DisCont:
Self-Supervised Visual Attribute Disentanglement using Context Vectors

Sarthak Bhagat[1] , Vishaal Udandarao[1] , Shagun Uppal[1]

[1] IIIT-Delhi, Equal Contribution

Disentangling

► Disentangling visual attributes leads to task-agnostic representations which can enhance efficiency and performance of deep models.

► 'Impossibility result': unsupervised disentanglement is impossible without explicit inductive biases on the models and data used.

Contrastive Learning Paradigms

► Exploit the spatial and structural inductive biases prevalent in visual datasets for unsupervised disentanglement.

► E.g. data with color and position attributes

color transformation -> position remains same and vice versa.

Contributions

► Propose a self-supervised method *DisCont* to simultaneously disentangle multiple underlying visual attributes by effectively introducing inductive biases in images via data augmentations.

► Highlight the utility of leveraging composite stochastic transformations for learning richer disentangled representations.

► Present the idea of 'Attribute Context Vectors' to capture and utilize intra-attribute variations in an extensive manner.

► Impose an attribute clustering objective that is commonly used in distance metric learning literature, and show that it further promotes attribute disentanglement.

Algorithm 1 Mask and Augmented Batch generation

Input: A batch of images x, the set of positive transformations $\mathcal{T}_p$, the set of negative transformations $\mathcal{T}_n$, number of feature attributes k

Output: The augmented batch x_{aug}, the mask m

Initialize $m = [0, 0, ..., 0]_k$, $x_{aug} = x$

for $i = 1$ **to** k **do**
    $p \sim Bernoulli(0.5)$
    if $p = 1$ **then**
        $m[i] = 1$
        $x_{aug} = n_i(x_{aug})$
    end if
end for

for $i = 1$ **to** k **do**
    $p \sim Bernoulli(0.5)$
    if $p = 1$ **then**
        $x_{aug} = p_i(x_{aug})$
    end if
end for
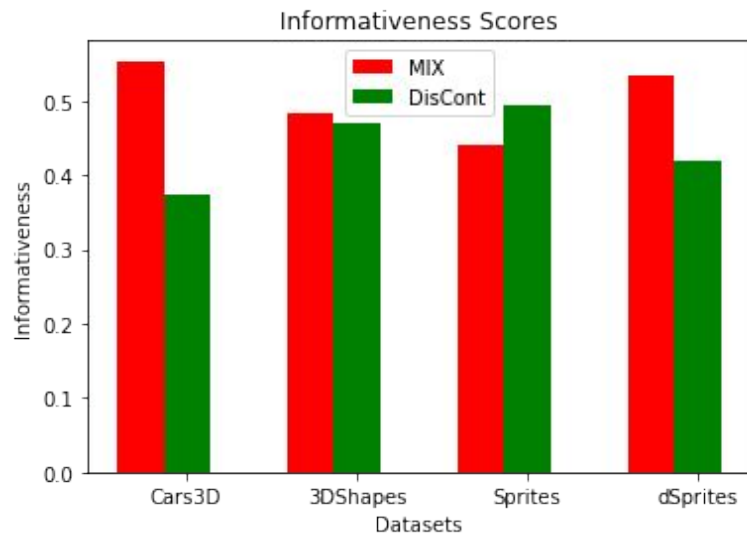
return x_{aug}, m
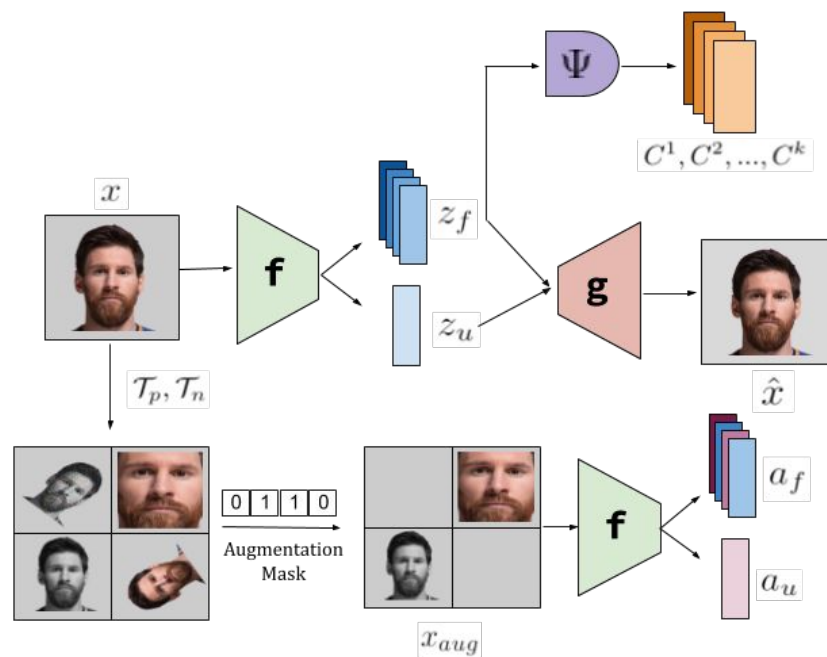
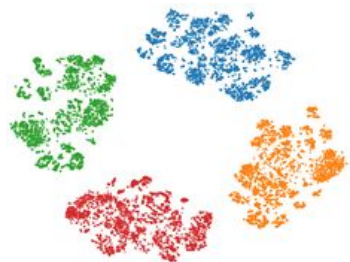
Figure 1: *DisCont* Training

Informativeness Score

$$\mathbb{I}(x, z_{f,i}) = \int_x \int_{z_f} p_{\mathcal{D}}(x) f(z_{f,i}|x) \log \left(\frac{f(z_{f,i}|x)}{f(z_{f,i})} \right) dx \, dz_f$$

Informativeness Scores

Latent Visualization
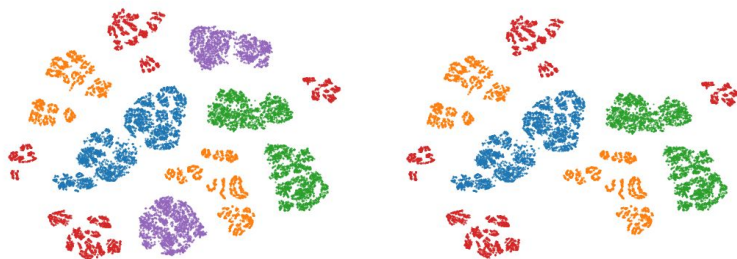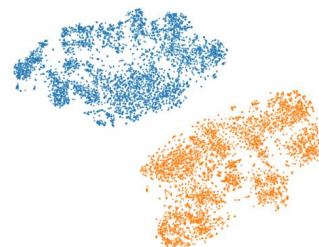

dSprites

Cars3D

3DShapes

Attribute Transfer
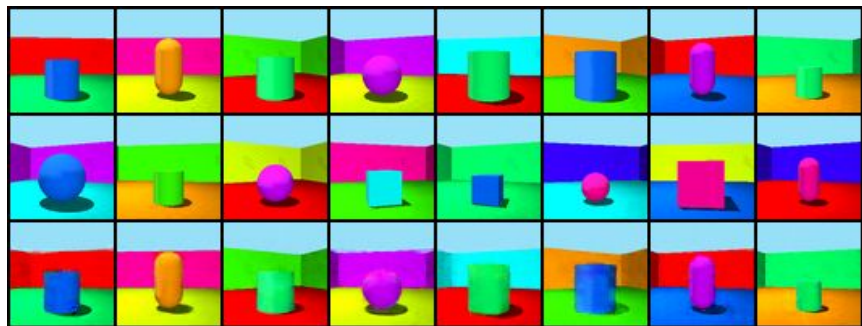
3DShapes, Orientation
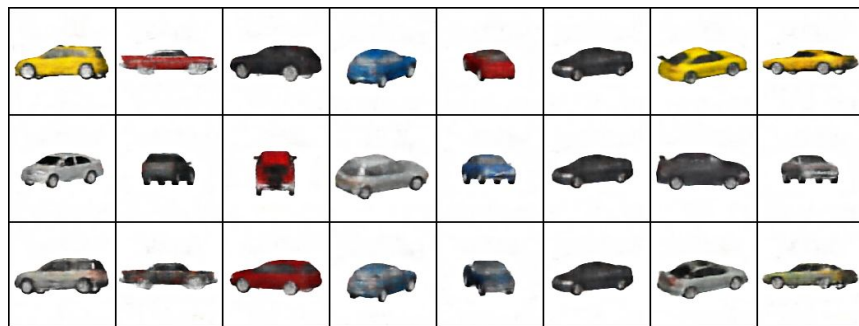
Cars3D, Color

Sprites, Hair Color

For more details, please check our paper and code:
DisCont: Self-Supervised Visual Attribute Disentanglement using Context Vectors
https://github.com/sarthak268/DisCont