Facial Beauty Prediction

Speaker：Cheng Ping Hsieh (b04901020@ntu.edu.tw)

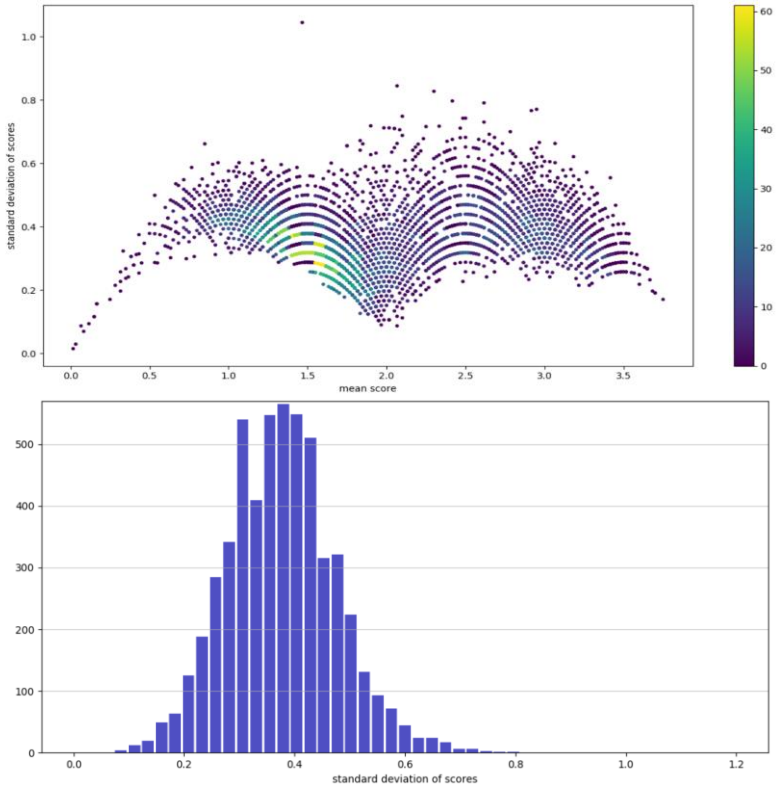
Advisor：Prof. Winston H. Hsu

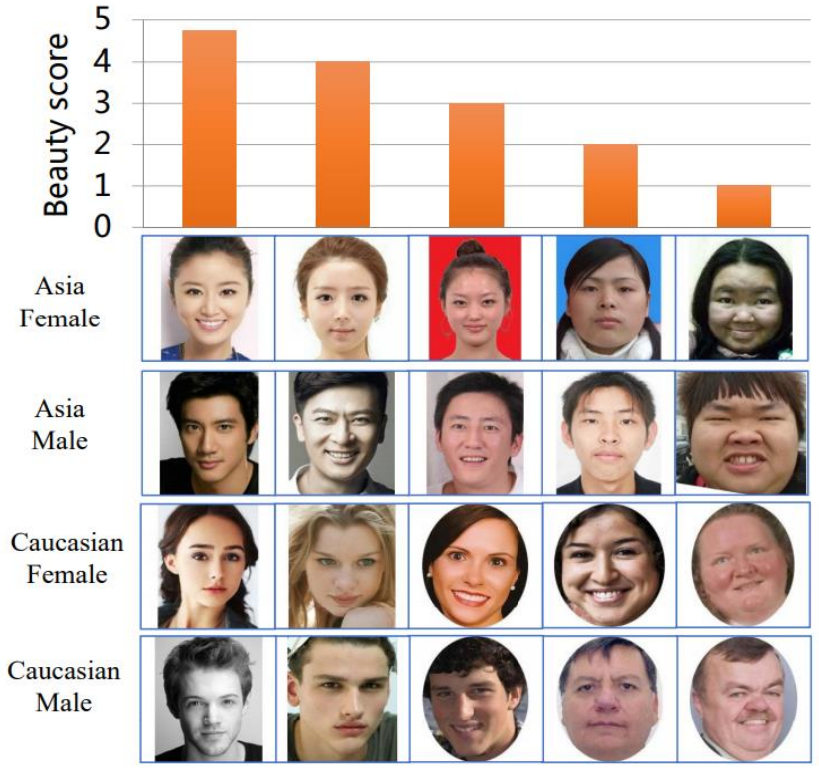
Outline

- Dataset
- Proposed Method
- Evaluation
- Experiment Result
- Discussion
- Conclusion
- Live Demo
- Reference

Dataset

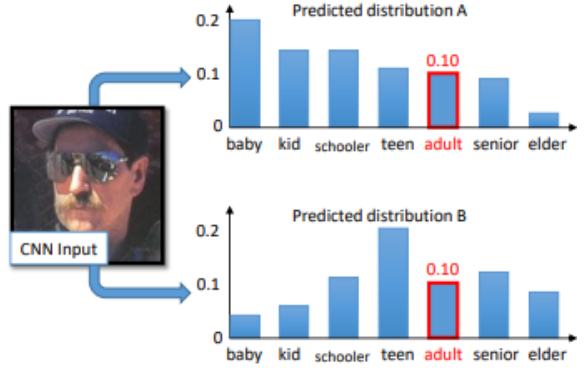
- Dataset：SCUT-FBP5500[1]
- Image number：5500；Labelers：60
- Face Property：Asian/Caucasian; Male/Female
- Extreme values and Median values have small variance





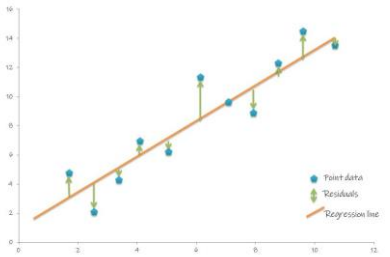
Proposed Method

- Pre-trained ImageNet Models
- Earth Mover Distance-based Loss[2][3]
  - Distribution of beauty score
  - Consider Class Relationship



$$\mathrm{E}_E(\mathbf{p},\mathbf{t}) = \sum_{i=1}^{C} \left(\mathrm{CDF}_i(\mathbf{p}) - \mathrm{CDF}_i(\mathbf{t})\right)^2$$

- Mean Square Error
  - Mean value of beauty score
  - Regression



$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2$$

- Multi-BinaryCrossEntropyLoss
  - Distribution of beauty score
  - Classification

$$-\frac{1}{C} * \sum_i y[i] * \log((1+\exp(-x[i]))^{-1}) + (1-y[i]) * \log\left(\frac{\exp(-x[i])}{(1+\exp(-x[i]))}\right)$$

Evaluation

- Metrics - 5 Cross Validation (80% train and 20% test)
  - Pearson correlation coefficient (PC)
  - Mean absolute error (MAE)
  - Root mean square error (RMSE)

PC	1	2	3	4	5	Average
AlexNet	0.8667	0.8645	0.8615	0.8678	0.8566	0.8634
ResNet-18	0.8847	0.8792	0.8929	0.8932	0.9004	0.89
ResNeXt-50	**0.8985**	**0.8932**	**0.9016**	**0.899**	**0.9064**	**0.8997**
MAE	1	2	3	4	5	Average
AlexNet	0.2633	0.2605	0.2681	0.2609	0.2728	0.2651
ResNet-18	0.248	0.2459	0.243	0.2383	0.2383	0.2419
ResNeXt-50	**0.2306**	**0.2285**	**0.226**	**0.2349**	**0.2258**	**0.2291**
RMSE	1	2	3	4	5	Average
AlexNet	0.3408	0.3449	0.3538	0.3438	0.3576	0.3481
ResNet-18	0.3258	0.3286	0.3184	0.3107	0.2994	0.3166
ResNeXt-50	**0.3025**	**0.3084**	**0.3016**	**0.3044**	**0.2918**	**0.3017**

Experiment Result-1

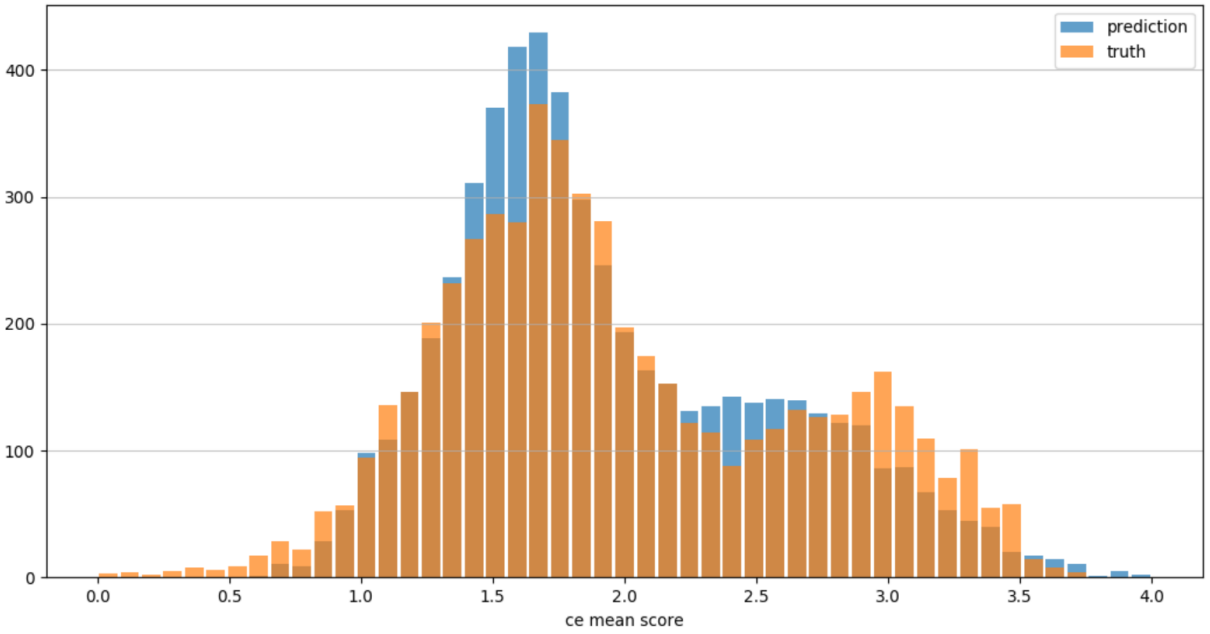
- Complexed model has higher performance
- EMD Loss has the highest performance with overall metrics
- Squeeze Net for future application

EMD > MSE > CE > Base

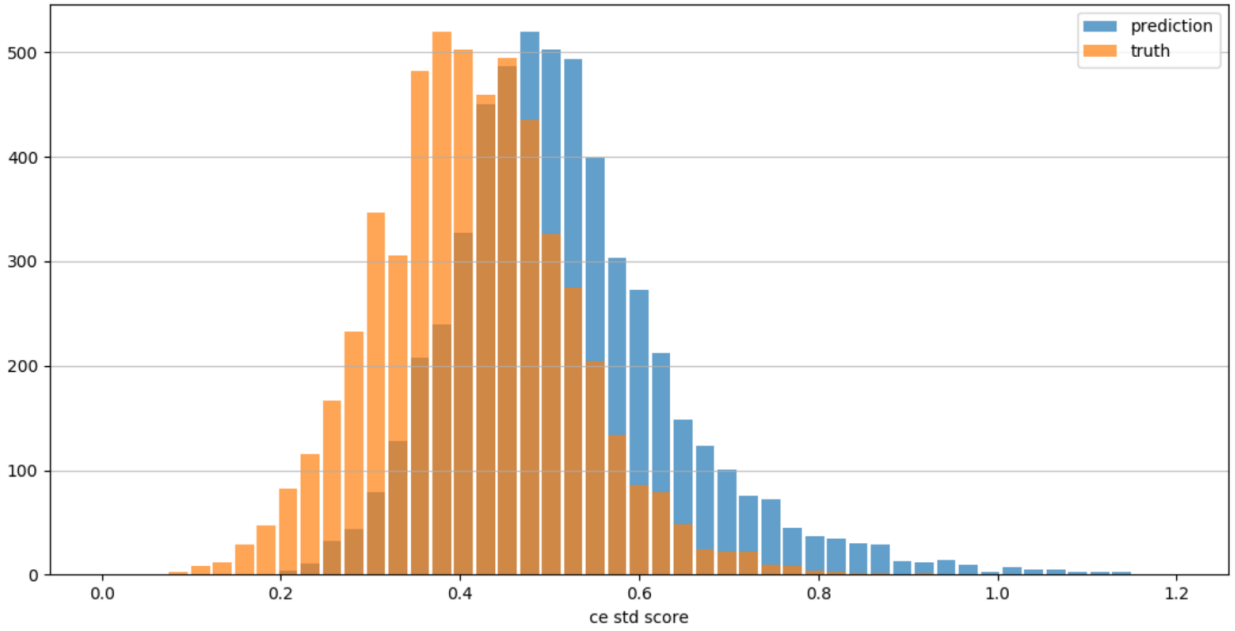
CV x 5	Model	Baseline	EMD Loss	MSE Loss	BCE Loss
PC ↑	Alex	0.8634	0.88852	0.85924	0.8668
	Squeeze		0.89904	0.89992	0.88144
	Resnet18	0.89	0.91668	0.92044	0.91554
	Resnet50	0.8997	0.9203	0.92082	0.9168
MAE ↓	Alex	0.2651	0.23992	0.27066	0.26294
	Squeeze		0.22752	0.22962	0.24772
	Resnet18	0.2419	0.2072	0.20742	0.2094
	Resnet50	0.2291	0.20268	0.20434	0.21368
RMSE ↓	Alex	0.3481	0.3172	0.35268	0.34856
	Squeeze		0.3012	0.30126	0.32792
	Resnet18	0.3166	0.27728	0.27352	0.27836
	Resnet50	0.3017	0.2698	0.26888	0.28184

Experiment Result-2

- Easily predict high frequency mean score
- Model performance in different frequency

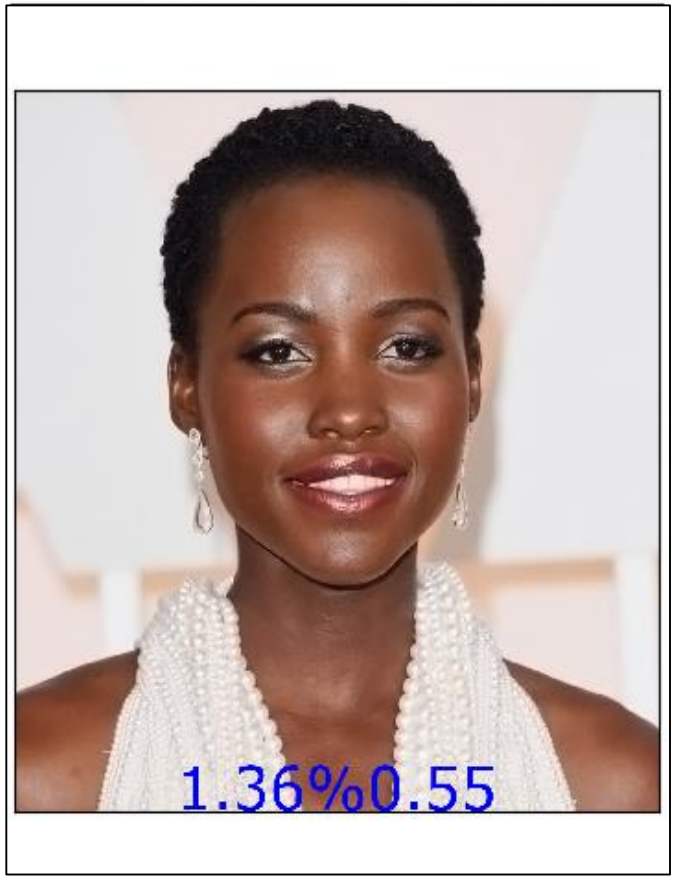


- Easily predict high frequency std score
- CE Loss has higher standard deviation

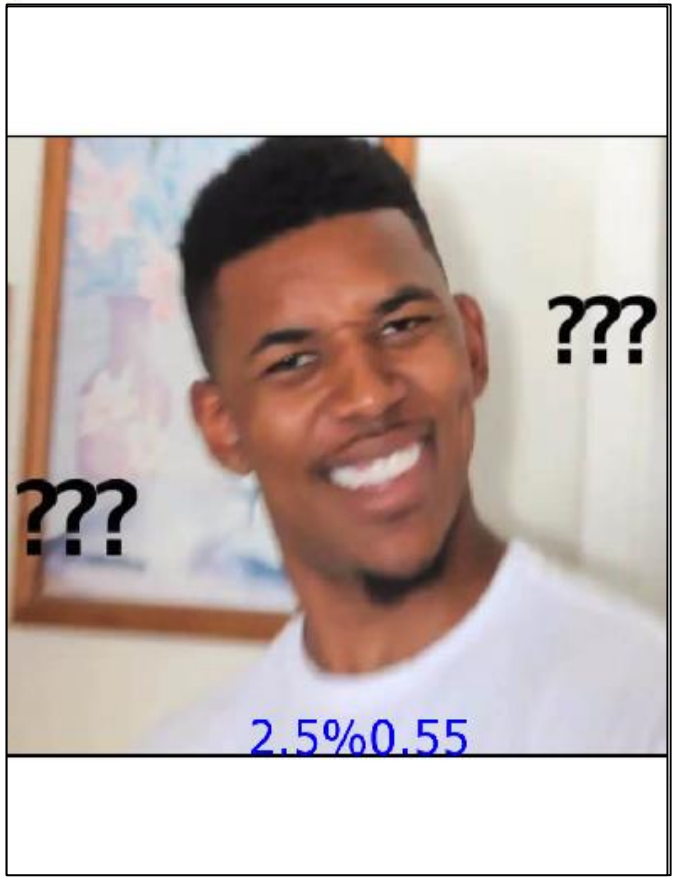


Experiment Result-3

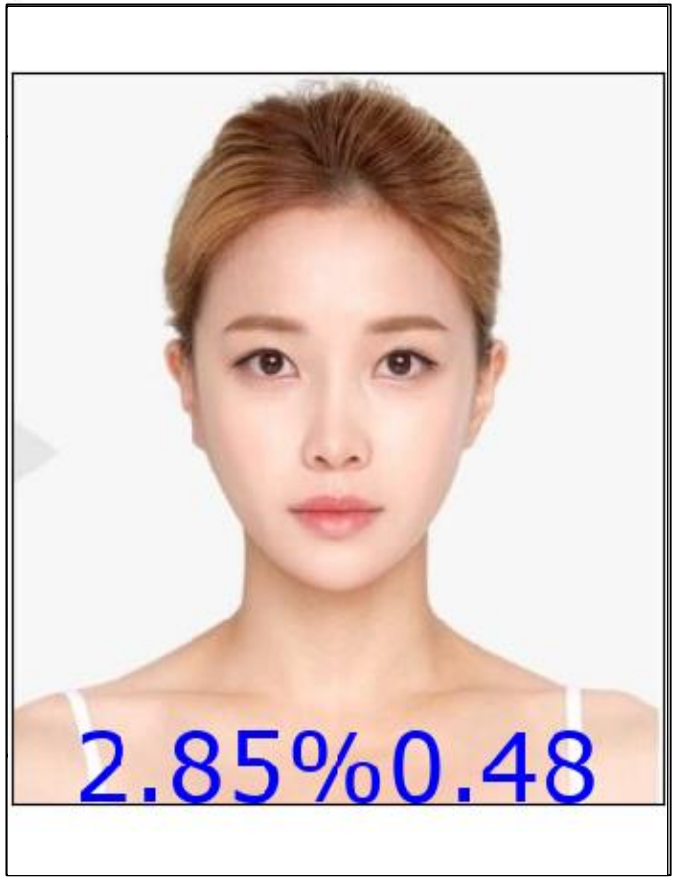
Experiment Result-3



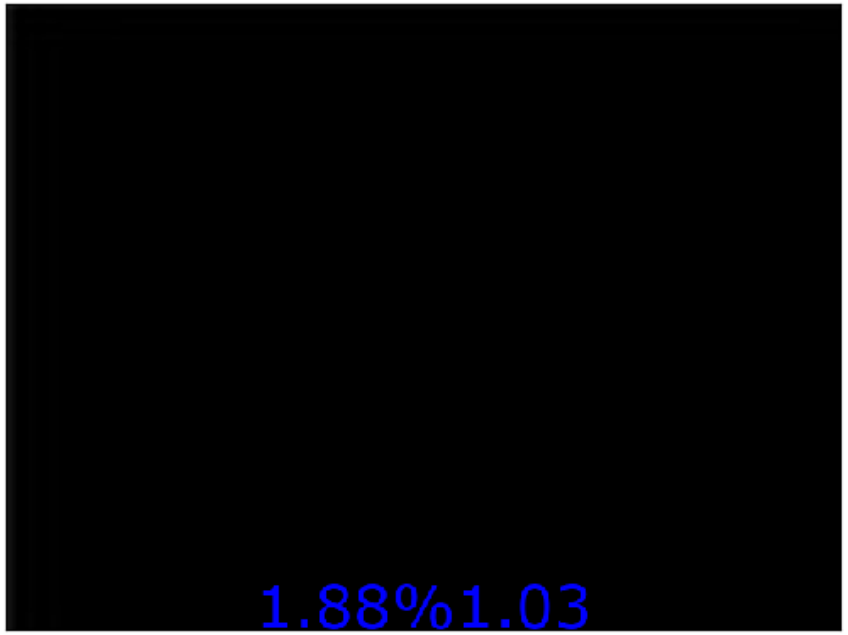
Experiment Result-3

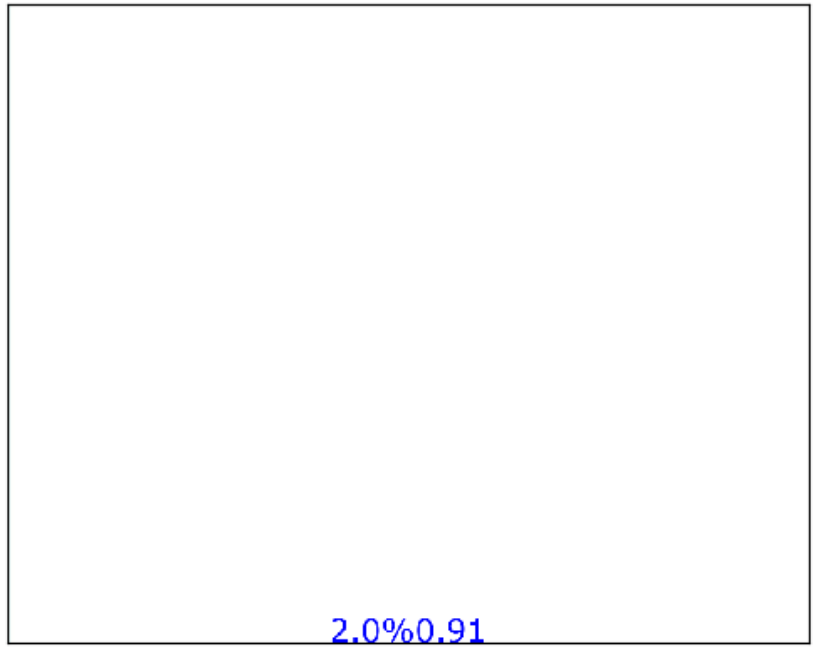


Experiment Result-3

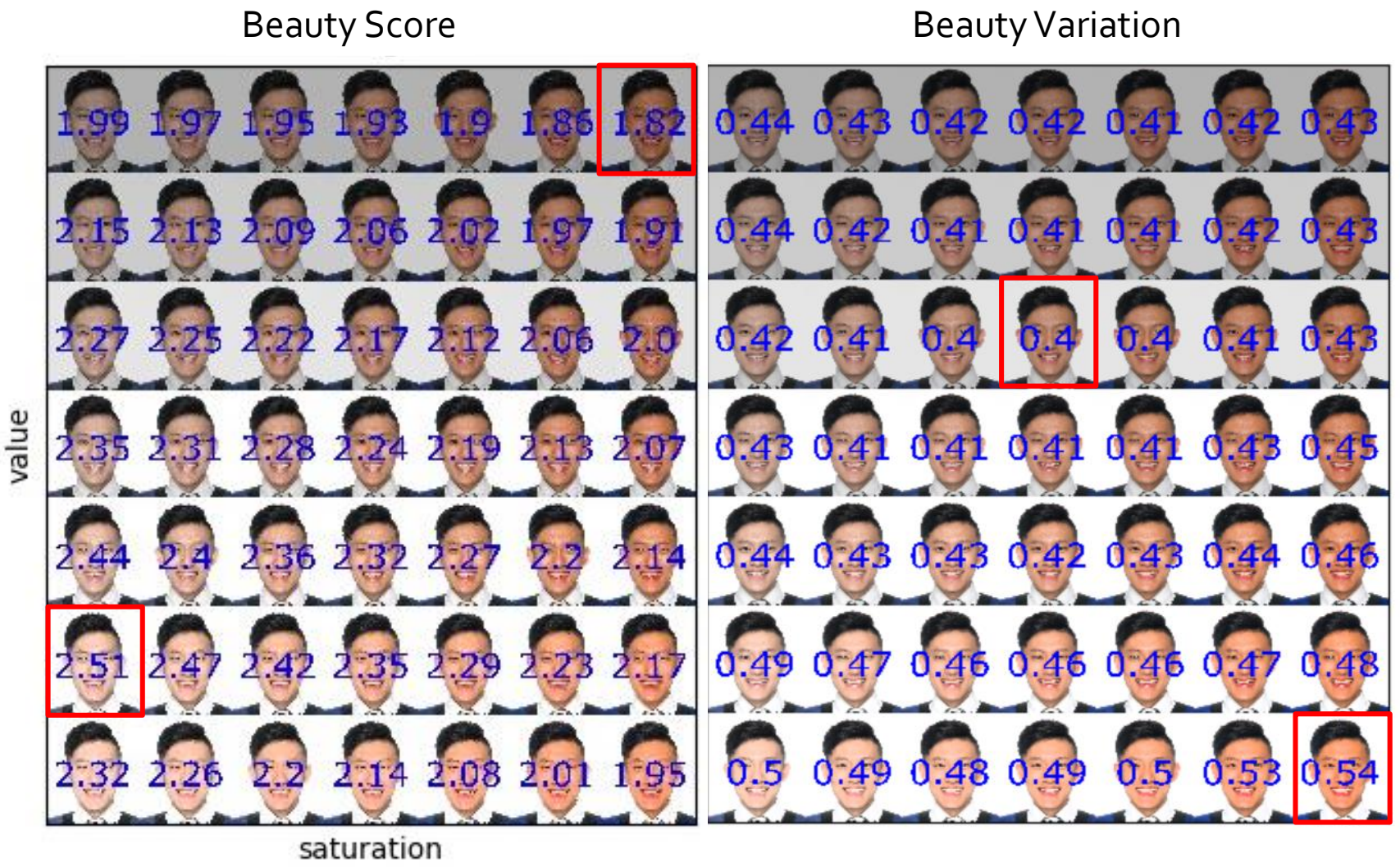


Experiment Result-3



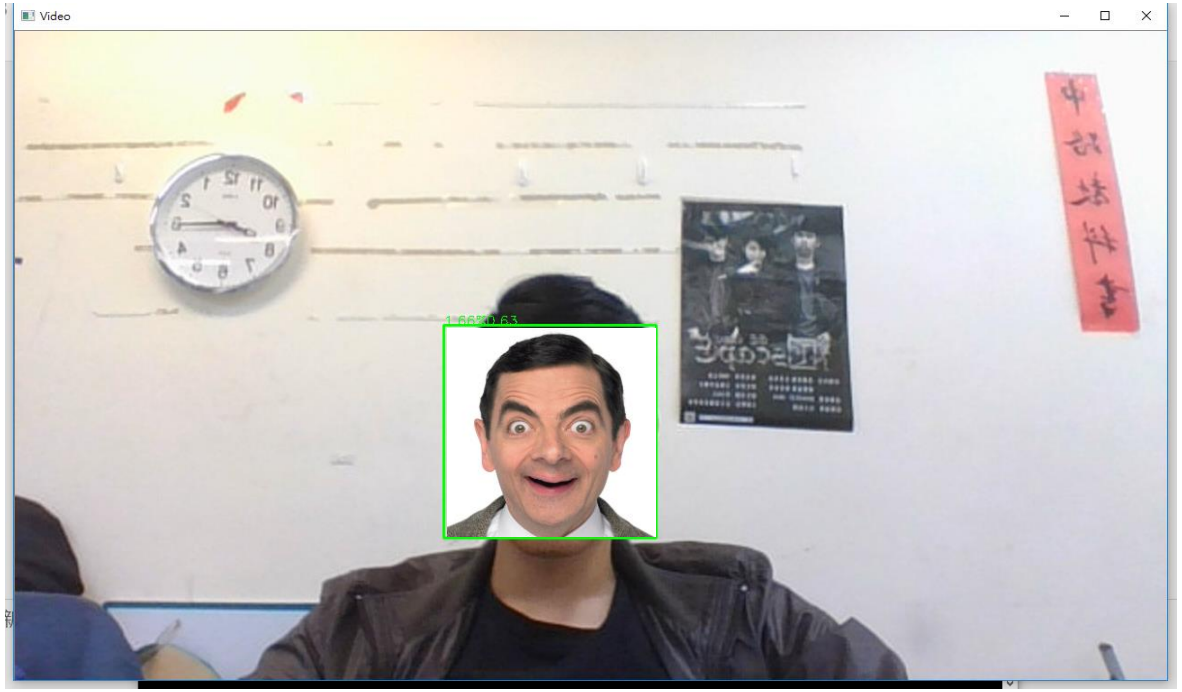


Discussion


Beauty Score
Beauty Variation
value
saturation
1.99 1.97 1.95 1.93 1.9 1.86 1.82
2.15 2.13 2.09 2.06 2.02 1.97 1.91
2.27 2.25 2.22 2.17 2.12 2.06 2.0
2.35 2.31 2.28 2.24 2.19 2.13 2.07
2.44 2.4 2.36 2.32 2.27 2.2 2.14
2.51 2.47 2.42 2.35 2.29 2.23 2.17
2.32 2.26 2.2 2.14 2.08 2.01 1.95
0.44 0.43 0.42 0.42 0.41 0.42 0.43
0.44 0.42 0.41 0.41 0.41 0.42 0.43
0.42 0.41 0.4 0.4 0.4 0.41 0.43
0.43 0.41 0.41 0.41 0.41 0.43 0.45
0.44 0.43 0.43 0.42 0.43 0.44 0.46
0.49 0.47 0.46 0.46 0.46 0.47 0.48
0.5 0.49 0.48 0.49 0.5 0.53 0.54

Conclusion

- EMD Loss has higher performance
- Complexed Model has higher performance
- Model is sensitive to light and skin color (data augmentation)
- Model has tendency to predict high frequency value



Reference

- [1] Liang, Lingyu, et al. "SCUT-FBP5500: A Diverse Benchmark Dataset for Multi-Paradigm Facial Beauty Prediction." *arXiv preprint arXiv:1801.06345* (2018).
- [2] Talebi, Hossein, and Peyman Milanfar. "Nima: Neural image assessment." *IEEE Transactions on Image Processing* 27.8 (2018): 3998-4011.
- [3] Hou, Le, Chen-Ping Yu, and Dimitris Samaras. "Squared Earth Mover's Distance-based Loss for Training Deep Neural Networks." *arXiv preprint arXiv:1611.05916* (2016).