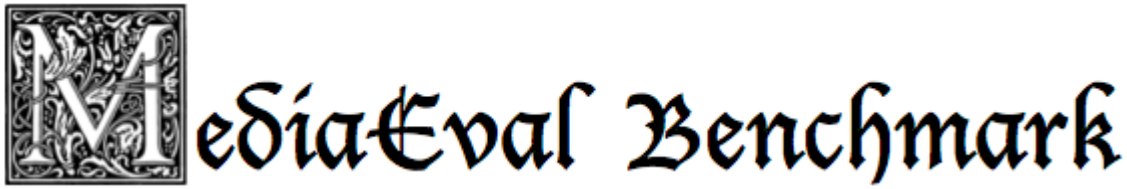


ChaLearn Looking at People: datasets and competitions

Sergio Escalera Guerrero
ChaLearn, Computer Vision Center (UAB) and University of Barcelona
Head of Human Pose Recovery and Behavior Analysis Group
Isabelle Guyon
Jordi Gonzàlez
Xavier Baró

17/10/2014

sergio@maia.ub.es

INDEX

ChaLearn LAP
1.History

ChaLearn http://www.chalearn.org/

Mission:

Machine Learning is the science of building hardware or software that can achieve tasks by learning from examples. The examples often come as {input, output} pairs. Given new inputs a trained machine can make predictions of the unknown output.

Examples of machine learning tasks include:

- automatic reading of handwriting
- assisted medical diagnosis
- automatic text classification (classification of web pages; spam filtering)
- financial predictions

We organize challenges to stimulate research in this field. **The web sites of past challenges remain open** for post-challenge submission as ever-going benchmarks.

ChaLearn is a tax-exempt organization under section 501(c)(3) of the US IRS code. DLN: 17053090370022.

ChaLearn Looking at people (multimedia datasets, http://gesture.chalearn.org/)

ChaLearn Looking at People Challenges and Workshops

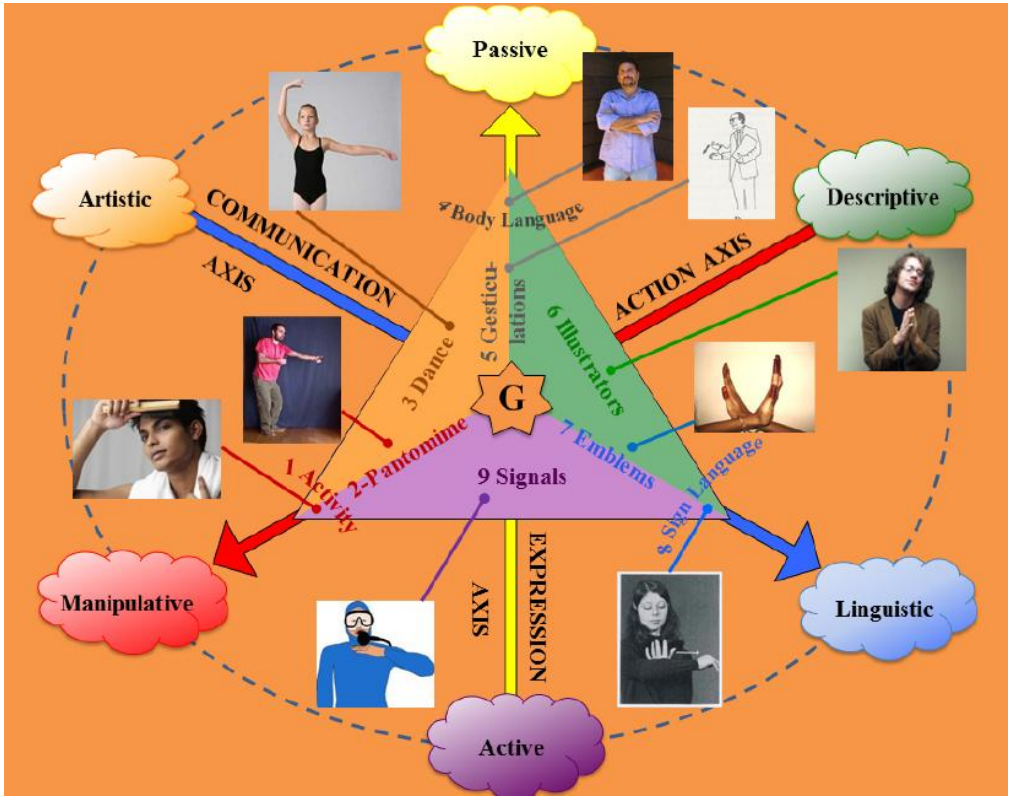
CVPR 2011	Workshop and Challenge on Multi-modal Sign Language Recognition
CVPR 2012	Workshop and Challenge on Multi-modal Sign Language Recognition
ICPR 2012	Workshop and Challenge on Multi-modal Sign Language Recognition
ICMI 2013	Workshop and Challenge on Multi-modal gesture recognition
ECCV 2014	Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting
CVPR 2015	Workshop and Challenge on human pose recovery, action/interaction spotting, age estimation, cultural event recognition
ICCV 2015	Workshop and Challenge on age estimation and cultural event recognition
CVPR 2016	Workshop and Challenge on Multi-modal speed interviews analysis

And so on! Let us know about your opinion! sergio@maia.ub.es

ChaLearn LAP
2.Datasets

ChaLearn Looking at People Challenges and Workshops

CVPR 2011	Workshop and Challenge on Multi-modal Sign Language Recognition
CVPR 2012	Workshop and Challenge on Multi-modal Sign Language Recognition
ICPR 2012	Workshop and Challenge on Multi-modal Sign Language Recognition



Microsoft Research

ChaLearn LAP
2.Datasets

ChaLearn Looking at People Challenges and Workshops

CVPR 2011	Workshop and Challenge on Multi-modal Sign Language Recognition
CVPR 2012	Workshop and Challenge on Multi-modal Sign Language Recognition
ICPR 2012	Workshop and Challenge on Multi-modal Sign Language Recognition

Microsoft Research

1. Body language gestures (like scratching your head, crossing your arms).
2. Gesticulations performed to accompany speech.
3. Illustrators (like Italian gestures).
4. Emblems (like Indian Mudras).
5. Signs (from sign languages for the deaf).
6. Signals (like referee signals, diving signals, or Marshalling signals to guide machinery or vehicle).
7. Actions (like drinking or writing).
8. Pantomimes (gestures made to mimic actions).
9. Dance postures.

Evaluation metric: levenstein edition distance

ChaLearn Looking at People Challenges and Workshops

ICMI 2013 Workshop and Challenge on Multi-modal gesture recognition

Multi-modal ChaLearn Gesture Recognition Challenge and Workshop

http://gesture.chalearn.org/
sunai.uoc.edu/chalearn

Web of the competition
Data

The challenge features a **quantitative evaluation** of automatic gesture recognition from a multi-modal dataset recorded with **Kinect** (providing RGB images of face and body, depth images of face and body, skeleton information, joint orientation and audio sources), **including 13,858 Italian gestures from near 30 users.**

The emphasis of this edition of the competition will be on multi-modal automatic learning of a **vocabulary of 20 types of Italian anthropological/cultural gestures performed by different users**, with the aim of **performing user independent continuous gesture recognition combined with audio information**.

ChaLearn Looking at People Challenges and Workshops

ICMI 2013 Workshop and Challenge on Multi-modal gesture recognition

Gesture categories (1/2)

(1) *Vattene*

(2) *Viene qui*

(3) *Perfetto*

(4) *E un furbo*

(5) *Che due palle*

(6) *Che vuoi*

(7) *Vanno d'accordo*

(8) *Sei pazzo*

(9) *Cos hai combinato*

(10) *Nonme me friega niente*

ChaLearn Looking at People Challenges and Workshops

ICMI 2013 Workshop and Challenge on Multi-modal gesture recognition

Gesture categories (2/2)


(11) *Ok*


(12) *Cosa ti farei*


(13) *Basta*


(14) *Le vuoi prendere*


(15) *Non ce ne piu*


(16) *Ho fame*


(17) *Tanto tempo fa*


(18) *Buonissimo*


(19) *Si sono messi d'accordo*

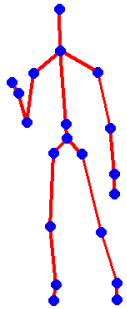

(20) *Sono stufo*

ChaLearn Looking at People Challenges and Workshops

ICMI 2013 Workshop and Challenge on Multi-modal gesture recognition

Data and modalities



- Framerate 20FPS
- RGB: 640x480
- Depth: 640x480
- Audio: Kinect 20 michropone array
- Users: 27
- Italians: 81%

- Total number of sequences: 956 ϵ [1,2] min.
- Total number of gestures: 13,858
- Total number of frames: 1.720.800
- Noisy gestures

Data structure information: *S. Escalera, J. Gonzàlez, X. Baró, M. Reyes, O. Lopes, I. Guyon, V. Athistos, H.J. Escalante, "Multi-modal Gesture Recognition Challenge 2013: Dataset and Results", ICMI 2013*.

ChaLearn Looking at People Challenges and Workshops

ICMI 2013 Workshop and Challenge on Multi-modal gesture recognition

Easy and challenging aspects of the data.

Easy
Fixed camera
Near frontal view acquisition
Within a sequence the same user
Gestures performed mostly by arms and hands
Camera framing upper body
Several available modalities: audio, skeletal model, user mask, depth, and RGB
Several instances of each gesture for training
Single person present in the visual field

Challenging
Within each sequence:
Continuous gestures without a resting pose
Many gesture instances are present
Distracter gestures out of the vocabulary may be present in terms of both gesture and audio
Between sequences:
High inter and intra-class variabilities of gestures in terms of both gesture and audio
Variations in background, clothing, skin color, lighting, temperature, resolution
Some parts of the body may be occluded
Different Italian dialects

Evaluation metric: levenstein edition distance

ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

•**Track 1: Human Pose Recovery:** More than 8,000 frames of continuous RGB sequences are recorded and labeled with the objective of performing human pose recovery by means of recognizing more than 120,000 human limbs of different people.

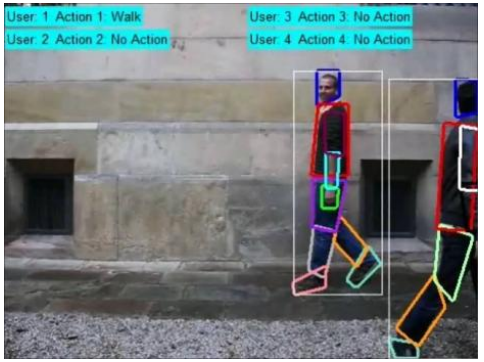
•**Track 2: Action/Interaction Recognition:** 235 performances of 11 action/interaction categories are recorded and manually labeled in continuous RGB sequences of different people performing natural isolated and collaborative behaviors.

•**Track 3: Gesture Recognition:** The gestures are drawn from a vocabulary of Italian sign gesture categories. The emphasis of this third track is on multi-modal automatic learning of a set of 20 gestures performed by several different users, with the aim of performing user independent continuous gesture spotting.

ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

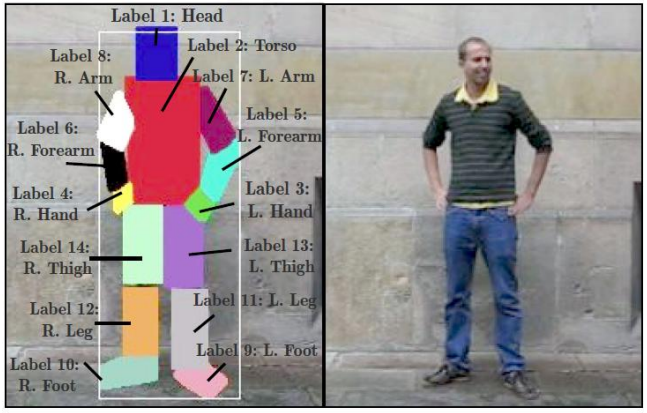
•**Track 1: Human Pose Recovery:** More than 8,000 frames of continuous RGB sequences are recorded and labeled with the objective of performing human pose recovery by means of recognizing more than 120,000 human limbs of different people.



Training frames	Validation frames	Test frames	Sequence duration	FPS
4,000	2,000	2,236	1-2 min	15
Modalities	**Num. of users**	**Limbs per body**	**Labeled frames**	**Labeled limbs**
RGB	14	14	8,234	124,761

Human pose recovery data characteristics.

•**9 videos** (RGB sequences) and a total of **14** different **actors**. Stationary camera with the same static background.
•**15 fps** rate, resolution **480x360** in BMP file format.
• For each actor **14 limbs** (if not occluded) were manually tagged: Head, Torso, R-L Upper-arm, R-L Lower-arm, R-L Hand, R-L Upper-leg, R-L Lower-leg, and R-L Foot.
• **Limbs are manually labeled** using binary masks and the minimum bounding box containing each subject is defined.



• The actors appear in a **wide range of different poses** and **performing different actions/gestures** which vary the visual appearance of human limbs. So there is a large variability of human poses, self-occlusions and many variations in clothing and skin color.

ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

•**Track 1: Human Pose Recovery:** More than 8,000 frames of continuous RGB sequences are recorded and labeled with the objective of performing human pose recovery by means of recognizing more than 120,000 human limbs of different people.

Overlap evaluation

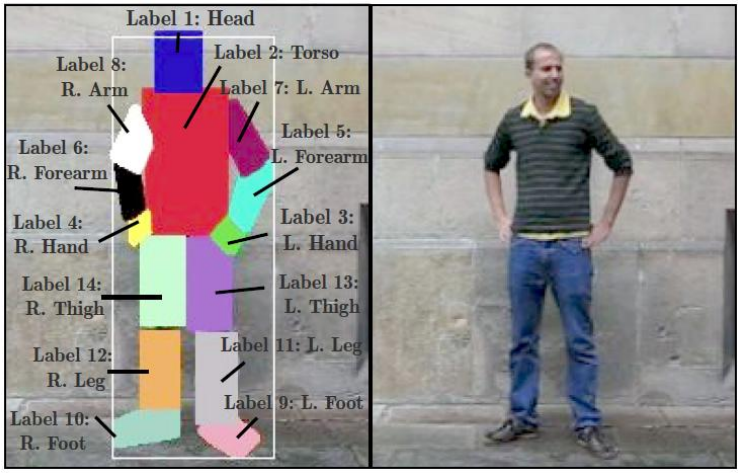
$$J_{i,n} = \frac{A_{i,n} \bigcap B_{i,n}}{A_{i,n} \bigcup B_{i,n}},$$

$$H_{i,n} = \begin{cases} 1 & \text{if } \frac{A_n \bigcap B_n}{A_n \bigcup B_n} \geq 0.5 \\ 0 & \text{otherwise} \end{cases}$$

$A_{i,head}$ $B_{i,head}$

$$J_{i,head} = \frac{A_{i,head} \cap B_{i,head}}{A_{i,head} \cup B_{i,head}} = 0.82$$

$$J_{i,head} > 0.5 \longrightarrow HR_{i,head} = 1$$



ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

•**Track 2: Action/Interaction Recognition:** 235 performances of 11 action/interaction categories are recorded and manually labeled in continuous RGB sequences of different people performing natural isolated and collaborative behaviors.

Training actions	Validation actions	Test actions	Sequence duration	FPS
150	90	95	9× 1-2 min	15

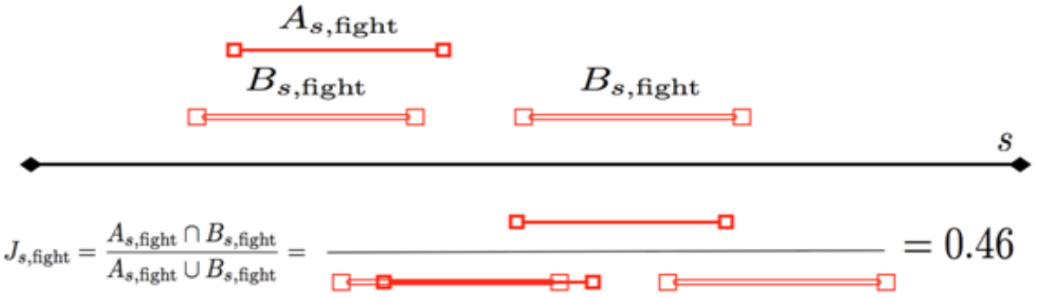
Modalities	Num. of users	Action categories	interaction categories	Labeled sequences
RGB	14	7	4	235

Action and interaction data characteristics.

- **235 action/interaction** samples performed by **14 actors**.
- Large **difference in length** about the performed actions and interactions.
- Several **distracter actions** out of the 11 categories are also present.
- **11 action categories, containing isolated and collaborative actions**: Wave, Point, Clap, Crouch, Jump, Walk, Run, Shake Hands, Hug, Kiss, Fight. There is a high intra-class variability among action samples.

Overlap evaluation

$$J_{s,n} = \frac{A_{s,n} \bigcap B_{s,n}}{A_{s,n} \bigcup B_{s,n}},$$



ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

Track 2: Action/Interaction Recognition

Action categories


Wave


Point


Clap


Crouch


Jump


Walk

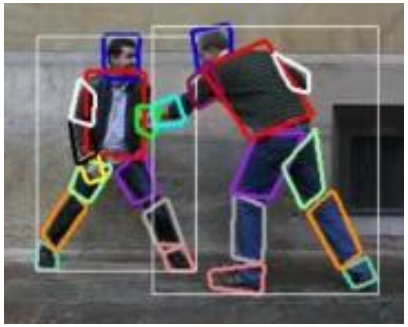

Run

Interaction categories


Shake Hands


Hug


Kiss


Fight

ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

Track 3: Gesture Recognition: The gestures are drawn from a vocabulary of Italian sign gesture categories. The emphasis of this third track is on multi-modal automatic learning of a set of 20 gestures performed by several different users, with the aim of performing user independent continuous gesture spotting.

Training seq.	Validation seq.	Test seq.	Sequence duration	FPS
393 (7,754 gestures)	287 (3,362 gestures)	276 (2,742 gestures)	1-2 min	20

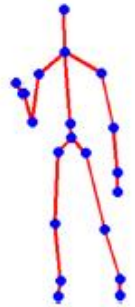
Modalities	Num. of users	Gesture categories	Labeled sequences	Labeled frames
RGB, Depth, User mask, Skeleton	27	20	13,858	1,720,800

Main characteristics of the *Montalbano* gesture dataset.

RGB

Depth

User mask



Skeletal model

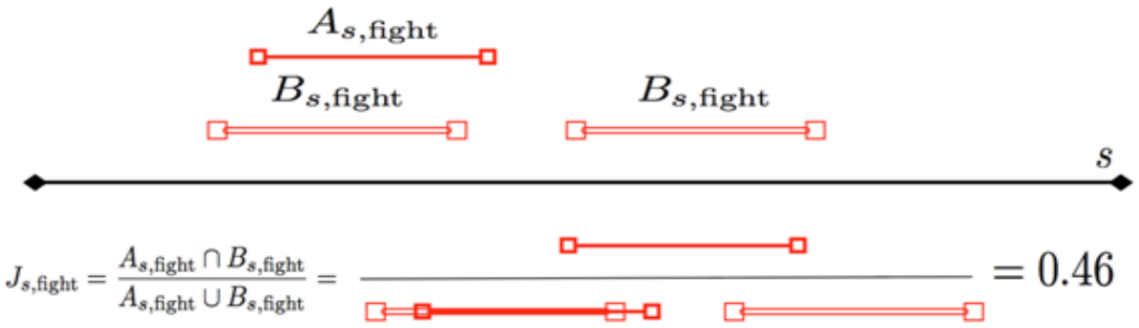
ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

- **Largest dataset** in the literature with a large duration of each individual performance showing **no resting poses and self-occlusions**.
- There is **no information about the number of gestures to spot** within each sequence, and **several distracter gestures** (out of the vocabulary) are present.
- **High intra-class variability** of gesture samples and **low inter-class variability** for some gesture categories.

Overlap evaluation

$$J_{s,n} = \frac{A_{s,n} \bigcap B_{s,n}}{A_{s,n} \bigcup B_{s,n}},$$



ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

- State of the art comparison

	Labeling at pixel precision	Number of limbs	Number of labeled limbs	Number of frames	Full body	Limb annotation	Gesture-action annotation	Number of gestures-actions	Number of gest-act. samples
Montalbano[8]	No	16	27 532 800	1 720 800	Yes	Yes	Yes	20	13 858
HuPBA $8K+$ [7]	Yes	14	124 761	8 234	Yes	Yes	Yes	11	235
LEEDS SPORTS[4]	No	14	28 000	2 000	Yes	Yes	No	-	-
UIUC people[10]	No	14	18 186	1 299	Yes	Yes	No	-	-
Pascal VOC[2]	Yes	5	8 500	1 218	Yes	Yes	No	-	-
BUFFY[3]	No	6	4 488	748	No	Yes	No	-	-
PARSE[11]	No	10	3 050	305	Yes	Yes	No	-	-
MPII Pose[12]	Yes	14	-	40 522	Yes	Yes	Yes	20	491
FLIC[13]	No	29	-	5 003	No	No	No	-	-
H3D[14]	No	19	-	2 000	No	No	No	-	-
Actions[15]	No	-	-	-	Yes	No	Yes	6	600
HW[5]	-	-	-	-	-	No	Yes	8	430

Comparison of public dataset characteristics.

ChaLearn LAP data sets, public available at:
http://sunai.uoc.edu/chalearnLAP/
ChaLearn LAP challenges and news:
http://gesture.chalearn.org/

ChaLearn LAP
3.Results

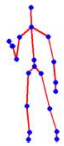
ChaLearn Looking at People Challenges and Workshops

ICMI 2013 Workshop and Challenge on Multi-modal gesture recognition


RGB

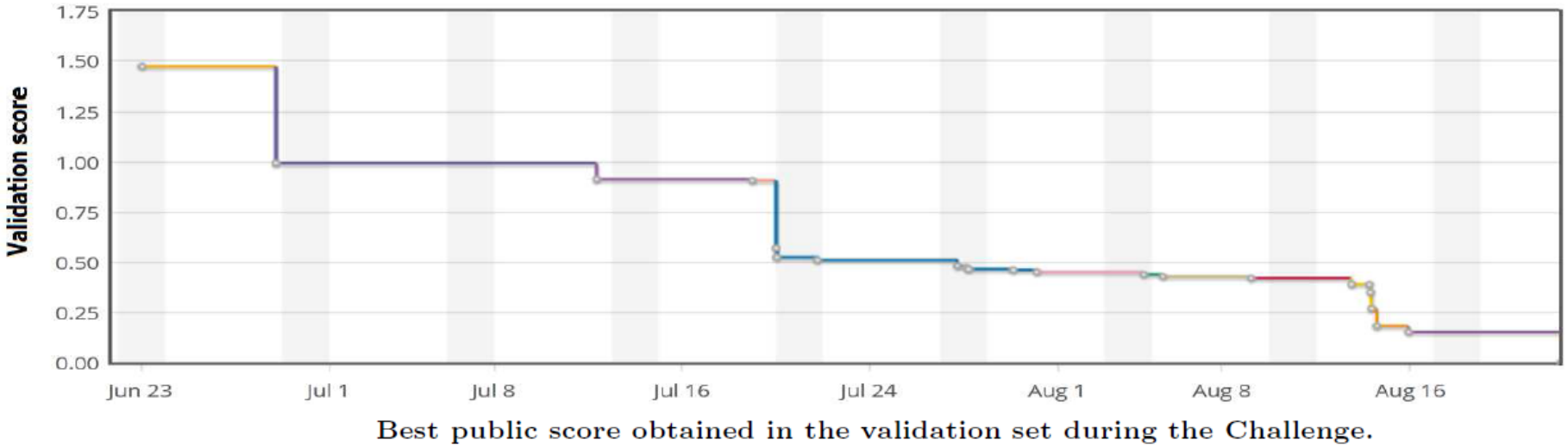

Depth


User mask


Skeletal model

+Audio

Levenstein score



Best public score obtained in the validation set during the Challenge.

kaggle

ChaLearn Looking at People Challenges and Workshops
ICMI 2013 Workshop and Challenge on Multi-modal gesture recognition

- Participation

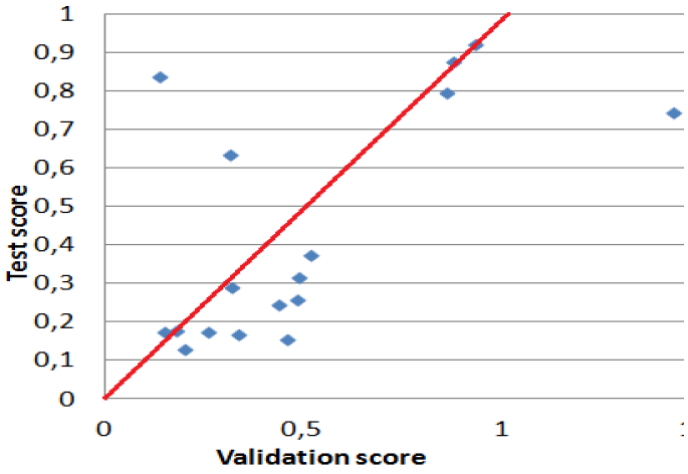
• The challenge attracted high level of participation, with a total of **54 teams and near 300 total number of entries**.
• Finally, **17 teams successfully submitted their prediction in final test set, while providing also their code for verification and summarizing their method by means of a fact sheet questionnaire**.
• After verifying the codes and results of the participants, the final scores of the top rank participants on both validation and test sets were made public.
• In the end, **the final error rate on the test data set was around 12%.**

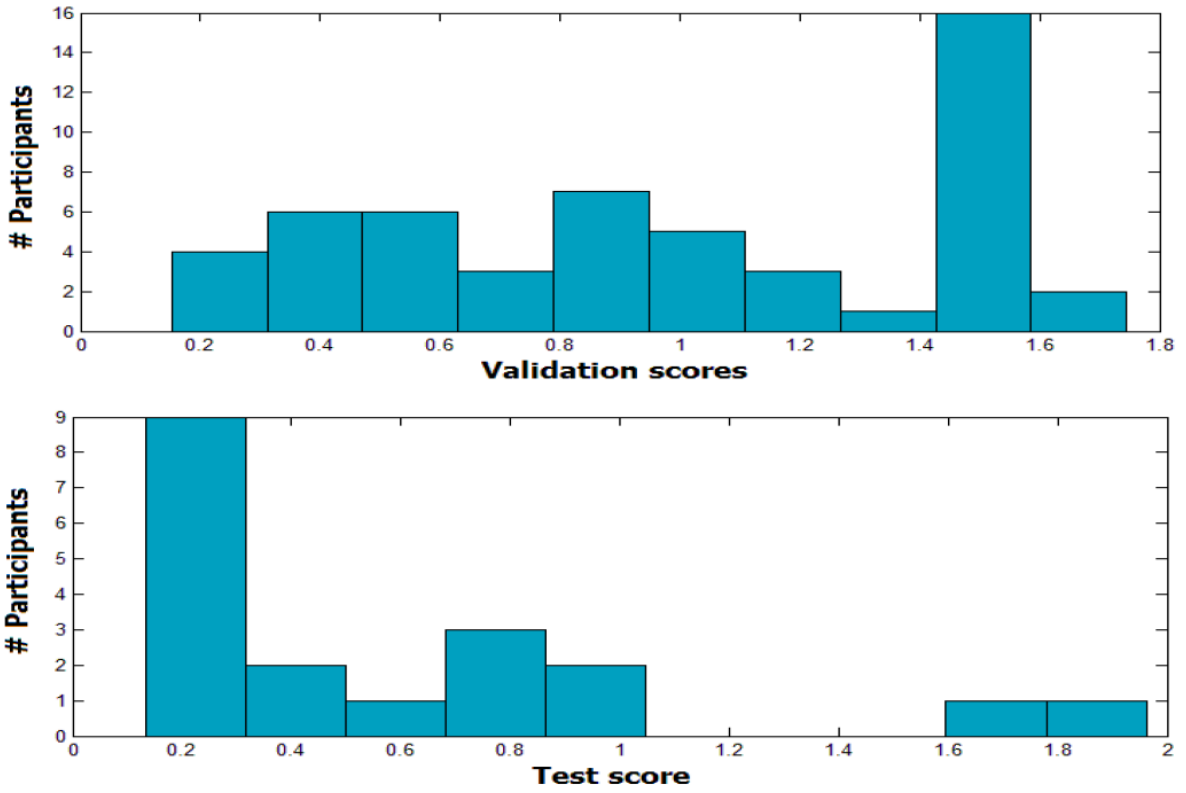
Top rank results on validation and test sets.

TEAM	Validation score	Test score
IVA MM	0.20137	**0.12756**
WWEIGHT	0.46163	**0.15387**
ET	0.33611	**0.16813**
MmM	0.25996	0.17215
PPTK	0.15199	0.17325
LRS	0.18114	0.17727
MMDL	0.43992	0.24452
TELEPOINTS	0.48543	0.25841
CSI MM	0.32124	0.28911
SUMO	0.49137	0.31652
GURU	0.51844	0.37281
AURINKO	0.31529	0.63304
STEVENWUDI	1.43427	0.74415
JACKSPARROW	0.86050	0.79313
JOEWAN	0.13653	0.83772
MILAN KOVAC	0.87835	0.87463
IAMKHADER	0.93397	0.92069

ChaLearn Looking at People Challenges and Workshops

ICMI 2013 Workshop and Challenge on Multi-modal gesture recognition





Validation and test scores histograms.

ChaLearn Looking at People Challenges and Workshops

ICMI 2013 Workshop and Challenge on Multi-modal gesture recognition

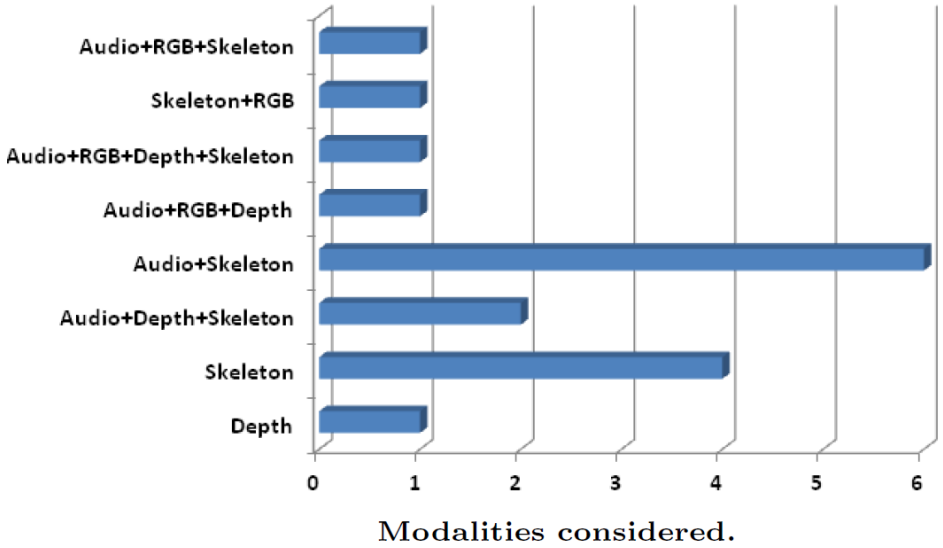
Team methods and results. Early and late refer to early and late fusion of features/classifier outputs. HMM: Hidden Markov Models. KNN: Nearest Neighbor. RF: Random Forest. Tree: Decision Trees. ADA: Adaboost variants. SVM: Support Vector Machines. Fisher: Fisher Linear Discriminant Analysis. GMM: Gaussian Mixture Models. NN: Neural Networks. DGM: Deep Boltzmann Machines. LR: Logistic Regression. DP: Dynamic Programming. ELM: Extreme Learning Machines.

TEAM	Test score	Rank position	Modalities	Segmentation	Fusion	Classifier
IVA MM	0.12756	1	Audio,Skeleton	Audio	None	HMM,DP,KNN
WWEIGHT	0.15387	2	Audio,Skeleton	Audio	Late	RF,KNN
ET	0.16813	3	Audio,Skeleton	Audio	Late	Tree,RF,ADA
MmM	0.17215	4	Audio,RGB+Depth	Audio	Late	SVM,Fisher,GMM,KNN
PPTK	0.17325	5	Skeleton,RGB,Depth	Sliding windows	Late	GMM,HMM
LRS	0.17727	6	Audio,Skeleton,Depth	Sliding windows	Early	NN
MMDL	0.24452	7	Audio,Skeleton	Sliding windows	Late	DGM+LR
TELEPOINTS	0.25841	8	Audio,Skeleton,RGB	Audio,Skeleton	Late	HMM,SVM
CSI MM	0.28911	9	Audio,Skeleton	Audio	Early	HMM
SUMO	0.31652	10	Skeleton	Sliding windows	None	RF
GURU	0.37281	11	Audio,Skeleton,Depth	DP	Late	DP,RF,HMM
AURINKO	0.63304	12	Skeleton,RGB	Skeleton	Late	ELM
STEVENWUDI	0.74415	13	Audio,Skeleton	Sliding windows	Early	DNN,HMM
JACKSPARROW	0.79313	14	Skeleton	Sliding windows	None	NN
JOEWAN	0.83772	15	Skeleton	Sliding windows	None	KNN
MILAN KOVAC	0.87463	16	Skeleton	Sliding windows	None	NN
IAMKHADER	0.92069	17	Depth	Sliding windows	None	RF

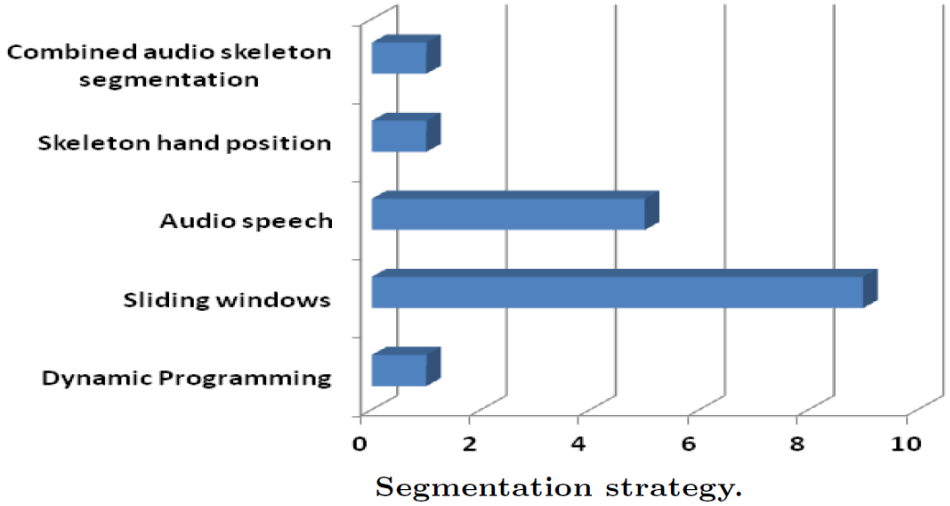
Data structure information: *S. Escalera, J. Gonzàlez, X. Baró, M. Reyes, O. Lopes, I. Guyon, V. Athistos, H.J. Escalante, "Multi-modal Gesture Recognition Challenge 2013: Dataset and Results", ICMI 2013*.

ChaLearn Looking at People Challenges and Workshops

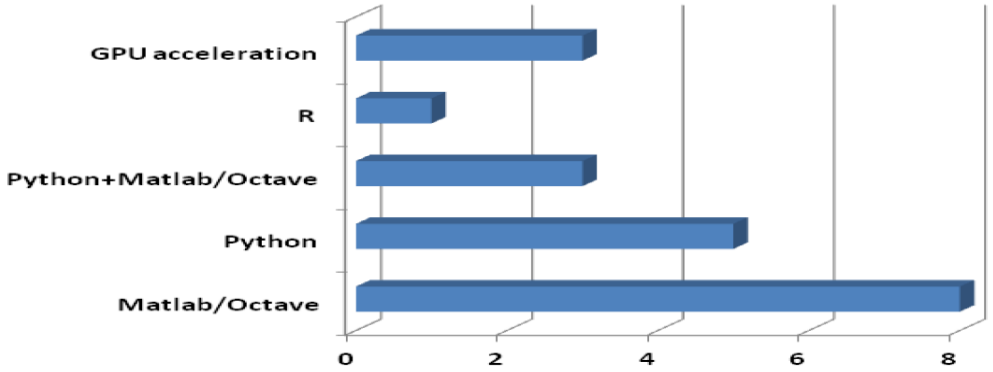
ICMI 2013 Workshop and Challenge on Multi-modal gesture recognition



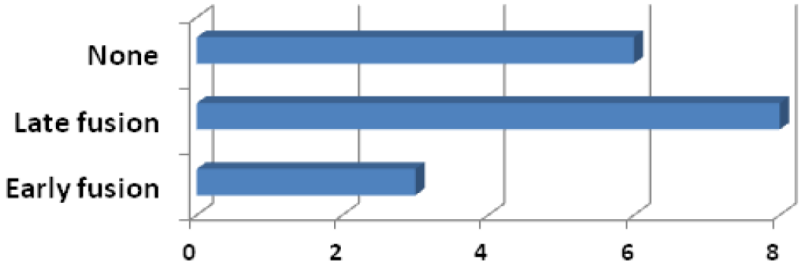
Modalities considered.



Segmentation strategy.



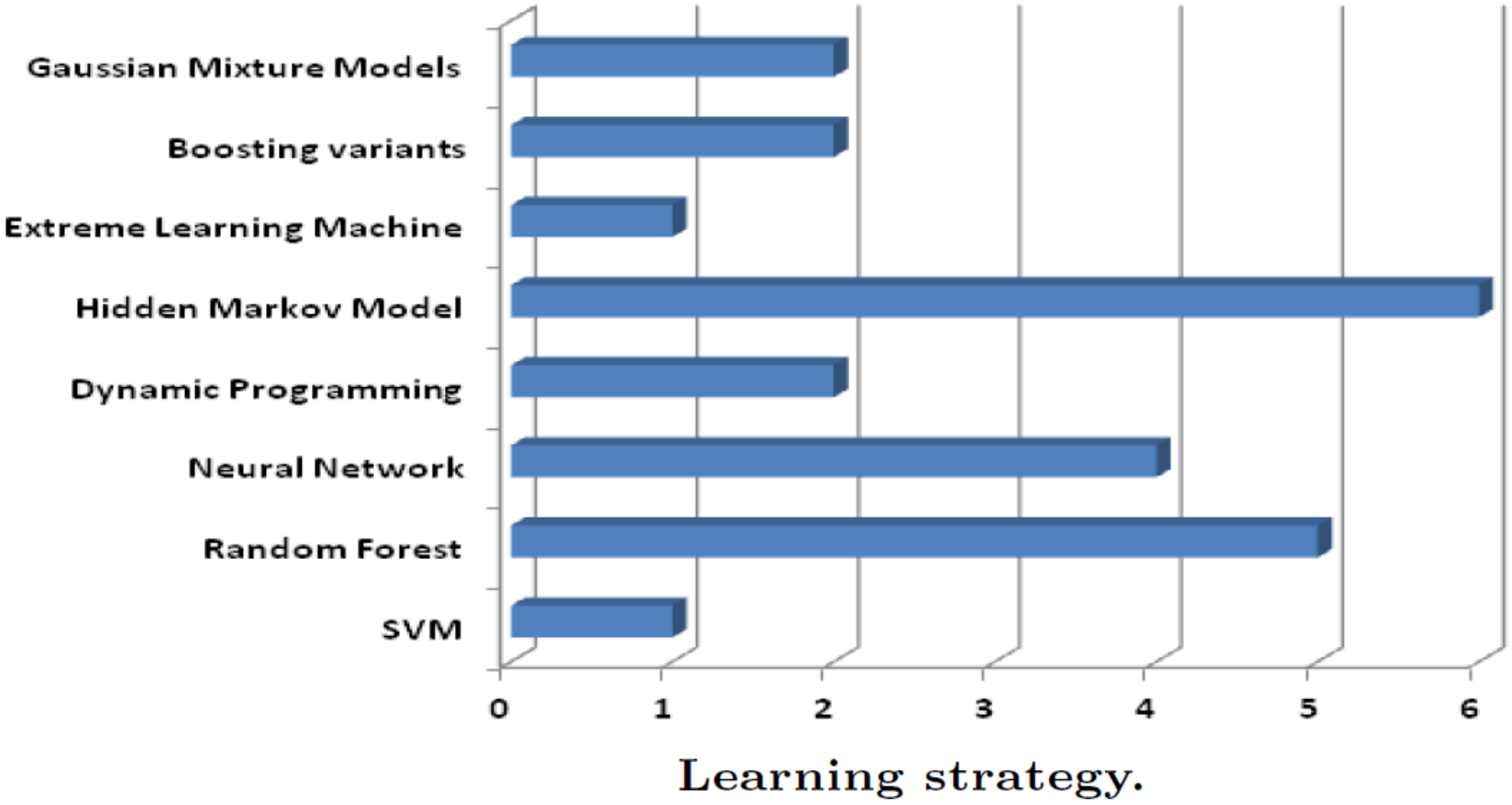
Programming language.



Fusion strategy.

ChaLearn Looking at People Challenges and Workshops

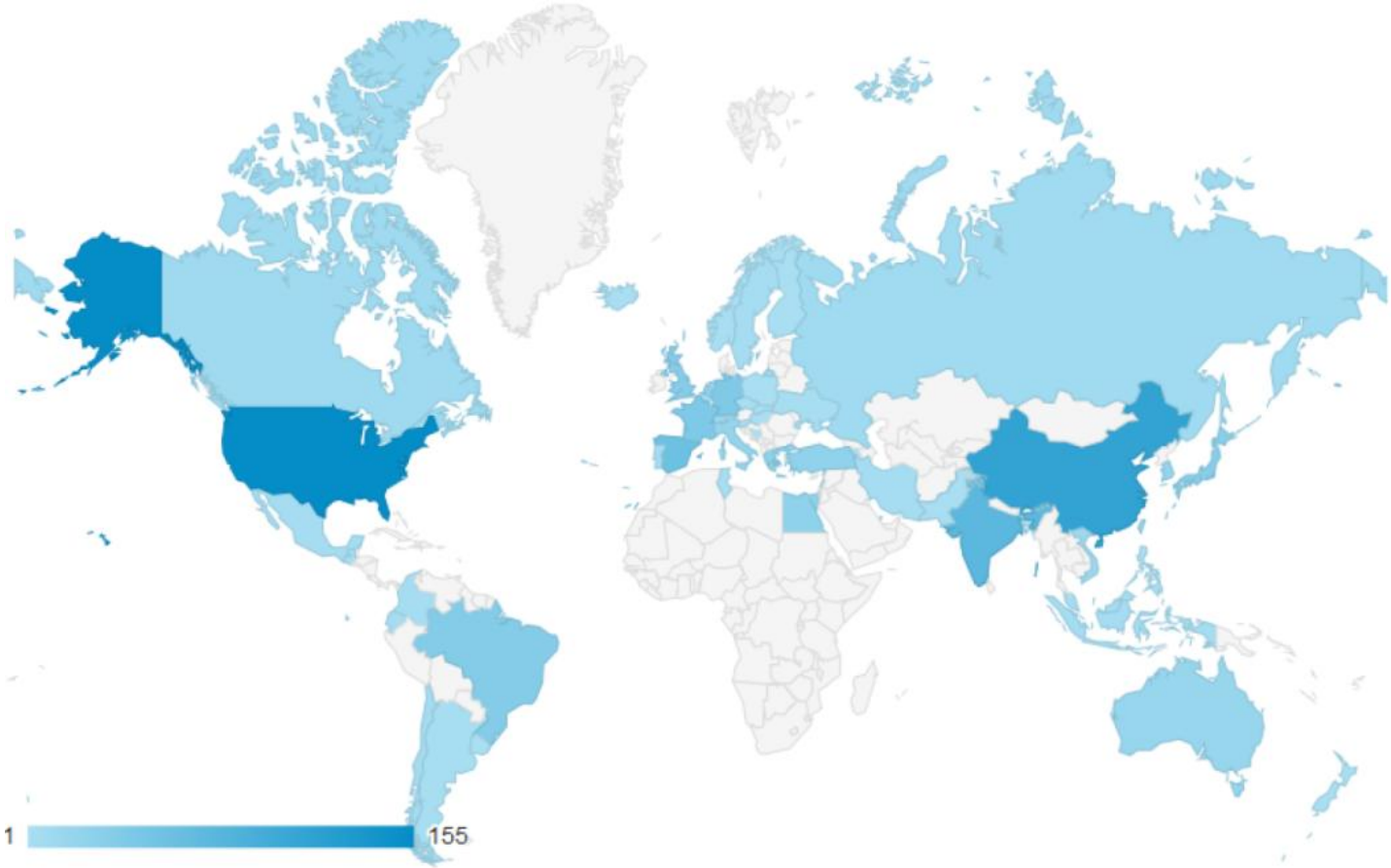
ICMI 2013 Workshop and Challenge on Multi-modal gesture recognition



ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

- **Connectivity:** During the Challenge period, the download page had a total of 2.895 visits from 920 different users of 59 countries.

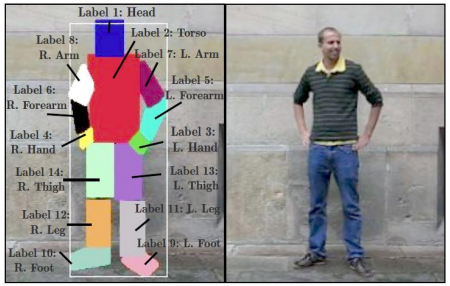


ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

- **Track1 results**

CodaLab



Team	Accuracy	Rank position	Features	Pose model
ZJU	**0.194144**	1	HOG	tree structure
Seawolf Vision	**0.182097**	2	HOG	tree structure

Track 1 Pose Recovery results.

Both winner participants applied a similar approach based on [*].

- Mixture of templates for each part. This method incorporates the co-occurrence relations, appearance and deformation into a model represented by an objective function of pose configurations. Model is tree-structured, and optimization is conducted via dynamic programming.

[*] Yang, Y., Ramanan, D.: Articulated human detection with flexible mixtures of parts. IEEE TPAMI (2013)

ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

- **Track2 results**

Team name	Accuracy	Rank	Features
CUHK-SWJTU	**0.507173**	1	Improved trajectories [*]
ADSC	**0.501164**	2	Improved trajectories [*]
SBUVIS	**0.441405**	3	Improved trajectories [*]
DonkeyBurger	0.342192	4	MHI, STIP
UC-T2	0.121565	5	Improved trajectories [*]
MindLAB	0.008383	6	MBF

Team name	Dimension reduction	Clustering	Classifier	Temporal coherence	Gesture representation
CUHK-SWJTU	PCA	-	SVM	Sliding windows	Fisher Vector
ADSC	-	-	SVM	Sliding windows	-
SBUVIS	-	-	SVM	Sliding windows	-
DonkeyBurger	-	Kmeans	Sparse code	Sliding windows	-
UC-T2	PCA	-	Kmeans	Sliding windows	Fisher Vector
MindLAB	-	Kmeans	RF	Sliding windows	BoW

* Wang, H., Schmid, C.: Action recognition with improved trajectories. ICCV (2013)

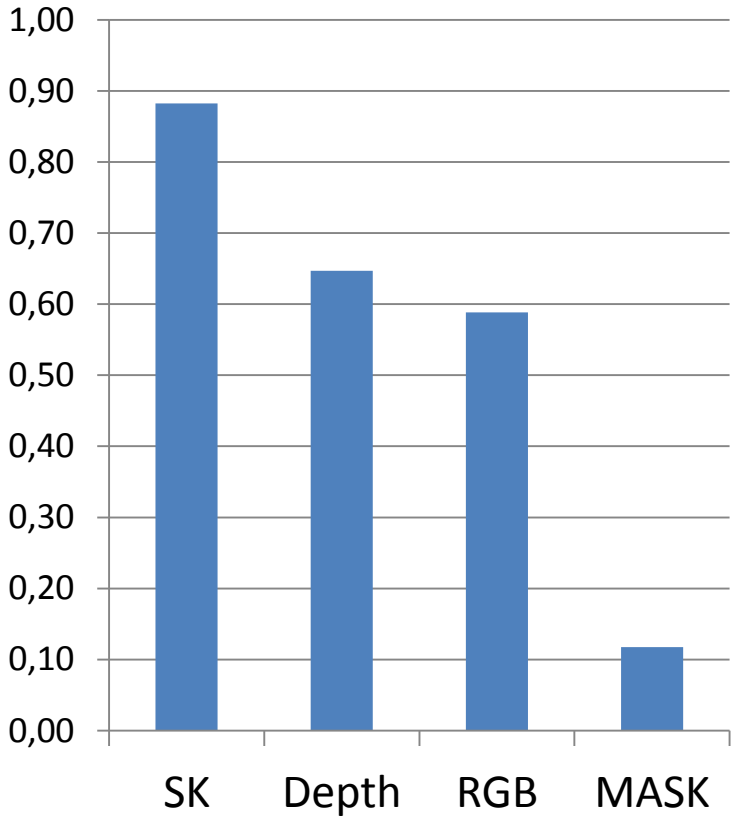
ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

- **Track3 results**

Team	Accuracy	Rank	Modalities
LIRIS	**0.849987**	1	SK, Depth, RGB
CraSPN	**0.833904**	2	SK, Depth, RGB
JY	**0.826799**	3	SK, RGB
CUHK-SWJTU	0.791933	4	RGB
Lpigou	0.788804	5	Depth, RGB
stevenwudi	0.787310	6	SK, depth
Ismar	0.746632	7	SK
Quads	0.745449	8	SK
Telepoints	0.688778	9	SK, Depth, RGB
TUM-fortiss	0.648979	10	SK, Depth, RGB
CSU-SCM	0.597177	11	Skeleton, Depth, mask
iva.mm	0.556251	12	Skeleton, RGB, depth
Terrier	0.539025	13	Skeleton
Team Netherlands	0.430709	14	Skeleton, Depth, RGB
VecsRel	0.408012	15	Skeleton, Depth, RGB
Samgest	0.391613	16	Skeleton, Depth, RGB, mask
YNL	0.270600	17	Skeleton

Percentage of methods using each independent modality

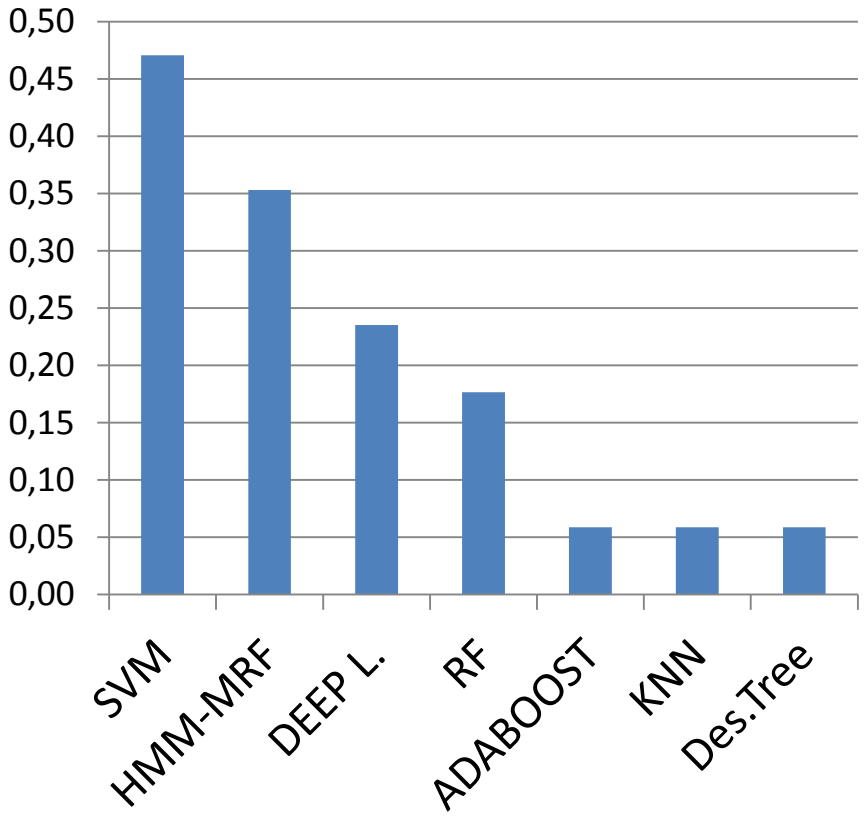


ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

- **Track3 results** **Percentage of methods using each gesture classification strategy**

Team	Gesture representation	Classifier
LIRIS	-	DNN
CraSPN	BoW	Adaboost
JY	-	MRF, KNN
CUHK-SWJTU	Fisher Vector, VLAD	SVM
Lpigou	-	CNN
stevenwudi	-	HMM, DNN
Ismar	-	RF
Quads	Fisher Vector	SVM
Telepoints	-	SVM
TUM-fortiss	-	RF, SVM
CSU-SCM	2DMTM	SVM, HMM
iva.mm	BoW	SVM, HMM
Terrier	-	RF
Team Netherlands	-	SVM, RT
VecsRel	-	DNN
Samgest	-	HMM
YNL	Fisher Vector	HMM, SVM



ChaLearn Looking at People Challenges and Workshops

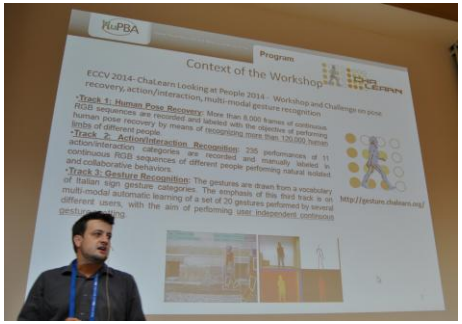
ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting

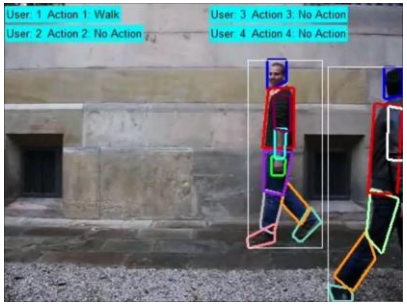
Team	Features	Fusion	Temp. segmentation	Dimension reduction
LIRIS	RAW, SK joints	Early	Joints motion	-
CraSPN	HOG, SK	Early	Sliding windows	-
JY	SK, HOG	Late	MRF	PCA
CUHK-SWJTU	Improved trajectories	-	Joints motion	PCA
Lpigou	RAW, SK joints	Early	Sliding windows	Max-pooling CNN
stevenwudi	RAW	Late	Sliding windows	-
Ismar	SK	-	Sliding windows	-
Quads	SK quads	-	Sliding windows	-
Telepoints	STIPS, SK	Late	Joints motion	-
TUM-fortiss	STIPS	Late	Joints motion	-
CSU-SCM	HOG, Skeleton	Late	Sliding windows	-
iva.mm	Skeleton, HOG	Late	Sliding windows	-
Terrier	Skeleton	-	Sliding windows	-
Team Netherlands	MHI	Early	DTW	Preserving projections
VecsRel	RAW, skeleton joints	Late	DTW	-
Samgest	Skeleton, blobs, moments	Late	Sliding windows	-
YNL	Skeleton	-	Sliding windows	-

For more details of the challenge and the results: *Sergio Escalera, Xavier Baró, Jordi Gonzàlez, Miguel Ángel Bautista, Meysam Madadi, Miguel Reyes, Víctor Ponce-López, Hugo J. Escalante, Jamie Shotton, Isabelle Guyon, ChaLearn Looking at People Challenge 2014: Dataset and Results, ChaLearn Looking at People, European Conference on Computer Vision, 2014.*

ChaLearn Looking at People Challenges and Workshops

ECCV 2014 Workshop and Challenge on multi-modal gesture spotting, human pose recovery, action/interaction spotting





•JOURNAL OF MACHINE LEARNING RESEARCH, SPECIAL TOPIC ON GESTURE RECOGNITION

JMLR Journal of Machine Learning Research

•IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENTE SPECIAL ISSUE

Call for Papers IEEE Transactions on Pattern Analysis and Machine Intelligence

Special Issue on Multimodal Human Pose Recovery and Behavior Analysis – M²HuPBA

DEADLINE 1 DECEMBER 2014

Invited speakers ICMI2013-ECCV2014

Professor Leonid Sigal, Disney Research

Professor Cristian Sminchisescu, Lund University

Professor Richard Bowden, University of Surrey

Professor Adrian Hilton University of Surrey

Professor Antonis Argyros, Univ. of Crete

Carol Neidle, Boston Linguistics

Diitris Metaxas, CBIM

Fernando de la Torre, CMU

Jeffrey Cohn, CMU

Upcoming ChaLearn Looking at People competitions:

- Track (2015): Age recognition **CVPR 2015**
- Track (2015): Cultural event recognition **ICCV 2015**
- Track (2016): Social signal and affective computing in interview scenarios

ChaLearn LAP
5.Sponsors/org.

Organizers

Universitat de Barcelona

Microsoft Research

Computer Vision Center

Universitat Autònoma de Barcelona

Universitat Oberta de Catalunya
www.uoc.edu

INAOE

Sponsors

kaggle

amazon

visada

PASCAL2
Pattern Analysis, Statistical Modelling and Computational Learning

Microsoft Research

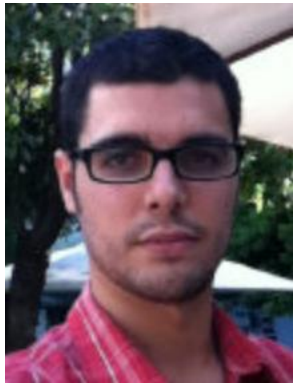
Universitat de Barcelona

2014 organizers


Sergio Escalera


Jordi Gonzàlez


Xavier Báró


Miguel Reyes


Víctor Ponce


Meysam Madadi


M. Ángel Bautista


Isabelle Guyon


Hugo J. Escalante


Jamie Shotton

ChaLearn LAP challenges and news:

http://gesture.chalearn.org/

Organization of ChaLean Looking at People requires:

-Good ideas to solve real problems focused on humans
-Collecting data
-Labeling tools
-Dissemination and repositories
-Baseline designs based on state of the art approaches
-Online platform for the competition
-Sponsoring
-Presentation of the results in a relevant events
-Organization of special issues and challenge report documents, making competition data public for the scientific community

For each competition many organizers contribute. Our plan is to perform yearly challenges.
Feel free to contact us if you want to be included in our ChaLearn LAP mailing list or collaborate in some aspect propose ideas related to ChaLearn Looking at People competitions:

sergio@maia.ub.es

Thank you

- **Thank you and hope to see you in our next event!**