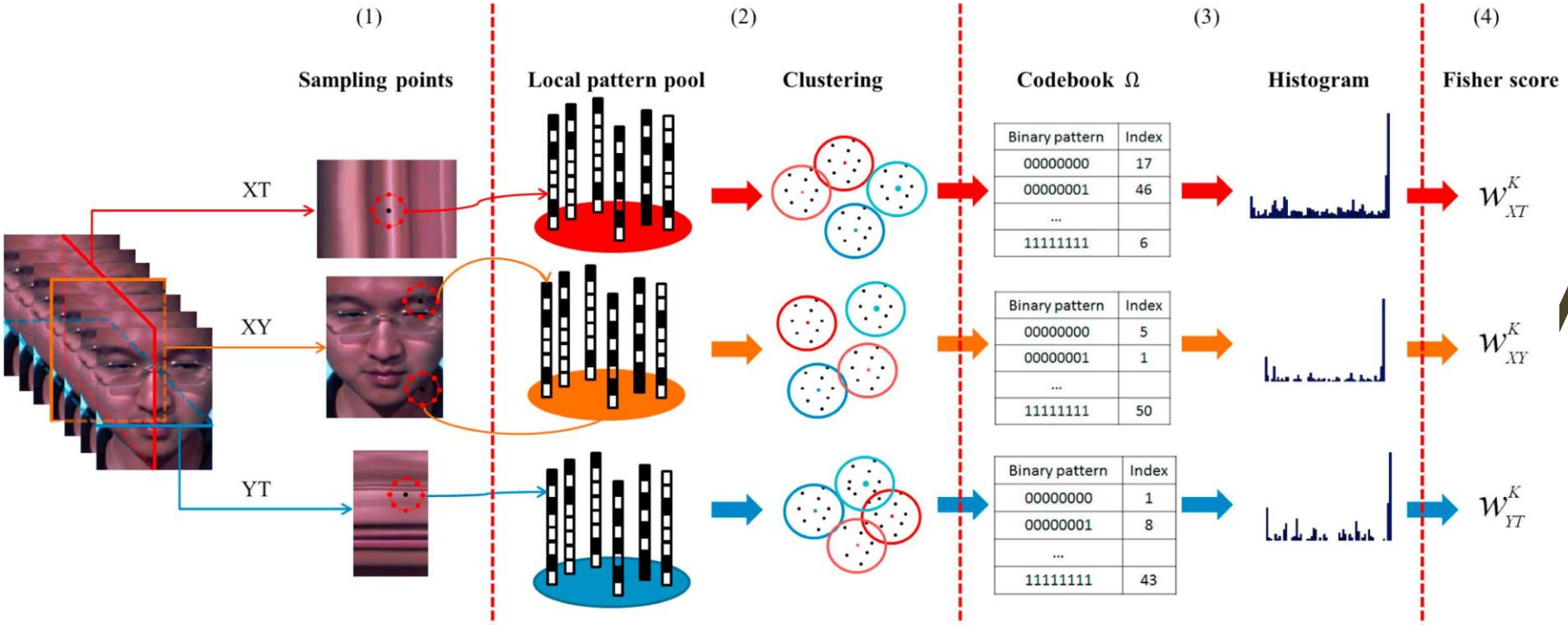
Table 5
Confusion matrix of 7-class facial expression recognition using SVM (RBF)

	Anger (%)	Disgust (%)	Fear (%)	Joy (%)	Sadness (%)	Surprise (%)	Neutral (%)
Anger	**85.0**	2.7	0	0	4.8	0	7.5
Disgust	0	**97.5**	2.5	0	0	0	0
Fear	0	2.0	**68.0**	22.0	1.0	0	7.0
Joy	0	0	0.7	**94.7**	1.1	0	3.5
Sadness	8.6	0	0	0	**69.5**	2.3	19.6
Surprise	0	0	1.3	0	0	**98.2**	0.5
Neutral	1.6	0.4	0	1.6	6.0	0.4	**90.0**

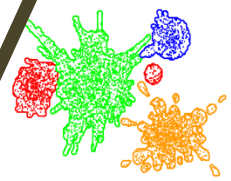
[Shan2009]



LBP temporal: LBP - TOP



[Huang2015-2]



LBP temporal: LBP - SIP

Reducere de informatie redundanta din LBP-TOP

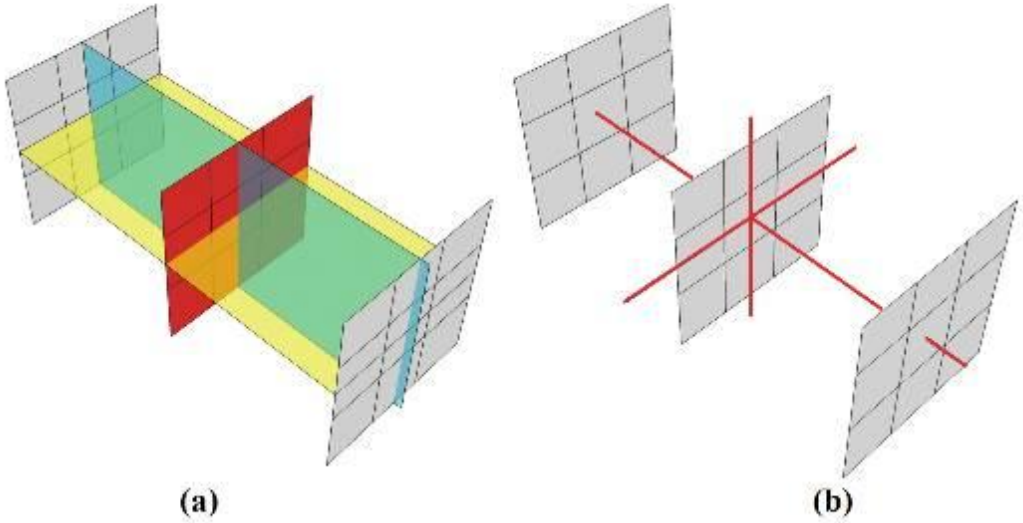
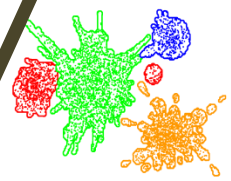


Fig. 1. (a) Three orthogonal planes in the LBP-TOP computation, and (b) Three intersecting lines crossing over the center pixel, formed by the three orthogonal planes, are considered to obtain the six distinct neighbour points surrounding the center point.

SVM pentru clasificare de microexpresii

[Wang2015]



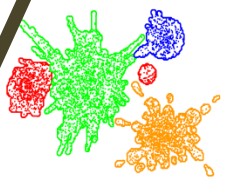
LBP - SIP

Table 1. Comparison of LBP-TOP and LBP-SIP on the CASMEII and SMIC datasets using SVM and RBF kernels for SVM

	CASMEII		SMIC	
	LBP-TOP (%)	LBP-SIP (%)	LBP-TOP (%)	LBP-SIP (%)
Linear	62.75	63.56	60.98	64.02
RBF	65.99	66.40	60.98	62.80

Rezultate pentru recunoastere de microexpresii.

[Wang2015]



LBP+Proiectii integrale

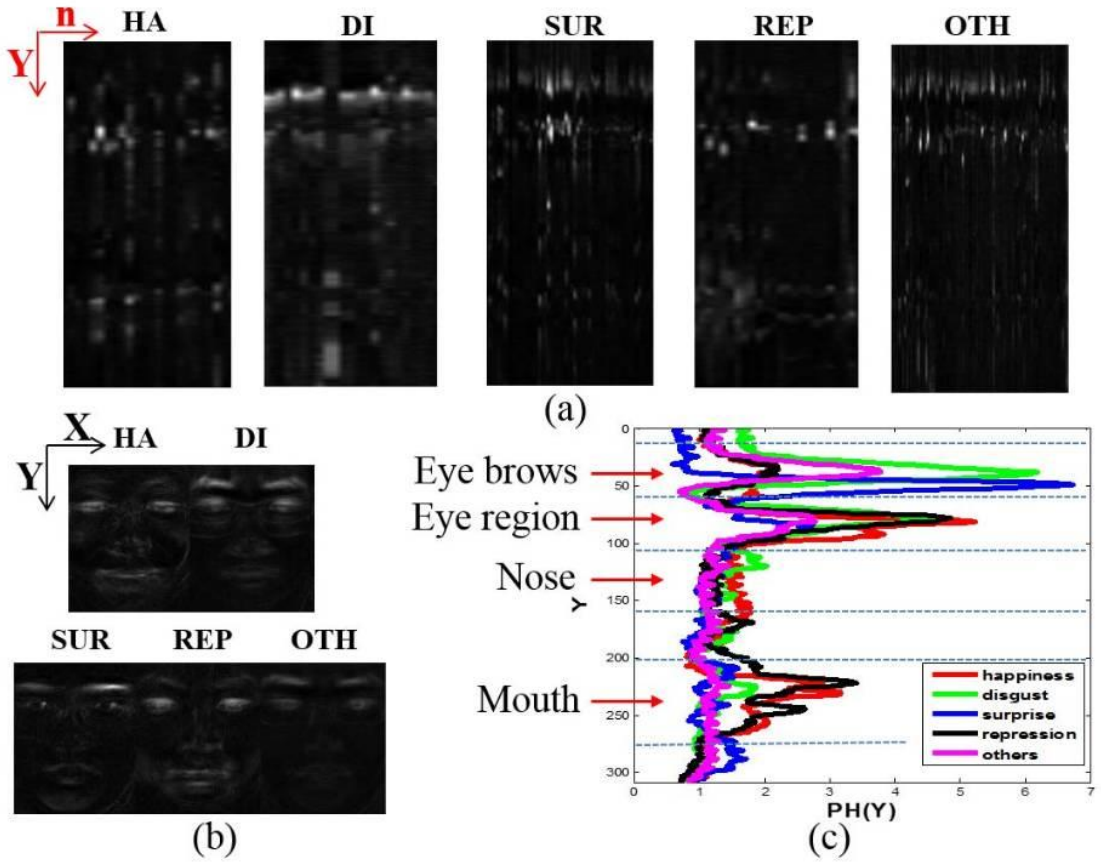
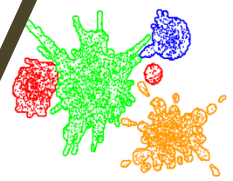


Figure 2: Horizontal integral projection based on difference images. (a) The horizontal integral projections of 247 difference images, where n-axis describes the number of facial images and each projection is represented as a single column. (b) The mean face of difference images for happiness (HA), disgust (DI), surprise (SUR), repression (REP) and others (OTH). (c) The horizontal integral projections of difference images in (b).

Diferenta intre fata cu expresie si cea neutra

[Huang2015-1]



LBP+Proiectii integrale

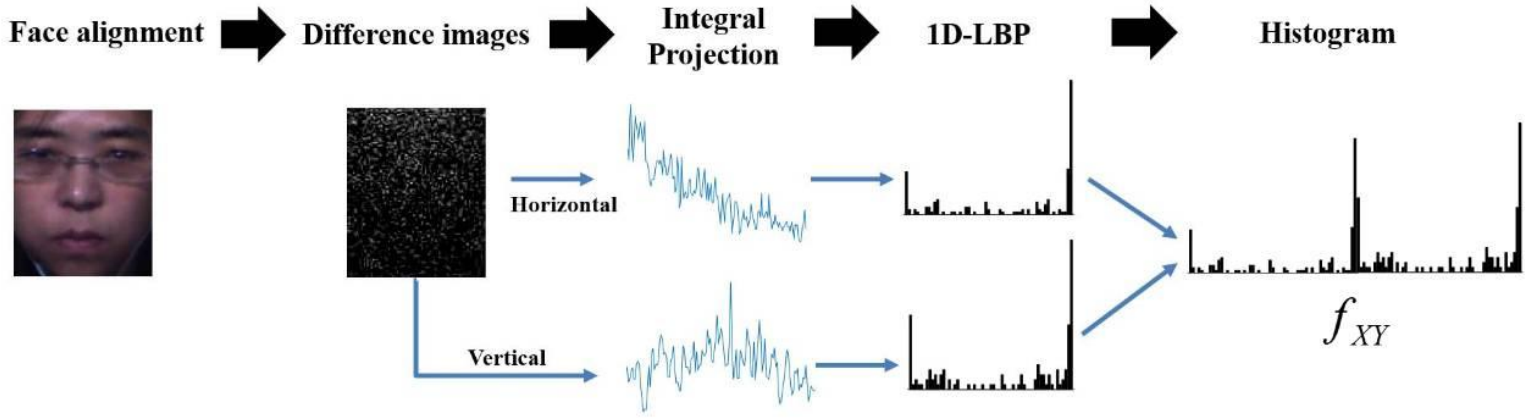


Figure 3: Procedure of encoding difference-image based integral projection on spatial domain.

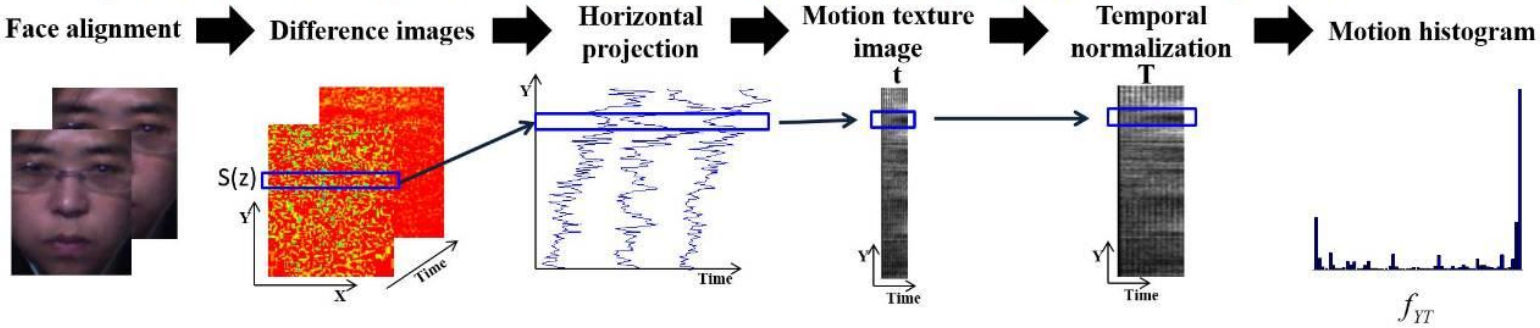
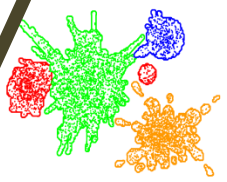


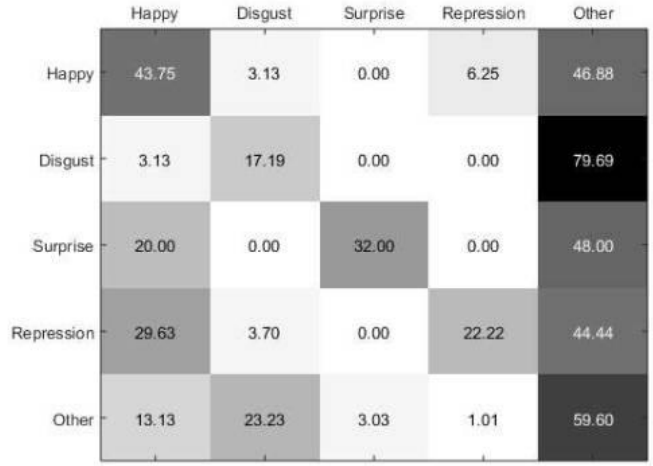
Figure 4: Motion histogram along horizontal direction: The blue rectangle represents the variation of horizontal projection at one position of Y-axis direction, where t and T are the original and temporal normalization time lengths, respectively.

SVM pentru clasificare

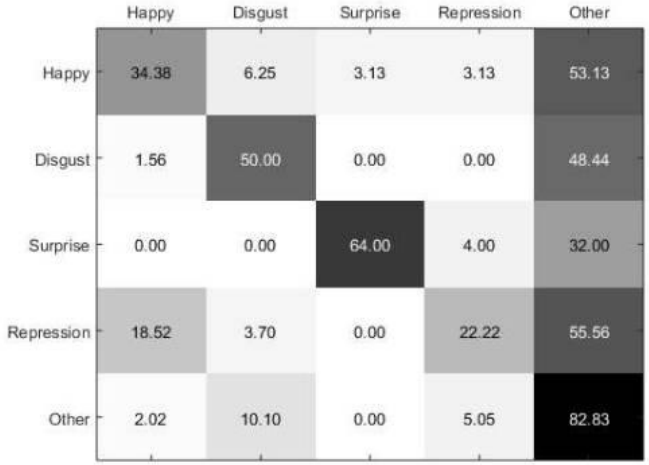
[Huang2015-1]



LBP+Proiectii integrale



(a)



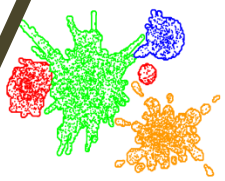
(b)

Table 1: Micro-expression recognition accuracies in CASME2. All methods are based on our implementation, where the values in bracket for LBP-TOP mean the radius for XY, XT and YT planes, for SVM represents kernel function.

Methods	Accuracy (%)
TIM15+LBP-TOP(3,3,3)+SVM(ChiSquare)	39.68
LBP-TOP(3,3,3)+SVM(ChiSquare) [29]	36.03
LBP-TOP(3,3,3)+SVM(linear) [27]	38.87
TIM15+LBP-SIP(3,3,3)+SVM(ChiSquare) [25]	40.08
TIM15+LOCP-TOP(3,3,3)+SVM(ChiSquare) [2]	42.11
STLBP-IP	**59.51**

Rezultate pentru recunoastere de microexpresii.

[Huang2015-1]



Gabor

[Gu2012]



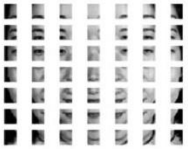
Original Image

Preprocess



Cropped Image

Partition



Local Blocks

Gabor Filters

Gabor Response of Patch 1

Gabor Response of Patch 2

…

Gabor Response of Patch M

Scale 1, Orientation 1
…
Scale 1, Orientation 8
Scale 2, Orientation 1
…
Scale 2, Orientation 8
Scale 3, Orientation 1
…
Scale 3, Orientation 8

Gabor Output

Radial Encoding

Local Feature of Patch 1, Scale 1
…
Local Feature of Patch 1, Scale 3
Local Feature of Patch 2, Scale 1
…
Local Feature of Patch 2, Scale 3
…
Local Feature of Patch M , Scale 1
…
Local Feature of Patch M, Scale 3

Local Feature Representation

PCA + FLD + KNN

Classifier output of Patch 1, Scale 1
…
Classifier output of Patch 1, Scale 3
Classifier output of Patch 2, Scale 1
…
Classifier output of Patch 2, Scale 3
…
Classifier output of Patch M , Scale 1
…
Classifier output of Patch M, Scale 3

Local Classifier Output

Classifier Combination

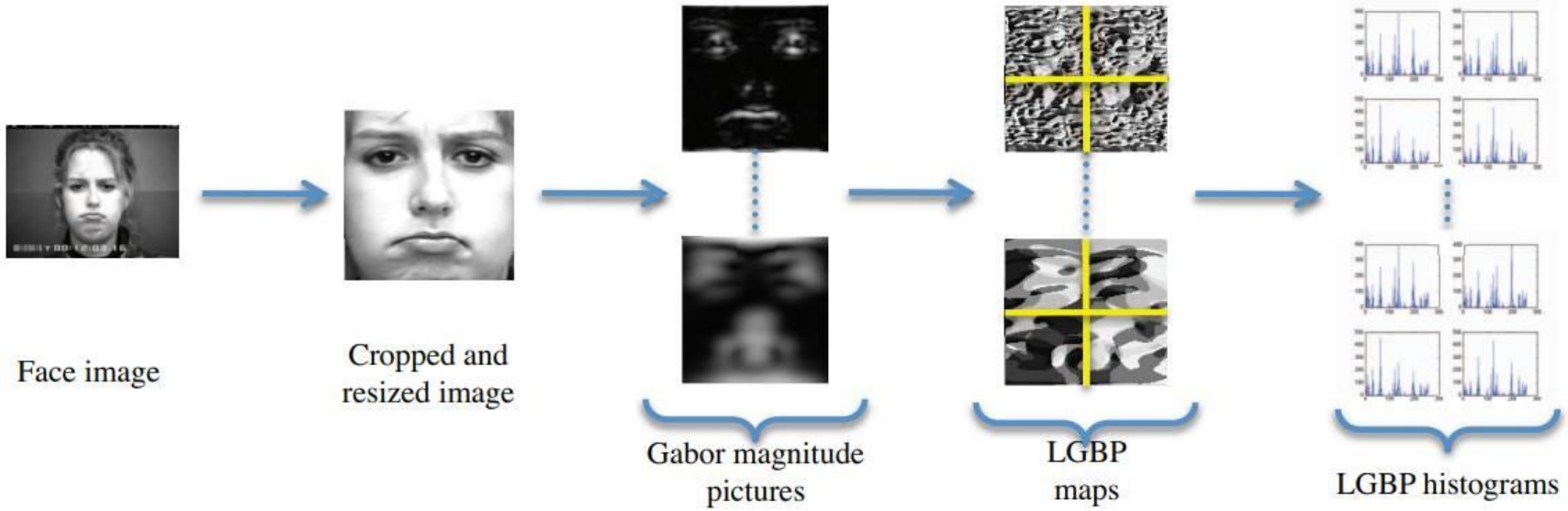
Final Decision of Expression Label

Table 1. Recognition rates (%) on JAFFE and CK for different no. of local blocks.

	25 LBs	49 LBs	81 LBs
JAFFE	84.03	87.32	87.32
CK	84.75	89.07	88.80

Gabor+LBP

[Senechal2011]



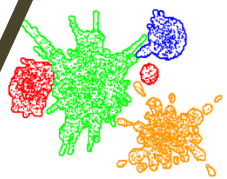
SVM pentru clasificare

Folosit la recunoastere de AU

Analiza holistica

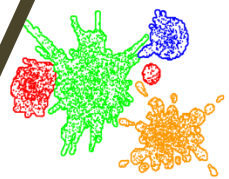
- Imaginile originale sunt considerate a fi unitare.
- Clasificatorul este cel care decide ce este relevant
- Configuratia trasaturilor sau formele sunt dificil de parametrizat

- Solutii:
  - PCA – Principal Component Analysis
  - LFA - Local Feature Analysis
  - LDA – Linear Discriminant Analysis
  - ICA – Independent Component Analysis



PCA pentru emotii

- Portretul este divizat in:
  - Partea superioara
  - Partea inferioara
- Se aplica PCA pe fiecare pentru extragerea dimensiunilor celor mai relevante
- Dimensiunea este redusa la ~ 30 componente
- Clasificarea este realizata pe baza:
  - Celui mai apropiat vecin
  - Template Matching (maximizarea similaritatii)



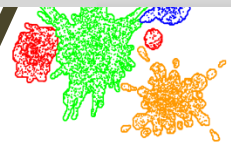
PCA pentru emotii

- Displaying samples along first two principal components shows the validity of the representation for clustering expression.
- First principal components codes large variations of the whole face while second principal component mostly codes opening of the mouth.



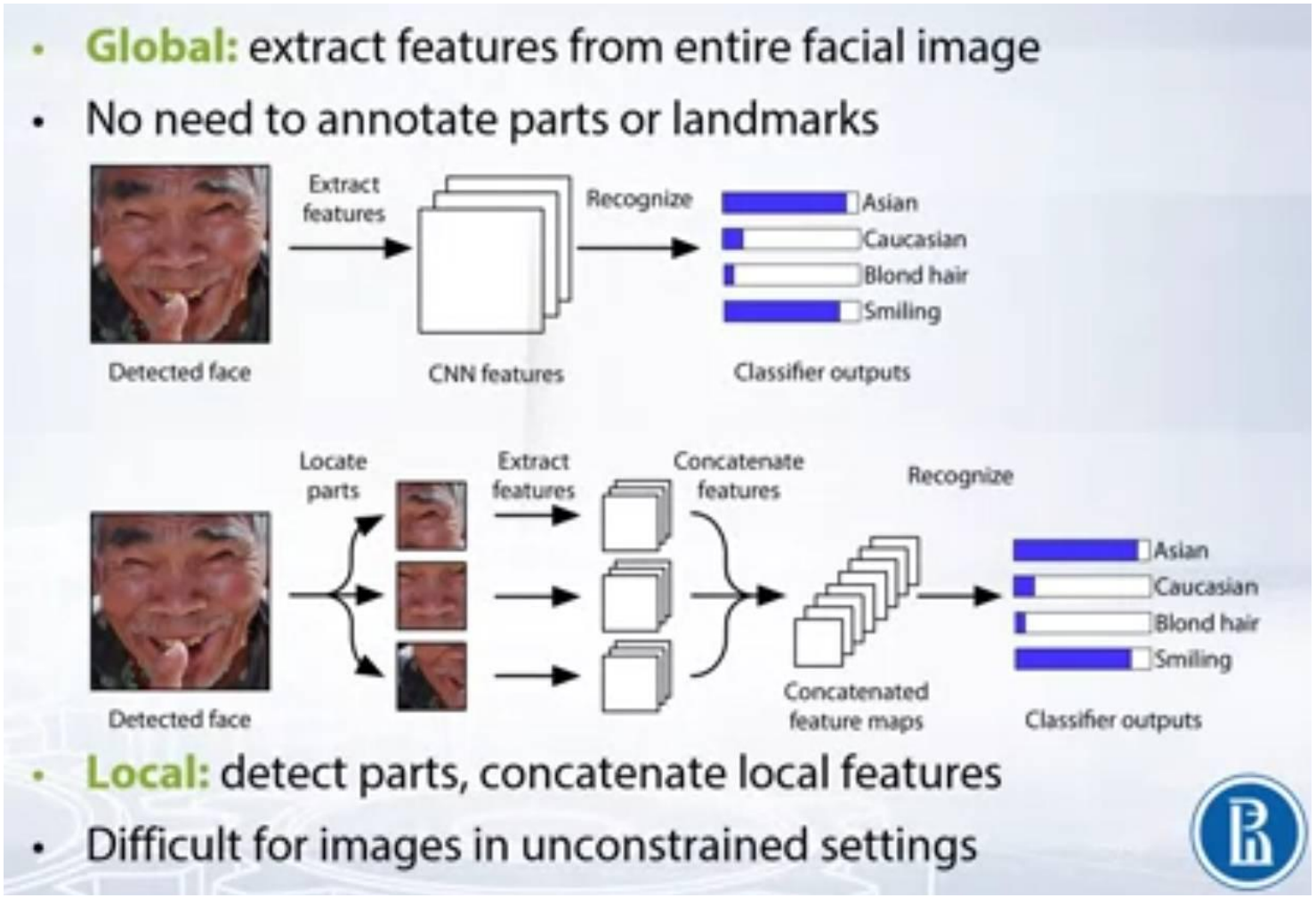
Data along first two principal components.

[Corneanu2015]



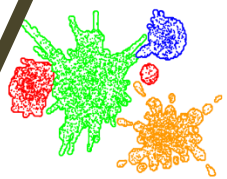
Trasaturi invatate

- Pe toata fata sau pe zone de interes



Trasaturi invatate

- Problema:
  - Nu exista destule date etichetate
  - Etichetarea este complicata
- Solutii posibile:
  - Retele pre-antrenate in probleme de analiza faciala (recunoastere de fete)
  - Invatare semi-supervizata
  - Sintetizare de noi imagini

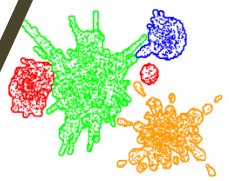


Importanta pre-procesarii

Table 1. Preprocessing steps tuning for the CK+ database: (a) no pre-processing; (b) just cropping; (c) just rotation correction; (d) cropping and rotation correction; (e) only intensity normalization; (f) both normalizations; (g) spatial normalization and synthetic samples; and (h) both normalizations and synthetic samples.

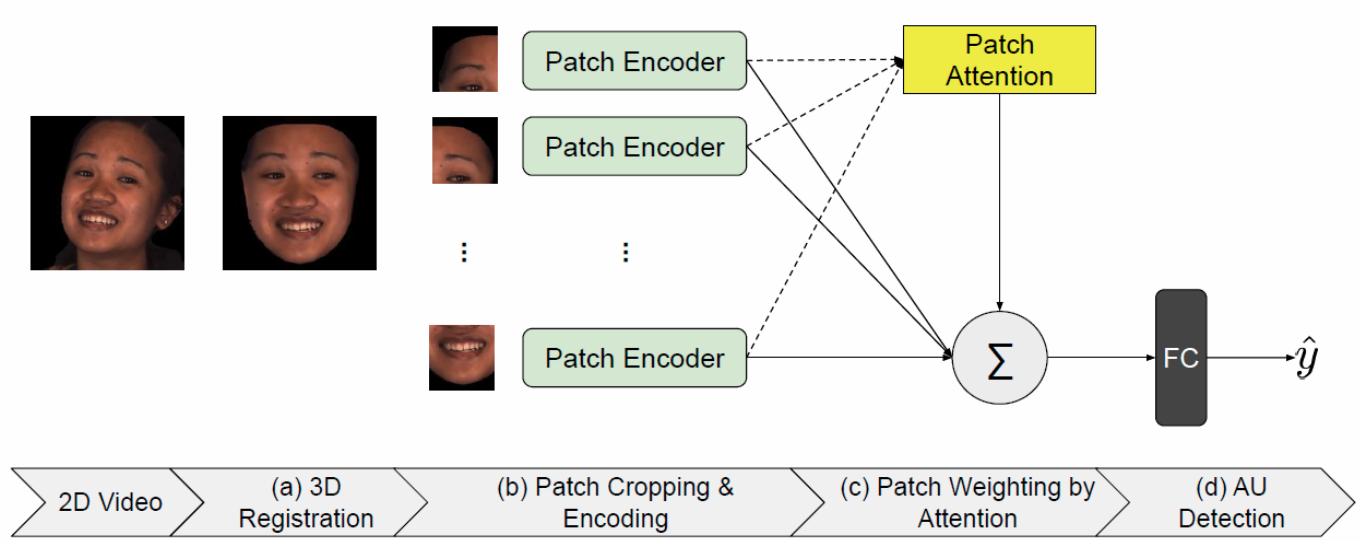
Preprocessing Step	Angry (%)	Disgust (%)	Fear (%)	Happy (%)	Sad (%)	Surprise (%)	Average (%)
(a)	28.10	51.23	17.91	70.68	20.99	77.52	53.57
(b)	68.60	79.01	23.37	86.39	23.46	87.16	71.67
(c)	17.17	79.09	00.00	48.92	05.05	91.25	61.55
(d)	81.82	90.74	73.13	95.81	66.67	94.50	87.86
(e)	27.27	52.94	08.22	79.10	18.29	85.54	57.00
(f)	78.51	93.21	53.73	95.29	75.31	93.12	86.67
(g)	86.05	88.30	69.33	96.60	77.11	95.34	87.10
(h)	**79.34**	**94.44**	**73.13**	**99.48**	**72.84**	**94.94**	**89.76**

[Lopes2017]



Exemplu: Trasaturi invatate

[Onal Ertugrul2019]



- **3D registration: Zface** (László A Jeni, Jeffrey F Cohn, and Takeo Kanade. Dense 3d face alignment from 2d video for real-time use. Image and Vision Computing, 58:13–24, 2017).
- Crop local patches that contain the same facial parts across frames and that are informative for detection of specific AUs – **Done manually**
- Encode patches with individual CNNs and obtain local representations.
- Patch Weighting by **Sigmoidal Attention Mechanism** - weight the importance of patches in detecting specific AUs. The network itself learns (1) encoding of patches; (2) the degree of attention to those patches to maximize AU detection.

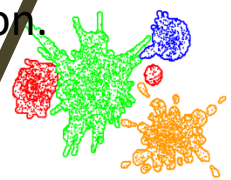
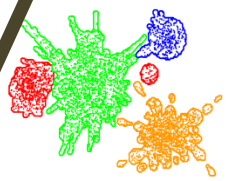


Table 3: Performances (F1-scores) obtained using single patch CNN, holistic face CNN, PAttNet with softmax attention and our proposed PAttNet approach. The best results are shown in bold for both single patch CNNs (left) and other methods (right). The second best results are shown underlined for single patch CNNs.

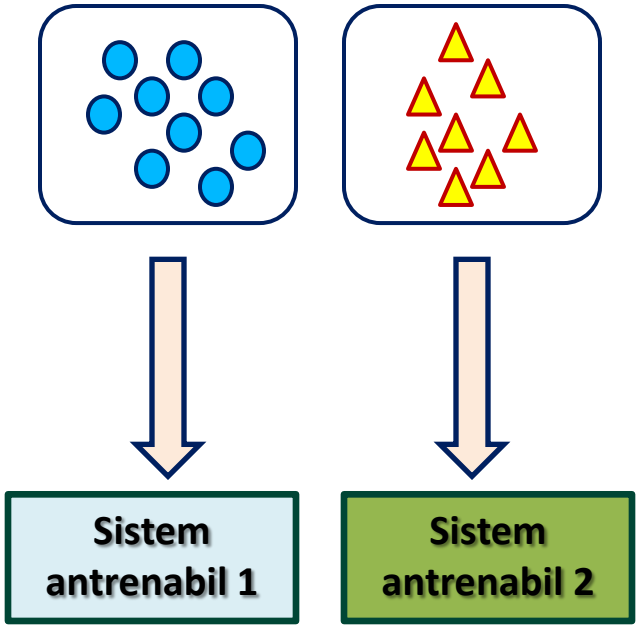
AU	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	Holistic	PAttNet Softmax	PAttNet Sigmoid
1	**39.5**	37.2	38.1	38.5	36.8	26.5	23.4	22.7	21.9	45.3	37.24	**46.1**
2	**36.3**	31.9	35.5	27.3	32.7	21.5	14.5	19.6	18.9	37.2	28.4	**41.4**
4	41.9	44.4	44.7	**46.1**	43.3	27.8	33.3	33.0	31.8	55.5	41.5	**57.1**
6	67.3	69.0	70.7	71.0	71.1	73.0	**73.2**	72.9	72.1	76.3	73.4	**77.9**
7	72.9	**75.6**	71.2	75.5	73.9	70.3	69.2	66.7	70.2	74.0	69.8	**76.2**
10	73.7	75.2	77.6	78.6	78.7	**80.6**	80.3	79.3	78.5	82.2	79.3	**83.8**
12	74.1	75.1	78.2	78.5	81.2	82.5	**83.9**	83.7	83.2	87.3	82.0	**88.4**
14	55.2	52.3	56.4	55.6	54.8	57.7	**61.4**	59.1	61.0	65.6	58.7	**66.5**
15	28.1	26.5	28.7	28.6	26.8	42.0	41.4	**45.7**	43.1	**51.6**	32.7	51.2
17	37.9	37.8	38.1	42.9	40.3	55.0	**58.1**	57.0	58.0	61.3	58.5	**61.6**
23	24.3	22.6	24.8	26.6	23.4	38.7	**42.9**	40.0	38.3	**47.1**	39.8	44.1
24	18.3	18.8	13.8	17.9	16.1	47.5	**52.4**	41.7	51.5	50.8	49.2	**57.3**

Rezultate pe BP4D

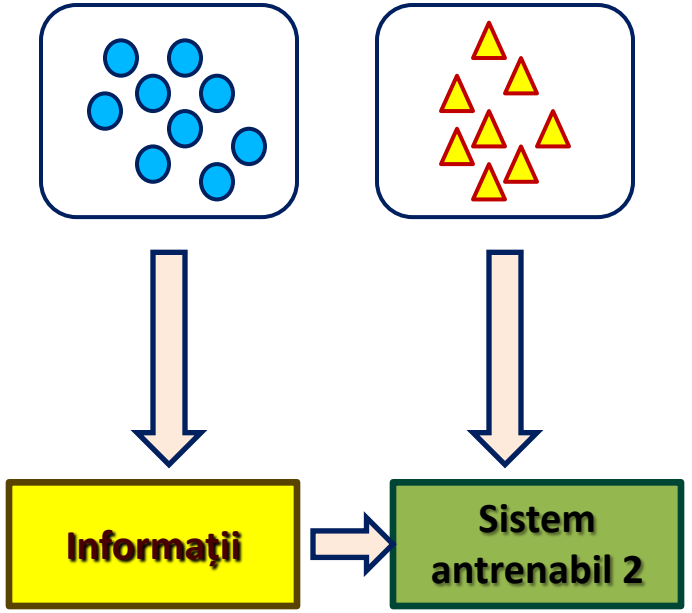
[Onal Ertugrul2019]



Transfer learning



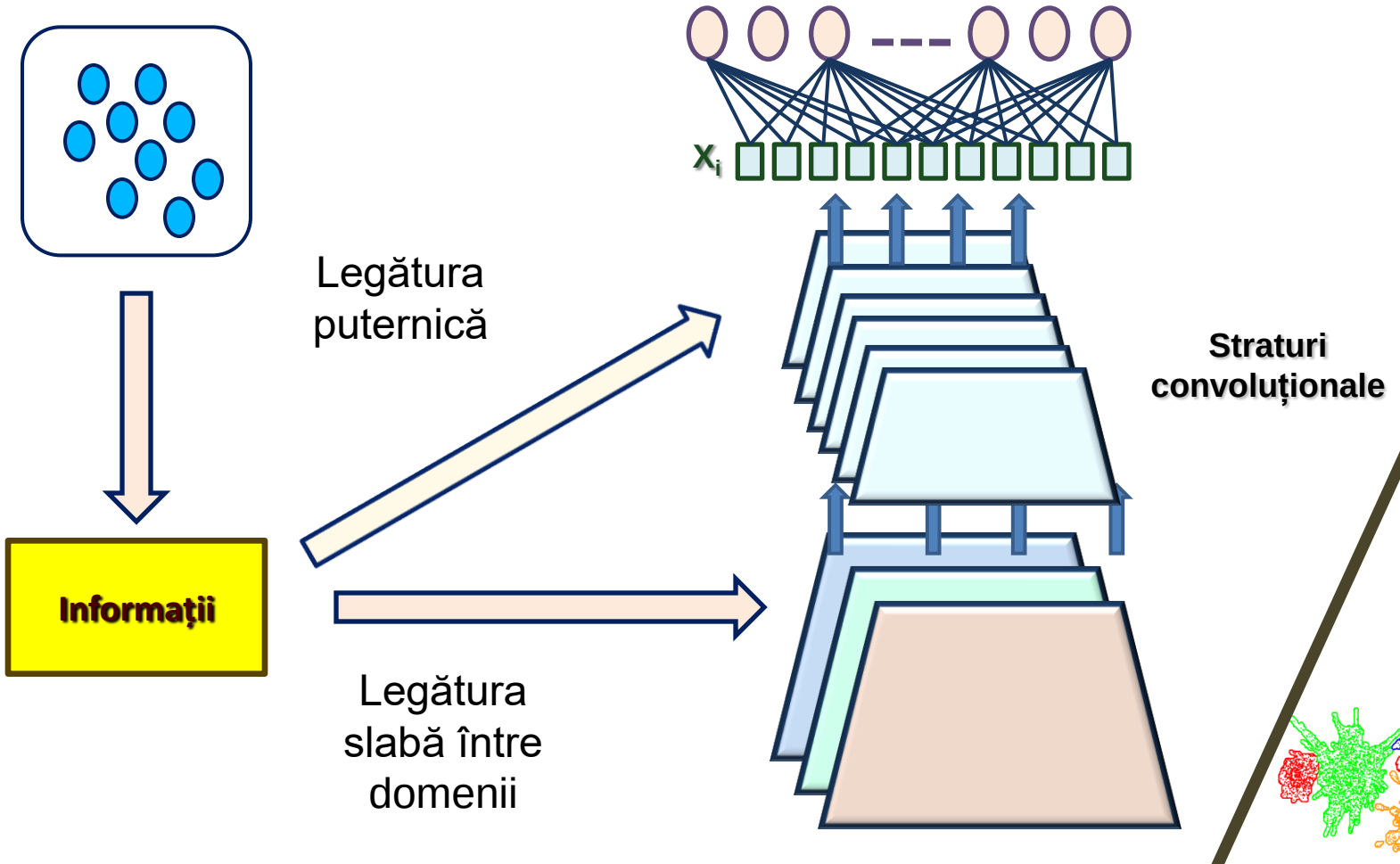
Machine learning clasic



Transfer learning

Transfer learning

Învățare prin transfer pt CNN-uri:



Exemplu: Trasaturi invatate

[Ding2017]

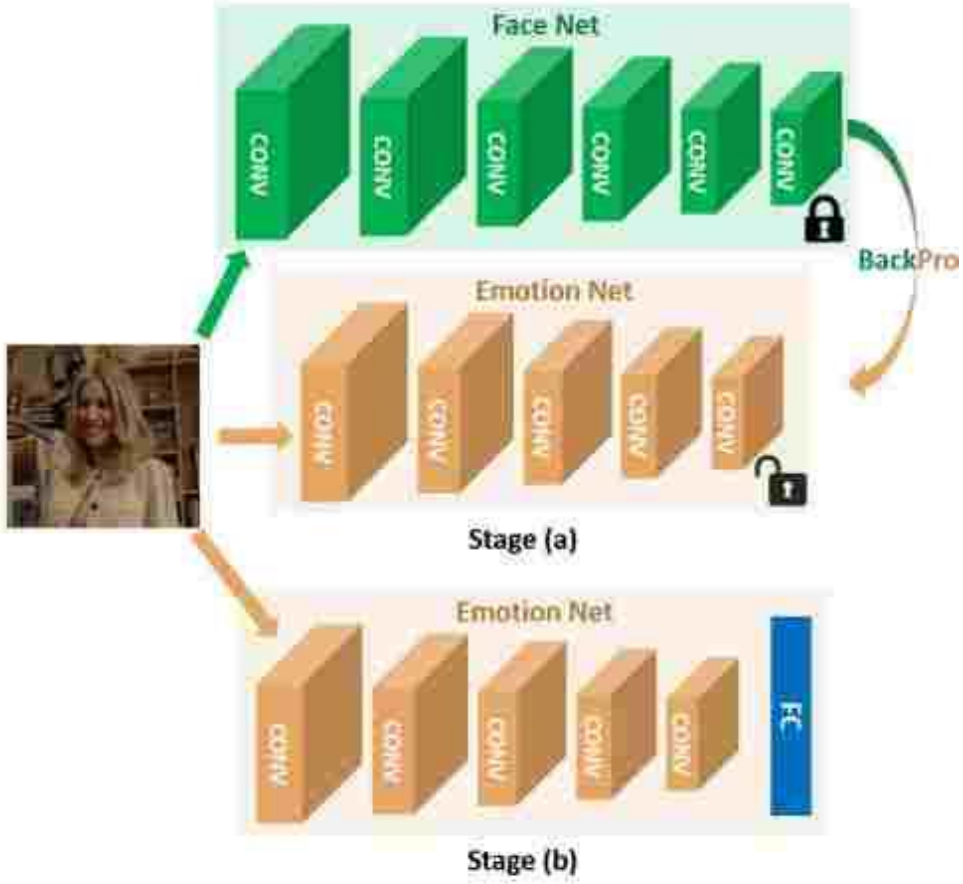


Fig. 2. Two-stage Training Algorithm. In stage (a), the face net is frozen and provides supervision for the expression net. The regression loss is backpropped only to the expression net. The convolutional layers are trained in this stage. In stage (b), the randomly initialized fully-connected layers are attached to the trained convolutional blocks. The whole network is trained jointly with cross-entropy loss. The face net is normally much deeper than the expression net.

Sintetizare de imagini

[Ding2017]

- GAN antrenat sa sintetizeze fete neutre
- Informatia din straturile intermediare se foloseste pentru clasificarea emotiilor

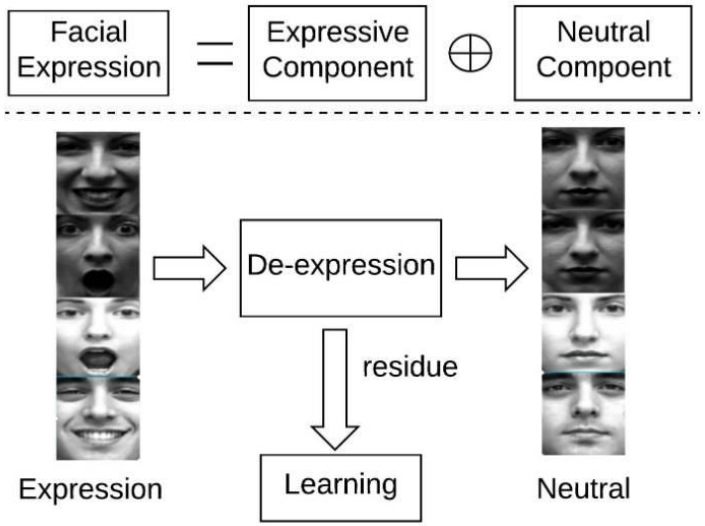


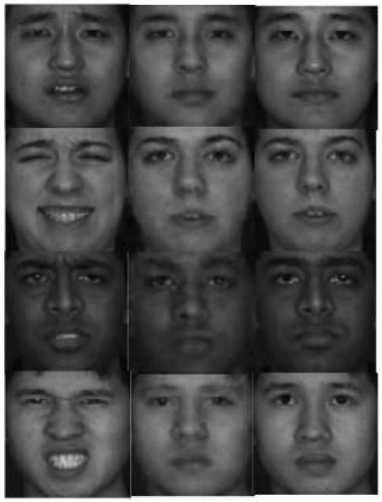
Figure 1. Illustration of our proposed method - De-expression Residue Learning (DeRL). A facial expression is the combination of a neutral face image and the expressive component. Our proposed method recognizes facial expression by learning the residual expressive component in the generative model.



CK+



Oulu-CASIA



BU-4DFE



BP4D+

Table 1. Average accuracy on the CK+ database for seven expressions classification.

Method	Setting	Accuracy
LBP-TOP [34]	sequence-based	88.99
HOG 3D [14]	sequence-based	91.44
3DCNN [17]	sequence-based	85.9
STM-Explet [18]	sequence-based	94.19
IACNN [21]	image-based	95.37
DTAGN [12]	sequence-based	97.25
CNN (baseline)	image-based	89.50
DeRF (Ours)	image-based	**97.30**

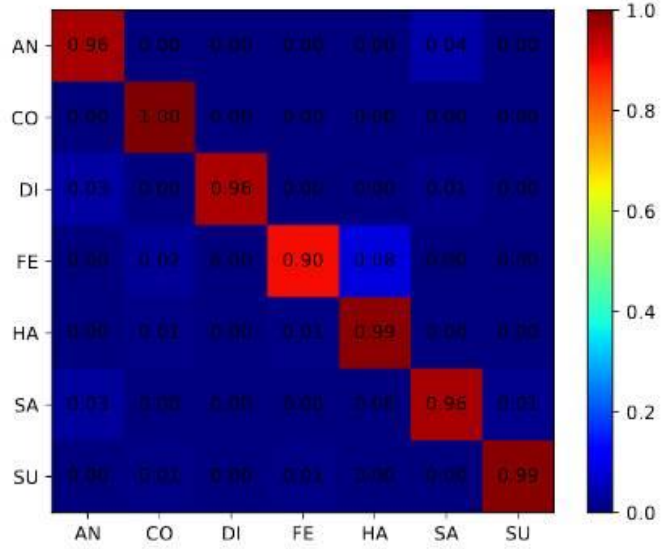


Figure 5. Confusion matrix on CK+

Table 2. Average accuracy on the Oulu-CASIA database for six expressions classification.

Method	Setting	Accuracy
LBP-TOP [34]	sequence-based	68.13
HOG 3D [14]	sequence-based	70.63
STM-Explet [18]	sequence-based	74.59
Atlases [9]	sequence-based	75.52
DTAGN-Joint [12]	sequence-based	81.46
FN2EN [6]	image-based	87.71
PPDN [35]	image-based	84.59
CNN (baseline)	image-based	72.92
DeRL (Ours)	image-based	**88.0**

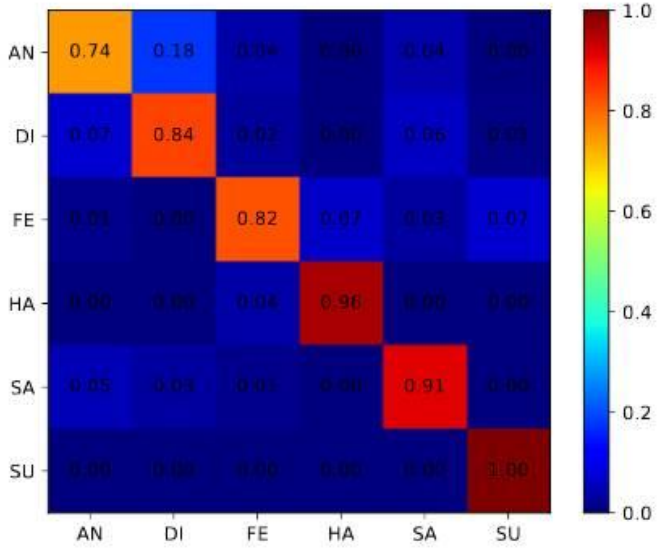


Figure 6. Confusion matrix on Oulu-CASIA

Table 3. Average accuracy on the MMI database for six expressions classification.

Method	Setting	Accuracy
LBP-TOP [34]	sequence-based	59.51
HOG 3D [14]	sequence-based	60.89
STM-Explet [18]	sequence-based	**75.12**
DTAGN-Joint [12]	sequence-based	70.24
IACNN [21]	image-based	71.55
CNN (baseline)	image-based	57.00
DeRL (Ours)	image-based	**73.23**

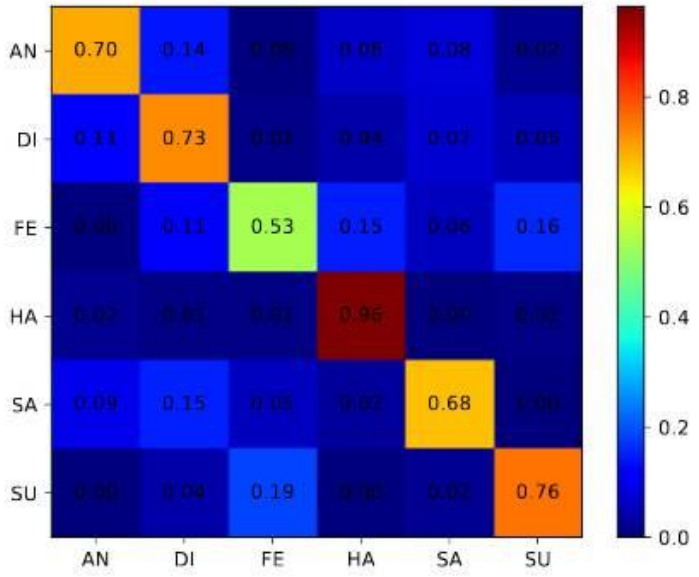


Figure 7. Confusion matrix on MMI

Table 4. Average accuracy on the BU-3DFE database for six expressions classification.

Method	Setting	Accuracy
Wang et al.[26]	3D	61.79
Berretti et al.[3]	3D	77.54
Yang et al.[27]	3D	84.80
Lo et al.[16]	2D image + 3D	**86.32**
Lopes [19]	image-based	72.89
CNN (baseline)	image-based	73.2
DeRL (Ours)	image-based	**84.17**

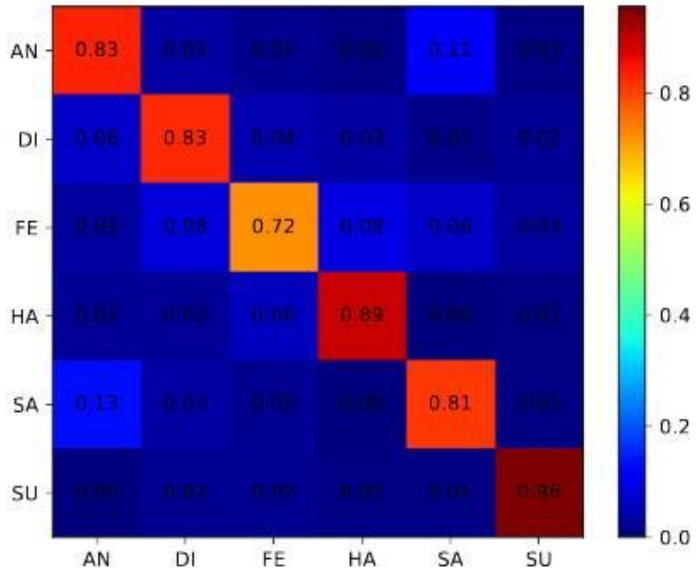
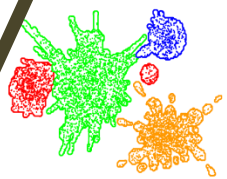


Figure 8. Confusion matrix on BU-3DFE

Referinţe bibliografice

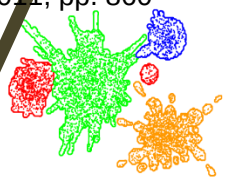
FACS, Generale:

- [Scherer2011] Scherer, Klaus & Clark-Polner, Elizabeth & Mortillaro, Marcello. (2011). In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. International journal of psychology : Journal international de psychologie. 46. 401-35.
- [Ekman1982] Ekman, P., "*Methods for measuring facial action*". In K. R. Scherer and P. Ekman,Eds."*Handbook of methods in nonverbal behavior research*" pp. 45–135, New York: Cambridge University Press, 1982.
- [Ekman2005] P. Ekman and E.L. Rosenberg. What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the FACS. Oxford Scholarship, 2005
- [Bartlett1999] M. Bartlet, J. Hager, P. Ekman, T. Sejnowski ," *Measuring facial expressions by computer image analysis*", *Psychophysiology, Vol. 36* , pp. 253–263. Cambridge University Press, 1999.
- [Fasel2003] B. Fasel, J. Luettin, "*Automatic Facial Expression Analysis: A Survey*" Pattern Recognition, Vol. 36, Issue 1, pp. 259-275, 2003
- [Zucker2007] Ursula Zucker," *Facial Expression Recognition - A Comparison Between Humans and Algorithms",* http://www9.cs.tum.edu/people/wimmerm/lehre/sep_zu, 2007
- [Pantic2000] M. Pantic and L. J. M. Rothkrantz, "Expert system for automatic analysis of facial expressions," Image Vis. Comput., vol. 18, no. 11, pp. 881–905, Aug. 2000.
- [Susskind2007] J Susskind, et al, Human and computer recognition of facial expressions of emotion, Neuropsychologia 45 (1) (2007) 152–162
- [Dodge2017] S. Dodge, L. Karam, Can the early human visual system compete with deep neural networks?, ICCV, 2017, pp. 2798–2804



Referinţe bibliografice

- [Onal Ertugrul2019] I. Onal Ertugrul, L. A. Jeni, J. F. Cohn, *PAttNet*: *Patch-attentive deep network* for *action unit detection*. British Machine Vision Conference (BMVC), 2019.
- G. Zhao and M. Pietikainen, “Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 6, pp. 915–928, 2007.
- Caifeng Shan, Shaogang Gong, Peter W. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study", Image and Vision Computing, Volume 27, Issue 6, 2009, Pages 803-816,
- I. Kotsia and I. Pitas, "Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines," in *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 172-187, Jan. 2007.
- [Liu2006] WeiFeng Liu and ZengFu Wang, "Facial Expression Recognition Based on Fusion of Multiple Gabor Features," *18th International Conference on Pattern Recognition (ICPR'06)*, Hong Kong, 2006, pp. 536-539
- H. Ding, S. K. Zhou and R. Chellappa, "FaceNet2ExpNet: Regularizing a Deep Face Recognition Net for Expression Recognition," *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, Washington, DC, 2017, pp. 118-126.
- André Teixeira Lopes, Edilson de Aguiar, Alberto F. De Souza, Thiago Oliveira-Santos, "Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order", Pattern Recognition, Volume 61, 2017, Pages 610-628
- [Huang2015-1] X. Huang, S. Wang, G. Zhao and M. Piteikäinen, "Facial Micro-Expression Recognition Using Spatiotemporal Local Binary Pattern with Integral Projection," *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, Santiago, 2015, pp. 1-9.
- [Huang2015-2]Huang, Xiaohua & Zhao, Guoying & Hong, Xiaopeng & Zheng, Wenming & Pietikäinen, Matti. (2015). Spontaneous Facial Micro-expression Analysis using Spatiotemporal Completed Local Quantized Patterns. Neurocomputing
- [Wang2015]Y. Wang, J. See, R. C.-W. Phan, Y.-H. Oh, “LBP with Six Intersection Points: Reducing Redundant Information in LBP-TOP for Micro-expression Recognition,” in Asian Conference on computer Vision, pp. 525-537, 2015.
- [Gu2012] Wenfei Gu, Cheng Xiang, Y.V. Venkatesh, Dong Huang, Hai Lin, "Facial expression recognition using radial encoding of local Gabor features and classifier synthesis", Pattern Recognition, Volume 45, Issue 1, 2012, Pages 80-91,
- [Yang2018] H. Yang, U. Ciftci and L. Yin, "Facial Expression Recognition by De-expression Residue Learning," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, 2018, pp. 2168-2177.
- [Senechal2011] T. Senechal, V. Rapp, H. Salam, R. Seguier, K. Bailly and L. Prevost, "Combining AAM coefficients with LGBP histograms in the multi-kernel SVM framework to detect facial action units," *Face and Gesture 2011*, Santa Barbara, CA, 2011, pp. 860-865.



Multumesc pentru atentie!

In ceea ce priveste intrebarile… va ascult

