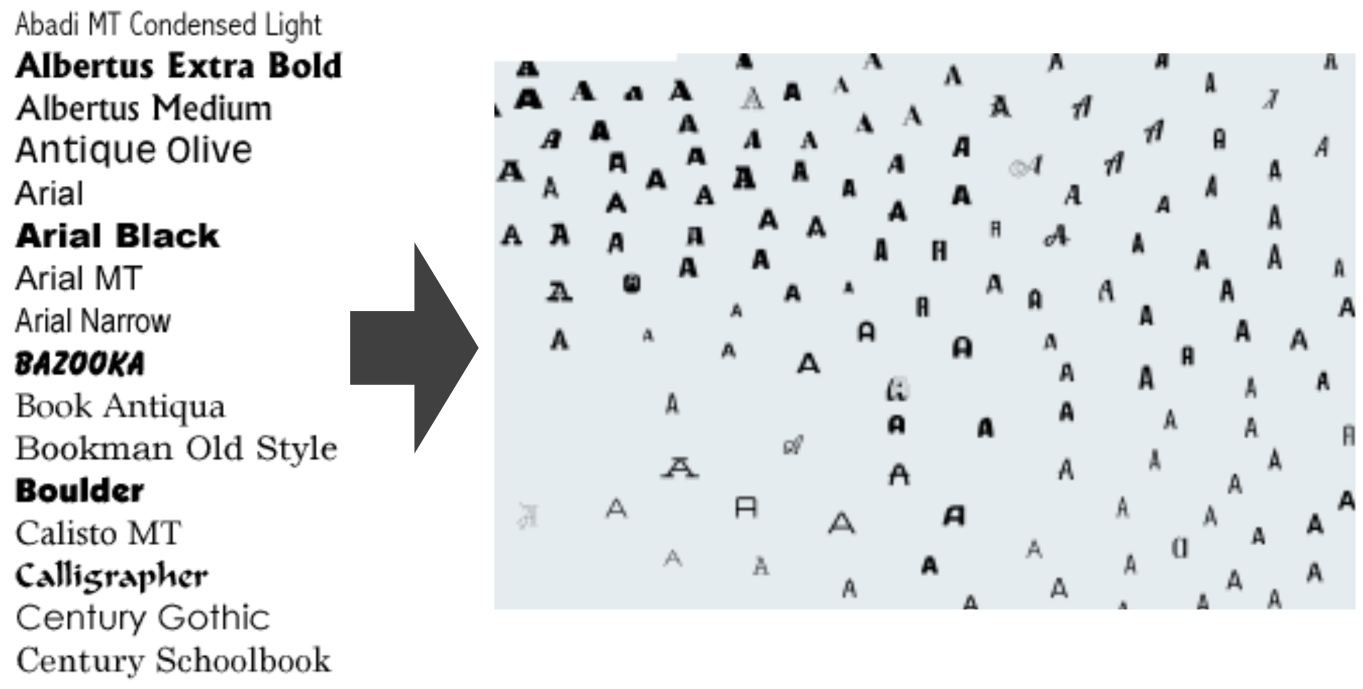
# Font Representation Learning via Paired-glyph Matching

Junho Cho[1,2], Kyuewang Lee[1], Jin Young Choi[1]
Department of ECE, ASRI, Seoul National University[1]
Samsung Advanced Institute of Technology, Samsung Electronics[2]

BMVC 2022

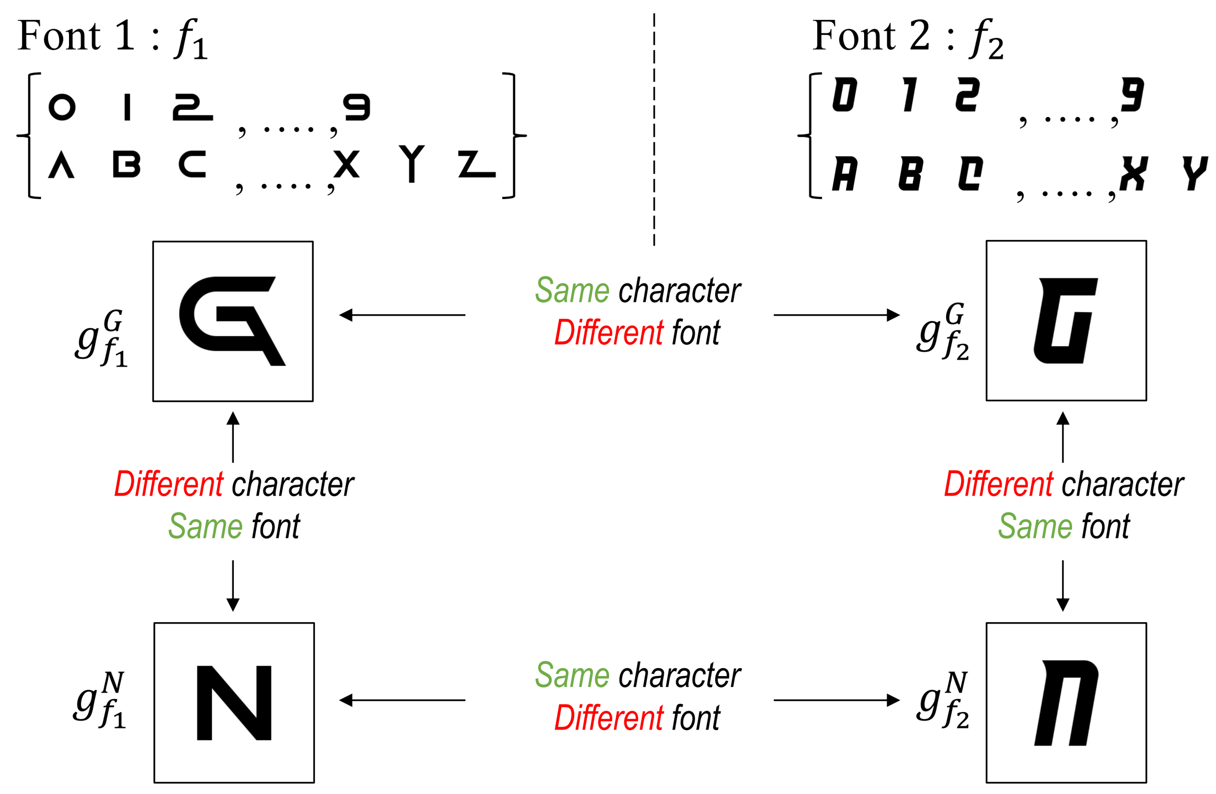
## Introduction



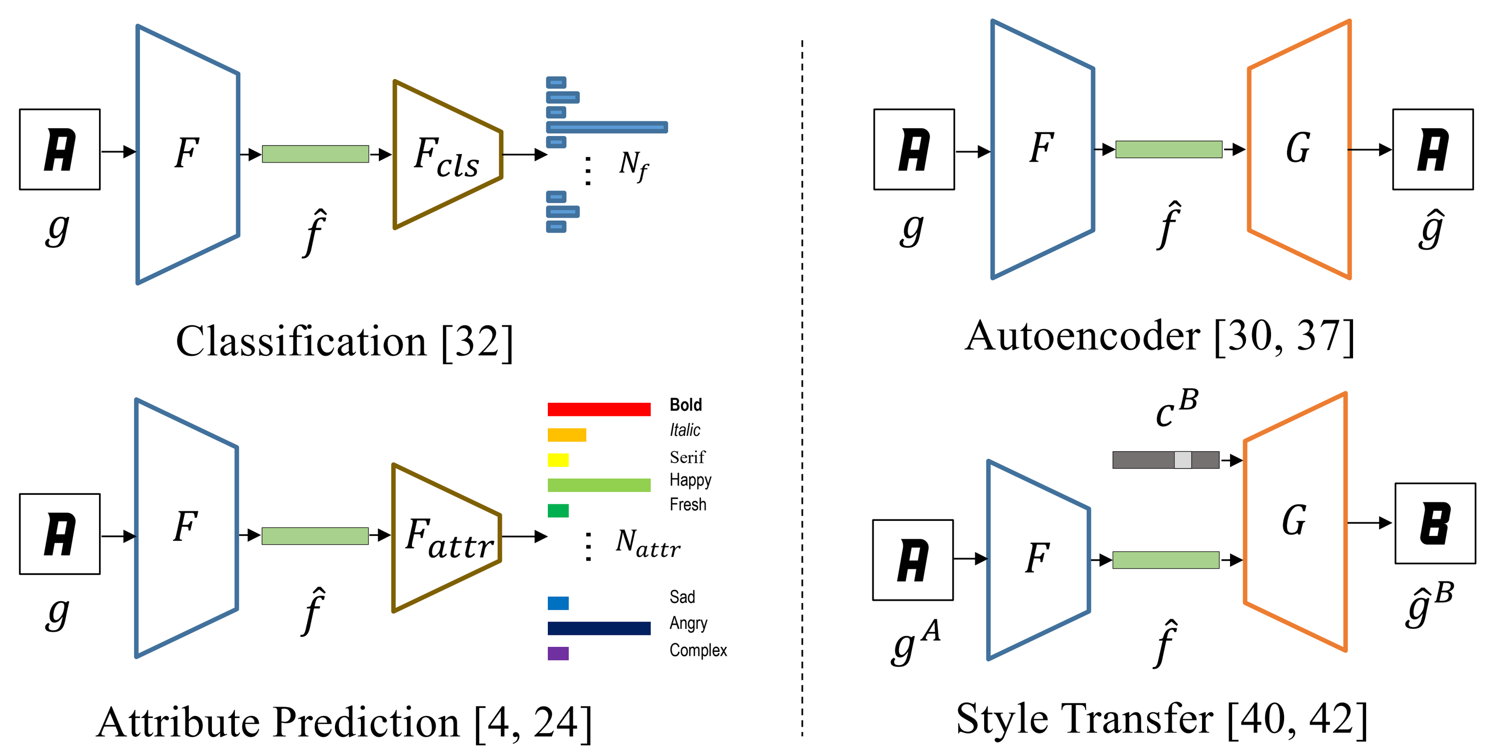

We aim to learn ***font representation*** to explore vast font styles.

## Preliminary

- A glyph ($g$) is an image form of a character that has a specific style in a font.
- A font ($f$) is a matching style set of glyphs.




## Previous Method




## Reference

[4] Tianlang Chen, Zhaowen Wang, N. Xu, Hailin Jin, and Jiebo Luo. Large-scale tag- based font retrieval with generative feature learning. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 9115–9124, 2019.

[5] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A sim- ple framework for contrastive learning of visual representations. *arXiv preprint arXiv:2002.05709*, 2020.

[24] Peter O'Donovan, Janis Libeks, Aseem Agarwala, and Aaron Hertzmann. Exploratory Font Selection Using Crowdsourced Attributes. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 33(4), 2014.

[28] Nikita Srivatsan, Jonathan Barron, Dan Klein, and Taylor Berg-Kirkpatrick. A deep factorization of style and structure in fonts. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 2195–2205, Hong Kong, China, November 2019.

[30] Shancheng Tang, Puyue Zhang, Xiongxiong Chen, Hanbo Wang, and Ming Chen. A word representation method based on glyph of chinese character. In *2020 International Conference on Intelligent Transportation, Big Data Smart City (ICITBS)*, pages 954– 957, 2020. doi: 10.1109/ICITBS49701.2020.00212.

[32] Chris Tensmeyer, Daniel Saunders, and Tony Martinez. Convolutional neural net- works for font classification. In *2017 14th IAPR International Conference on Doc- ument Analysis and Recognition (ICDAR)*, volume 01, pages 985–990, 2017. doi: 10.1109/ICDAR.2017.164.

[37] Chen Wang, Yani Zhu, Zhangyi Shen, Dong Wang, Guohua Wu, and Ye Yao. Font transfer based on parallel auto-encoder for glyph perturbation via strokes moving. In Yongxuan Lai, Tian Wang, Min Jiang, Guangquan Xu, Wei Liang, and Aniello Cas- tiglione, editors, *Algorithms and Architectures for Parallel Processing*, pages 586–602, Cham, 2022. Springer International Publishing. ISBN 978-3-030-95388-1.

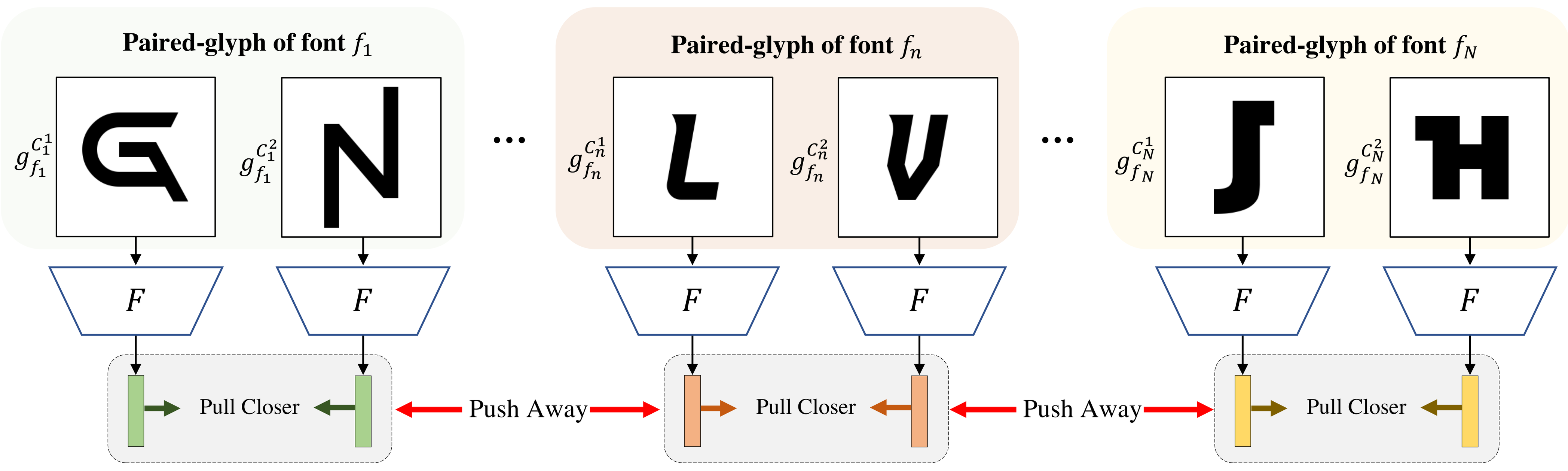
[40] Yangchen Xie, Xinyuan Chen, Li Sun, and Yue Lu. Dg-font: Deformable generative networks for unsupervised font generation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2021.

[42] Yexun Zhang, Ya Zhang, and Wenbin Cai. Separating style and content for generalized style transfer. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, 2018.

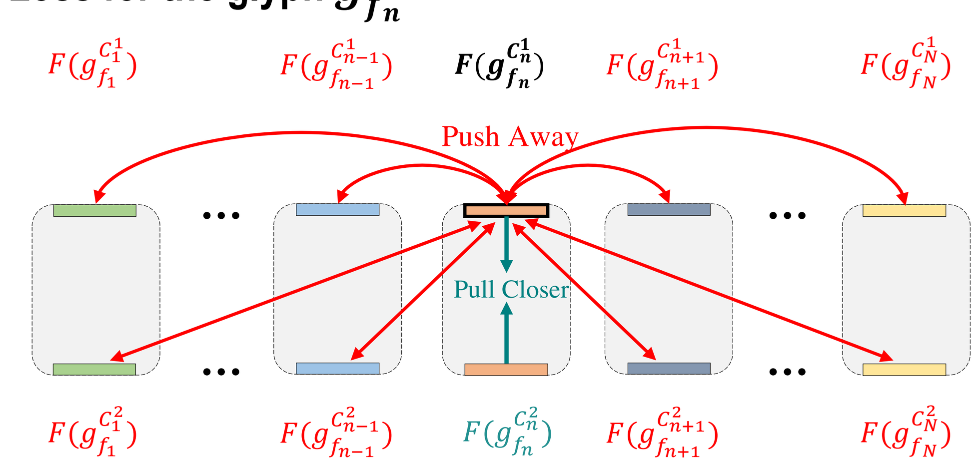
| Methods | Captials64 valset $\text{MACC}_{Ret}(\mathbb{C}_{\text{A-Z}})$ | Captials64 testset $\text{MACC}_{Ret}(\mathbb{C}_{\text{A-Z}})$ | O'Donovan $L_1$-error |
|---|---|---|---|
| $\hat{f}$ of Paired-$g$ Matching | **61.38** | **62.66** | **0.09589** |
| $\hat{f}$ of Classification [32] | 55.27 | 56.31 | 0.1275 |
| $\hat{f}$ of Style Transfer [40, 42] | 32.22 | 32.53 | 0.1217 |
| $\hat{f}$ of Autoencoder [30, 37] | 13.60 | 14.16 | 0.1312 |
| $\hat{f}$ of Srivatsan *et al.* [28] | 11.72 | 11.56 | 0.1097 |

Performance evaluation (MACC and Font attribute prediction) on the Capitals64 dataset. Font attribute prediction is evaluated on O'Donovan dataset with $L1$ error.

## Proposed Method



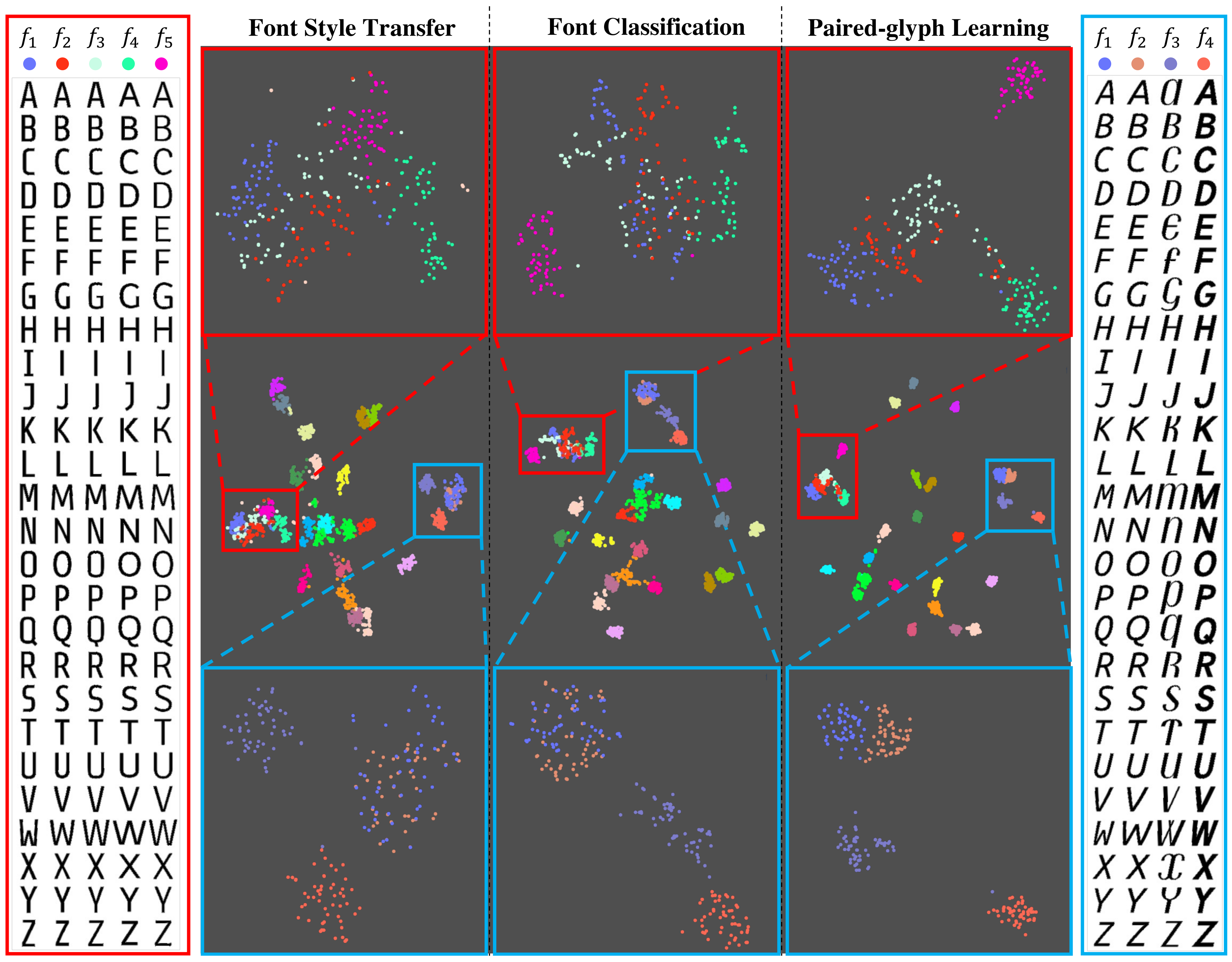

Loss for the glyph $g_{f_n}^{c_n^1}$

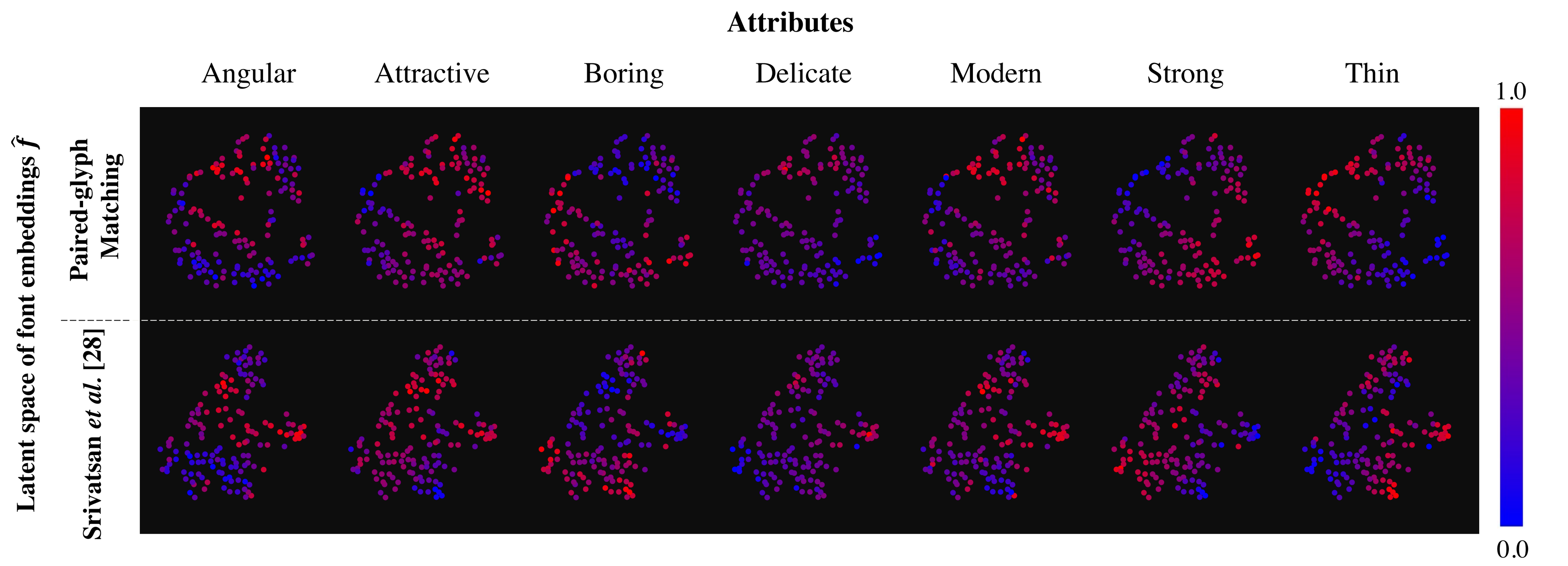



$$\ell_{n_1} = -\log \frac{\texttt{pos-sim}(f_n)}{\texttt{pos-sim}(f_n) + \sum_{m=1,m\neq n}^{N} \texttt{neg-sim}_1(f_n, f_{m\neq n})}$$
$$= \log\left(1 + \frac{\sum_{m=1,m\neq n}^{N} \texttt{neg-sim}_1(f_n, f_{m\neq n})}{\texttt{pos-sim}(f_n)}\right).$$

The loss is derived from "the normalized temperature-scaled cross entropy loss" [5].

## Result



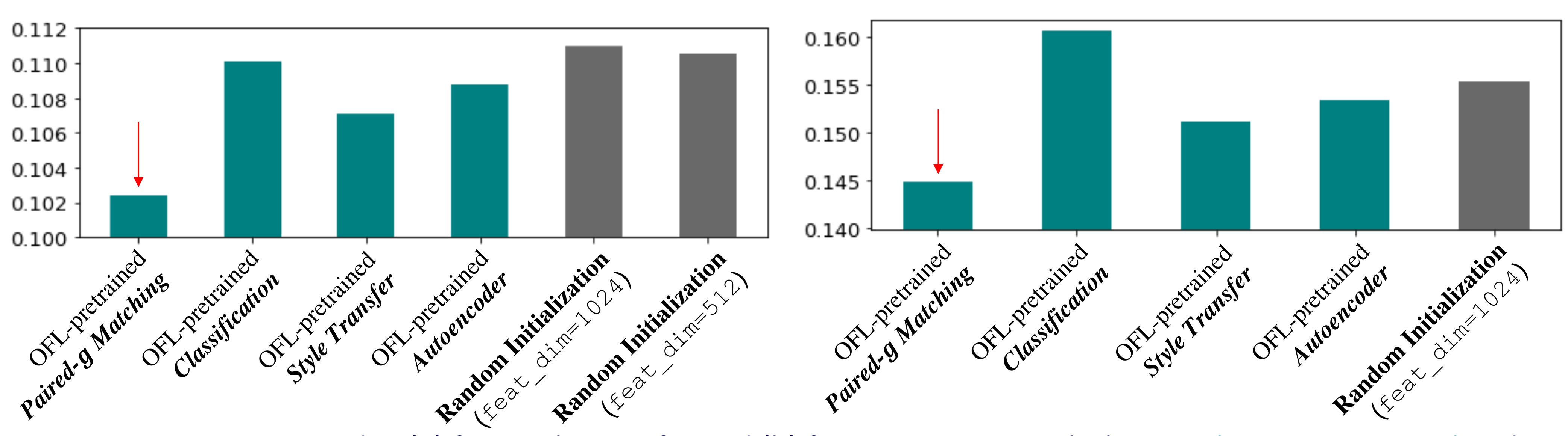




Image L1-error measured in (a) font style transfer and (b) font generation with the OFL dataset pretrained and random initialization.