Design an age estimation algorithm based on multi-task Neural Process

李禹陈 11712116
罗运升 11712608
叶致景 11712923

Southern University of Science and Technology

Content

Introduction
What's this project want to solve

Model
Outlook of the Neural Process

Method
What we have done

Plan
Our Future Work

01 Introduction

What's this project want to solve

Definition

What is the meaning of age estimation?

-According to a single image, implement an end-to-end system

-to estimate the approximate age of this person.

02

Literature review

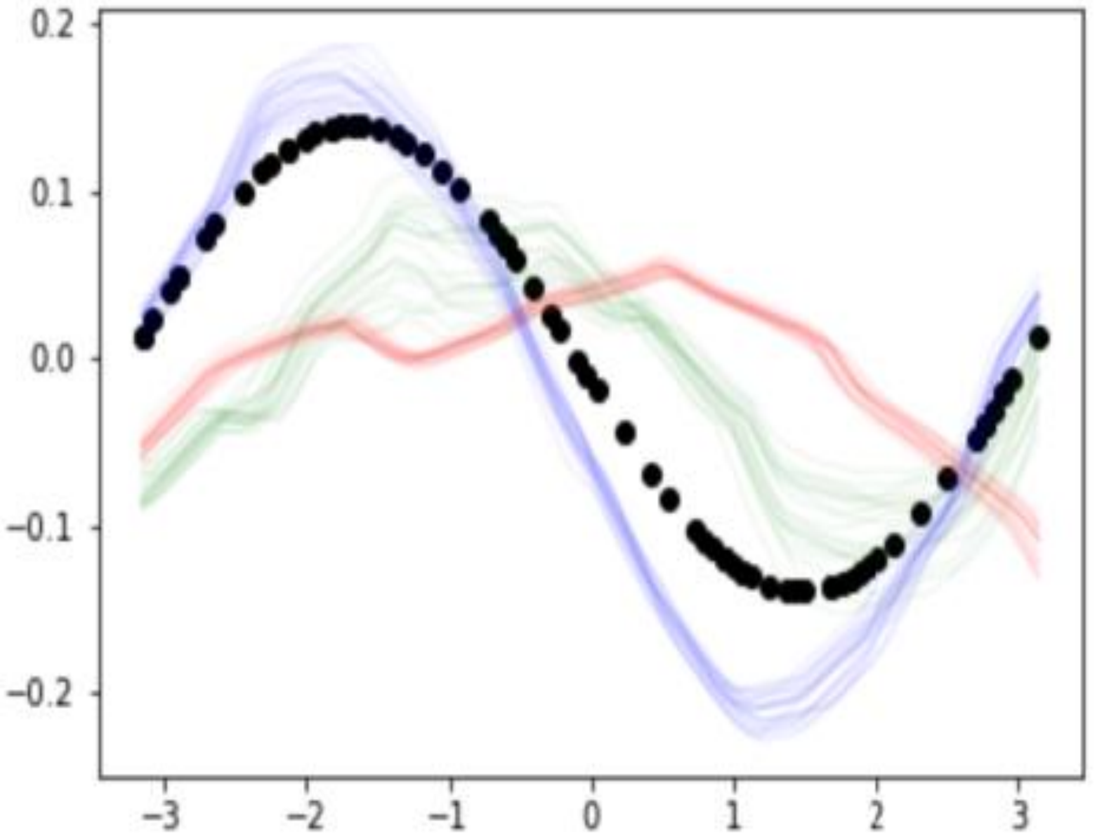
Outlook of the Neural Process

Neural Processes

- Neural Processes (NPs) are a neural network (NN) based probabilistic model which can represent a distribution over functions:

Test scenario: NP One-Dimension Regression

https://github.com/LuckyOne09/neural-processes/blob/master/experiments/example-1d.ipynb

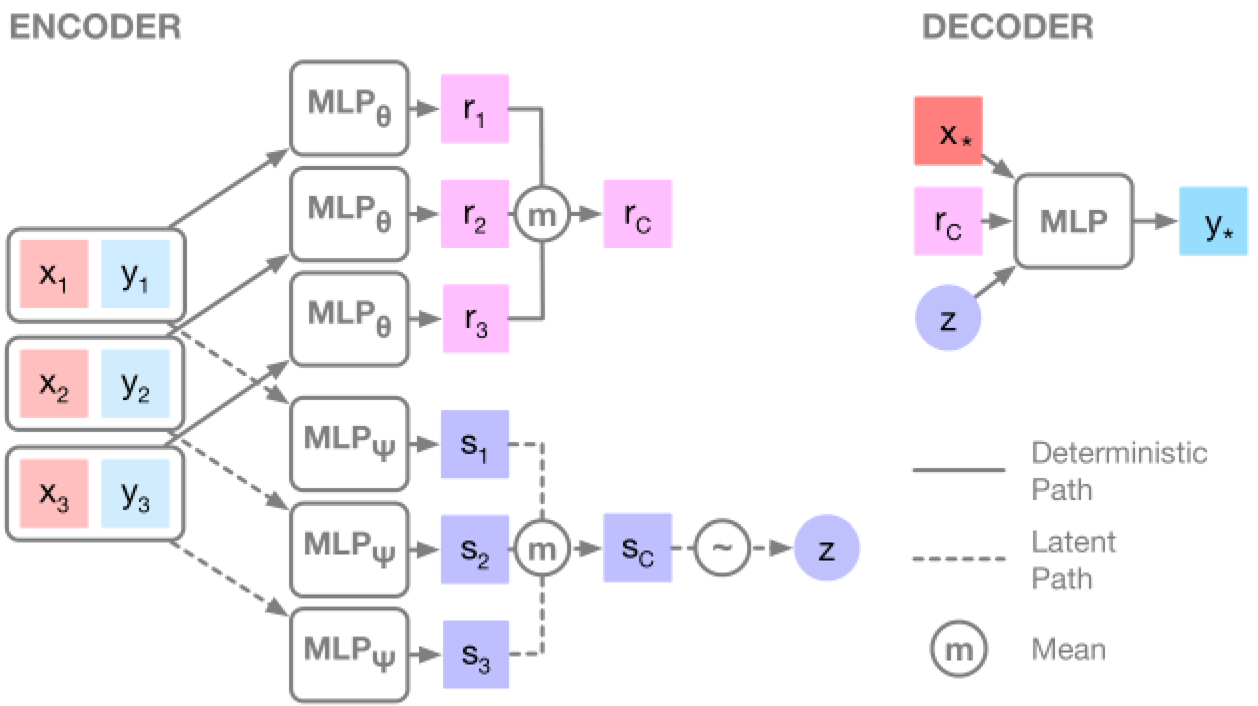


Neural Process for a simple class of 1D functions. The functions are defined by f(x) = a * sin(x - b) where a and b and randomly sampled.

Red line: context points: 2
Green line: context points: 20
Blue line: context points: 50

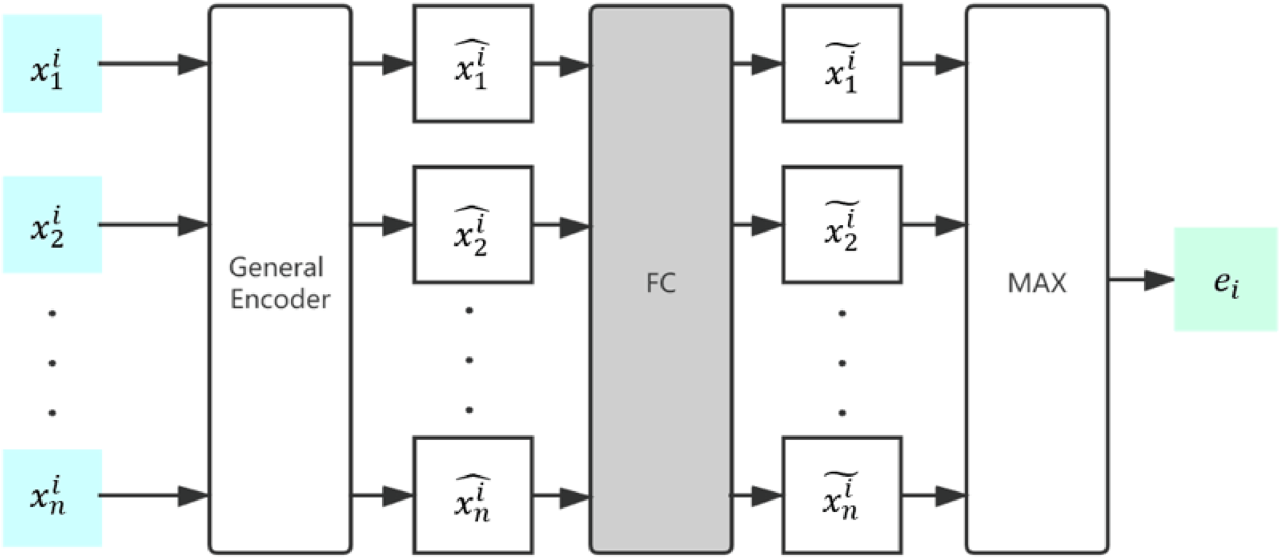
Black line: context points

NEURAL PROCESS



Method 03

WHAT WE HAVE DONE

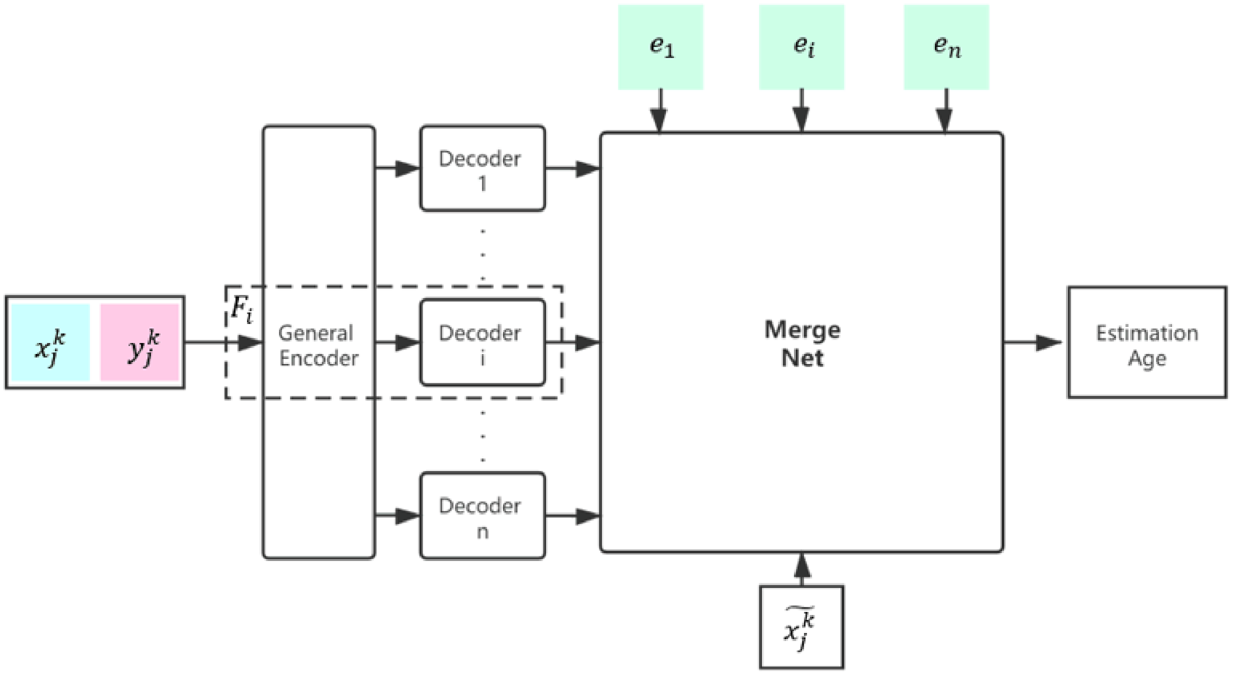


Shared Encoder: $\widehat{x_j^i} = E(x_i^j)$

FC: A fully connected layer we defined $\widetilde{x_j^i} = \sigma(W\widehat{x_j^i})$

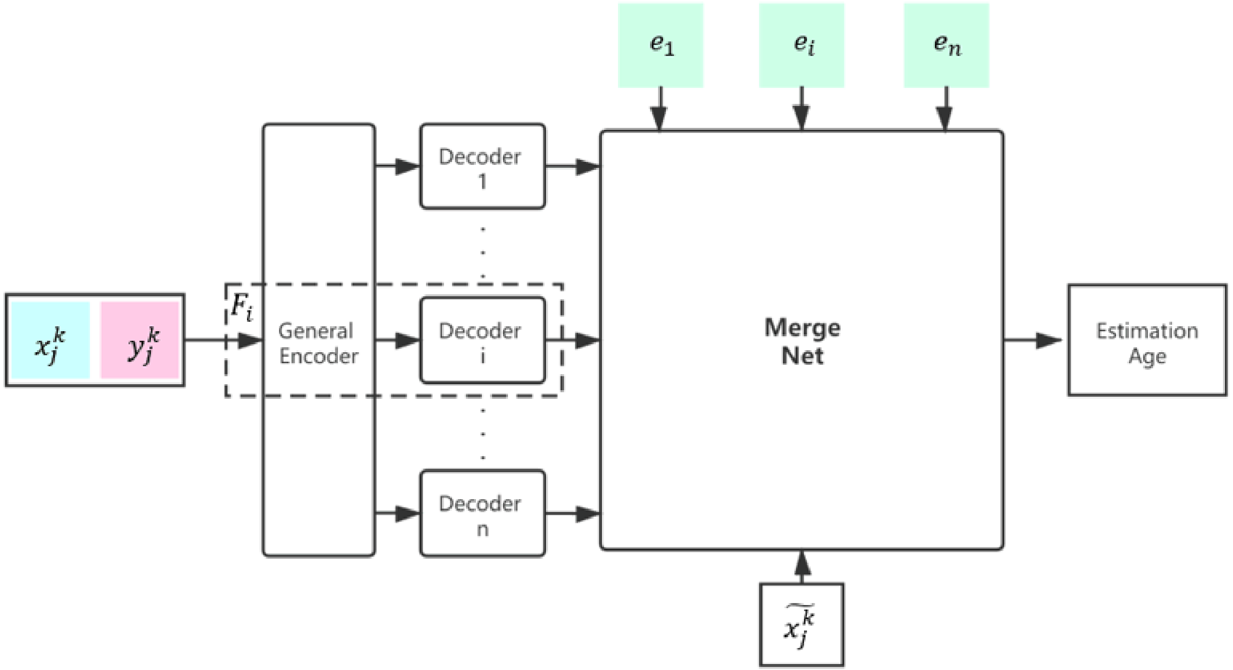
MAX: To get subject embedding. The max operator is conducted elementwisely

$$e_i = \max_j(\widetilde{x_j^i})$$



We define the similarity between pth subject and $\widetilde{x_j^k}$ as $s_{p,k,j}$

$$s_{p,k,j} = \frac{\exp(e_p^T \widetilde{x_j^k})}{\sum_{q \neq k} \exp(e_p^T \widetilde{x_j^k})}$$



The Estimation Age: $\sum_{\substack{p=1 \\ p \neq k}}^{n} s_{p,k,j} F_p(x_j^k)$

Actually we are here to train the FC we defined

The objective function to learn parameters:

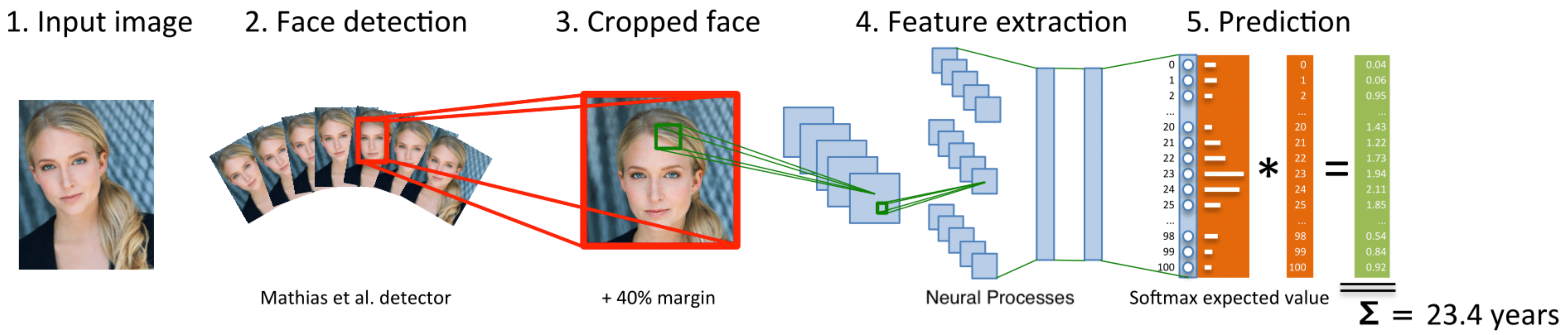
$$\min_{W} \sum_{k=1}^{n} \sum_{j=1}^{n_k} \left(y_j^k - \sum_{\substack{p=1 \\ p \neq \mathrm{k}}}^{n} s_{p,k,j} \, F_p(\widehat{x_j^k}) \right)$$

04

Result

WHAT WE HAVE REALIZED

Pipline


1. Input image
2. Face detection
3. Cropped face
4. Feature extraction
5. Prediction
0 1 2 ... 20 21 22 23 24 25 ... 98 99 100
*
0 1 2 ... 20 21 22 23 24 25 ... 98 99 100
=
0.04 0.06 0.95 ... 1.43 1.22 1.73 1.94 2.11 1.85 ... 0.54 0.84 0.92
Mathias et al. detector
+ 40% margin
Neural Processes
Softmax expected value
Σ = 23.4 years

Pre - processes: Face Detection

DPM ECCV2014

Brightness adjustment and contrast adjustment

Brightness adjustment

were the color and brightness of the image. To encode invariance to varying color contrast between images, we introduced brightness adjustment with a random scale factor α per image, sampled from a uniform distribution over $[-0.3, 0.3]$, through equation 1,

$$y = (x - \text{mean}) \times (1 + \alpha) \tag{1}$$

Contrast adjustment

and contrast adjustment with a random scale factor β per image, sampled from a uniform distribution over $[-0.2, 0.2]$, using equation 2.

$$y = (x - \text{mean}) \times (\beta) \tag{2}$$

Pre - processes: Face Cropping

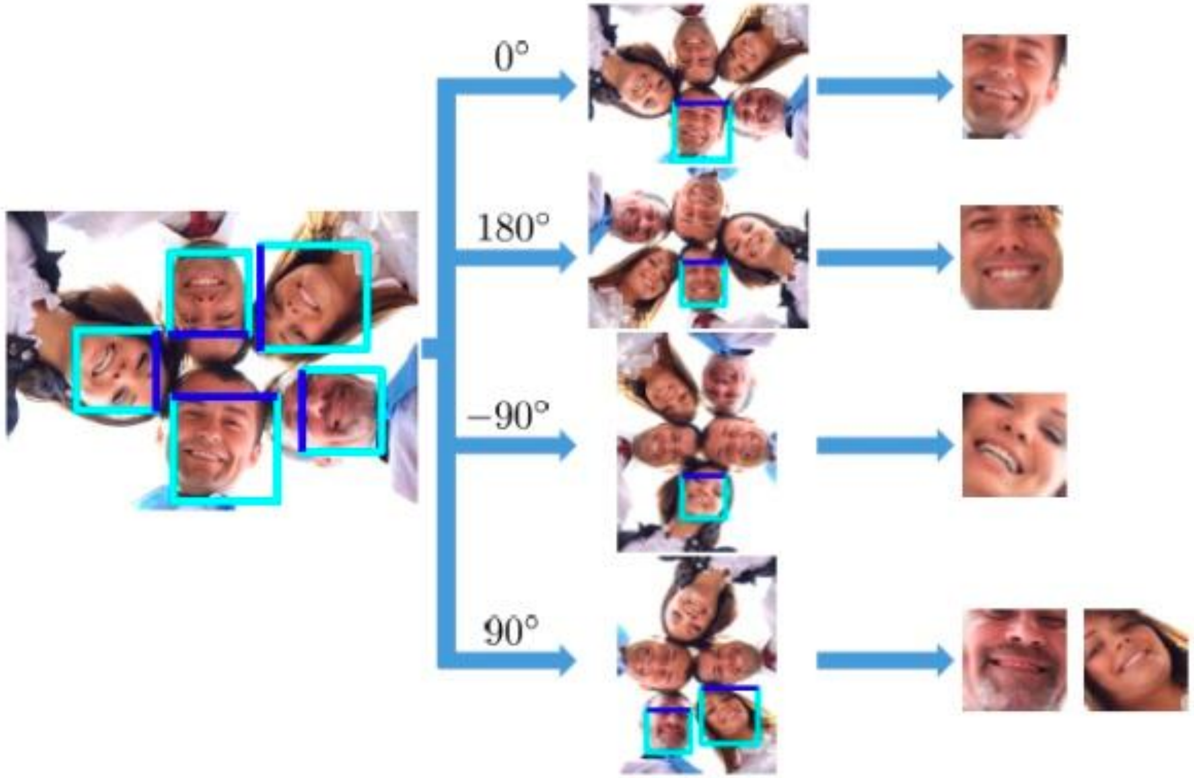
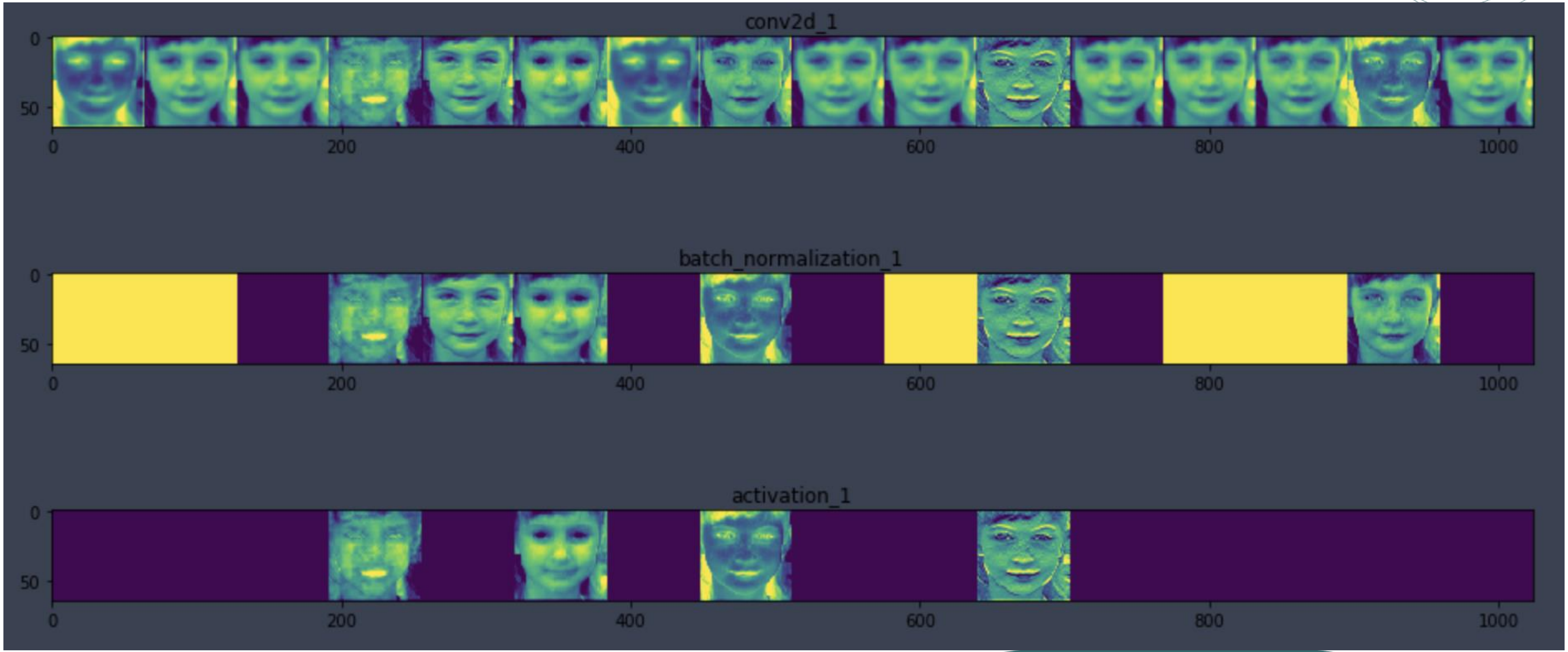
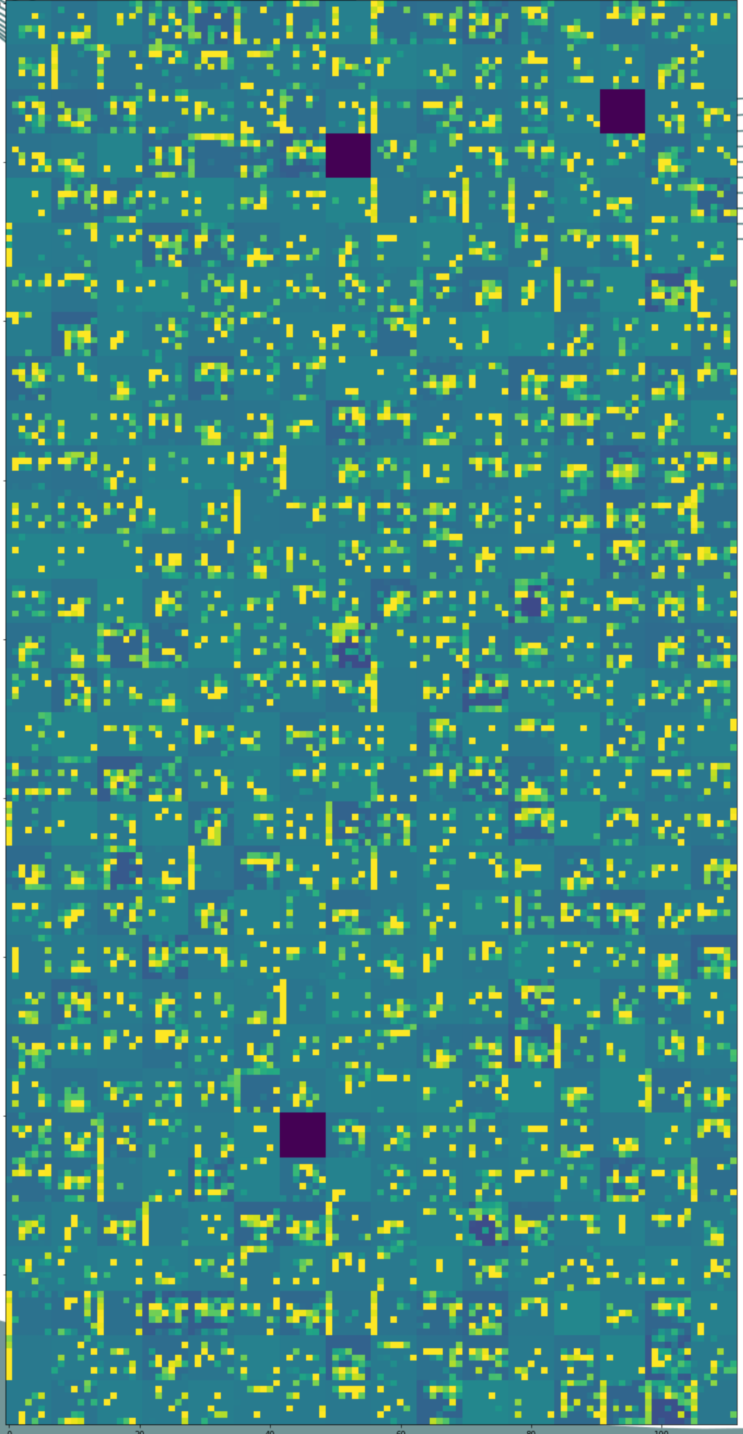


Figure 5. Rotate the original image by $-90°$, $90°$, and $180°$ to get
image left, image right, and image down. And the windows with

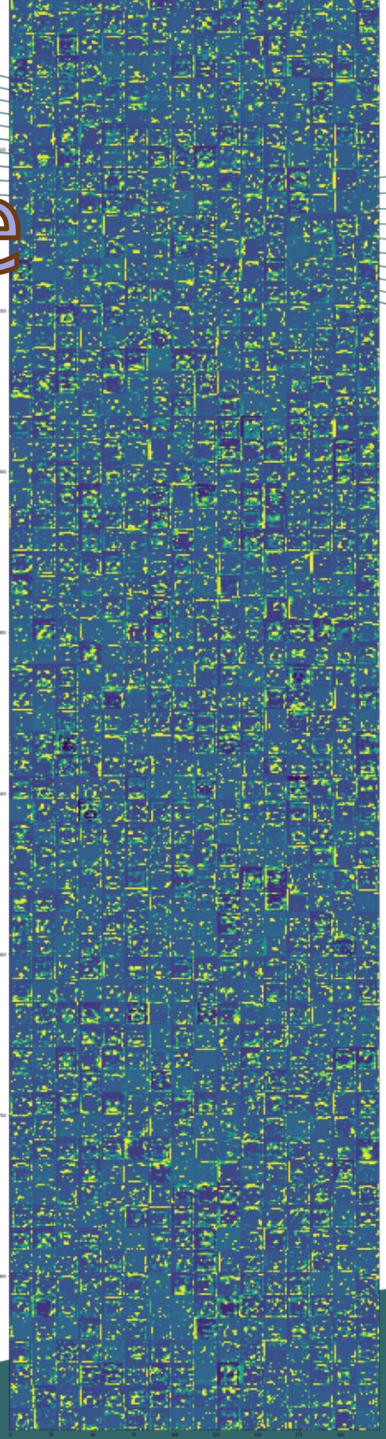
Layer Visualization



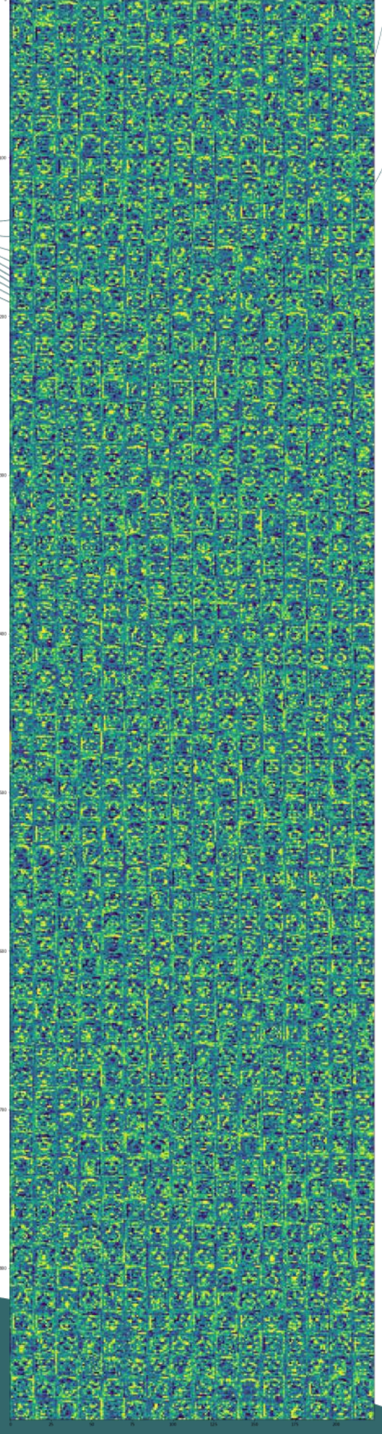
Feature Extractio
n



Abstract



Abstract

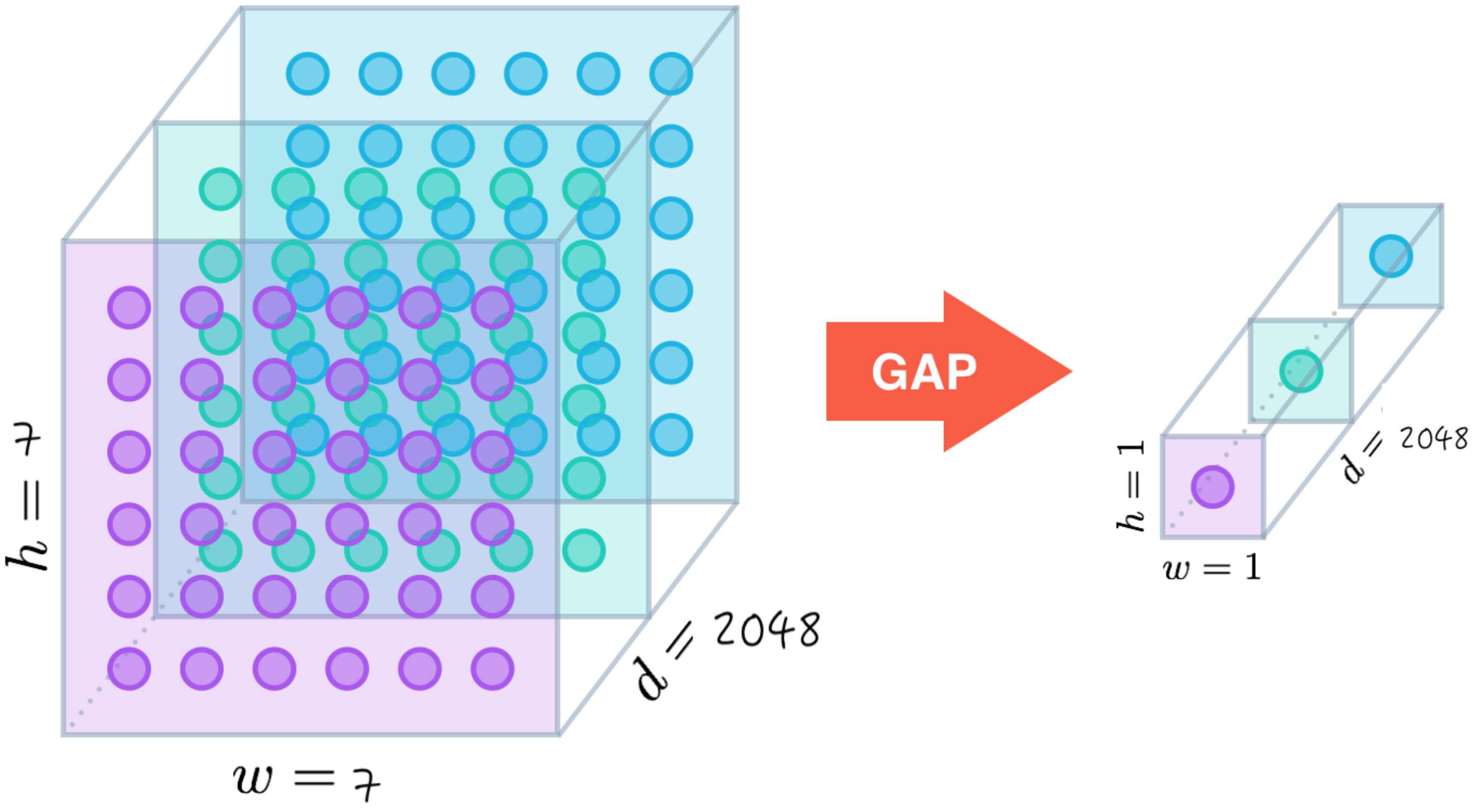


Model Summary

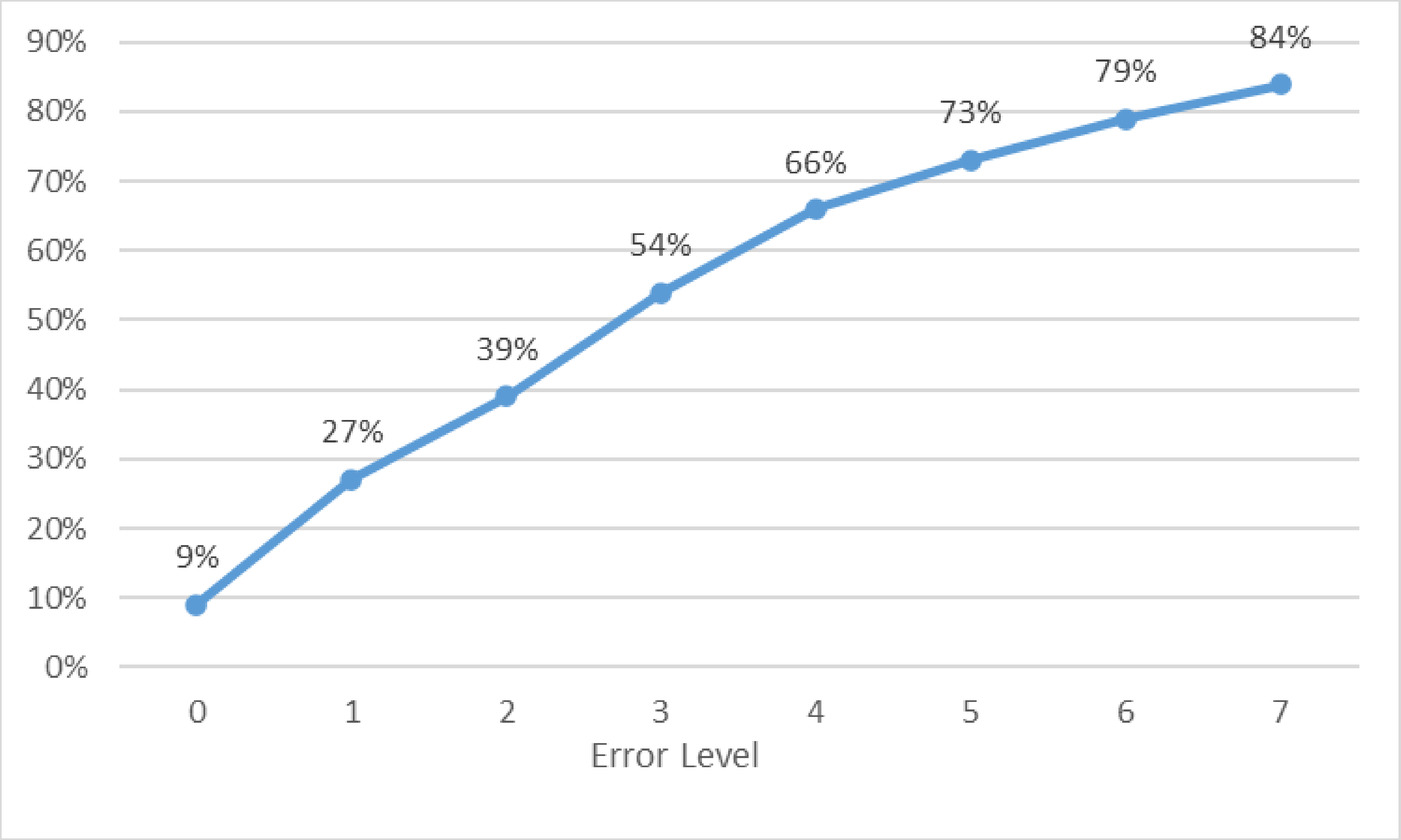
```
activation_879 (Activation)     (None, 7, 7, 2048)   0           add_287[0][0]
________________________________________________________________________________
res5c_branch2a (Conv2D)         (None, 7, 7, 512)    1049088     activation_879[0][0]
________________________________________________________________________________
bn5c_branch2a (BatchNormalizati (None, 7, 7, 512)    2048        res5c_branch2a[0][0]
________________________________________________________________________________
activation_880 (Activation)     (None, 7, 7, 512)    0           bn5c_branch2a[0][0]
________________________________________________________________________________
res5c_branch2b (Conv2D)         (None, 7, 7, 512)    2359808     activation_880[0][0]
________________________________________________________________________________
bn5c_branch2b (BatchNormalizati (None, 7, 7, 512)    2048        res5c_branch2b[0][0]
________________________________________________________________________________
activation_881 (Activation)     (None, 7, 7, 512)    0           bn5c_branch2b[0][0]
________________________________________________________________________________
res5c_branch2c (Conv2D)         (None, 7, 7, 2048)   1050624     activation_881[0][0]
________________________________________________________________________________
bn5c_branch2c (BatchNormalizati (None, 7, 7, 2048)   8192        res5c_branch2c[0][0]
________________________________________________________________________________
add_288 (Add)                   (None, 7, 7, 2048)   0           bn5c_branch2c[0][0]
                                                                 activation_879[0][0]
________________________________________________________________________________
activation_882 (Activation)     (None, 7, 7, 2048)   0           add_288[0][0]
________________________________________________________________________________
global_average_pooling2d_18 (Gl (None, 2048)         0           activation_882[0][0]
================================================================================
Total params: 23,587,712
Trainable params: 23,534,592
Non-trainable params: 53,120
```

Global Average Pooling

activation_49 (Activation)   (None, 7, 7, 2048)   0
global_average_pooling2d_1   (None, 2048)   0



Single Task Estimation



Some Prediction Result

The Correct Example

THe Fail Example

Reference

- 1.Garnelo, M., Schwarz, J., Rosenbaum, D., Viola, F., Rezende, D. J., Eslami, S. M., & Teh, Y. W. (2018). Neural processes. arXiv preprint arXiv:1807.01622.
- 2. Garnelo, M., Rosenbaum, D., Maddison, C. J., Ramalho, T., Saxton, D., Shanahan, M., ... & Eslami, S. M. (2018). Conditional neural processes. arXiv preprint arXiv:1807.01613.
- 3. Kim, H., Mnih, A., Schwarz, J., Garnelo, M., Eslami, A., Rosenbaum, D., ... & Teh, Y. W. (2019). Attentive neural processes. arXiv preprint arXiv:1901.05761.
- 4. Zhang, Y., & Yeung, D. Y. (2010, June). Multi-task warped gaussian process for personalized age estimation. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (pp. 2622-2629). IEEE.
- 5. Rothe, R., Timofte, R., & Van Gool, L. (2015). Dex: Deep expectation of apparent age from a single image. In Proceedings of the IEEE International Conference on Computer Vision Workshops (pp. 10-15).