ChaLearn Looking at People Challenge 2015: Dataset and Results

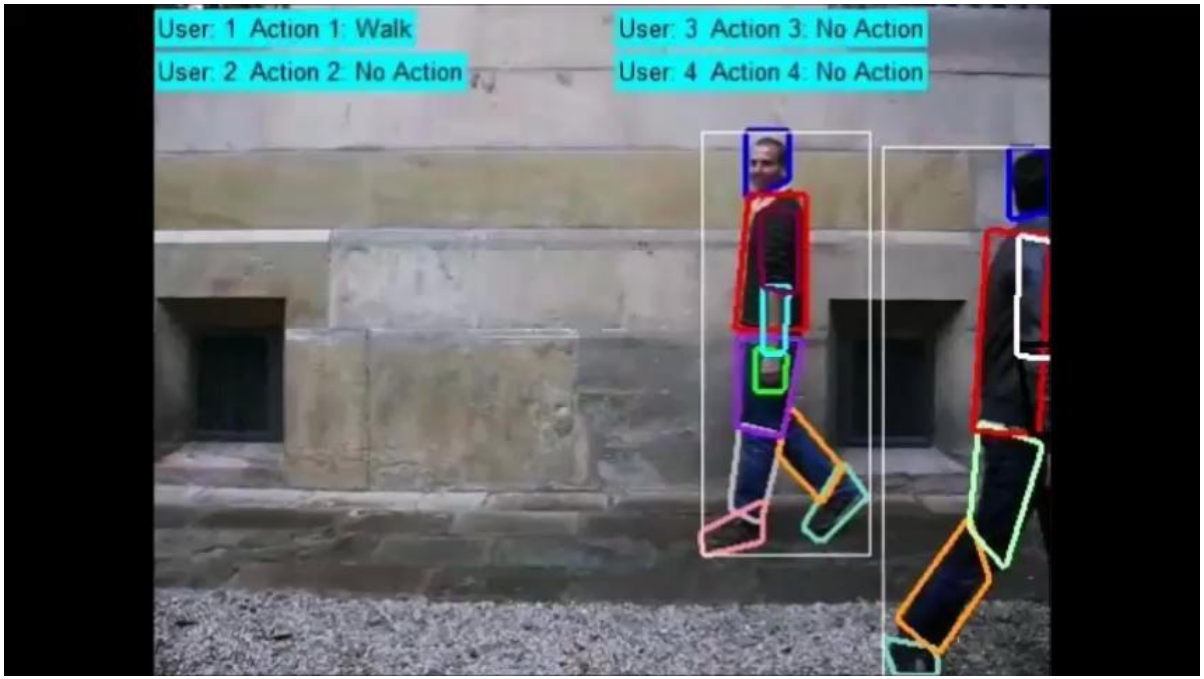
http://gesture.chalearn.org/

Xavier Baró, UOC & CVC
Jordi Gonzàlez, UAB & CVC
Junior Fabian, UAB & CVC
Miguel Ángel Bautista, UB & CVC
Marc Oliu, UB & CVC
Hugo J. Escalante, INAOE & ChaLearn
Isabelle Guyon, ChaLearn
Sergio Escalera, UB & CVC & ChaLearn

http://gesture.chalearn.org/

Challenge on multi-modal gesture recognition and cultural event recognition

•**Track on 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.
(Japan), among others.



http://gesture.chalearn.org/

Challenge on multi-modal gesture recognition and cultural event recognition

•**Track on Cultural Event Recognition:** More than 10,000 images corresponding to 50 different cultural event categories will be considered. Examples of cultural events will be Carnival (Brasil, Italy, USA), Oktoberfest (Germany), San Fermin (Spain), Maha-Kumbh-Mela (India) and Aoi-Matsuri (Japan), among others.

•**Track on 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.

http://gesture.chalearn.org/

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

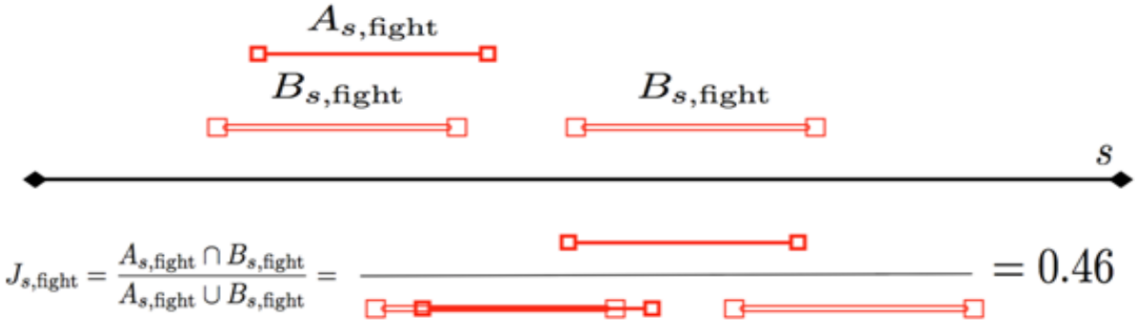
Modalities	Num. of users	Action categories	interaction categories	Labeled sequences
RGB	14	7	4	235

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

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



•**Track on 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.

http://gesture.chalearn.org/

Action categories


Wave


Point


Clap


Crouch


Jump


Walk

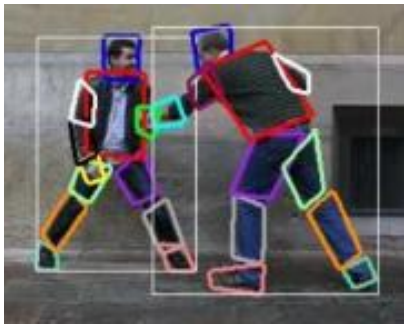

Run

Interaction categories


Shake Hands


Hug


Kiss


Fight

http://gesture.chalearn.org/

•**Track on Cultural Event Recognition:** More than 10,000 images corresponding to 50 different cultural event categories will be considered.

Dataset	#Images	#Classes	Year
Action Classification Dataset [8]	5,023	10	2010
Social Event Dataset [11]	160,000	149	2012
Event Identification Dataset [1]	594,000	24,900	2010
Cultural Event Dataset	11,776	50	2015

- **First dataset** on cultural events
- **10.000 images** corresponding to **50 cultural events**.
- **Person related** events.
- High intra and inter-class variability.
- **Different cues** can be exploited like garments, human poses, crowds analysis, objects and background scene.

http://gesture.chalearn.org/

•**Track on Cultural Event Recognition:** More than 10,000 images corresponding to 50 different cultural event categories will be considered.

Inter-class variability

http://gesture.chalearn.org/

•**Track on Cultural Event Recognition:** More than 10,000 images corresponding to 50 different cultural event categories will be considered.

Inter-class variability

Carnival of Dunkerque

Carnival of Rio

Carnival of Venice

Carnival of Helsinki

Nothing Hill Carnival

Carnival of Quebec

http://gesture.chalearn.org/

•**Track on Cultural Event Recognition:** More than 10,000 images corresponding to 50 different cultural event categories will be considered.

Inter-class variability

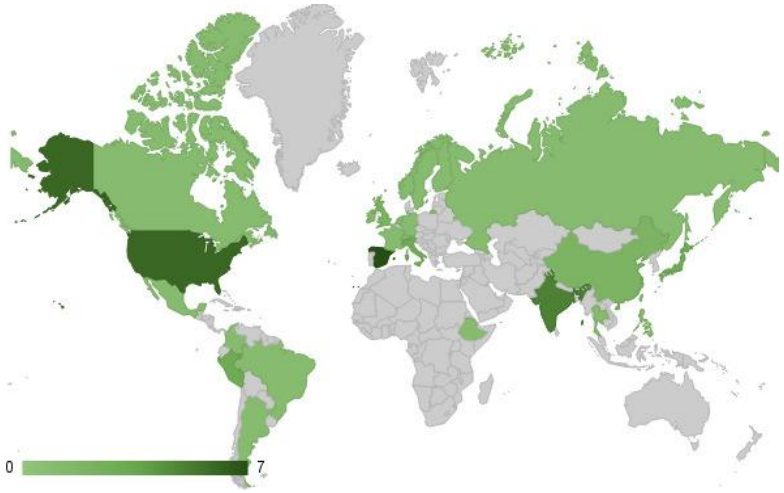
Quebec Winter Carnival

Harbin Ice and Snow Festival

http://gesture.chalearn.org/

•**Track on Cultural Event Recognition:** More than 10,000 images corresponding to 50 different cultural event categories will be considered.

Cultural Event	Country	#Images
1. Annual Buffalo Roundup	USA	334
2. Ati-atihan	Philippines	357
3. Ballon Fiesta	USA	382
4. Basel Fasnacht	Switzerland	310
5. Boston Marathon	USA	271
6. Bud Billiken	USA	335
7. Buenos Aires Tango Festival	Argentina	261
8. Carnival of Dunkerque	France	389
9. Carnival of Venice	Italy	455
10. Carnival of Rio	Brazil	419
11. Castellers	Spain	536
12. Chinese New Year	China	296
13. Correfocs	Catalonia	551
14. Desert Festival of Jaisalmer	India	298
15. Desfile de Silleteros	Colombia	286
16. Día de los Muertos	Mexico	298
17. Diada de Sant Jordi	Catalonia	299
18. Diwali Festival of Lights	India	361
19. Falles	Spain	649
20. Festa del Renaixement Tortosa	Catalonia	299
21. Festival de la Marinera	Peru	478
22. Festival of the Sun	Peru	514
23. Fiesta de la Candelaria	Peru	300
24. Gion matsuri	Japan	282
25. Harbin Ice and Snow Festival	China	415
26. Heiva	Tahiti	286
27. Helsinki Samba Carnival	Finland	257
28. Holi Festival	India	553
29. Infiorata di Genzano	Italy	354
30. La Tomatina	Spain	349
31. Lewes Bonfire	England	267
32. Macys Thanksgiving	USA	335
33. Maslenitsa	Russia	271



34. Midsommar	Sweden	323
35. Notting hill carnival	England	383
36. Obon Festival	Japan	304
37. Oktoberfest	Germany	509
38. Onbashira Festival	Japan	247
39. Pingxi Lantern Festival	Taiwan	253
40. Pushkar Camel Festival	India	433
41. Quebec Winter Carnival	Canada	329
42. Queens Day	Netherlands	316
43. Rath Yatra	India	369
44. SandFest	USA	237
45. San Fermin	Spain	418
46. Songkran Water Festival	Thailand	398
47. St Patrick's Day	Ireland	320
48. The Battle of the Oranges	Italy	276
49. Timkat	Ethiopia	425
50. Viking Festival	Norway	262

http://gesture.chalearn.org/

•**Track on Cultural Event Recognition:** More than 10,000 images corresponding to 50 different cultural event categories will be considered.

Average Precision evaluation

For each image, participants submit their confidence for all the categories.

1. Precision/recall curve computed with precision monotonically decreasing.
2. AP is computed by numeric integration, using the trapezoidal rule.

$$\int_a^b f(x)dx \approx (b-a)\left[\frac{f(a)+f(b)}{2}\right]$$

http://gesture.chalearn.org/

Competition schedule

The challenge was managed using the Microsoft Codalab platform. The schedule of the competition was as follows:

- **December 1st, 2014:** Beginning of the quantitative competition, release of development and validation data.
- **February 15th, 2015**: Release of encrypted final evaluation data and validation labels. Participants can start training their methods with the whole data set.
- **March 13th, 2015:** Release of final evaluation data decryption key. Participants start predicting the results on the final evaluation data.
- **March 20th, 2015:** End of the quantitative competition. Deadline for submitting the predictions over the final evaluation data. Deadline for code submission. The organizers start the code verification by running it on the final evaluation data.
- **March 25th, 2015:** Deadline for submitting the fact sheets.
- **March 27th, 2015:** Release of the verification results to the participants for review. Top ranked participants are invited to follow the workshop submission guide for inclusion at CVPR 2015 ChaLearn Looking at People workshop proceedings.

http://gesture.chalearn.org/

Participation

- We created a different competition for each track, having the specific information and leaderboard.
- **A total of 116 users has been registered in the Codalab platform:**
  - **62 for action/interaction track**
  - **54 for cultural event track**
- All these users were able to access the data for the Developing stage, and submit their predictions for this stage. For the final evaluation stage, a team registration was mandatory, and a total of **8 teams were successfully registered:**
  - **2 for action/interaction track**
  - **6 for cultural event track**
- **Only registered teams had access to the data for the last stage.**
- The data was downloadable from the Codalab platform.

http://gesture.chalearn.org/

Track on Action/interaction results

Action/Interaction Track								
Rank	Team name	Score	Features	Dimension reduction	Clustering	Classification	Temporal coherence	Action representation
1	MMLAB	0.5385	IDT [19]	PCA	-	SVM	-	Fisher Vector
2	FIKIE	0.5239	IDT	PCA	-	HMM	Appearance+Kalman filter	-

- Both methods are based on the Improved Dense Trajectories
- PCA is used for dimension reduction
- The first team uses fisher vectors for action representation
- The second team uses tracking with Kalman Filters
- Generative vs Discriminative classifiers
  - Both strategies have been used.

http://gesture.chalearn.org/

Track on Action/interaction results

- In the case of **action/interaction** RGB data sequences, Improved Dense Trajectories are the standard **action description**.
- No general rule for classifiers, generative and discriminative models used with similar results.
- Stalled methodologies. From last challenge only fine tune has been performed, with a performance increment of just 3%.

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

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

http://gesture.chalearn.org/

Track on Cultural event recognition Results

Cultural Event Track				
Rank	Team name	Score	Features	Classification
1	MMLAB	0.855	Multiple CNN	Late weighted fusion of CNNs predictions.
2	UPC-ST	0.767	Multiple CNN	SVM and late weighted fusion.
3	MIPAL_SNU	0.735	Discriminant regions [18] + CNNs	Entropy + Mean Probabilities of all patches
4	SBU_CS	0.610	CNN-M [2]	SPM [10] based on LSSVM [16]
5	MasterBlaster	0.58	CNN	SVM, KNN, LR and One Vs Rest
6	Nyx	0.319	Selective-search approach [17] + CNN	Late fusion AdaBoost

- All the teams are using at least on CNN
  - Pre-trained CNNs
- Many late-fusion strategies
  - From the final layer of the CNN
  - Use fine-tuned features as input to classifiers

http://gesture.chalearn.org/

Track on Cultural event recognition Results

- In the case of **Cultural Event Recognition**, all teams use only CNN for description.
- Not enough images for CNN training, pre-trained CNNs used.
- Different methodologies for CNN fusing.
  - Ad-hoc methodologies addressed to solve the problem
- No new methodologies applied
  - No specific methods to take advantage of the different available cues.
- 85% of average precision obtained. There is still room for improvement.

http://gesture.chalearn.org/

Track on Cultural event recognition Results

- Hard classes


Chinese New Year


Falles


Infiorata Genzano


Maslenitza


Nothin Hill Carn.

- Easy classes

Boston Marathon


Carnaval of Venice


Desf. Silleteros


Oktoberfest

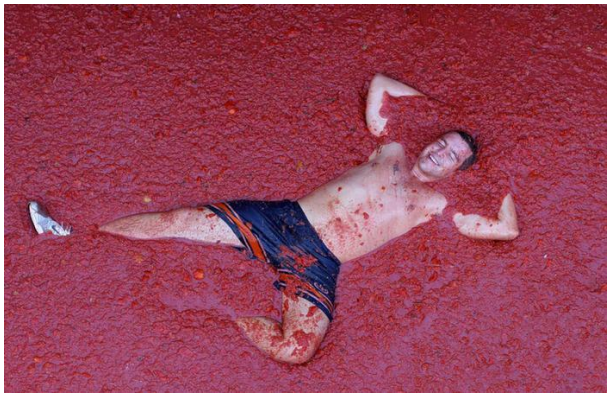

Batle of Oranges

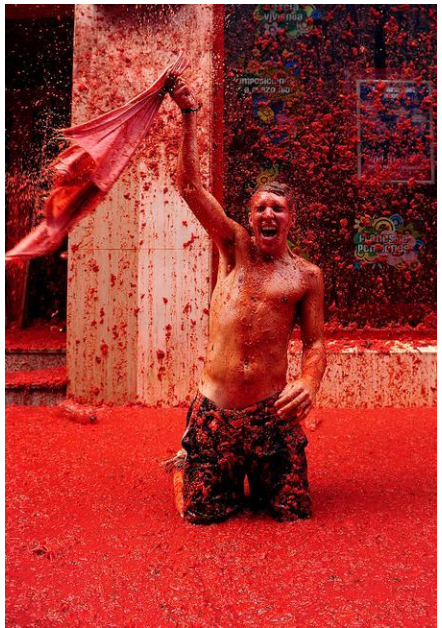
Track on Cultural event recognition Results

http://gesture.chalearn.org/

- No colour cue used may be the reason for bad results on classes like Tomatina








Organizers

Sponsors

http://gesture.chalearn.org/

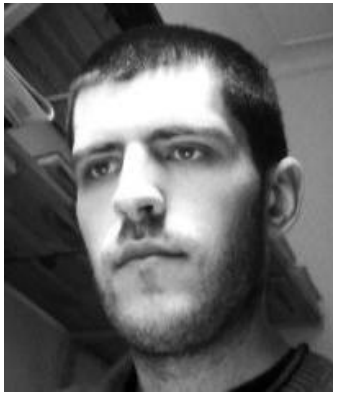
Organizers


Sergio Escalera


Jordi Gonzàlez


Xavier Báró


Junior Fabian


Marc Oliu


M. Ángel Bautista


Isabelle Guyon


Hugo J. Escalante

http://gesture.chalearn.org/

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

ChaLearn LAP challenges and news:
http://gesture.chalearn.org/

Send us en email if you want to be included in our ChaLearn LAP mailing list:
mmgesture@chalearn.org