Historical Context-based Style Classification of Painting Images via Label Distribution Learning

Jufeng Yang[1], Liyi Chen[1], Le Zhang[2], Xiaoxiao Sun[1], Dongyu She[1], Shao-Ping Lu[1], Ming-Ming Cheng[1]

[1]Nankai University, [2]Advanced Digital Sciences Center

http://cv.nankai.edu.cn/ chenliyi1995@126.com

1. Motivation

The evolution of painting style is both **continuous**, in a sense that new styles may inherit, develop or even mutate from their predecessors and **multi-modal** because of various issues such as the visual appearance, the birthplace, the origin time and the art movement. We propose to synthesize historical knowledge into the image label via the **label distribution learning** and encapsulate into a **multi-task learning framework** to assist visual feature learning.

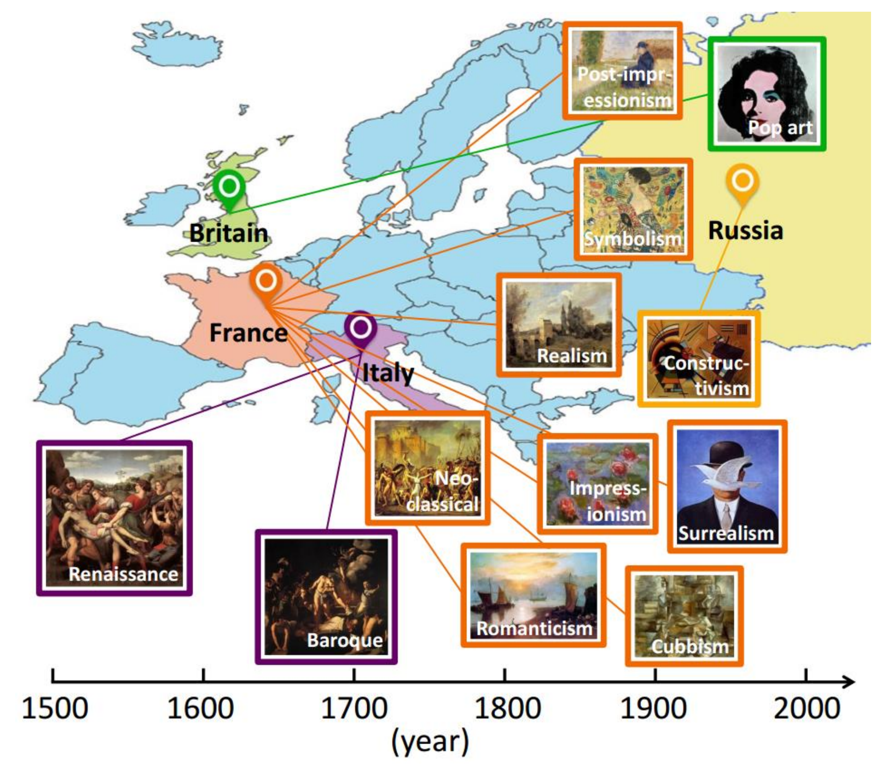

Fig.1 Style examples on Painting91 dataset.

2. Method

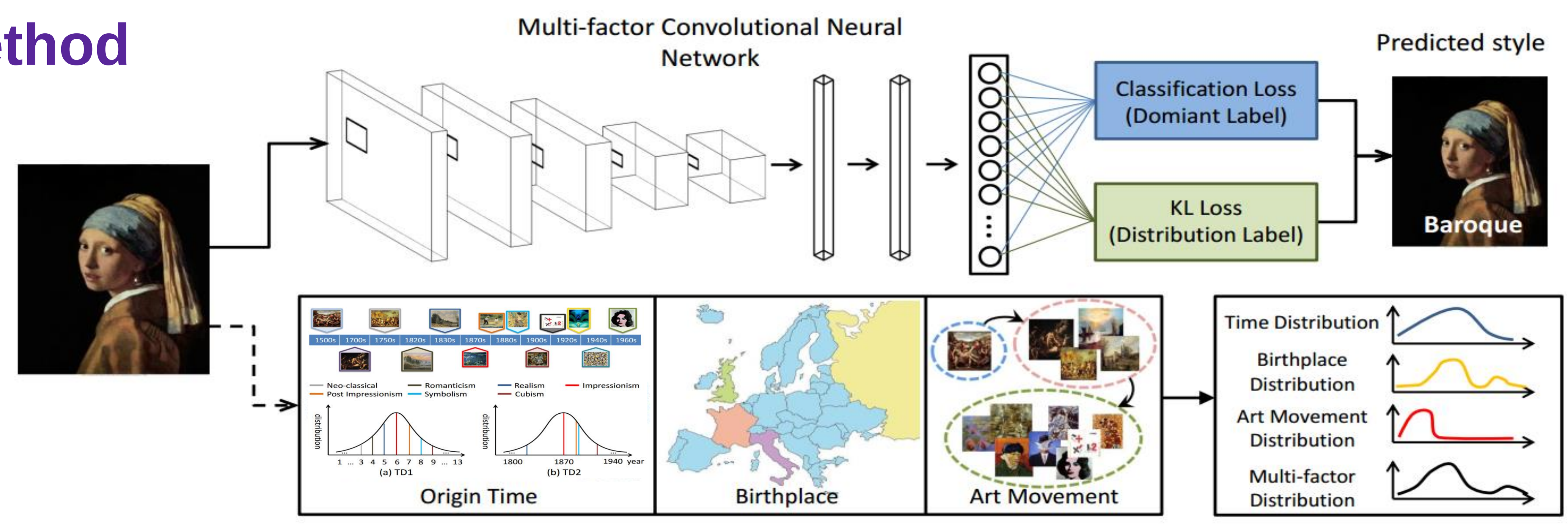

Fig.2 The illustration of the proposed method.

➢ Label Distribution:

The label distribution is generated based on three kinds of historical information:

a. Time

$$\boldsymbol{t1}_i = \frac{f(T_i, T_y, \sigma)}{\sum_{k=1}^{c} f(T_k, T_y, \sigma)} \quad (1)$$

b. Birthplace

$$\boldsymbol{b}_i = \begin{cases} 1, & i = y \\ \frac{\beta}{n_b}, & B_i = B_y, i \neq y \\ 0, & otherwise \end{cases} \quad (2)$$

c. Art Movement

$$\boldsymbol{a}_i = \begin{cases} 1, & i = y \\ \frac{\alpha}{n_a}, & A_i = A_y, i \neq y \\ 0, & otherwise \end{cases} \quad (3)$$

d. Multi-factor Distribution

$$\boldsymbol{l} = \eta \times \boldsymbol{t1} + (1-\eta) \times \boldsymbol{t2} + \boldsymbol{b} + \boldsymbol{a} \quad (4)$$

➢ Optimization:

$$L = \lambda \underbrace{\boxed{L_{sty}(x, y)}}_{\text{classification loss}} + (1-\lambda) \underbrace{\boxed{L_{dis}(x, \boldsymbol{l})}}_{\text{KL loss}} \quad (5)$$

- The classification loss calculates the loss of the ground truth and predicted style.
- The KL loss calculates the loss of the distribution of generated and predicted.

3. Performance

Method	Painting91	OilPainting	Pandora
VGGNet [44]	72.89%	64.24%	70.52%
Khan F. S. *et al.* [23]	62.20%	-	-
Condorovici *et al.* [6]	-	-	37.90%
Florea *et al.* [9]	-	-	54.70%
CMFFV [37]	67.43%	-	-
MSCNN1 [34]	69.67%	55.24%	70.32%
MSCNN2 [34]	70.96%	57.92%	69.75%
CNN F4 [33]	69.21%	58.47%	70.47%
Peng K. C. *et al.* [35]	71.05%	-	-
Gram [5]	71.86%	60.61%	-
Gram-Cov [5]	72.41%	60.72%	-
Gram dot Cos [5]	73.59%	63.33%	-
SCMFA [38]	73.16%	-	-
Anwer R. M. *et al.* [1]	74.80%	-	-
Ours	**77.76%**	**70.59%**	**73.28%**

Fig.3 Classification accuracy of different baseline methods, including traditional and deep methods.

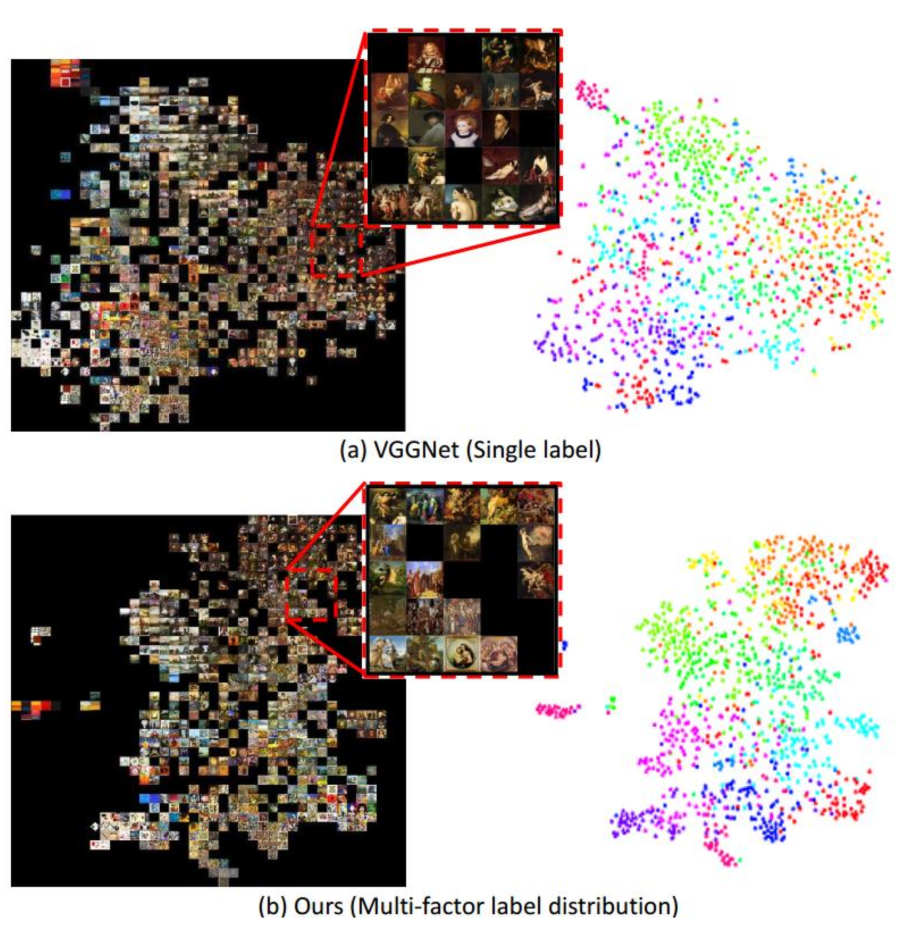

Fig.4 Comparison of VGG-Net and our method on the Painting-91 dataset.

4. Conclusion

- ➢ The three historical context-based information has been encoded as a soft label to assist visual feature learning in CNN.
- ➢ Experimental results demonstrate that our proposed method performs favorably against the state-of-the-art approaches on various datasets.