Fine-Grained Visual Comparisons

Densifying Supervision
Generate synthetic images exhibiting subtle differences
- "fill in" the sparsely sampled regions to enhance fine-grained supervision
- pre-trained Attribute2Image [Tan et al. ’17] image generation engine
- attribute-conditioned generation of synthetic identities \( i_j = (\gamma_j, \epsilon_j) \)

Semantic “jittering” to augment real training images
- high-level changes that modify underlying meaning

Our Idea
Densify the attribute space using synthetic image pairs to improve supervision for fine-grained learning.

Experimental Results
Observation: The synthetic image pairs successfully densify the supervision. Given a novel pair, the nearest neighbors consist of both real and synthetic pairs, suggesting their combined importance.