ChaLearn Looking at People 2015: Apparent Age and Cultural Event Recognition datasets and results

http://gesture.chalearn.org/

Sergio Escalera, Computer Vision Center (UAB) & University of Barcelona, Spain, sergio@maia.ub.es

Junior J. Fabian, Computer Vision Center, jfabian@cvc.uab.es

Pablo Pardo, University of Barcelona, ppardoga7@gmail.com

Xavier Baró, Universitat Oberta de Catalunya and CVC, Spain, xbaro@uoc.edu

Jordi Gonzàlez, Universitat Autònoma de Barcelona and CVC, Spain, lopoal@gmail.com

Hugo Jair Escalante, INAOE, Puebla, Mexico, hugo.jair@gmail.com

Dusan Misevic, Center for Research and Interdisciplinarity

Ulrich Steiner, Max-Planck Odense Center

Isabelle Guyon, ChaLearn, USA and Université Paris-Saclay, France *guyon@chalearn.org*

Marc Oliu Simón, University of Barcelona, moliusimon@gmail.com

Context of the Challenge

ChaLearn LAP Challenges, Workshops, and Special Issues

CVPR 2011 - Workshop and Challenge on Gesture Recognition

CVPR 2012 - Workshop and Challenge on Gesture Recognition

ICPR 2012 - Workshop and Challenge on Gesture Recognition

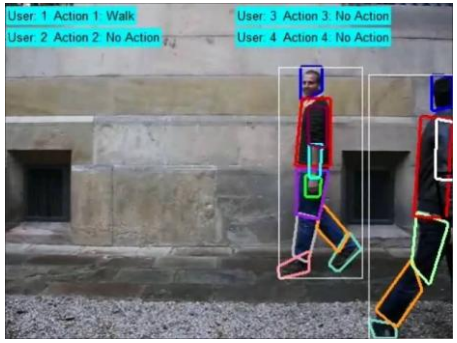
ICMI 2013 - Workshop and Challenge on Gesture Recognition

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

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

ICCV 2015 - ChaLearn Looking at People 2015 – Apparent age and cultural event Recognition

JMLR special issue 2011-2014, TPAMI special issue 2015-2016, IJCV special issue on LAP 2016



http://gesture.chalearn.org/

Competitions Schedule

- **June 15, 2015:** Beginning of the competition, release of development and validation data.
- **August 30, 2015:** Release of the encrypted final evaluation data and validation labels. Participants can start training their methods with the whole data set.
- **September 7, 2015:** Release of the decryption key for the final evaluation data. Participants can start predicting the results on the final evaluation labels. Deadline for code submission.
- **September 12, 2015:** End of the competition. Deadline for submitting the predictions over the final evaluation data. The organisers start the code verification by running it on the final evaluation data.
- **September 15, 2015:** Deadline for submitting the fact sheets.
- **September 19, 2015:** Release of the verification results to the participants for review.

The CodaLab Team

Percy Liang is an assistant professor of Computer Science at Stanford University. His primary research areas are machine learning and natural language processing. He leads the development of CodaLab in close collaboration with Microsoft Research and the rest of the community.

Isabelle Guyon is an independent consultant, specializing in statistical data analysis, pattern recognition and machine learning. Isabelle served as an advisor in the development of the CodaLab competition platform and pioneered the implementation of several challenges on Codalab.

Sergio Escalera leads the Human Analysis group (HuPBA) at the University of Barcelona and the Computer Vision Center. He is one of the directors of ChaLearn. He pioneered the implementation of several Computer Vision challenges on Codalab within the field of Looking at People.

Evelyne Viegas is a Director at Microsoft Research responsible for the outreach artificial intelligence program. She leads the CodaLab project working in collaboration with Isabelle Guyon, Percy Liang and the machine learning and artificial intelligence communities.

CodaLab

Accelerating reproducible computational research.

https://competitions.codalab.org/

https://github.com/codalab/

http://gesture.chalearn.org/

ChaLearn Looking at People 2015 - Track 2: Cultural Event Recognition

Organized by jfabianarteaga

For this track more than 28000 images are labeled with the objective of performing automatic cultural event recognition from 100 ...

Jun 15, 2015-Sep 12, 2015

74 participants

USD **$6,000** reward

ChaLearn Looking at People 2015 - Track 1: Age Estimation

Organized by xbaro

For this track nearly 5000 images were collectively labeled with the objective of performing automatic apparent age estimation from RGB ...

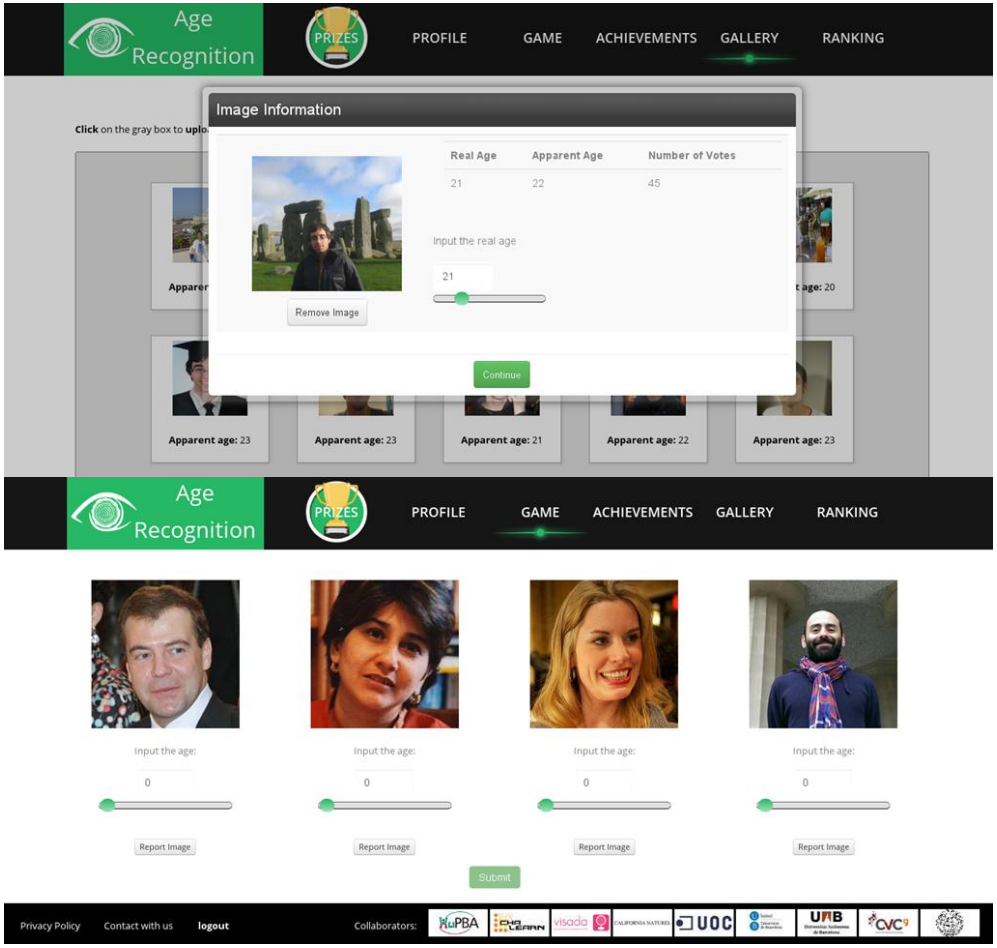
Jun 15, 2015-Sep 12, 2015

135 participants

USD **$6,000** reward

Age Estimation Track

Age estimation is a difficult task which requires the automatic detection and interpretation of facial features. We have designed an application using the Facebook API for the collaborative harversting and labeling by the community in a gamified fashion (http://sunai.uoc.edu:8005/).

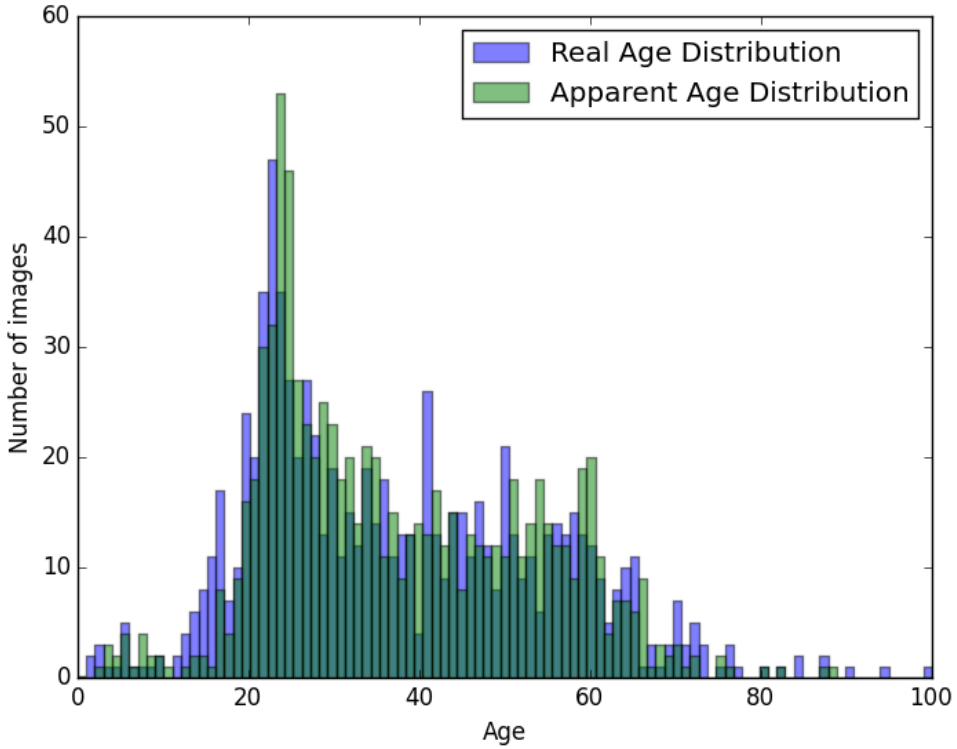


Dataset and evaluation metric

The dataset consists of 4691 face images,

- **Train Set**: 2476 images
- **Validation Set**: 1136 images
- **Test Set**: 1079 images

The images were labelled by hundreds of people with the apparent age, so for each image we have the mean and standard deviation of its apparent age.



The evaluation metric was the following:

$$\epsilon = 1 - e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where x is the prediction and μ and σ are the mean and std. of the apparent age.

Rank	Team	Development	Test
1	**CVL ETHZ**	**0.295116**	**0.264975**
2	**ICT-VIPL**	**0.292297**	**0.270685**
3	AgeSeer	0.327321	0.287266
3	**WVU CVL**	**0.316289**	**0.294835**
4	SEU-NJU	0.380615	0.305763
5	UMD	-	0.373352
6	Enjuto	0.370656	0.37439
7	Sungbin Choi	-	0.420554
8	Lab219A	0.477079	0.499181
9	Bogazici	0.483337	0.524055
10	Notts CVLab	-	0.594248

More than 100 registered teams

10 teams submitted results and provided code at test stage

9 of 10 methods based on CNN

Methods Used by the Winners

- **First place (CVL ETHZ)**. First **face detection and alignment**. Pretrain **20 VGG-16** deep networks using Imagenet, **fine-tune** on faces from IMDB and Wikipedia. Fine-tune again with the challenge dataset. The network was trained for **classification with 101 classes** (0-100 years). The **final prediction is the average** over the 20 nets of the expected value of the last layer.

- **Second place (ICT-VIPL)**. **Pretrain two** 22 layer large-scale deep network based on **GoogleNet**, one for multiclass face classification and the other for age estimation using a face database. **Fine-tune** on large real age dataset. Fine-tune on the challenge apparent age data. And finally, **Ensemble Learning**, where the final age estimation output is the **fusion of 8 deep neural networks**.

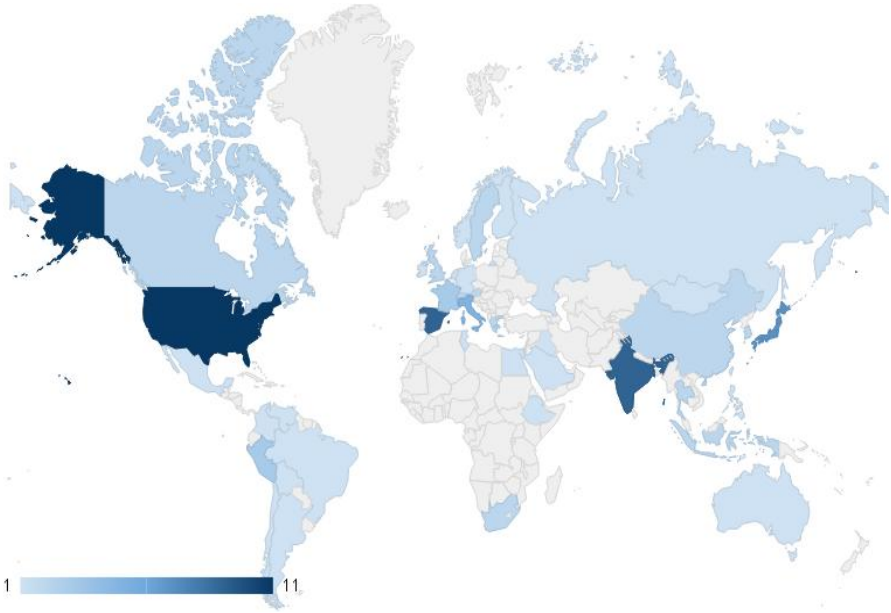
- **Third place (WVU CVL)**. **GoogleNet** trained with 240k extra images. **Fine-tune** the net using the Challenge data augmented and the features are extracted. Then the images are classified **into ten different groups and finally Random Forests and SVR are used to estimate the age**. Score level **fusion** is applied to get the final prediction.

Cultural Event Recognition

The Cultural Event Recognition challenge aims to investigate the performance of recognition methods based on several cues like garments, human poses, objects, background, etc. To this end, our dataset contains significant variability in terms of clothes, actions, illumination, localization and context.
This is the second round for this track. We have significantly incremented the number of images and classes

Dataset and evaluation metric

The dataset is composed of 28, 705 images corresponding to different world-wide cultural events.
Classes: 99 + non-class
Train Set: (50%) → 14, 332 images
Validation Set: (20%) → 5, 704 images
Test Set: (30%) → 8, 669 images



Evaluation metric

- Participants were asked to submit for each image their confidence for each one of the categories (events).

- Participants submissions were evaluated using the average precision (AP), inspired in the metric used for PASCAL challenges.

- Finally, to compare the results of the teams we use the mean average precision (mAP) to obtain the final score.

Rank	Team	Development	Test
1	**VIPL-ICT-CAS**	**0.783**	**0.854**
2	**FV**	**0.770**	**0.851**
3	**MMLAB**	**0.717**	**0.847**
4	NU&C	0.387	0.824
5	CVL_ETHZ	0.662	0.798
6	SSTK	0.740	0.770
7	MIPAL_SNU	0.801	0.763
8	ESB	0.729	0.758
9	Sungbin Choi	-	0.624
10	UPC-STP	0.503	0.588

More than 50 registered teams

10 teams submitted results and provided code at test stage

All methods are based on CNN

Methods Used by the Winners

•**First place (VIPL-ICT-CAS)**. Combination of visual features extracted from deep convolutional neural networks (**VGGNet and GoogLeNet**) and adapt them to the task by performing event-specific **fine-tuning on both global and local images**. In recognition stage, it is employed two kinds of linear classifiers, **Logistic Regression (LR) and Linear Discriminant Analysis (LDA)** on image features from the different deep models and decision **scores are fused**.

•**Second place (FV)**. Images of all different scales are feed into a **pre-trained CNN model**. In each scale, the corresponding rescaled image is encoded. Then, vectors are merged into a single vector by average pooling. In particular VGG16, VGG19 and Place-CNN models were employed. Finally, one test image is then represented by five equally instances, and at the test stage, the prediction scores of these five instances, obtained by **a logistic regression and softmax prediction, are averaged to get the final score.**

•**Third place (MMLAB)**. The approach uses a deep architecture to perform event recognition by extracting visual information from the perspectives of object and scene. **Based on OS-CNN**, it is presented an effective image representation, by extracting the activations of fully connected layers and convolutional layers. Average pooling is applied to aggregate the activations of fully connected layers. Then, **Fisher vector encodes those convolutional layers, and an SVM is used for classification**.

Thank you!

¿Questions and/or suggestions related to the competitions and future events?

Main sponsor

Microsoft

Sponsors

amazon Computer Vision

Google

facebook

Disney Research

NVIDIA

IJCV special issue on LAP 2016
Deadline: 15th February

ChaLearn LAP workshop CVPR 2016 and 2nd round of apparent age competition

Winners

Apparent age

Position	Team	Team members
1	CVL_ETHZ	Rasmus Rothe, Radu Timofte, Luc Van Gool
2	ICT-VIPL	Xin Liu, Shaoxin Li, Meina Kan, Jie Zhang, Shuzhe Wu, Hu Han, Shiguang shan
3	WVU_CVL	Yu Zhu, Yan Li, Guowang Mu, Guodong Guo

Cultural event

Position	Team	Team members
1	VIPL-ICT-CAS	Mengyi Liu, Xin Liu, Yan Li, Shiguang Shan, Xilin Chen
2	FV	Xiu-Shen Wei, Bin-Bin Gao, Jianxin Wu
3	MMLAB	Limin Wang, Zhe Wang, Sheng Guo, Yu Qiao

Best workshop paper by NVIDIA

DEX: Deep EXpectation of apparent age from a single image,
Rasmus Rothe, Radu Timofte, Luc Van Gool

You are invited to submit an improved version of your work to IJCV LAP SI deadline 15th February