IEEE International Conference on Image Processing
25-28 October 2020, United Arab Emirates

Kernelized dense layers for facial expression recognition

Presented by : M.Amine Mahmoudi

M.Amine Mahmoudi[1,3], Aladine Chetouani[2], Fatma Boufera[1], Hedi Tabia[3]

[1]Mustapha Stambouli University of Mascara, Algeria
[2]PRISME laboratory, Université d'Orléans, Orléans, France
[3]IBISC laboratory, Université Paris-Saclay, Paris, France

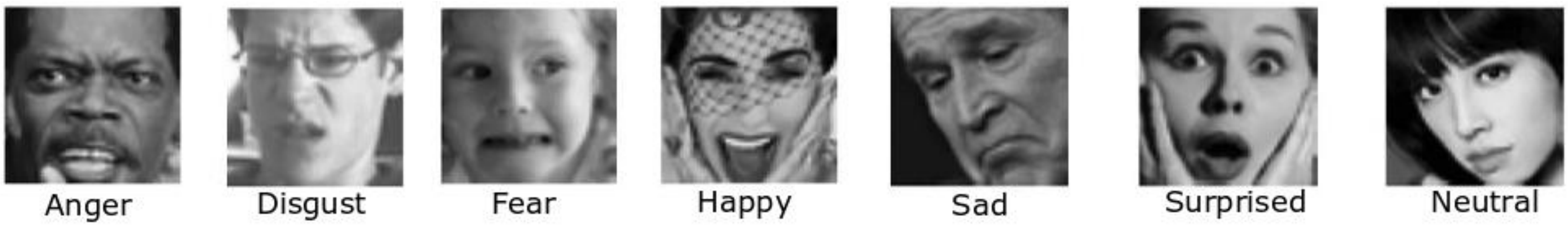
(Machine Learning for Recognition in Images and Videos II)

Outline

- Facial expression recognition
- Fine-grained recognition
- Improving convolutional neural networks
- Kernelized dense layers
- Experiments and results
- Conclusion and perspectives

Facial expression recognition

- **Objective**
  - Classifying human emotions given facial images as one of seven basic emotions: happy, sad, fear, disgust, anger, surprise and neutral.



- **Challenge**
  - Small inter-class differences;
  - Large intra-class differences;

Fine-grained recognition

- Going beyond classical image classification.
- It consist of discriminating categories with only small subtle visual differences.
- Many fine-grained datasets emerged for this purpose (e.g. CUB Bird, Stanford Car…etc)

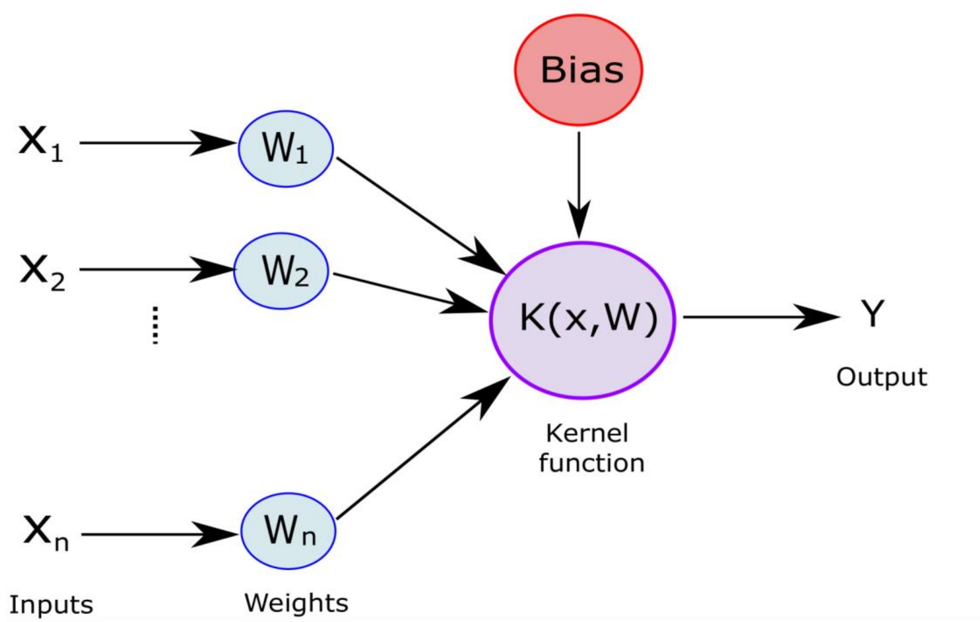
Improving convolutional neural networks (1/2)

- Many researchers tried to improve different levels of CNNs by:
- Replacing convolution
  (e.g. Kervolutional neural networks, Wang et al, 2019);
- Enhancing pooling
  (e.g. Universal Pooling, Hyun et al, 2019);
- Dimensionality expansion
  (e.g. Bilinear pooling, Kernel pooling)

Improving convolutional neural networks (2/2)

- The majority of these works uses kernel functions.
- Their focus is more on improving the process of fully connected layers.
- We build upon that and propose to use kernel functions on fully connected layers.

Kernelized dense layers



- It is similar to a classical neuron.
- It applies higher degree kernel function.

Datasets (1/3)

- FER2013

- ICML 2013 Challenges in Representation Learning.
- 28.709 images for training.
- 3.589 images for validation.
- 3.589 images for test.

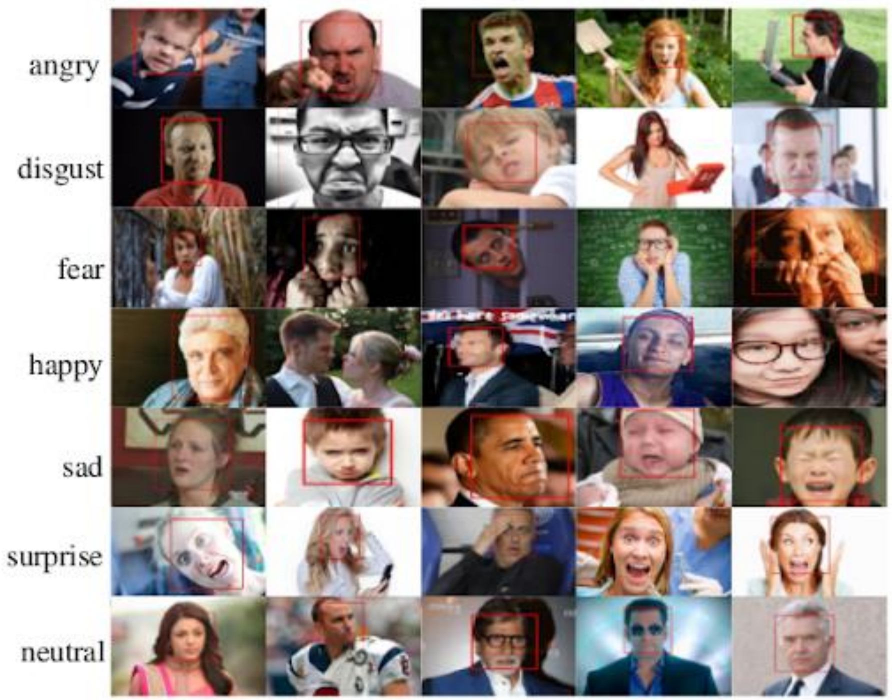
Datasets (2/3)

- RAF-DB

- Real-world Affective Face DataBase.
- It contains 29.672 images.

Datasets (3/3)

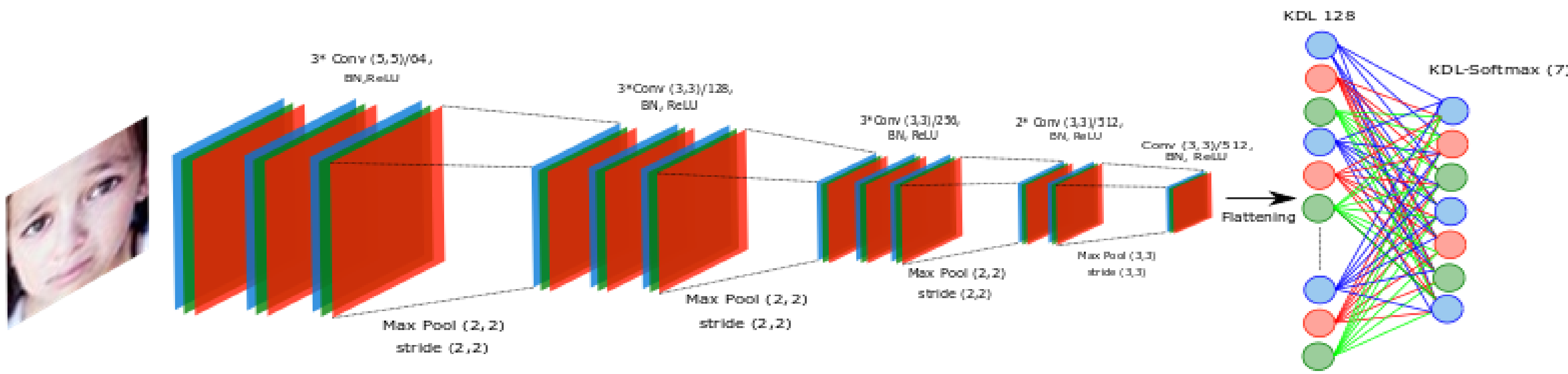
- ExpW



  - EXPression in-the-Wild dataset.
  - It contains 91.793 images.

Experiments and results (1/4)

- We built a CNN with two Kernelized dense layers.



- We used three kernel functions for our experiments:
  - A linear kernel = fully connected layers.
  - Polynomial kernel of second and third degree.

- Many other kernels can be used (e.g. L1 norm, L2 norm, Gaussian RBF , Laplacian, Abel...etc.)

Experiments and results (2/4)

- Second-order polynomial kernel increases for about 0.7% for FER2013, 0.25% for ExpW and 0.6% for RAF-DB.
- Third-order polynomial kernel increases further the overall accuracy for about 1.15% for FER2013, 0.75% for ExpW and 1% for RAF-DB.

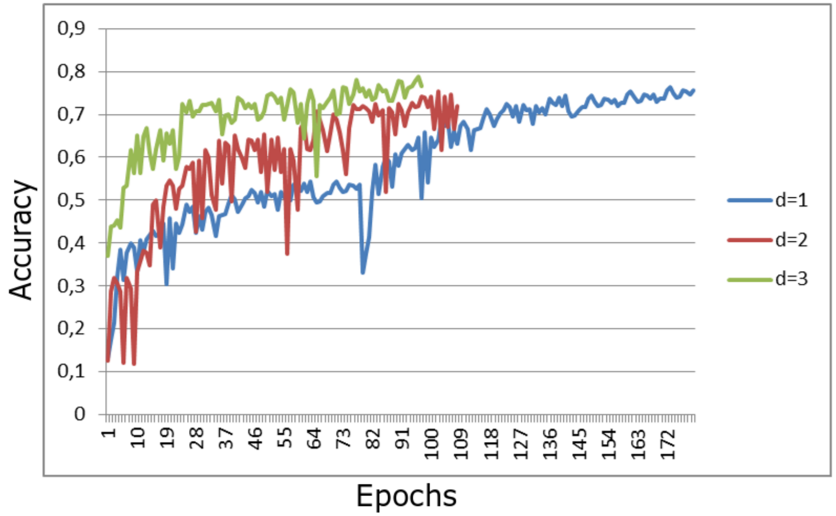
Models	**Dataset**		
	FER2013	***ExpW***	***RAF-DB***
Base-Model-FC (Base-Model-KDL, n=1)	70.13%	75.91%	87.05%
Base-Model-KDL (n=2)	70.85%	76.13%	87.64%
Base-Model-KDL (n=3)	**71.28%**	**76.64%**	**88.02%**

Table 1: Accuracy rates of the proposed approach with different kernel functions.

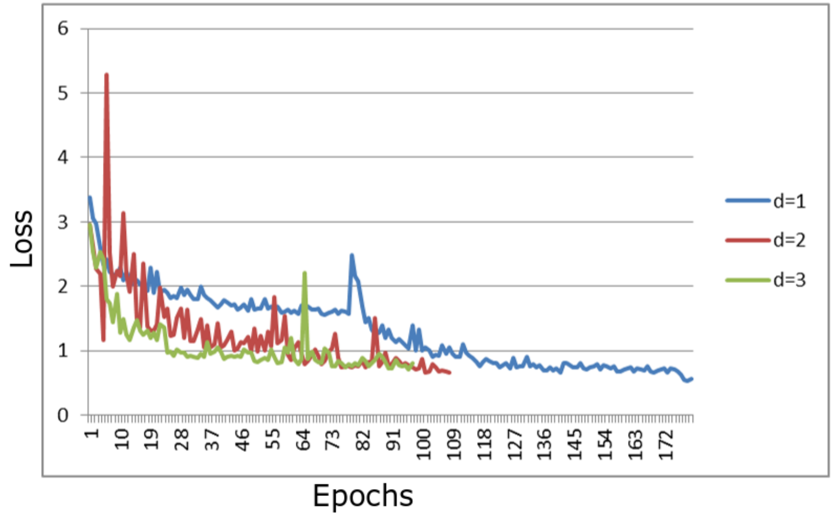
Experiments and results (3/4)

- Due to the use of early stopping, the learning process is interrupted as soon as the model reaches its max capacity.

- The higher degree is the kernel function, the fast it converge.



(a) Validation accuracy



(b) Validation loss

Validation accuracy and validation loss on ExpW with the three kernel configurations.

Experiments and results (4/4)

Models	Dataset		
	FER2013	***ExpW***	***RAF-DB***
Base-Model-FC (Base-Model-KDL, n=1)	70.13%	75.91%	87.05%
Base-Model-KDL (n=2)	70.85%	76.13%	87.64%
Base-Model-KDL (n=3)	71.28%	**76.64%**	**88.02%**
Tang et al. [12]	71.16%	–	–
Guo et al. [13]	71.33%	–	–
Kim et al. [1]	**73.73%**	–	–
Bishay et al. [14]	–	**73.1%**	–
Lian et al. [15]	–	71.9 %	–
Acharya et al. [7]	–	–	**87%**
S Li et al. [16]	–	–	74.2%
Z.Liu et al. [17]	–	–	73.19%

Table 2: Accuracy rate of the proposed approach and state-of-the-art approach

Conclusion and perspectives

- Using higher order kernel function for FC layer enhances the discriminative power of CNN.
- It also improves the convergence speed of the network.
- The ability of capturing high order information that are crucial for fine-grained classification tasks such as the FER.
- As future work, other kernel functions will be considered, compared and combined

mohamed.mahmoudi@univ-mascara.dz