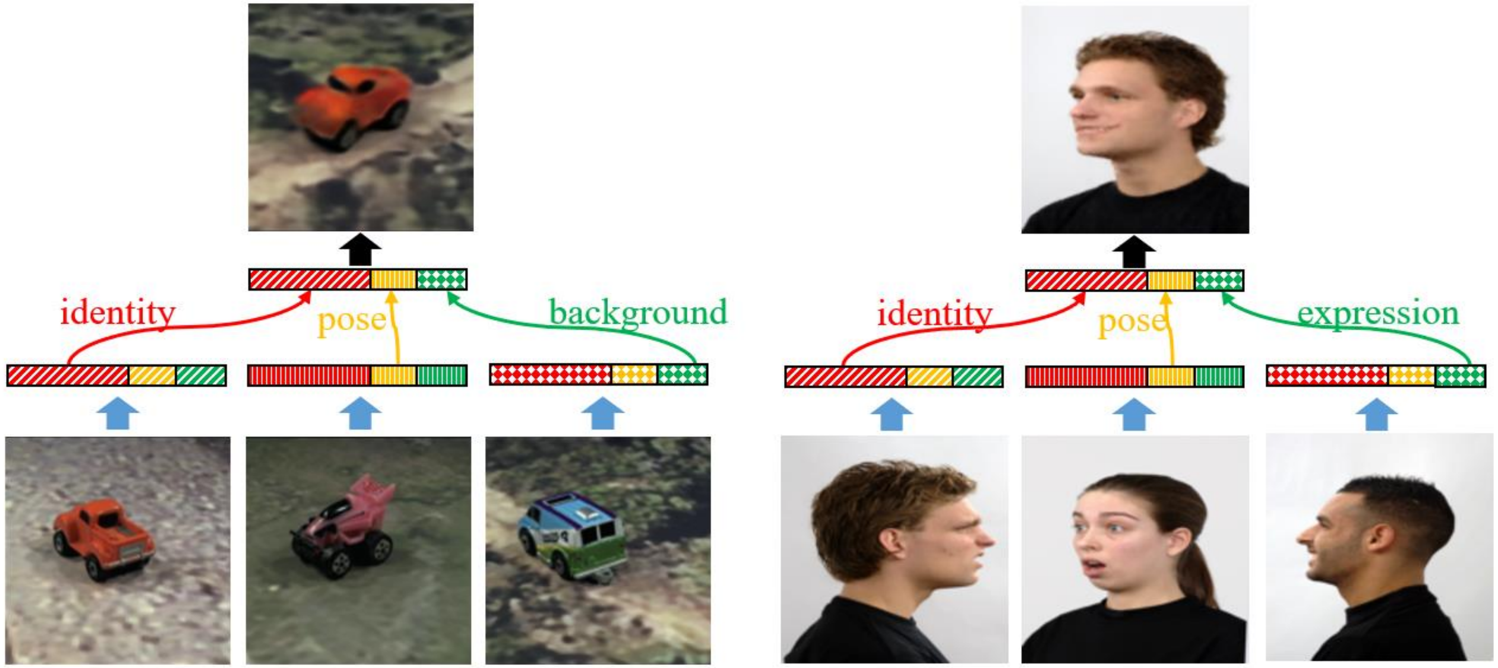


Zero-shot Synthesis with Group-Supervised Learning

ICLR 2021

Yunhao Ge, Sami Abu-El-Haija, Gan Xin, and Laurent Itti

USC University of Southern California




ICLR

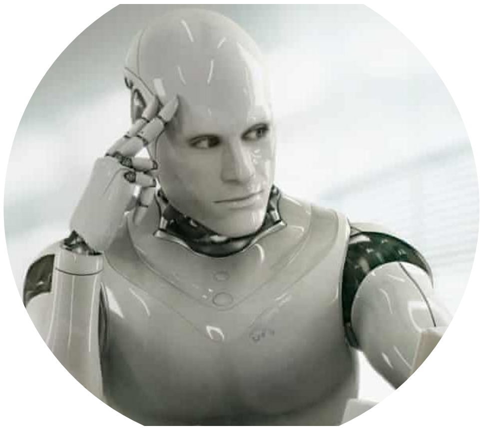
Agenda

- Motivation
    - Visual Cognition: Human -->Machines
- Problem Statement and Approach
- Experiments
    - Qualitative results
    - Quantitative results

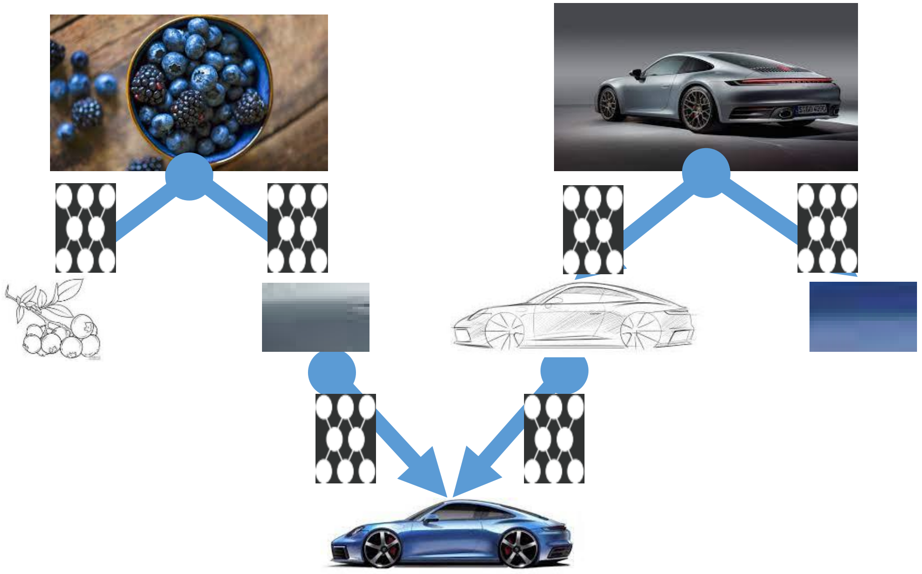
"Envision" a *novel* visual object

Group-Supervised Learning

Zero-shot synthesis

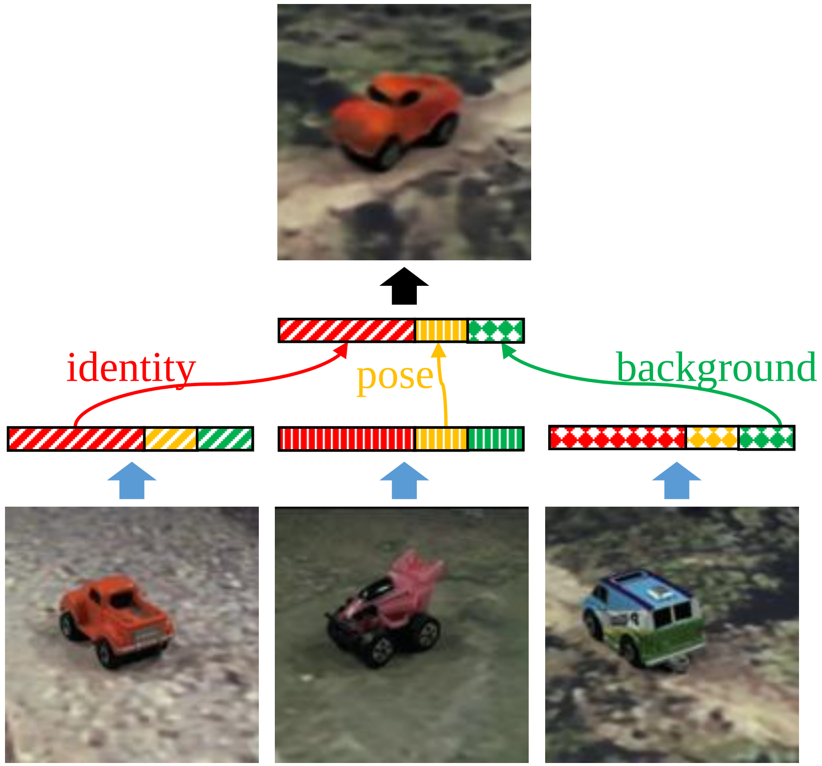


Knowledge Factorization[1]



Controllable
Disentangled
Representation Learning

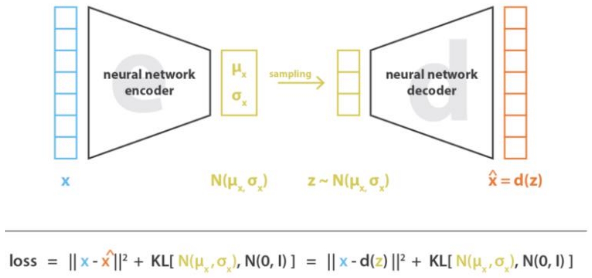
[1] Logothetis et al., 1995.



Group-Supervised learning (GSL) allows us to **decompose** inputs into a **disentangled representation** with **swappable** components, that can be recombined to synthesize new samples.

Controllable Disentangled Representation Learning

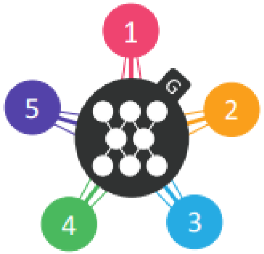
Unsupervised

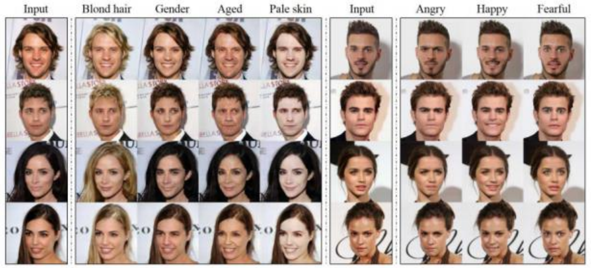


VAE DP Kingma, et al. 2016

☹ Controllable

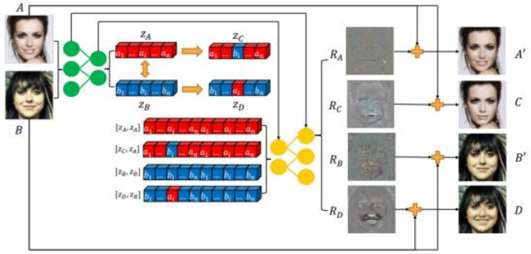
Supervised

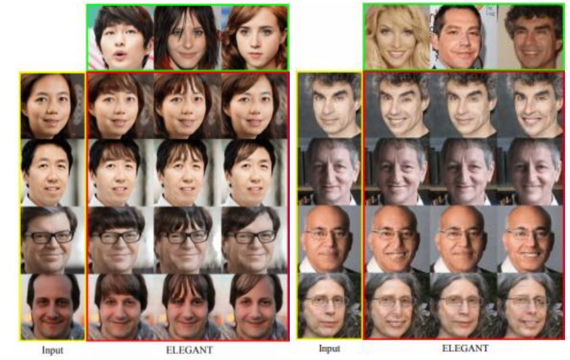




StarGAN Choi, Yunjey, et al. 2018

Supervised

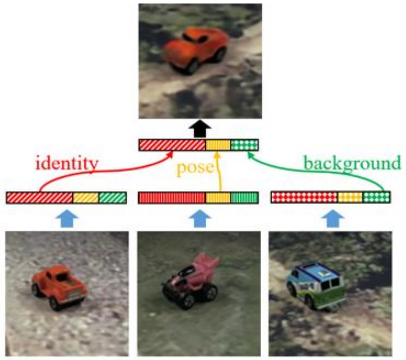




ELEGANT Xiao, T. et al. 2018

😀 Controllable
☹ Global semantical consistency
☹ Easy implement and training

Supervised



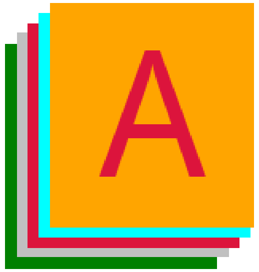
Group-Supervised Learning (Ours)

😀 Controllable
😀 Global semantical consistency
😀 Easy implement and training

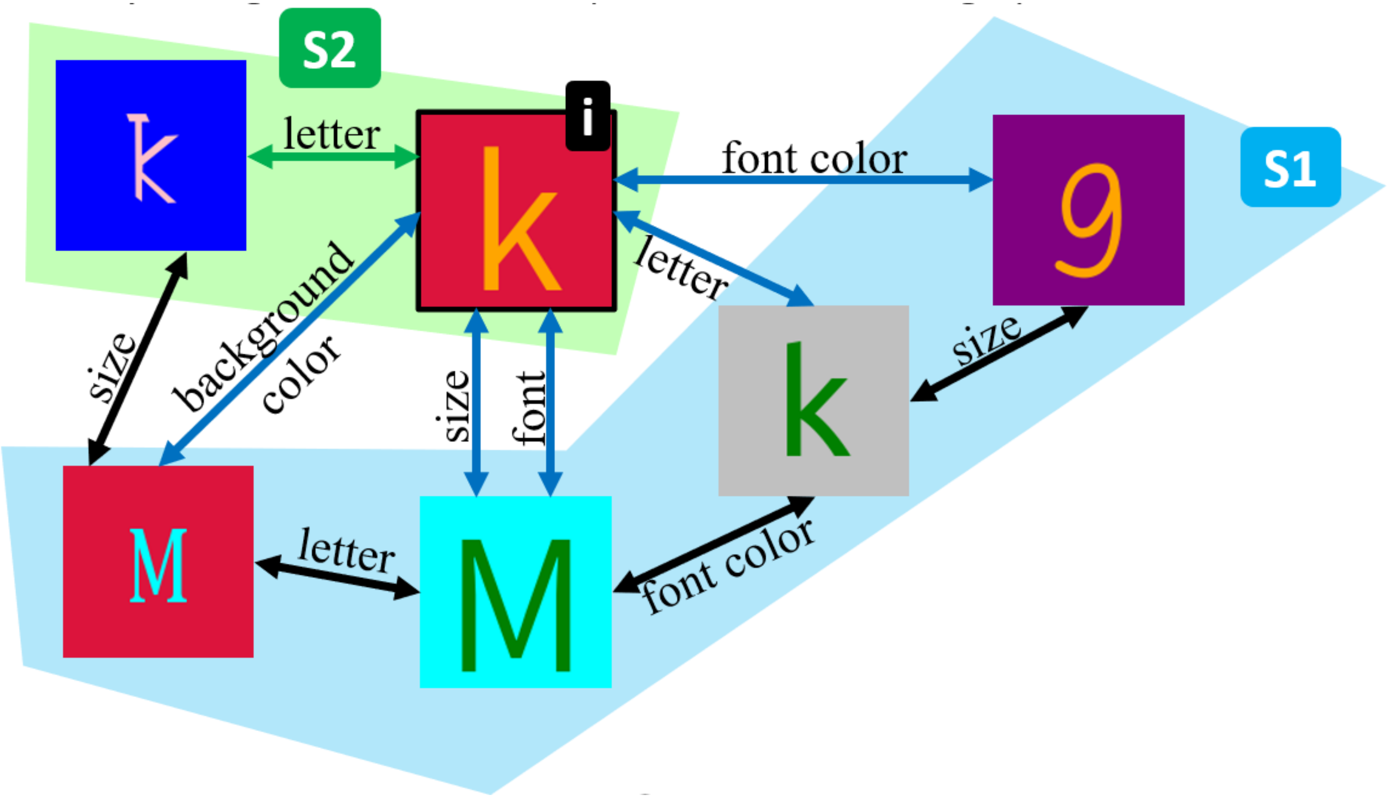
Agenda

- Motivation
  - Visual Cognition: Human -->Machines
- Problem Statement and Approach
- Experiments
  - Qualitative results
  - Quantitative results

Given: Dataset --> Multi-graph

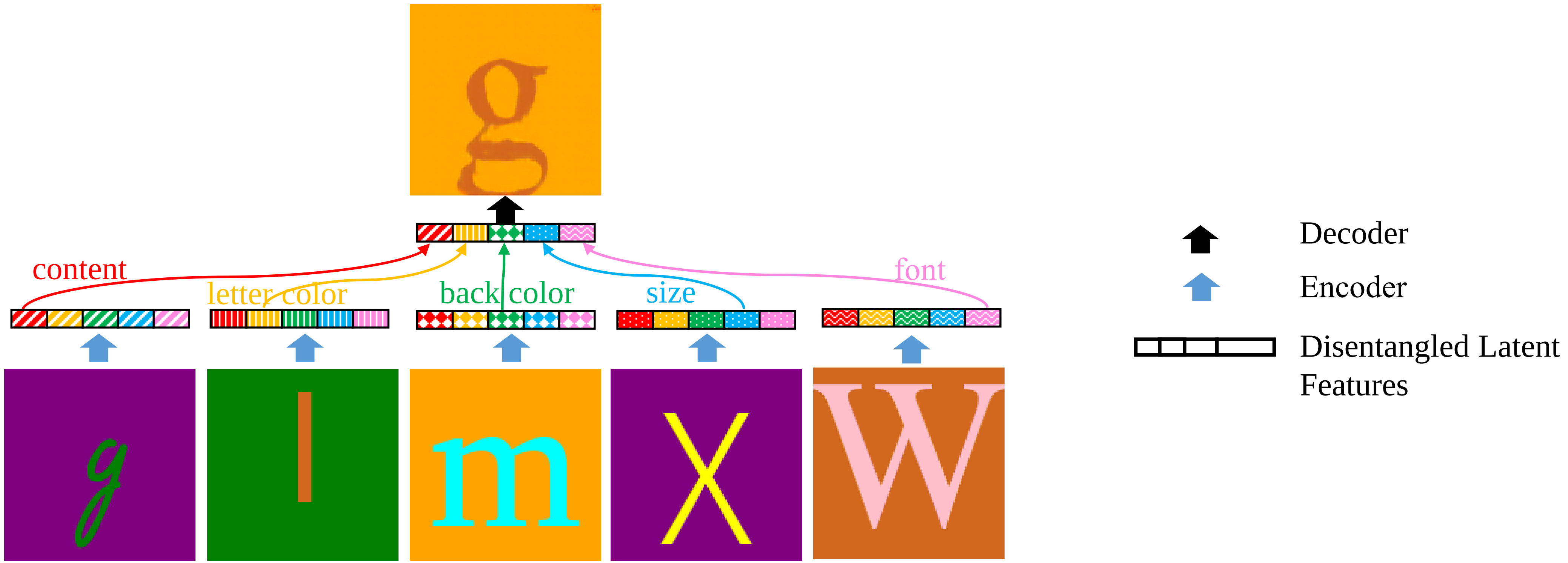


Letter : A
Size : large
Font color : red
Background color : orange
Font : Abadi



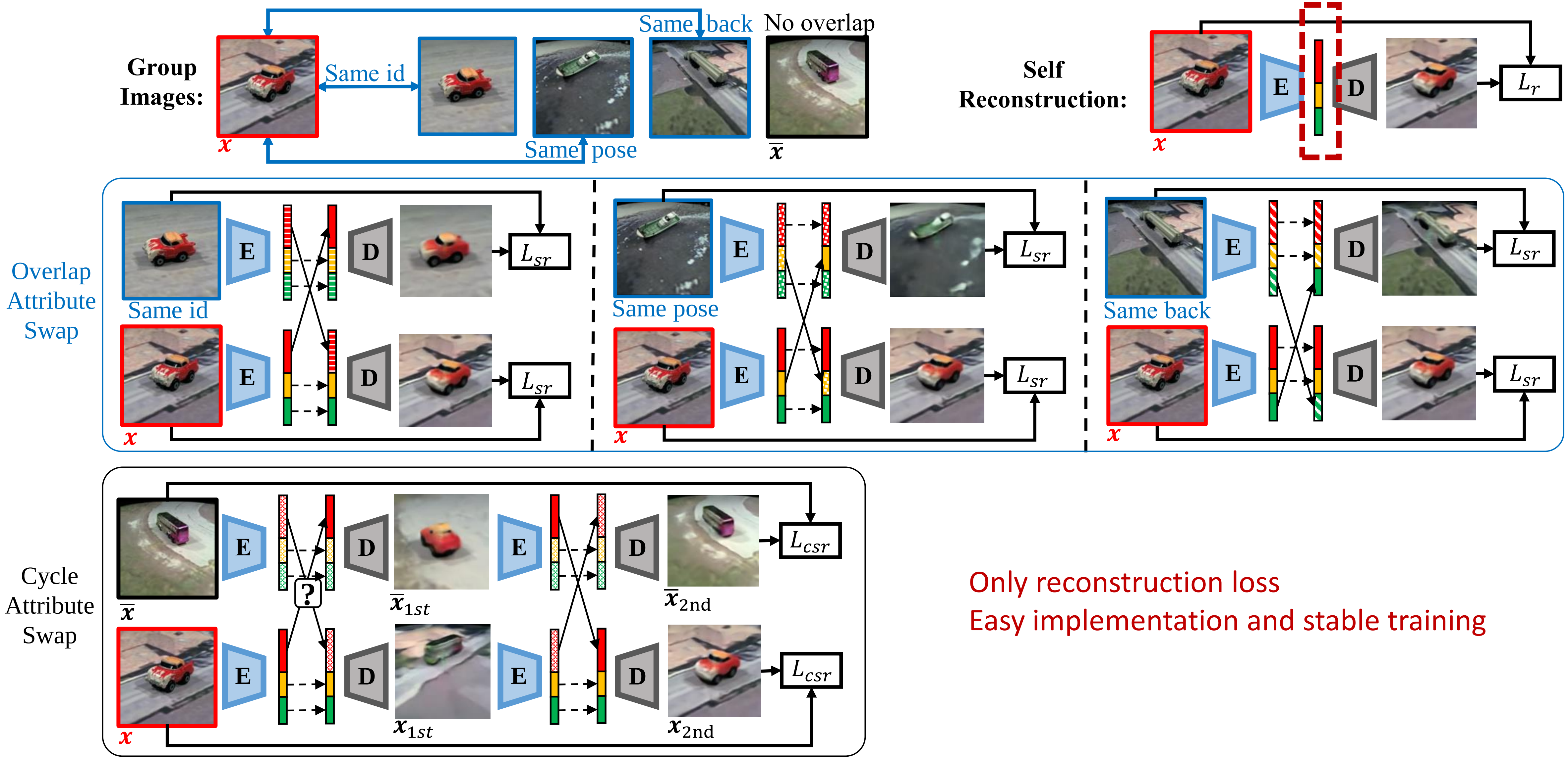
Fonts: http://ilab.usc.edu/datasets/fonts

Goal: Controllable Synthesis by Disentangle Representation Learning



Group-Supervised Learning

Controllable Disentanglement: (1) Predefine partition (2) Mine the similarity by attribute swap.



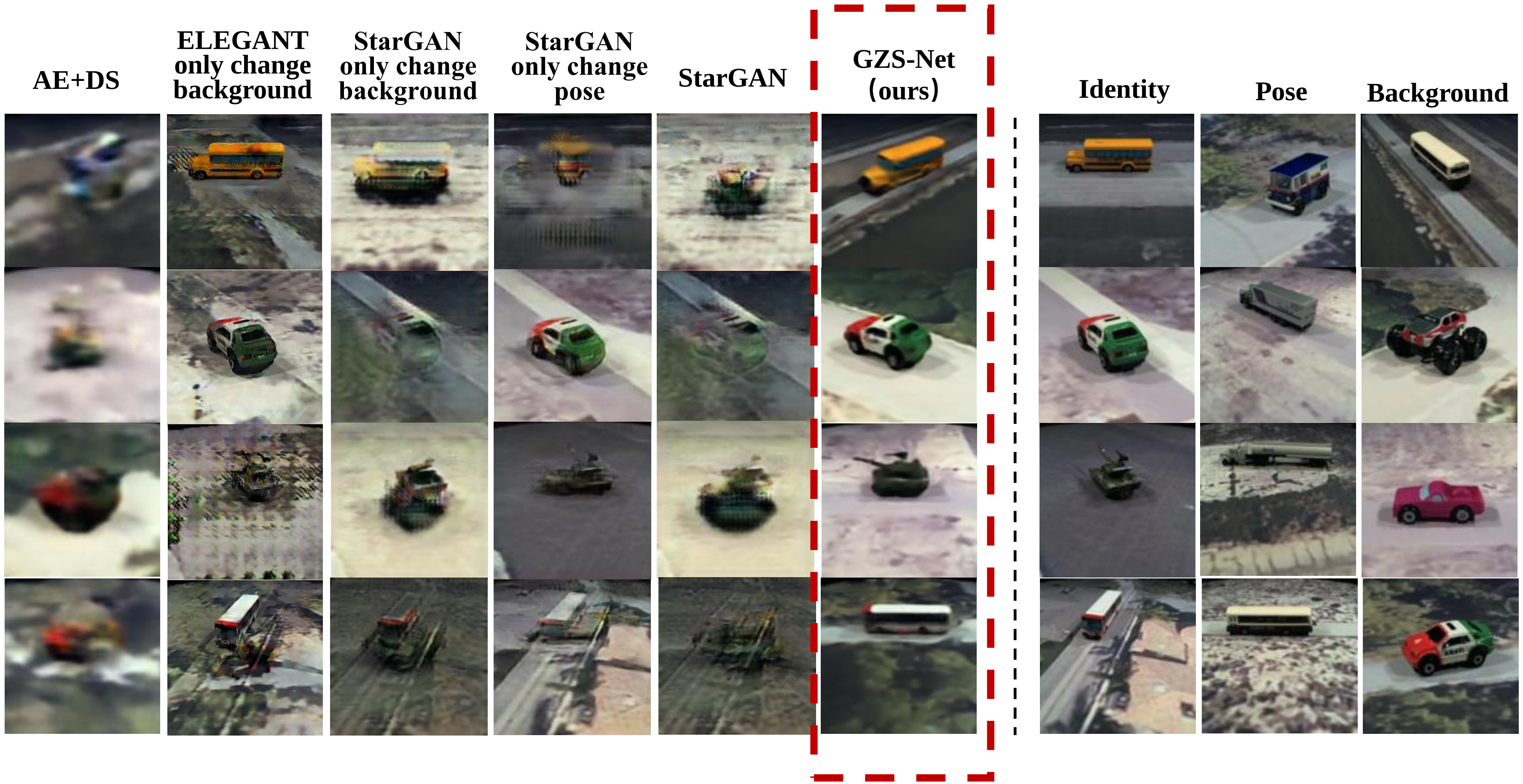
Only reconstruction loss

Easy implementation and stable training

Agenda

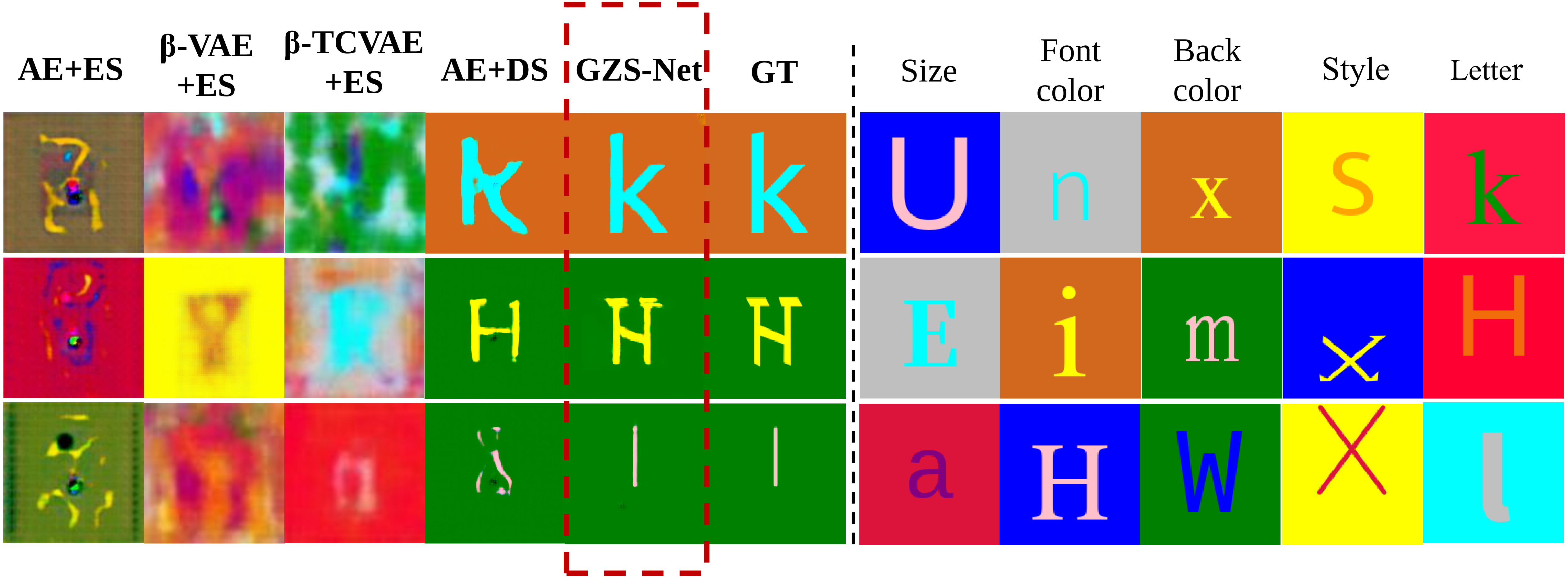
- Motivation
  - Visual Cognition: Human -->Machines
- Problem Statement and Approach
- Experiments
  - Qualitative results
  - Quantitative results

Qualitative Results 1 --- iLab-20M [1]



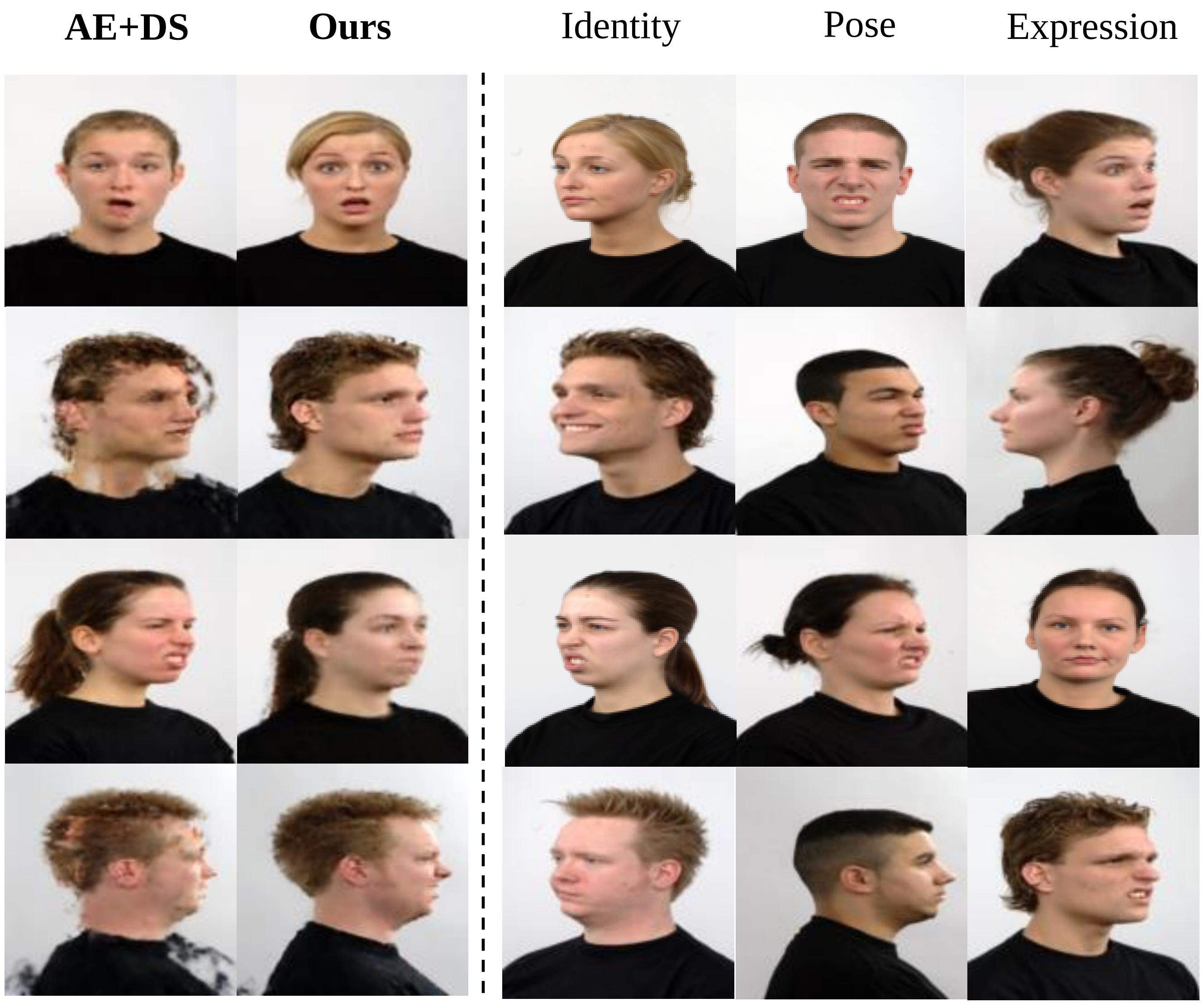
[1] ilab-20m: A large-scale controlled object dataset to investigate deep learning. Borji, A., Izadi, S. and Itti, L., CVPR 2016

Qualitative Results 2 --- Fonts



Fonts: http://ilab.usc.edu/datasets/fonts

Qualitative Results 3 --- RaFD [1]



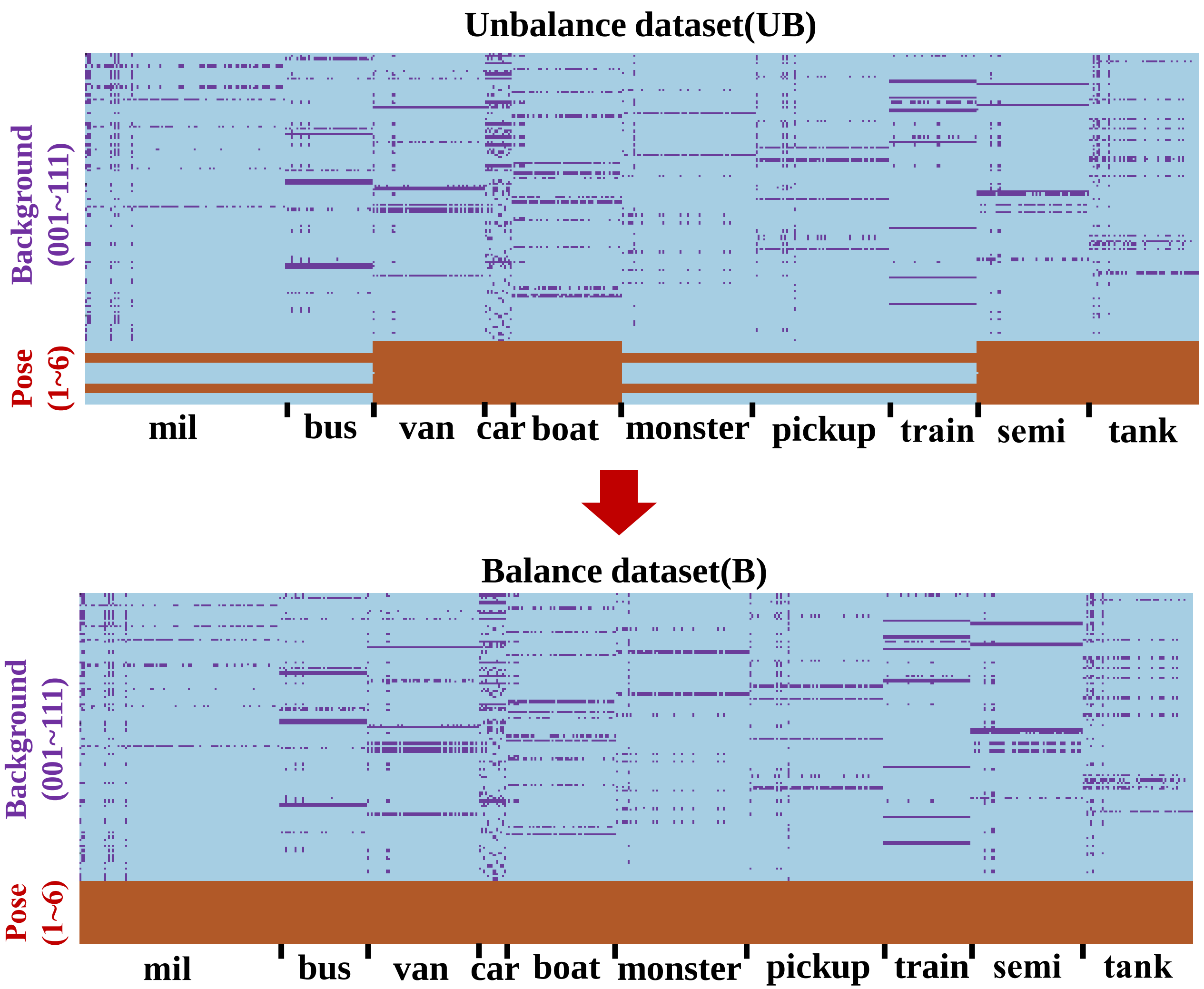
[1] Langner, Oliver, et al. "Presentation and validation of the Radboud Faces Database." Cognition and emotion 24.8 (2010): 1377-1388.

Quantitative Results 1 --- Disentanglement analysis

Table 1: Disentangled representation analysis. Diagonals are bolded.

	GZS-Net					Auto-encoder					AE + DS					β-VAE + ES					β-TCVAE + ES				
$\mathcal{A}$ ($\lvert\mathcal{A}\rvert$)	**C**	**S**	**FC**	**BC**	**St**	**C**	**S**	**FC**	**BC**	**St**	**C**	**S**	**FC**	**BC**	**St**	**C**	**S**	**FC**	**BC**	**St**	**C**	**S**	**FC**	**BC**	**St**
Content (52)	**.99**	.92	.11	.13	.30	**.48**	.60	.71	.92	.06	**.99**	.72	.22	.20	.25	**.02**	.35	.11	.19	.01	**.1**	.39	.13	.11	.01
Size (3)	.78	**1.0**	.11	.15	.36	.45	**.61**	.77	.96	.07	.54	**1.0**	.19	.23	.25	.02	**.38**	.29	.11	.01	.02	**.47**	.18	.19	.01
FontColor (10)	.70	.88	**1.0**	.16	.23	.48	.60	**.67**	.95	.06	.19	.64	**1.0**	.66	.20	.02	.33	**.42**	.11	.01	.02	.35	**.21**	.13	.01
BackColor (10)	.53	.78	.21	**1.0**	.15	.53	.63	.64	**.93**	.08	.32	.65	.29	**1.0**	.25	.02	.34	.11	**.86**	.01	.03	.40	.24	**.75**	.01
Style (100)	.70	.93	.12	.12	**.63**	.49	.60	.70	.94	**.06**	.38	.29	.20	.20	**.65**	.02	.33	.10	.11	**.02**	.02	.33	.10	.08	**.01**

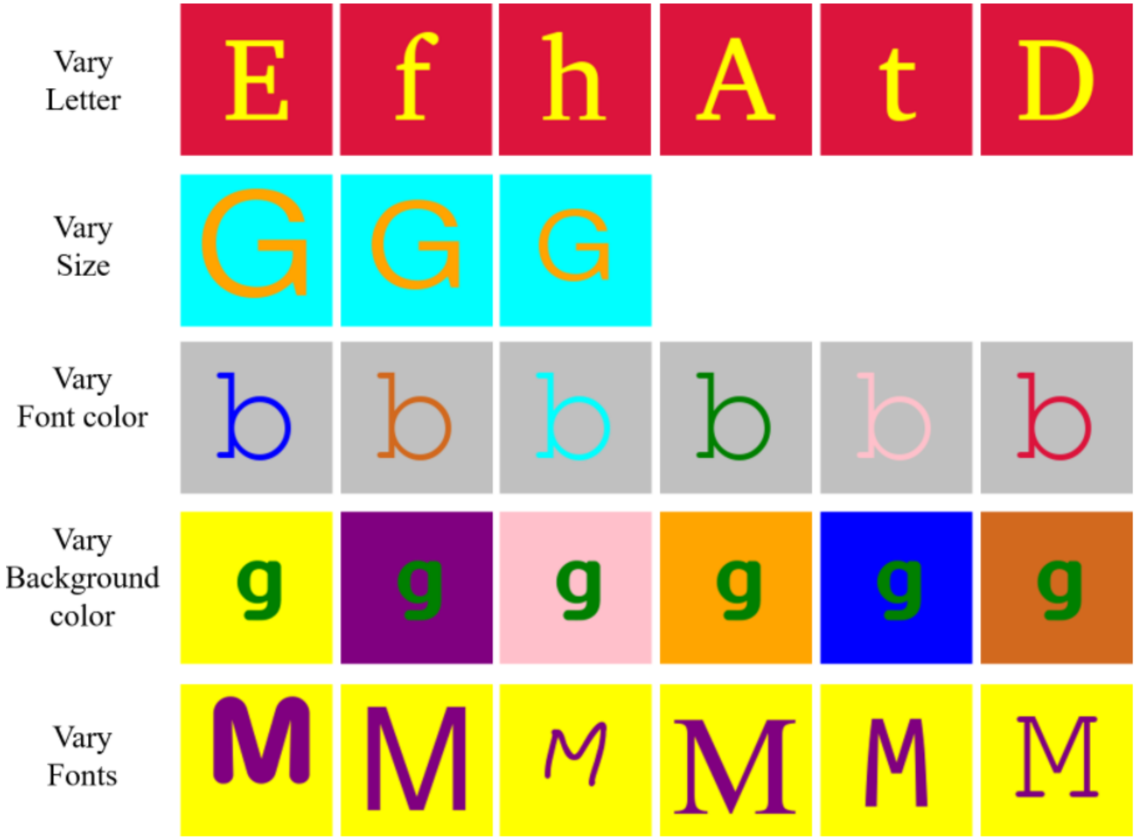
Quantitative Results 2 --- Data Augmentation

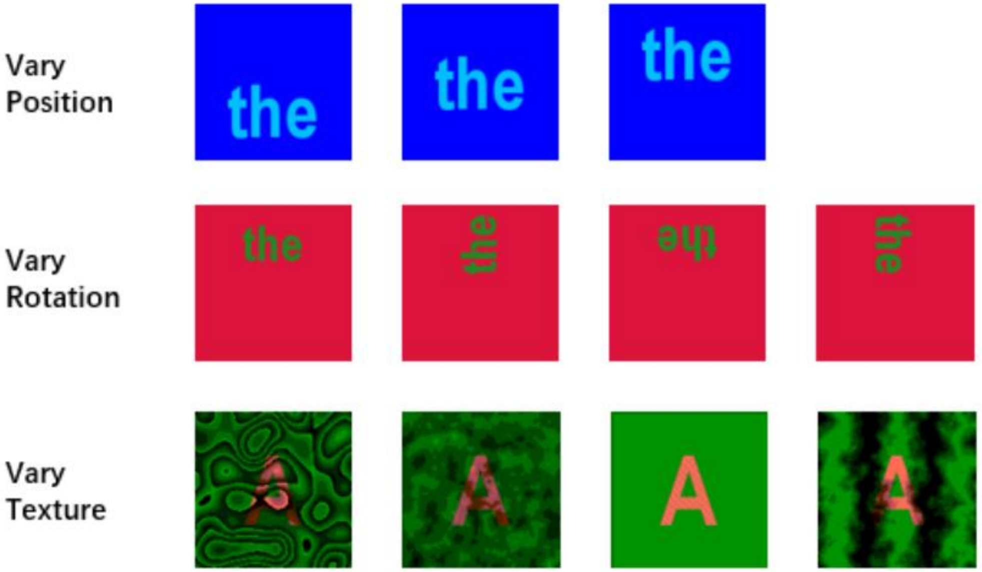


Dataset	UB	B	S-B	A-UB	Test
Source	real	real	GZS-Net synthesized	Traditional augmented	real
Numbe	25149	37417	37395	37395	37469
Overall Accuracy on Test	56.5	64.4	**63.5**	56.6	

Font dataset

Primary motivation : allows fast testing and idea iteration, on disentangled representation learning and zero-shot synthesis.





Fonts: http://ilab.usc.edu/datasets/fonts

Group-Supervise Learning

Paper: https://arxiv.org/pdf/2009.06586.pdf

Code: **https://github.com/gyhandy/Group-Supervised-Learning**

Website: