ChaLearn Looking at People 2015 new competitions: Age Estimation and Cultural Event Recognition

Sergio Escalera, UB & CVC & ChaLearn
Jordi Gonzàlez, UAB & CVC
Xavier Baró, UOC & CVC
Pablo Pardo, UB
Junior Fabian, UAB & CVC
Marc Oliu, UB & CVC
Hugo J. Escalante, INAOE & ChaLearn
Ivan Huerta, University of Venezia
Isabelle Guyon, ChaLearn

http://gesture.chalearn.org/
Index

• ChaLearn Looking at People events
• ChaLearn Looking at People Cultural Event Recognition
• ChaLearn Looking at People Age Recognition
• Ongoing and upcoming events
ChaLearn Gesture Recognition Challenges and Workshops

**CVPR 2011** - Workshop and Challenge on Gesture Recognition

**CVPR 2012** - Workshop and Challenge on Gesture Recognition

**ICPR 2012** - Workshop and Challenge on Gesture Recognition

**ICMI 2013** - Workshop and Challenge on Gesture Recognition

**ECCV 2014** – Workshop and Challenge on Human Pose, Action and Gesture Recognition

**Quantitative competition:**
- One-shot learning
- New depth-rgb data set
- Dictionaries among 5-8 gesture categories
- Leveinstein: recognizing list of sequences within each sequence

**Quantitative competition:**
- User independent multiple instance learning
- New depth-rgb-mask-skeleton-audio data set
- Dictionary of 20 gesture categories
- Leveinstein: recognizing list of gestures within each sequence
ECCV 2014 – Workshop and Challenge on Human Pose, Action and Gesture Recognition

CVPR 2015 - ChaLearn Looking at People 2015 - Action spotting and cultural event recognition

ICCV 2015 - ChaLearn Looking at People 2015 - Age recognition and extended cultural event recognition
Invited speakers: Antonis Argyros, Takeo Kanade, Deva Ramanan, Stan Sclaroff, Cordelia Schmid, Fernando de la Torre, Jeffrey Cohn, Tinne Tuytellers, Leonid Sigal, Larry Davis, among others. Special issues: JMLR, TPAMI, (IJCV LaP opened until February 2016)
Challenge on 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.

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

<table>
<thead>
<tr>
<th>Dataset</th>
<th>#Images</th>
<th>#Classes</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Action Classification Dataset [8]</td>
<td>5,023</td>
<td>10</td>
<td>2010</td>
</tr>
<tr>
<td>Event Identification Dataset [1]</td>
<td>594,000</td>
<td>24,900</td>
<td>2010</td>
</tr>
<tr>
<td><strong>Cultural Event Dataset</strong></td>
<td>11,776</td>
<td>50</td>
<td>2015</td>
</tr>
</tbody>
</table>

• **First dataset** on cultural events
• **10,000 images** corresponding to **50 cultural events**.
• **Person related** events.
• High intra and low inter-class variability.
• **Different cues** can be exploited like garments, human poses, crowds analysis, objects and background scene.
• **Track on Cultural Event Recognition**: More than 10,000 images corresponding to 50 different cultural event categories will be considered.

**Inter-class variability**

[Images showing various cultural events and activities]
• **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
• 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

Average Precision evaluation
For each image, participants submit their confidence for all the categories (Average Precision).
**Track on Cultural Event Recognition:** More than 10,000 images corresponding to 50 different cultural event categories will be considered.

<table>
<thead>
<tr>
<th>Cultural Event</th>
<th>Country</th>
<th>#Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Annual Buffalo Roundup</td>
<td>USA</td>
<td>334</td>
</tr>
<tr>
<td>2. Ati-atihan</td>
<td>Philippines</td>
<td>357</td>
</tr>
<tr>
<td>3. Ballon Fiesta</td>
<td>USA</td>
<td>382</td>
</tr>
<tr>
<td>4. Basel Fastnacht</td>
<td>Switzerland</td>
<td>310</td>
</tr>
<tr>
<td>5. Boston Marathon</td>
<td>USA</td>
<td>271</td>
</tr>
<tr>
<td>6. Bud Billiken</td>
<td>USA</td>
<td>335</td>
</tr>
<tr>
<td>7. Buenos Aires Tango Festival</td>
<td>Argentina</td>
<td>261</td>
</tr>
<tr>
<td>8. Carnival of Dunkerque</td>
<td>France</td>
<td>389</td>
</tr>
<tr>
<td>9. Carnival of Venice</td>
<td>Italy</td>
<td>455</td>
</tr>
<tr>
<td>10. Carnival of Rio</td>
<td>Brazil</td>
<td>419</td>
</tr>
<tr>
<td>11. Castellers</td>
<td>Spain</td>
<td>536</td>
</tr>
<tr>
<td>12. Chinese New Year</td>
<td>China</td>
<td>296</td>
</tr>
<tr>
<td>13. Correfocs</td>
<td>Catalonia</td>
<td>551</td>
</tr>
<tr>
<td>14. Desert Festival of Jaisalmer</td>
<td>India</td>
<td>298</td>
</tr>
<tr>
<td>15. Desfile de Silteiros</td>
<td>Colombia</td>
<td>286</td>
</tr>
<tr>
<td>16. Dia de los Muertos</td>
<td>Mexico</td>
<td>298</td>
</tr>
<tr>
<td>17. Diada de Sant Jordi</td>
<td>Catalonia</td>
<td>299</td>
</tr>
<tr>
<td>18. Diwali Festival of Lights</td>
<td>India</td>
<td>361</td>
</tr>
<tr>
<td>19. Falles</td>
<td>Spain</td>
<td>649</td>
</tr>
<tr>
<td>20. Festa del Renaiement Tortosa</td>
<td>Catalonia</td>
<td>299</td>
</tr>
<tr>
<td>21. Festival de la Marinera</td>
<td>Peru</td>
<td>478</td>
</tr>
<tr>
<td>22. Festival of the Sun</td>
<td>Peru</td>
<td>514</td>
</tr>
<tr>
<td>23. Fiesta de la Candeleria</td>
<td>Peru</td>
<td>300</td>
</tr>
<tr>
<td>24. Gion matsuri</td>
<td>Japan</td>
<td>282</td>
</tr>
<tr>
<td>25. Harbin Ice and Snow Festival</td>
<td>China</td>
<td>415</td>
</tr>
<tr>
<td>26. Helaiva</td>
<td>Tahiti</td>
<td>286</td>
</tr>
<tr>
<td>27. Helsinki Samba Carnival</td>
<td>Finland</td>
<td>257</td>
</tr>
<tr>
<td>28. Holi Festival</td>
<td>India</td>
<td>553</td>
</tr>
<tr>
<td>29. Infiorni di Genzano</td>
<td>Italy</td>
<td>354</td>
</tr>
<tr>
<td>30. La Tomatina</td>
<td>Spain</td>
<td>349</td>
</tr>
<tr>
<td>31. Lewes Bonfire</td>
<td>England</td>
<td>267</td>
</tr>
<tr>
<td>32. Macys Thanksgiving</td>
<td>USA</td>
<td>335</td>
</tr>
<tr>
<td>33. Maldentosa</td>
<td>Russia</td>
<td>271</td>
</tr>
<tr>
<td>34. Midsommar</td>
<td>Sweden</td>
<td>323</td>
</tr>
<tr>
<td>35. Notting hill carnival</td>
<td>England</td>
<td>383</td>
</tr>
<tr>
<td>36. Obon Festival</td>
<td>Japan</td>
<td>304</td>
</tr>
<tr>
<td>37. Oktoberfest</td>
<td>Germany</td>
<td>509</td>
</tr>
<tr>
<td>38. Onbashira Festival</td>
<td>Japan</td>
<td>247</td>
</tr>
<tr>
<td>39. Pingxi Lantern Festival</td>
<td>Taiwan</td>
<td>253</td>
</tr>
<tr>
<td>40. Pushkar Camel Festival</td>
<td>India</td>
<td>433</td>
</tr>
<tr>
<td>41. Quebec Winter Carnival</td>
<td>Canada</td>
<td>329</td>
</tr>
<tr>
<td>42. Queens Day</td>
<td>Netherlands</td>
<td>316</td>
</tr>
<tr>
<td>43. Rath Yatra</td>
<td>India</td>
<td>369</td>
</tr>
<tr>
<td>44. SandFest</td>
<td>USA</td>
<td>237</td>
</tr>
<tr>
<td>45. San Fermin</td>
<td>Spain</td>
<td>418</td>
</tr>
<tr>
<td>46. Songkran Water Festival</td>
<td>Thailand</td>
<td>398</td>
</tr>
<tr>
<td>47. St Patrick’s Day</td>
<td>Ireland</td>
<td>320</td>
</tr>
<tr>
<td>48. The Battle of the Oranges</td>
<td>Italy</td>
<td>276</td>
</tr>
<tr>
<td>49. Timkat</td>
<td>Ethiopia</td>
<td>425</td>
</tr>
<tr>
<td>50. Viking Festival</td>
<td>Norway</td>
<td>262</td>
</tr>
</tbody>
</table>
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.
Participation

- We created a different competition for each track, having the specific information and leaderboard.
- A total of 54 users has been registered in the Codalab platform for cultural event recognition.
- 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 6 teams were successfully registered.
- Only registered teams had access to the data for the last stage.
- The data was downloadable from the Codalab platform.
Track on Cultural event recognition Results

<table>
<thead>
<tr>
<th>Rank</th>
<th>Team name</th>
<th>Score</th>
<th>Features</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MMLAB</td>
<td>0.855</td>
<td>Multiple CNN</td>
<td>Late weighted fusion of CNNs predictions.</td>
</tr>
<tr>
<td>2</td>
<td>UPC-ST</td>
<td>0.767</td>
<td>Multiple CNN</td>
<td>SVM and late weighted fusion.</td>
</tr>
<tr>
<td>3</td>
<td>MIPAL_SNU</td>
<td>0.735</td>
<td>Discriminant regions [18] + CNNs</td>
<td>Entropy + Mean Probabilities of all patches</td>
</tr>
<tr>
<td>5</td>
<td>MasterBlaster</td>
<td>0.58</td>
<td>CNN</td>
<td>SVM, KNN, LR and One Vs Rest</td>
</tr>
<tr>
<td>6</td>
<td>Nyx</td>
<td>0.319</td>
<td>Selective-search approach [17] + CNN</td>
<td>Late fusion AdaBoost</td>
</tr>
</tbody>
</table>

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

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

[Images of people covered in red pigment, likely from a cultural event like Tomatina, with a link to http://gesture.chalearn.org/ for more information.]
Crowdsourcing application using Facebook API for uploading and voting apparent age

http://sunai.uoc.edu:8005
Profile: your uploaded photos, the real age, apparent age and number of votes
Gamification fashion: Ranking among colleagues and rest of participants. Winning points for voting near mean vote, number of voted images, number of images uploaded, etc.
Updated data

<table>
<thead>
<tr>
<th>Features</th>
<th>HuPBA¹</th>
<th>AgeGuess²</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Images</td>
<td>1506</td>
<td>3359</td>
<td>4865</td>
</tr>
<tr>
<td>Users</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>44</td>
<td>1828</td>
<td>1872</td>
</tr>
<tr>
<td>male</td>
<td>110</td>
<td>1143</td>
<td>1253</td>
</tr>
<tr>
<td>Total</td>
<td>154</td>
<td>2971</td>
<td>3125</td>
</tr>
<tr>
<td>Votes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>1753</td>
<td>75136</td>
<td>76889</td>
</tr>
<tr>
<td>male</td>
<td>14897</td>
<td>53117</td>
<td>68004</td>
</tr>
<tr>
<td>Total</td>
<td>16640</td>
<td>128253</td>
<td>144893</td>
</tr>
</tbody>
</table>

¹ HuPBA web application: http://sunai.uoc.edu:8005
² AgeGuess web application: http://www.ageguess.org/
Age recognition

http://gesture.chalearn.org/
Interesting findings: with the previous version of the apparent age data set. The baseline method (BIF features) showed better performance with apparent age than with real age.
Evaluation

Competition already started: join us!

https://www.codalab.org/competitions/4711

Schedule details:

http://gesture.chalearn.org/
Evaluation – competition open until middle september 2015
Performed baseline

- **Deep Learning Method:** based on Convolutional Neural Networks (CNN)
- **Face Detection:** OpenCV Viola & Jones implementation.
- **Grey Scale:** Transform RGB to grey scale.
- **Face Alignment:** Shape regressor using 68 facial landmarks by Shaoqing et al. ²

---

1 Bradski, G. published at *Dr. Dobb’s Journal of Software Tools* in 2000
2 Ren, Shaoqing and Cao, Xudong and Wei, Yichen and Sun, Jian, *Face Alignment at 3000 FPS via Regressing Local Binary Features*, CVPR 2014
Age recognition

Evaluation – competition open until middle september 2015
Performed baseline

<table>
<thead>
<tr>
<th>Layer</th>
<th>Input size</th>
<th>Output size</th>
<th>Filter size</th>
<th>Pooling size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conv1</td>
<td>200 × 200</td>
<td>190 × 190</td>
<td>10 × (11 × 11)</td>
<td>-</td>
</tr>
<tr>
<td>Pool1</td>
<td>190 × 190</td>
<td>95 × 95</td>
<td>-</td>
<td>(2, 2)</td>
</tr>
<tr>
<td>Conv2</td>
<td>95 × 95</td>
<td>89 × 89</td>
<td>20 × (7 × 7)</td>
<td>-</td>
</tr>
<tr>
<td>Pool2</td>
<td>89 × 89</td>
<td>44 × 44</td>
<td>-</td>
<td>(2, 2)</td>
</tr>
<tr>
<td>Conv3</td>
<td>44 × 44</td>
<td>40 × 40</td>
<td>40 × (5 × 5)</td>
<td>-</td>
</tr>
<tr>
<td>Pool3</td>
<td>40 × 40</td>
<td>20 × 20</td>
<td>-</td>
<td>(2, 2)</td>
</tr>
<tr>
<td>Full1</td>
<td>16,000</td>
<td>500</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Full2</td>
<td>500</td>
<td>200</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Full3</td>
<td>200</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

http://gesture.chalearn.org/
Evaluation – competition open until middle september 2015
Performed baseline

<table>
<thead>
<tr>
<th></th>
<th>User</th>
<th>Team Name</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Raducu</td>
<td></td>
<td>0.488526 (1)</td>
</tr>
<tr>
<td>2</td>
<td>mpopescu</td>
<td></td>
<td>0.542466 (2)</td>
</tr>
<tr>
<td>3</td>
<td>palm_seu</td>
<td></td>
<td>0.589424 (3)</td>
</tr>
<tr>
<td>4</td>
<td>all_about_faces</td>
<td></td>
<td>0.603691 (4)</td>
</tr>
<tr>
<td>5</td>
<td>cvl</td>
<td></td>
<td>0.608589 (5)</td>
</tr>
<tr>
<td>6</td>
<td>PabloPG</td>
<td></td>
<td>0.686668 (6)</td>
</tr>
</tbody>
</table>

Already outperformed by 4 participants on 12/7/2015 (in less than one month competition)
Competition still opened until middle september, join and have fun!

https://www.codalab.org/competitions/4711
Schedule details:
http://gesture.chalearn.org/
Context of the Workshops

Broader workshop scope focus on Looking at People:

• Gesture, posture, and sign recognition, analysis and synthesis
• Face recognition, analysis and synthesis
• Body motion analysis and synthesis, and action/interaction recognition and spotting
• Psychological and behavioral analysis
• Multi-modal strategies for gesture recognition and spotting
• Data sets and evaluation protocols
• Computer Vision applications of human pose recovery, gesture recognition and spotting
ChaLearn Looking at People Challenges and Workshops

ICCV 2015 Workshop and Challenge on action recognition, cultural event recognition, and apparent age recognition

<table>
<thead>
<tr>
<th>Total number of images</th>
<th>Number of countries for all the events</th>
<th>Number of images per category</th>
<th>Number of categories</th>
<th>Number of test images</th>
<th>Number of validation images</th>
<th>Number of training images</th>
</tr>
</thead>
<tbody>
<tr>
<td>28705</td>
<td>45</td>
<td>&gt;200</td>
<td>100</td>
<td>8669</td>
<td>5704</td>
<td>14332</td>
</tr>
</tbody>
</table>

**CULTURAL EVENT RECOGNITION**

- First database on cultural events.
- More than 25,000 images representing 90 different categories.
- High intra- and inter-class variability.
- For this type of images, different cues can be exploited like garments, human poses, crowds analysis, objects and background scene.
- The evaluation metric will be the recognition accuracy.
ChaLearn Looking at People Challenges and Workshops

ICCV 2015 Workshop and Challenge on action recognition, cultural event recognition, and **apparent age recognition**

<table>
<thead>
<tr>
<th>Range of labeled ages</th>
<th>Information from the labelers</th>
<th>Contains real age</th>
<th>Contains estimated age by the labelers</th>
<th>Number of labelers</th>
<th>Number of actors</th>
<th>Number of images</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-85</td>
<td>Nationality, age, and gender of the labelers</td>
<td>YES</td>
<td>YES</td>
<td>&gt; 3600</td>
<td>&gt;2000</td>
<td>5000</td>
</tr>
</tbody>
</table>

**AGE ESTIMATION**

• More than 5,000 faces from more than 2000 different people.
• Images with background.
• Non-controlled environments.
• Non-labeled faces neither landmarks, making the estimation problem even harder.
• One of the first datasets in the literature including estimated age labeled by many users to define the ground truth with the objective of estimating the age.
• The evaluation metric will be pondered by the mean and the variance of the labeling by the participants.
• The dataset also provides for each image the real age although not used for recognition (just for analysis purposes). In the same way for all the labelers we have their nationality, age, and gender, which will allow analyzing demographic and other interesting studies among the correlation of labelers.
ChaLearn Looking at people news

ICCV 2015 COMPETITIONS AND WORKSHOP!! STARTING 15th JUNE!


Economic prizes (1500$, 1000$, 500$ for three top positions), travel grants (500$ for each top 3 position), NVIDIA Titan X devices (for all three winners!), winner certificate
Best papers will be invited to a high impact factor SI on LAP (IJCV: special issue deadline on LAP beginning 2016). Best paper workshop award: NVIDIA Titan X. Call for papers opened.
ChaLearn LAP challenges and news: [http://gesture.chalearn.org/](http://gesture.chalearn.org/)

Organization of ChaLearn 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 propsoe ideas related to ChaLearn Looking at People competitions:
[lap@chalearn.org](mailto:lap@chalearn.org)
Thank you!

http://gesture.chalearn.org/
lap@chalearn.org