Aggregate Channel Features for newborn face detection in Neonatal Intensive Care Units

B. OLMI [1], C. MANFREDI [1], L. FRASSINETI [1,2], C. DANI [3], S. LORI [4], G. BERTINI [3], S. GABBANINI [4], A. LANATA' [1]

[1] Department of Information Engineering, Università degli Studi di Firenze, Firenze, Italy
[2] Department of Medical Biotechnologies, Università di Siena, Siena, Italy
[3] Department of Neurosciences, Psychology, Drug Research and Child Health, Division of Neonatology, Maternal and Child Department, AOU Careggi, Firenze, Italy
[4] Neurophysiology Unit, Neuro-Musculo-Skeletal Department, AOU Careggi Firenze, Italy

INTRODUCTION

An efficient face detector could be very helpful to point out possible neurological dysfunctions in Neonatal Intensive Care Units (NICUs) [1] such as seizure events, which show some facial clinical correlates (e.g. head movements, eye deviations, repetitive opening and closing of the eyelids, mouth movements). Thus, the analysis of video extracted features could provide new useful information for a more accurate and fast seizure assessment in newborns.

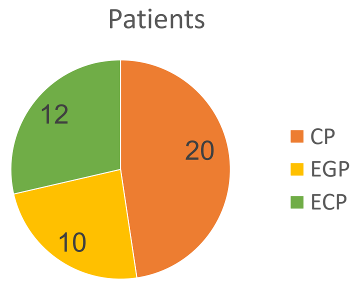
AIM

Over the years several studies introduced semi-automatic approaches. This study proposes a fully automated face detector for newborns in NICUs, based on the Aggregate Channel Feature (ACF) algorithm [2].

DATASET

Video recordings collected at the Neuro-physiopathology and Neonatology Clinical Units of AOU Careggi, Firenze, Italy:

- from 42 full-term newborns (gestational age: 38-41 weeks):
  - 20 control patients (**CP**),
  - 10 subjects showed electrographic-only seizures (**EGP**),
  - 12 subjects exhibited electro-clinical seizures (**ECP**).
- mean time duration of about 4 h (337631 frames).



METHODS

- **Labelling process**: The MATLAB “Video Labeler” tool was applied to create the ground-truth bounding boxes for the training step. The ROI bounding box containing the newborn's face was manually defined in the first frame of each considered video segment. Then, in the following frames, the ROIs were automatically built using the Point Tracker algorithm based on the Kanade-Lucas-Tomasi (KLT) algorithm.

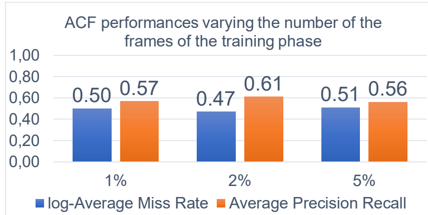
- **Training Design**: The training phase of the ACF Object Detector is regulated by several parameters. The system's performances were evaluated varying:
  - the number of stages for the iterative training process (NumStage) between 2 and 6;
  - the number of samples not containing the newborn face (NegativeSamplesFactor) between 2 and 5;
  - resize dimensions (ObjectTrainingSize) between [50x50] and [100x100];
  - the percentage number of frames, randomly extracted from the dataset, equal to 1%, 2% and 5%.

- **Validation Design: L**eave-One-Subject-Out (LOSO) cross-validation operation: at each iteration, the training set is defined by excluding the data of a patient and considering varying percentage of frames of the remaining patients; the test set is composed of all the video frames from that excluded patient

RESULTS

- The best results were obtained with:
  - NumStages = 4;
  - NegativeSamplesFactor = 2;
  - ObjectTrainingSize = [100x100]
  - Percentage number of frames randomly extracted from the dataset for the training step = 2%.

  The detector gave (mean ± standard error): log-Average Miss Rate (**l-AMR**) = 0.47 ± 0.05 and Average Precision Recall (**APR**) = 0.61 ± 0.05.



- The average performance for the three groups of newborns were also evaluated. The l-AMR and APR results suggest that the developed ACF detector provides better results when applied to the CP and the EGP groups than ECP.

Patients	Log-Average Miss Rate (mean ± standard error)	Average Precision Recall (mean ± standard error)
CP	0.40 ± 0.07	0.70 ± 0.06
EGP	0.31 ± 0.11	0.73 ± 0.11
ECP	0.73 ± 0.07	0.38 ± 0.07

- The statistically significant difference between the three considered groups for the l-AMR and APR values was tested with the Kruskal-Wallis test and the Tukey's HSD correction. The detector's performances were significantly different for newborns with seizures characterized by clear clinical correlations as compared to the control subjects and newborns with electrographic-only seizures.

Comparison	Log-Average Miss Rate p-value	Average Precision Recall p-value
CP vs. EGP	0.79	0.79
CP vs. ECP	0.02	0.031
EGP vs. ECP	0.011	0.017

CONCLUSIONS

We developed a promising automatic ACF-based system for newborns' face detection in NICUs. An efficient face detector for newborns could be very helpful in clinical practice. It could speed up clinical diagnosis and thus early intervention, by facilitating the extraction of quantitative features of facial motion related to pathological conditions such as pain or seizures [3].

REFERENCES

[1] S. Cabon et al., “Video and audio processing in paediatrics: a review”. Physiol Meas. 2019 Feb 26; vol. 40(2), 02TR02.
[2] P. Dollar et al., “Pedestrian Detection: An Evaluation of the State of the Art”, in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 4, pp. 743-761, April 2012.
[3] M. Pediaditis et al., “Model-free vision-based facial motion analysis in epilepsy”. 10th International Workshop on Biomedical Engineering, BioEng 2011, pp. 1-4.

ACKNOWLEDGEMENTS

This work was partially funded by MAECI PGR01276, Project MX18MO14

CONTACT INFORMATION

lorenzo.frassineti@student.unisi.it