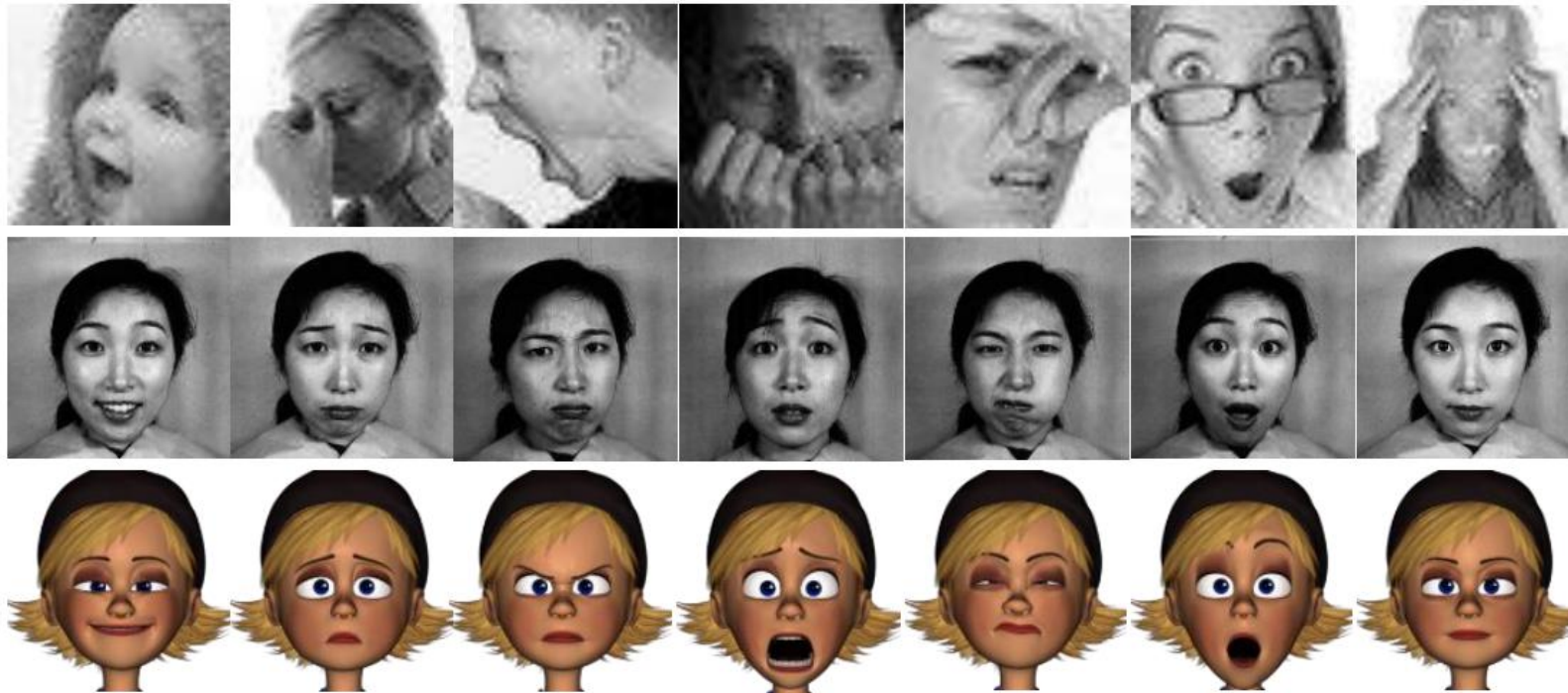# Facial Expression Recognition Using Attentional Convolutional Network

Ran Huo

Yin Jiang

Problem: Facial Expression Recoginition[1]

Problem: Facial Expression Recognition

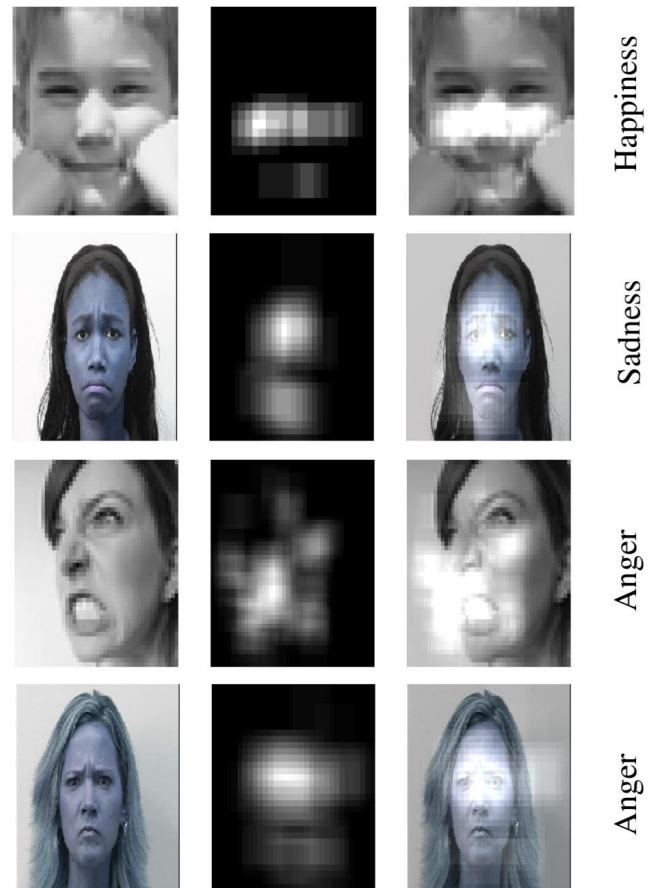7 Categries [2]

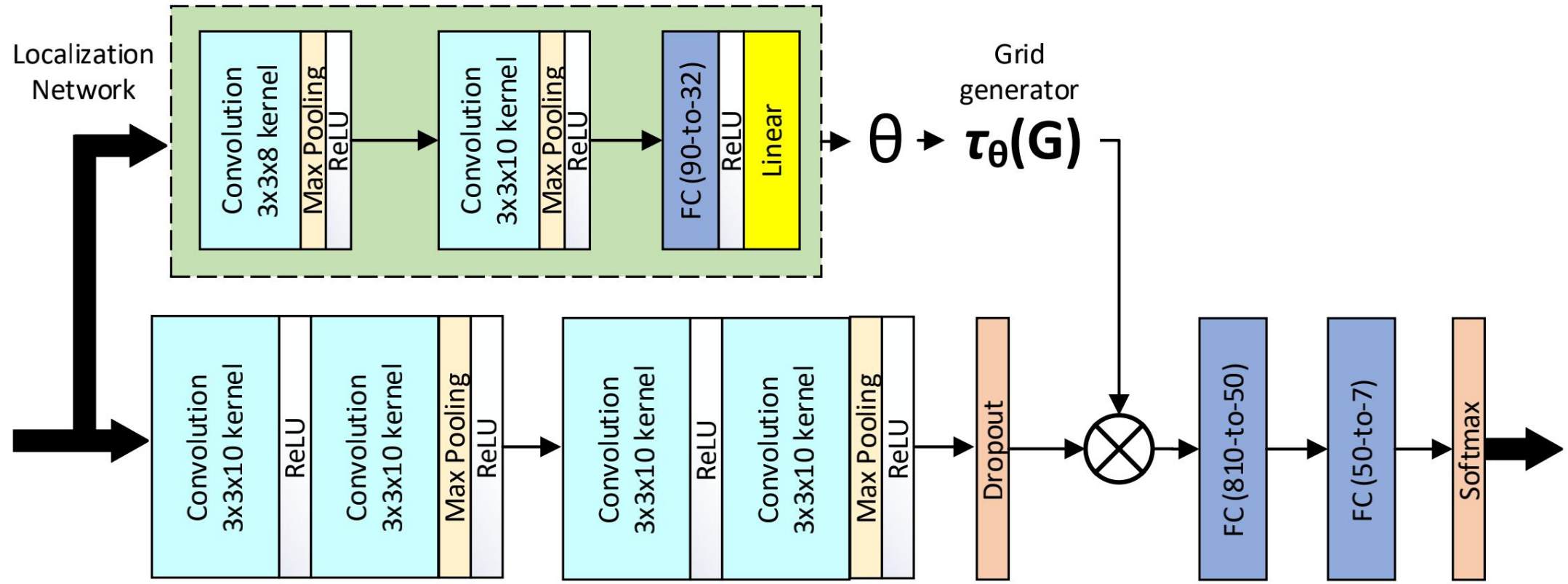- 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral.

Traditional models

- Hand-crafted feature extraction (HOG, SIFT)
- Classifier (SVM, random forest)

Solution: Attentional Convolutional Network [1]

Solution: Attentional Convolutional Network [1]

Solution: Attentional Convolutional Network

```
Model archticture:  Deep_Emotion(
  (conv1): Conv2d(1, 10, kernel_size=(3, 3), stride=(1, 1))
  (conv2): Conv2d(10, 10, kernel_size=(3, 3), stride=(1, 1))
  (pool2): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (conv3): Conv2d(10, 10, kernel_size=(3, 3), stride=(1, 1))
  (conv4): Conv2d(10, 10, kernel_size=(3, 3), stride=(1, 1))
  (pool4): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (norm): BatchNorm2d(10, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (fc1): Linear(in_features=810, out_features=50, bias=True)
  (fc2): Linear(in_features=50, out_features=7, bias=True)
  (localization): Sequential(
    (0): Conv2d(1, 8, kernel_size=(7, 7), stride=(1, 1))
    (1): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
    (2): ReLU(inplace=True)
    (3): Conv2d(8, 10, kernel_size=(5, 5), stride=(1, 1))
    (4): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
    (5): ReLU(inplace=True)
  )
  (fc_loc): Sequential(
    (0): Linear(in_features=640, out_features=32, bias=True)
    (1): ReLU(inplace=True)
    (2): Linear(in_features=32, out_features=6, bias=True)
  )
)
```

Data Set

- FER: facial expression recognition [4]
- 28,709 examples
- 48x48 pixel grayscale images of faces
- Facial expression categories: 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral.

Result

- result from paper [1]

TABLE I: Classification Accuracies on FER 2013 dataset

Method	Accuracy Rate
Bag of Words [52]	67.4%
VGG+SVM [53]	66.31%
GoogleNet [54]	65.2%
Mollahosseini et al [19]	66.4%
The proposed algorithm	70.02%

Result

- result from paper [1]

TABLE II: Classification Accuracy on FERG dataset

Method	Accuracy Rate
DeepExpr [2]	89.02%
Ensemble Multi-feature [49]	97%
Adversarial NN [48]	98.2%
The proposed algorithm	99.3%

Result

- result from our implementation
- 62% accuracy

```
Training Loss: 0.00808909    Validation Loss 0.01246600    Training Acuuarcy 61.385%    Validation Acuuarcy 44.920%
======================================Training Finished==============================================
```

Teamwork

- Ran Huo
- Model implementation 60%
- Project proposal
- Yin Jiang
- Model implementation 40%
- Final report and presentation

References

- [1] Minaee, S & Amirali, A. 2019. Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network.

- [2] Friesen, E., and P. Ekman. "Facial action coding system: a technique for the measurement of facial movement." Palo Alto, 1978.

- [3] Jaderberg, Max, Karen Simonyan, and Andrew Zisserman. "Spatial transformer networks." Advances in neural information processing systems, 2015.

- [4] Challenges in Representation Learning: Facial Expression Recognition Challenge https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data