Wearable ImageNet: Synthesizing Tileable Textures via Dataset Distillation Supplementary Material

George Cazenavette[1] Tongzhou Wang[2] Antonio Torralba[2] Alexei Efros[3] Jun-Yan Zhu[1]

[1]Carnegie Mellon University [2]Massachusetts Institute of Technology [3]UC Berkeley

A. Algorithm

Algorithm 1 Dataset Distillation via Trajectory Matching

Input: $\{\tau_i^*\}$: set of expert parameter trajectories trained on $\mathcal{D}_{\text{real}}$.
Input: M: # of updates between starting and target expert params.
Input: N: # of updates to student network per distillation step.
Input: $\mathcal{A}$: Differentiable augmentation function.
Input: $T^+ < T$: Maximum start epoch.

Initialize distilled data $\mathcal{D}_{\text{syn}} \sim \mathcal{N}(0, I)$
Initialize trainable learning rate $\alpha := \alpha_0$
for each distillation step... **do**
  ▷ Sample expert trajectory: $\tau^* \sim \{\tau_i^*\}$ with $\tau^* = \{\theta_t^*\}_0^T$
  ▷ Choose random start epoch, $t \leq T^+$
  ▷ Initialize student network with expert params:
    $\hat{\theta}_t := \theta_t^*$
  for $n = 0 \rightarrow N - 1$ **do**
    ▷ Sample a mini-batch of distilled images:
      $b_{t+n} \sim \mathcal{D}_{\text{syn}}$
    ▷ Take random crop with circular padding:
      $b_{t+n}^+ = \text{crop\_and\_pad}(b_{t+n})$
    ▷ Update student network w.r.t. classification loss:
      $\hat{\theta}_{t+n+1} = \hat{\theta}_{t+n} - \alpha \nabla \ell(\mathcal{A}(b_{t+n}^+); \hat{\theta}_{t+n})$
  end for
  ▷ Compute loss between ending student and expert params:
    $\mathcal{L} = \|\hat{\theta}_{t+N} - \theta_{t+M}^*\|_2^2 \;/\; \|\theta_t^* - \theta_{t+M}^*\|_2^2$
  ▷ Update $\mathcal{D}_{\text{syn}}$ and α with respect to $\mathcal{L}$
end for

Output: distilled data $\mathcal{D}_{\text{syn}}$ and learning rate α

B. Performance as Training Data

Despite not explicitly following the dataset distillation objective, our distilled textures still achieve surprising classification results. By training on only random patches of our ImageSquawk textures, we achieve %39.6 test accuracy. Likewise, we achieve %26.8 test accuracy by only training on random patches of ImageFruit.

C. More Visualizations

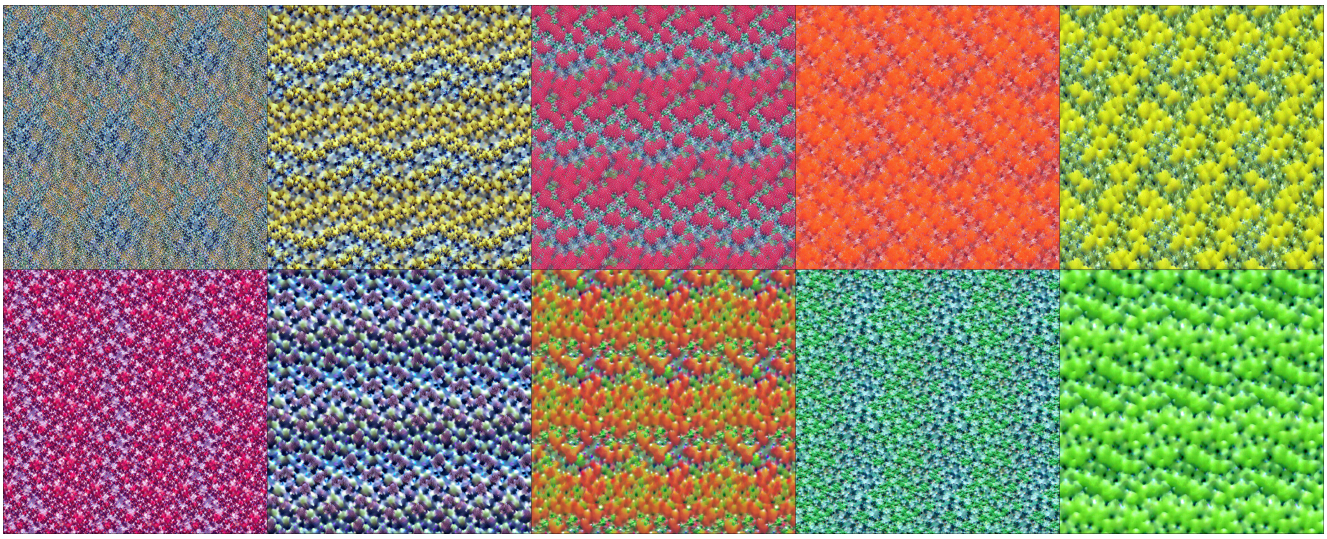

Figure 1. ImageFruit Distilled Textures Tiled 3x3