

ChaLearn Multi-modal Gesture Recognition Challenge

Evaluation process

The focus of the challenge is on “multiple instance, user independent learning” of gestures, which means learning to recognize gestures from several instances for each category performed by different users, drawn from a gesture vocabulary. A gesture vocabulary is a set of unique gestures, generally related to a particular task. In this challenge we will focus on the recognition of a vocabulary of 20 Italian cultural/anthropological signs.

During the challenge, organizers provide the participants with three different datasets:

- **Training:** This dataset was published on May 20th, and contains 393 labeled sequences with a total of 7754 gestures.
- **Validation:** This dataset was published on May 20th, and contains 287 sequences with a total of 3363 gestures. Initially it was published without the labels, and the labels were made available on August 1st.
- **Final evaluation:** This dataset was published on August 20th, after the end of the competition, and it contains a total of 276 sequences with 2742 gestures. The labels for this dataset are not provided to the participants.

Each sequence contains a single person performing a sequence with 8 to 15 gestures and their related oral expression. RGB, depth and audio information is provided for each sequence, together with additional information such as the skeleton and user mask. In addition to the 20 considered target gestures, some of the sequences contained some noisy gesture, that is, a gesture that not belong to the set of target gestures. Most of the individuals were native Italian people from different regions, but some non-native Italian speakers were added to the database in order to provide more variability.

For each sequence of the database, participants were asked to provide an ordered list of labels P corresponding to the recognized gestures. Although the labeled data contains the gestures at frame level, we did not consider frame level predictions on this competition. User predictions are compared with the corresponding list of ground truth labels GT . For each sequence, the so-called Levenshtein distance $L(P,GT)$ is computed, that is the minimum number of edit operations (substitution, insertion, or deletion) that one has to perform to go from P to GT (or vice versa). The Levenshtein distance is also known as "edit distance".

All the competition was held at Kaggle platform (www.kaggle.org), where participants were able to submit their predictions in order to be evaluated. The competition had two different phases.

The first phase, from May 20th to August 1st, participants had the labeled training data in order to learn their models, and predict the labels for the unlabeled validation data. The validation labels are updated to the Kaggle system, therefore participants predictions can be evaluated and the participants receive their score to each submitted prediction.

After this initial phase, on August 1st, the validation data labels were made available to all the participants, which had until the August 15th to learn their final models using the whole development data (training and validation). The final models were uploaded to the Kaggle platform.

Once final models were submitted, the organizers published the final evaluation data, and participants were asked to run their codes on this data and submit by the 24th August their predictions and the fact sheets explaining their models. The final predictions were managed with the Kaggle platform, where the final evaluation labels were prepared to be used in a blinded way, therefore, participants didn't know their scores and couldn't use them to tune their models.

The final evaluation started on 25th August, using the scores of the predictions submitted by participants over the final evaluation data. A board of judges composed of 4 doctors from different institutions was created in order to validate the final rankings:

- **Dr. Hugo Jair Escalante**
Computational Sciences Department of the National Institute of Astrophysics, Optics and Electronics (INAOE)
- **Dr. Jordi González**
Computer Sciences Department of the Autonomous University of Barcelona (UAB)
- **Dr. Sergio Escalera**
Applied Mathematics and Analysis Department of the University of Barcelona (UB)
- **Dr. Xavier Baró**
IT, Multimedia and Telecommunication department of the Open University of Catalunya (UOC)

From the total of 54 teams that participate on the challenge, only 19 submit the predictions for the final evaluation process, the rest were not considered. The judges checked all the predictions for those 19 teams following those steps:

1. **Analyze predictions:** All team's predictions were analyzed in order to validate the content and format. Two teams were discarded on this step: One team predicted a fix number of gesture 10 for all the samples. The second one, submitted the ground truth for validation dataset.
2. **Correct predictions:** For the final evaluation dataset, we detect that three samples had problems with the videos or audios, and participants couldn't predict their gestures. The team's predictions were modified in order to remove the effects of those errors, and obtaining the true score for all the teams. The ranking was not affected by this change.
3. **Model validation:** To validate the winners, the judges run the models of the 5 first teams and calculate the final scores. Some problems appear during this validation with one of the teams, that demands to contact with the member of the teams and ask help to run their code. Finally the code was verified.

Due to the problem with one of the models validation, the publication of the final scores was delayed some days. There was no changes on the final ranking due to the validation process. Regarding the prizes, some of the teams inform the organizers that wasn't able to go to the ICMI Workshop, and were informed that it supposed to lose the travel grant of the prize.

Prizes

After the conclusion of the ChaLearn Multimodal Gesture Recognition Challenge, the winners were:

Position	Team Name	Score	Members
1	IVA MM	0.12756	Jiaxiang Wu Jian Cheng Hangqing Lu cyzhao
2	WWEIGHT	0.15387	Ryan Keisler
3	ET	0.17105	Li Wei LiMin Wang




1st October 2013
Dr. Hugo Jair Escalante



1st October 2013
Dr. Jordi González



1st October 2013
Dr. Sergio Escalera



1st October 2013
Dr. Xavier Baró