Supplementary Material for Paper ID 2153

1. Visualization of synthetic images

Figures 1, 2, 3, 4, 5, and 6 show visualization of synthetic images generated through SRe^2L and SRe^2L + DiRe methods. The diversity introduced by addition of DiRe is clearly visible.

Figure 1. Visualization of synthetic images belonging to the ‘Great Dane’ class from the ImageNet-1K dataset. The top row images are condensed through SRe^2L and the bottom row images are condensed through SRe^2L + DiRe.

Figure 2. Visualization of synthetic images belonging to the ‘Fly’ class from the ImageNet-1K dataset. The top row images are condensed through SRe^2L and the bottom row images are condensed through SRe^2L + DiRe.

Figure 3. Visualization of synthetic images belonging to the ‘Ringlet Butterfly’ class from the ImageNet-1K dataset. The top row images are condensed through SRe^2L and the bottom row images are condensed through SRe^2L + DiRe.

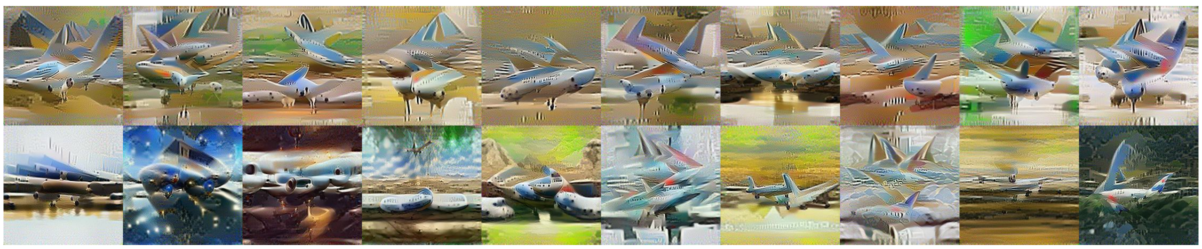

Figure 4. Visualization of synthetic images belonging to the ‘Airliner’ class from the ImageNet-1K dataset. The top row images are condensed through SRe^2L and the bottom row images are condensed through SRe^2L + DiRe.

Figure 5. Visualization of synthetic images belonging to the ‘church’ class from the ImageNet-1K dataset. The top row images are condensed through SRe^2L and the bottom row images are condensed through SRe^2L + DiRe.

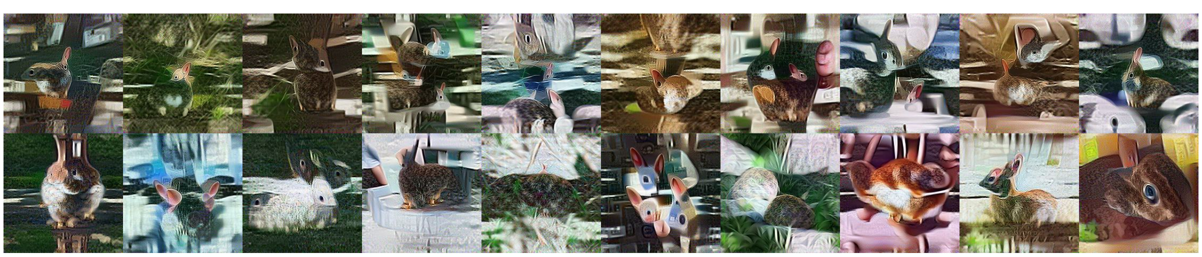

Figure 6. Visualization of synthetic images belonging to the ‘Rabbit’ class from the ImageNet-1K dataset. The top row images are condensed through SRe^2L, and the bottom row images are condensed through SRe^2L + DiRe.

Figure 7. Visualization of synthetic images belonging to the 'Dining Table' class from the ImageNet-1K dataset. The top row images are condensed through SRe^{2}L and the bottom row images are condensed through SRe^{2}L + DiRe.

2. Intuitive motivation for DiRe

In this section, we discuss some intuitive motivation for why DiRe, comprising of cosine diversity, cosine distribution matching, and Euclidean distribution matching leads to improved diversity in DC.

2.1. Notation

Let:
- $\mathcal{D}_{\text{real}} = \{x_i, y_i\}_{i=1}^{N}$: original dataset
- $\mathcal{D}_{\text{syn}} = \{\tilde{x}_j, \tilde{y}_j\}_{j=1}^{M}$: synthetic dataset, with $M \ll N$
- $h_\theta(\cdot)$: pretrained embedding function (e.g., ResNet avgpool output)
- $E_{\text{real}}, E_{\text{syn}}$: embedding matrices for real and synthetic data

Goal: Ensure $\mathcal{D}_{\text{syn}}$ generalizes similarly to $\mathcal{D}_{\text{real}}$ for training neural networks.

2.2. Cosine Diversity Loss

Let E_{syn}^c denote synthetic embeddings of class c.

Define the cosine diversity loss:

$$\mathcal{L}_{\text{cos-div}} = \sum_{i<j} \cos(E_i^c, E_j^c)$$

- The Cosine similarity of two orthogonal vectors is zero.
- Minimization of the sum of pairwise cosine similarities results in vectors being mutually orthogonal.
- Thus, it ensures vectors are geometrically spread out, resulting in higher diversity.

More formally, we can relate the minimization of pairwise cosine similarity to the maximization of the determinant of a corresponding Gram matrix. The determinant of the Gram matrix is related to the volume spanned by the vectors in the feature space. A larger determinant indicates that the vectors are more spread out, corresponding to higher diversity.

If the vectors are highly similar (high cosine similarities), they are nearly collinear and the Gram matrix becomes close to singular (low determinant). Conversely, if the vectors are orthogonal (cosine similarities close to 0), the Gram matrix approaches a diagonal matrix with a determinant of 1, indicating a maximal spread for the vectors.

2.3. Distribution Matching Losses

Define:

$$\mathcal{L}_{\text{cos-dm}} = \sum_{i,j} \left(1 - \cos(E_{\text{syn},i}^c, E_{\text{real},j}^c)\right)$$

$$\mathcal{L}_{\text{euc-dm}} = \sum_{i,j} \|E_{\text{syn},i}^c - E_{\text{real},j}^c\|_2$$

Minimizing both losses jointly aligns synthetic data with real data in:
- **Direction** (via cosine similarity)
- **Magnitude** (via Euclidean distance)

This encourages synthetic data to approximate the local geometry of the real data manifold.

2.4. Conclusion

Thus DiRe promotes diversity and alignment in embedding space. This:

- Maximizes class-wise dispersion.
- Aligns synthetic samples with the data manifold.

3. Results for Distribution Matching (DM) method

Comparison of test set accuracies on CIFAR-10 and CIFAR-100 with a ConvNet architecture for DM and DM + DiRe methods is given in Table 1. Addition of DiRe is able to enhance accuracy values obtained by DM for various IPC values considered.

Table 1. Comparison of accuracies on ConvNet architecture with DM Dataset Condensation method.

IPC	DM	DM + DiRe
	CIFAR-10	
10	48.9 ± 0.6	**51.6 ± 0.3**
50	63.0 ± 0.4	**64.5 ± 0.1**
	CIFAR-100	
10	29.7 ± 0.3	**31.8 ± 0.5**
50	43.6 ± 0.4	**44.9 ± 0.3**

4. Compute analysis

We carry out synthesis in a single 32 GB V100 GPU. The timing and memory requirement comparison for synthesizing IPC=10 images from ImageNet-1K dataset is provided in Table 2. It may be noted that, similar to SRe^2L, all the computation required to add DiRe can be parallelized across multiple GPUs, depending upon availability.

Table 2. Comparison of total time taken and GPU memory consumed for synthesizing IPC=10 images from ImageNet-1K dataset.

Method	Time (in Hrs.)	Memory (in GB)
SRe^2L	2.94	10.6
SRe^2L + DiRe	3.13	13.0

5. Intra-class cosine similarity for ILSVRC

Figure 8 shows that, for all classes in ImageNet-1K, DiRe lowers intra-class cosine similarity, which is an indicator of improving diversity.

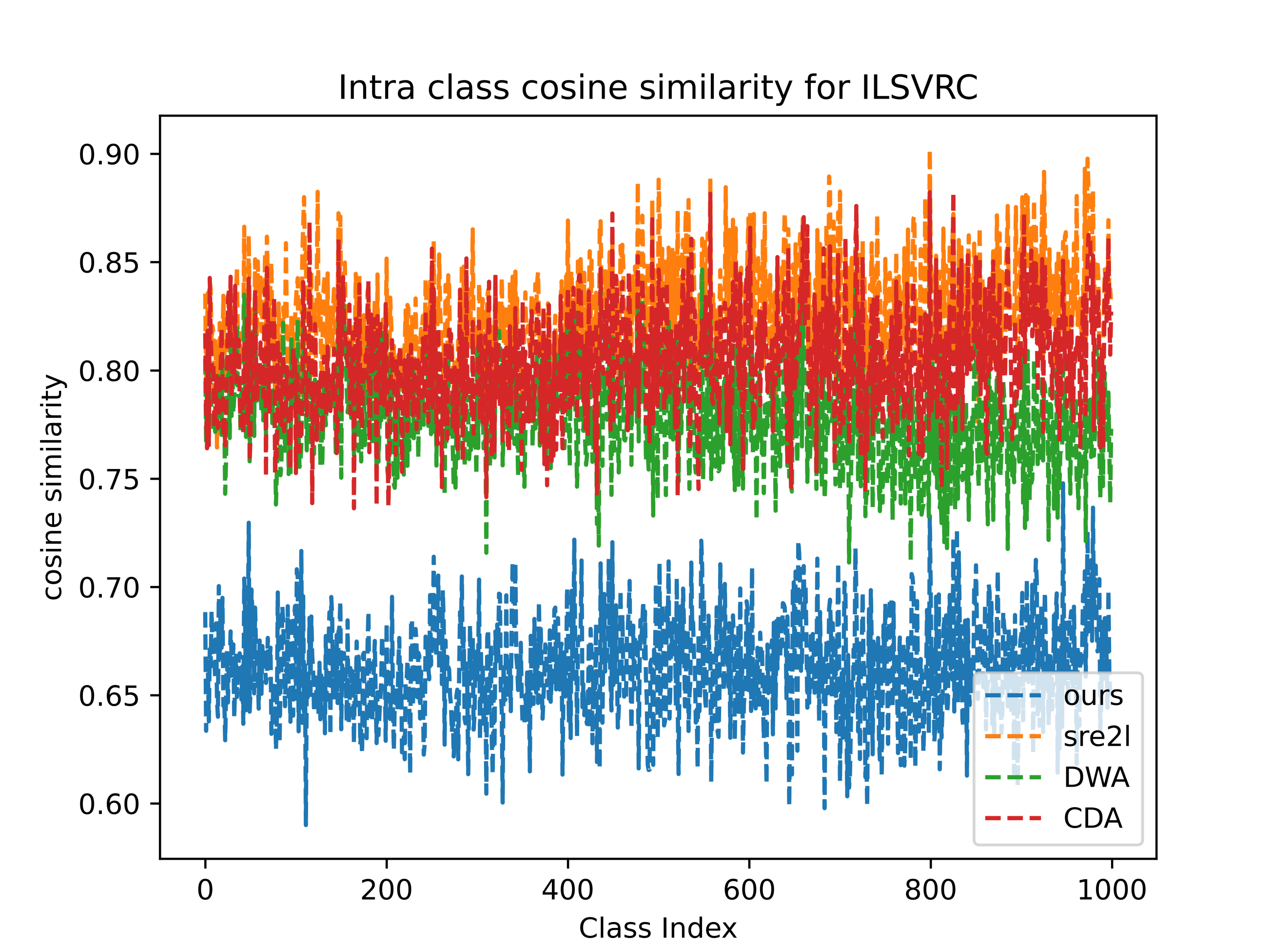

Figure 8. Intra-class cosine similarity of synthetic images generated from ImageNet-1K. We observe that DiRe lowers the cosine similarity across all classes, signifying improving diversity among the generated images.