Generative Cooperative Net for Image Generation and Data Augmentation

University of Tehran
School of Electrical and Computer Engineering

Deep Learning in Computer Vision and Speech Processing

Dr. R. Hosseini

Erfan Mirzaei, Mohammad Mahdi Mehmanchi, Sara Jahedazad

Summer 2021

Content:

Other Generative Models

- Autoregressive
- Generative Adversarial Networks
- Variational Auto-Encoder
- Flow-based Generative Model

Conventional Augmentation Methods

- Random Cropping
- Scale Variation
- Affine Transformation ,....

Limitations and Issues with Other Works and How This Paper Solves the Problems

- Supervised Methods → Generalization
- GAN →Limited Controllability Over the Desired Features
- Conditional GAN → Large Dataset

Main Contributions:

1. Can employ smaller dataset due to absence of need for learning the distribution of visual concept.

2. Deploying a classifier facilitates knowledge transformation, to solve supervised generative model problems.

Datasets:

- The Karolinska Directed Emotional Faces (KDEF)
- QMNIST

KDEF Preprocessing

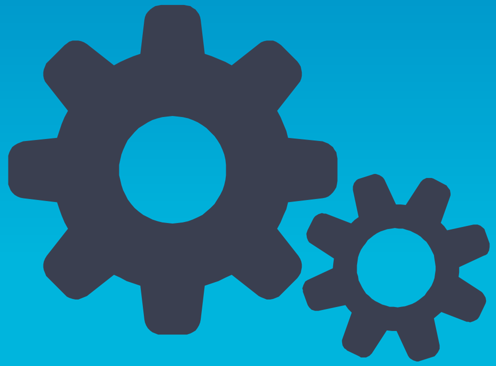
KDEF Dataset Consists of 4900, 562*762 RGB Pictures of Human Expressions.

Preprocessing Below was Applied:

- Face Detection and Resizing
- Mirroring
- Rotation (0, 90, 180, and 270 degrees)
- Train & Test Splitting

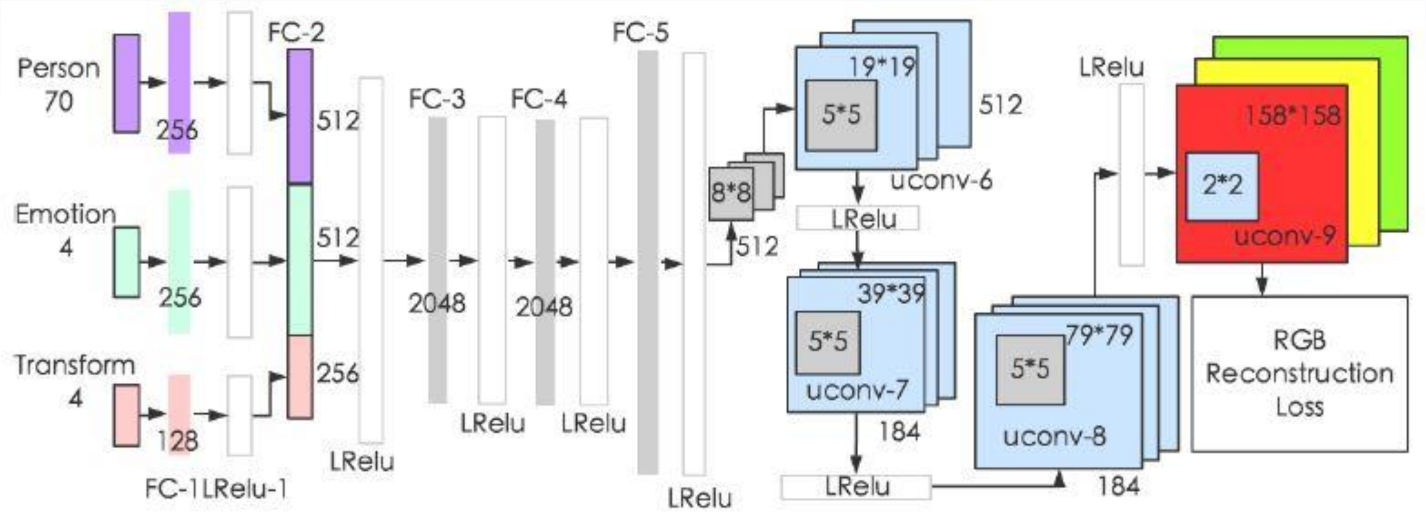
GCN Architecture:

- Generator
- Classifier



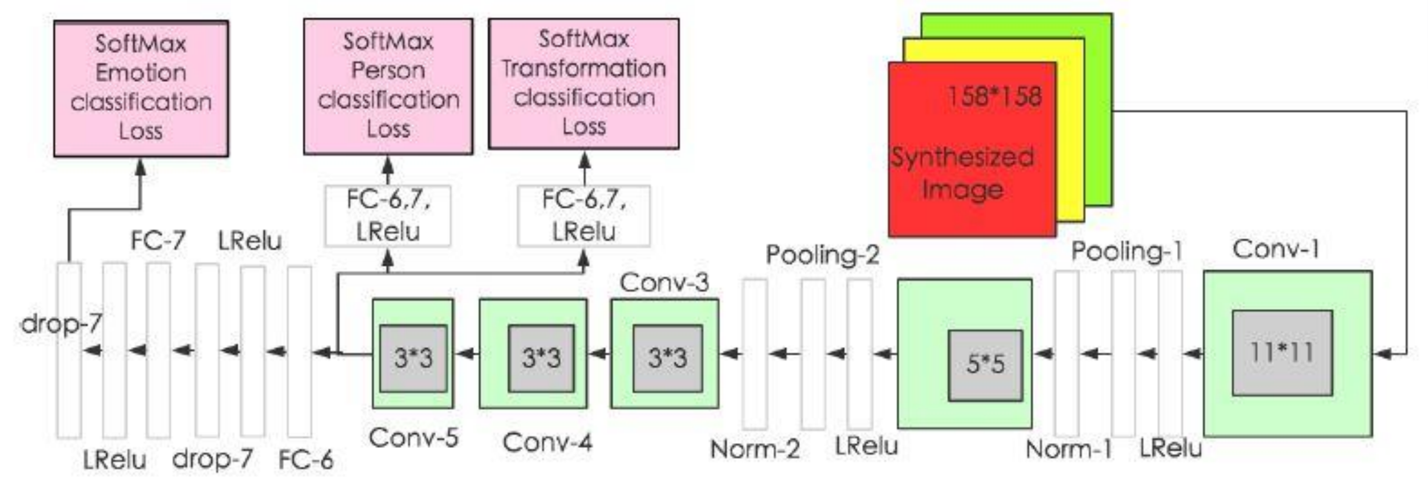
Generator

Inspired by Network Introduced by Dosovitskiy et al.



Classifier

Inspired by AlexNet.



Loss Function

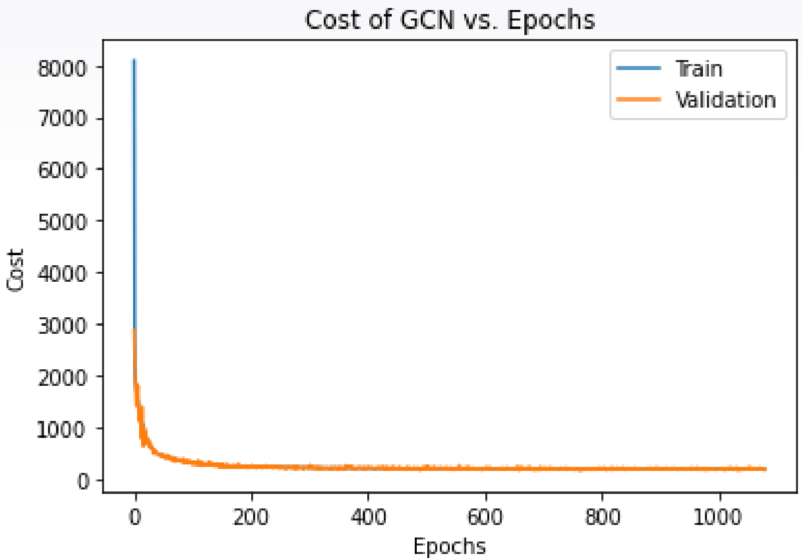
$$\frac{1}{N}\{\sum_{i=1}^{N}\left\|Pixel_r^i - Pixel_s(f_{i_1},\dots,f_{iM})^i\right\|^2 - [\sum_{i=1}^{N}\sum_{f=1}^{M}\sum_{j=1}^{K_f} Weight_f\{y_f^i = j\}\log\frac{e^{\theta_{f_j}^T x_f^i}}{\sum_{l=1}^{K_f} e^{\theta_{f_l}^T x_f^i}}]\}$$

Hyper Parameters:

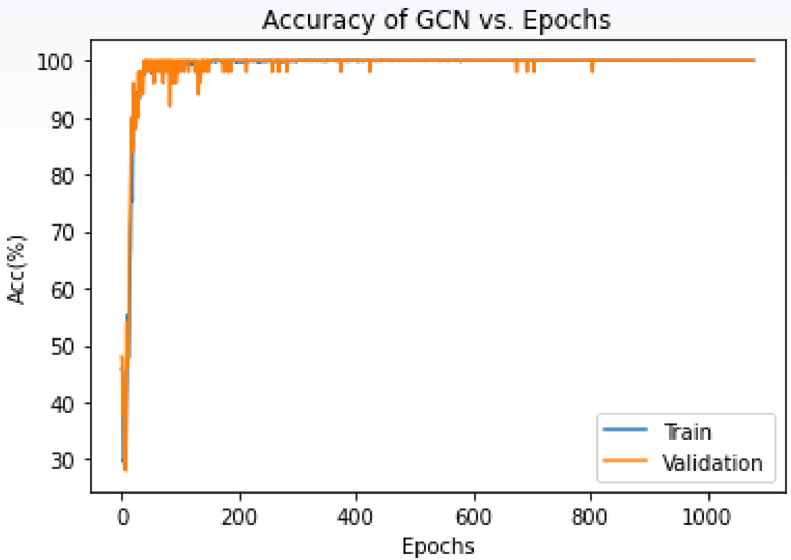
- **Optimizer: Adam**
- **Learning Rate: 0.0002**
- **(β_1, β_2) : (0.9,0.99)**
- **ϵ: 10^{-8}**
- **Batch Size: 64**
- **Epochs: 1100**

Image Generation Results

KDEF

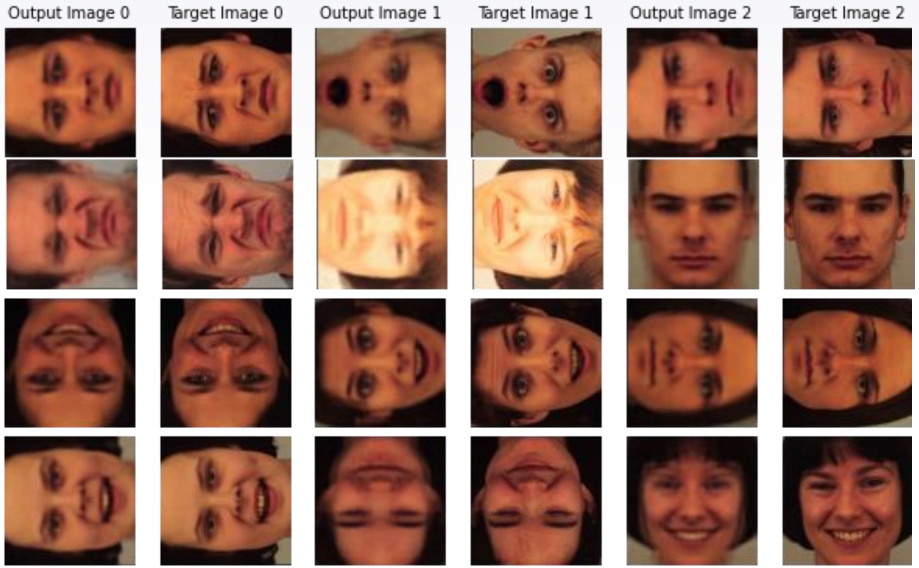


Cost of GCN vs. Epochs
Train
Validation
Cost
Epochs
0
1000
2000
3000
4000
5000
6000
7000
8000
0
200
400
600
800
1000



Accuracy of GCN vs. Epochs
Train
Validation
Acc(%)
Epochs
30
40
50
60
70
80
90
100
0
200
400
600
800
1000

KDEF



Training set faces reconstructed by our implementation

Training set faces reconstructed by the paper

KDEF

Low level features generation of the test set
Our implementation

Low level features generation of the test set
Original paper

KDEF

High level features generation of the test set
Our implementation

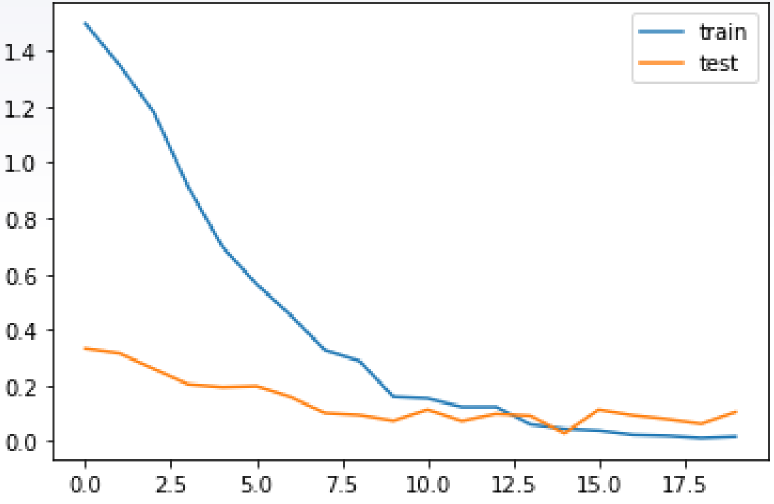
High level features generation of the test set
Original paper

Data Augmentation

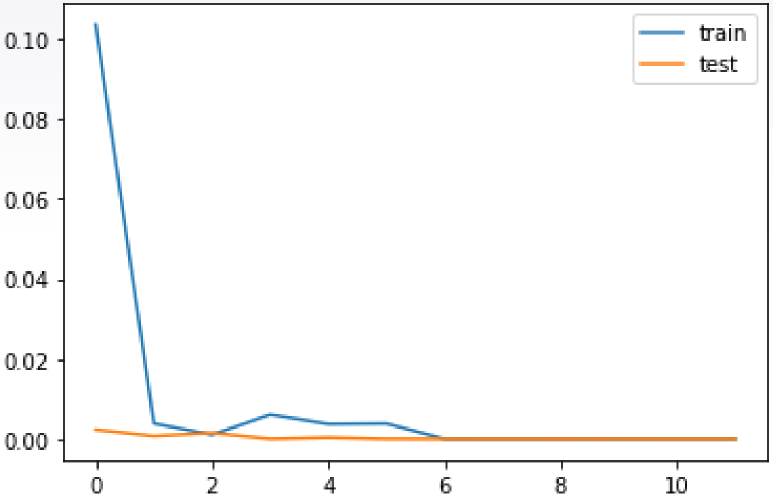
Results

Augmented images (from left: Surprised, Disgusted, Happy, Neutral)

Results



Loss function without augmented data



Loss function with augmented data

Results

Without augmented data	With augmented data
92%	100%

Accuracy on test images

QMNIST Preprocessing

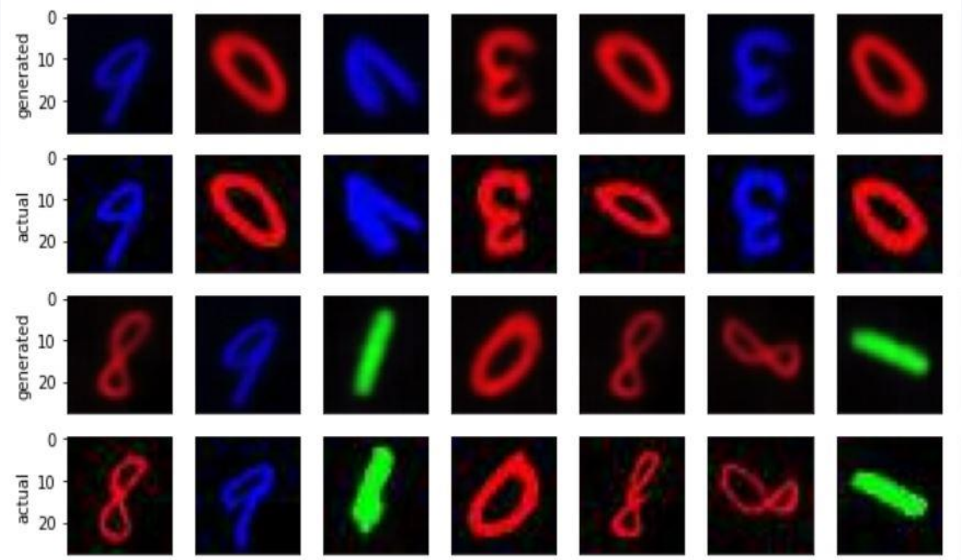
QMNIST Consists of 28*28 Black and White Pictures of Handwritten Digits Whose Writers are Known.

Preprocessing Below was Applied:

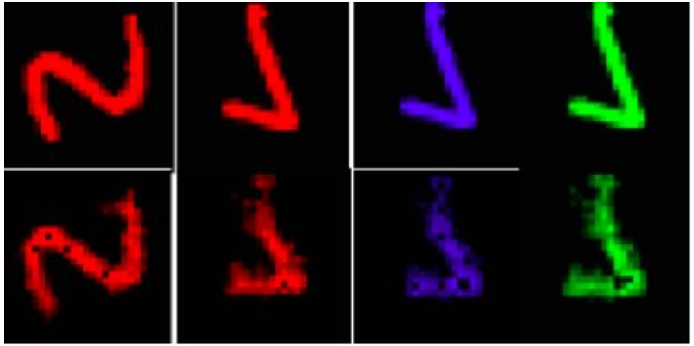
- Coloring the Pictures (to R, G and B)
- Rotation (0, 90, 180 and 270 Degrees)

** 10 Writers' Handwritten Digits Were Chosen. Either a Specific Rotation or Color of a Specific Digit was Taken Out for the Test Set.

QMNIST



Low level features generation of the test set
Our implementation



Low level features generation of the test set
Original paper

Conclusion and Suggestions

Conclusion

Generator

Takes One Hot Vectors as Inputs and Generates Images.

Classifier

Takes Images Generated by the Generator and Classifies Them.

Image Generation and Data Augmentation

The Trained Network is Utilized for Generating New Images, and This Method is Used as a Data Augmentation Approach to Train a More Accurate Network.

Critiques

- Information provided for data preprocessing and train/test splitting is vague
  - Unclear pre processing used for (Q)MNIST
  - Unclear test set for face generation
  - Used 5 emotions instead of 4 for the classification!!
- Network structure explanation is not exact:
  - Mismatch between network structure in the figure and text
  - No clear explanation for MNIST dataset network
- Learning rate decay seems <u>unreasonable</u>!!

THANKS!

Any questions?

You can contact us:

- erfunmirzaei@gmail.com
- mehmanchimahdi@gmail.com
- Sara.jahedazad@gmail.com

References

- Presentation template by SlidesCarnival
- Illustrations by Sergei Tikhonov
- Xu, Q., Qin, Z., & Wan, T. (2019, March). Generative cooperative net for image generation and data augmentation. In *International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making* (pp. 284-294). Springer, Cham.

References

- **Dosovitskiy, A., Tobias Springenberg, J., & Brox, T. (2015). Learning to generate chairs with convolutional neural networks. In** *Proceedings of the IEEE conference on computer vision and pattern recognition* **(pp. 1538-1546).**
- **Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks.** *Advances in neural information processing systems*, *25*, **1097-1105.**