Project specification - Theoretical research

MEi:CogSci 1st year semester project

General Project Information

Project Title	Emotion Recognition in the Wild: A Machine Learning Challenge
Author	Oswaldo Macedo
Supervisor	Martin Takac

Summary of Topic/Phenomenon

The use of emotion in robots and software had a growing interest since the nineties, currently with substantial improvement in face detection and faster processing computers it is becoming more efficient to detect emotions. My interest is in knowing whether it is possible to create a software that can precisely detect human emotions, therefore, my objective doing this project is to find the most notable state of the art psychological/computational models of emotions, compare them, and suggest some improvements or the possibility to combine them.

Learning Outcomes[1]

Subject specific

• An interdisciplinary perspective between computer science and psychology will allow me to combine ideas from different fields and merge them together.

Methodological

• Improvement of information searching skills.
• Experience with elaboration of theoretically oriented projects.

Generic/Instrumental

• Improve the ability to create and follow a project plan.

Systemic

• Interdisciplinary work & thinking
• Project-oriented work & organizational skill
• Critical evaluation of approaches & methods
• Quick orientation and navigation in computer science & psychology of emotions
• Change of viewpoint & perspectives

[1] as defined in the MEi:CogSci curriculum

• Phenomenon-oriented thinking
• Problem-solving abilities

Short Project Description

The main objective of the project is to review state of the art models of emotion recognition, analyze their solutions and suggest a new model combining previous solutions. Currently, artificial emotion recognition exceeds 90% precision in laboratory settings, nevertheless, emotion recognition in the wild is as low as 75.2% of precision for state of the art implementations. In this paper I mention convolutional neural networks studies and the constrained local neural fields solution for face detection in the wild.

Project Plan

Project Steps

1. Initial Literature Research (Overview)				**Total Working Hours (WH)/ECTS: 7.5 / 0.25**	
Working-package (WP)	**Start – End**	**WH / ECTS**	**Activities**	**Resources required**	**Mile-stones (M)**
WP 1.1 Literature 1	18.03 – 28.03	7.5/ 0.25	Search for literature related to the topic	Internet – scientific journals	M1.1

2. Formulating Research Question and Theses				**Total Working Hours (WH)/ECTS: 7.5 / 0.25**	
Working-package (WP)	**Start – End**	**WH / ECTS**	**Activities**	**Resources required**	**Mile-stones (M)**
WP 2.1 Research questions	29.03 – 27.04	7.5/ 0.25	Based on searched literature, propose a main objective and new questions to be solved by the research.	Selected papers	M2.1

3. Focussed Literature Research & Analysis (on Sub-Topics/Concepts)				**Total Working Hours (WH)/ECTS: 30 / 1**	
Working-package (WP)	**Start – End**	**WH / ECTS**	**Activities**	**Resources required**	**Mile-stones (M)**
WP 3.1 Literature 2	07.04 – 09.05	30/1	Research and analyze state of the art solutions of EmotiW challenge, convolutional neural networks, constrained local neural fields, and local binary patterns.	Internet – scientific journals	M3.1

4. Synthesis of Findings/Insights				**Total Working Hours (WH)/ECTS: 45 / 1.5**	
Working-package (WP)	**Start – End**	**WH / ECTS**	**Activities**	**Resources required**	**Mile-stones (M)**
WP 4.1 Findings	10.05 – 31.05	30/1	Combine knowledge from all the referenced articles and suggest a new approach.	Referenced articles	M4.1
WP 4.2 Reviews	25.05 - 31.05	15/0.5	Apply suggestions received by peer review authors.	Review document	M4.2

5. Project Documentation				**Total Working Hours (WH)/ECTS: 60 / 2**	
Working-package (WP)	**Start – End**	**WH / ECTS**	**Activities**	**Resources required**	**Mile-stones (M)**
WP 5.1 Project specification	14.03 – 08.06	15/0.5	Do Project specification	Literature, abstract	M5.1
WP 5.2 Abstract	30.04 - 31.05	30/1	Write abstract	Researched literature, peer review	M5.2
WP 5.3 Poster	15.04 – 11.06	15/0.5	Create poster	Abstract, researched literature, peer review	M5.3

Project Milestones

Milestone	**Result/"Product" and/or Deliverables**
M1.1	Research topic chosen.
M2.1	Main objective of the research and questions.
M3.1	Knowledge about new AI emotion recognition methodologies and implementations.
M4.1	New approach solution to the current research.
M4.2	Improved peer reviewed abstract.
M5.1	Project specification.
M5.2	Abstract
M5.3	Poster

Short Project Report [Conference Abstract]

Currently, artificial emotion recognition exceeds 90% precision in laboratory settings, nevertheless, emotion recognition in the wild is as low as 75.2% of precision for state of the art implementations. Emotion recognition in the wild is understood as the ability to recognize emotions from facial expressions in unconstrained environments where data can be noisy, illumination conditions and head pose may differ, and occlusion can be present [1].

The most popular methods for emotion recognition involve convolutional neural networks (CNN). One promising CNN-based solution is the multi-column deep neural network (MCDNN). It is a committee architecture, which means that there is a specific amount of independent CNN's with varied network architecture, input normalization and random weight initialization that processes the inputs [2]. However, MCDNN used several CNN's that by now are surpassed by more recent CNN architectures. One of the new architectures are residual networks that reduce the training error by creating a residual function from the initial inputs and outputs of the network. The other are "Inception" networks based on fire together wire together Hebbian theory and multi-scale processing.

As for face detection, constrained local neural field (CLNF) is an improved effective facial detection method for images with two main parts, the Local Neural Field patch expert which captures more complex information and exploits spatial relationships between pixels, and the Non-Uniform Regularized Mean-Shift which takes the patch expert reliabilities into account. The reliabilities are represented by a parameter that extracts vertex features and while there is less variance inside the patch more reliability is achieved [3].

Based on my current research, I propose to update MCDNN with residual networks and "Inception" networks because committee architectures were shown to be more precise in emotion recognition than individual networks and combine the outcome with CLNF which is specialized in face detection. CLNF allows preprocessing of 'in the wild' data that can further be processed by the updated MCDNN having as an outcome or faster processing times or major accuracy in emotion recognition than when using previous face detection methods.

References

[1] Dhall, A., Goecke, R., Gedeon, T., & Sebe, N. (2016). Emotion recognition in the wild. *Journal on Multimodal User Interfaces*, 10(2), 95–97. https://doi.org/10.1007/s12193-016-0213-z

[2] Kim, B.-K., Roh, J., Dong, S.-Y., & Lee, S.-Y. (2016). Hierarchical committee of deep convolutional neural networks for robust facial expression recognition. *Journal on Multimodal User Interfaces*, 10(2), 173–189. https://doi.org/10.1007/s12193-015-0209-0

[3] Baltrušaitis, T., Robinson, P., & Morency, L. P. (2013). Constrained local neural fields for robust facial landmark detection in the wild. *Proceedings of the IEEE International Conference on Computer Vision*, 354–361. https://doi.org/10.1109/ICCVW.2013.54