Batch Decorrelation for Active Metric Learning

Priyadarshini[1]

Joint work with Ritesh Goru[1]*, Siddhartha Chaudhuri*[1,2]*, and Subhasis Chaudhuri*[1]

[1]IIT Bombay, [2]Adobe Research

IJCAI-PRICAI 2020

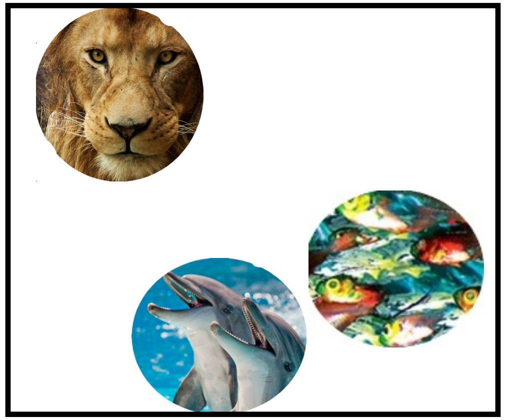
January 7, 2021

Metric Learning – Two Key Concerns

(1) Discrete class-based learning

Class-based model

Perceptual model
- concept



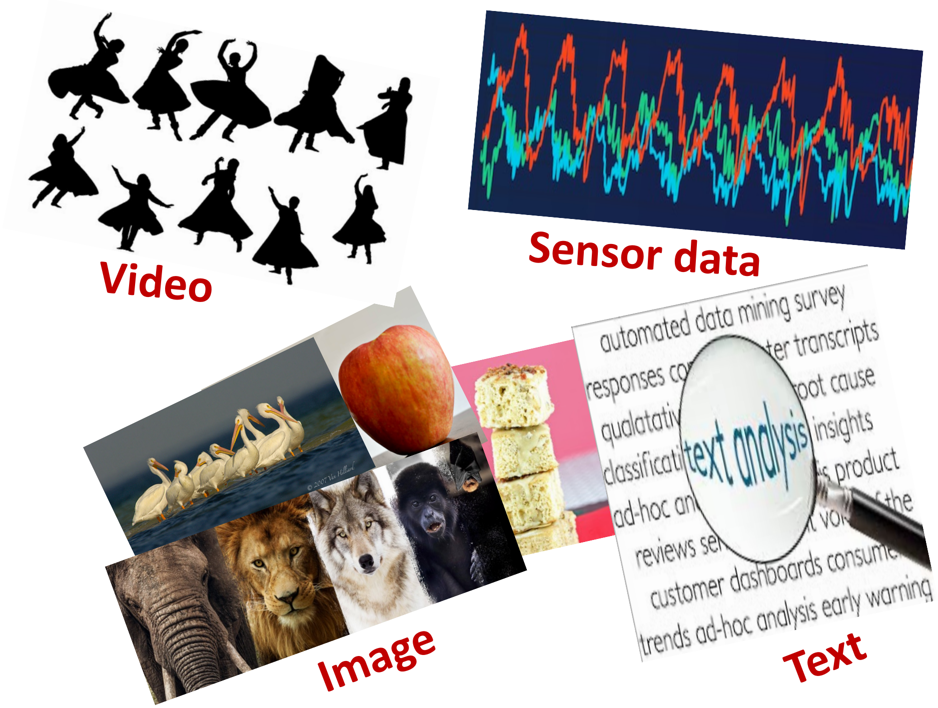
Perceptual model
- visual appearance

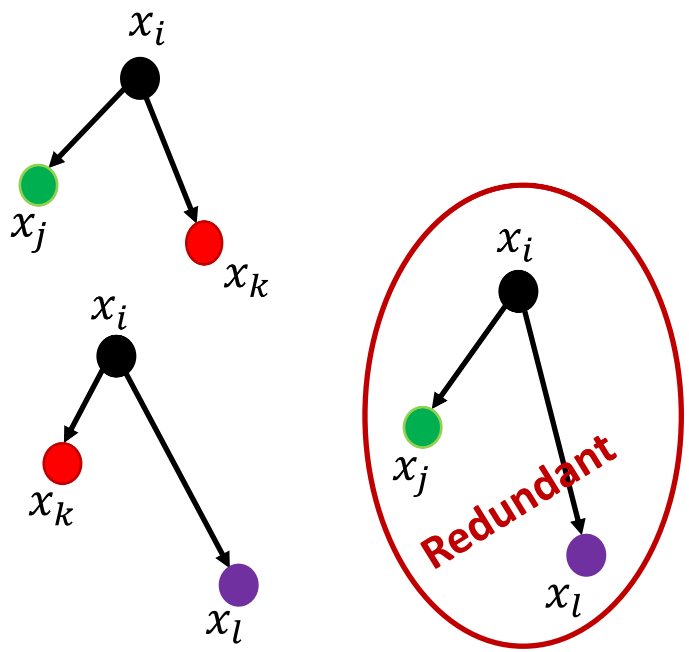
Perceptual metric trained on relative similarity comparisons between objects - Is object *"x"* more similar to object *"y"* or object *"z"* ?

Google Image

Metric Learning – Two Key Concerns

(2) Annotation-intensive



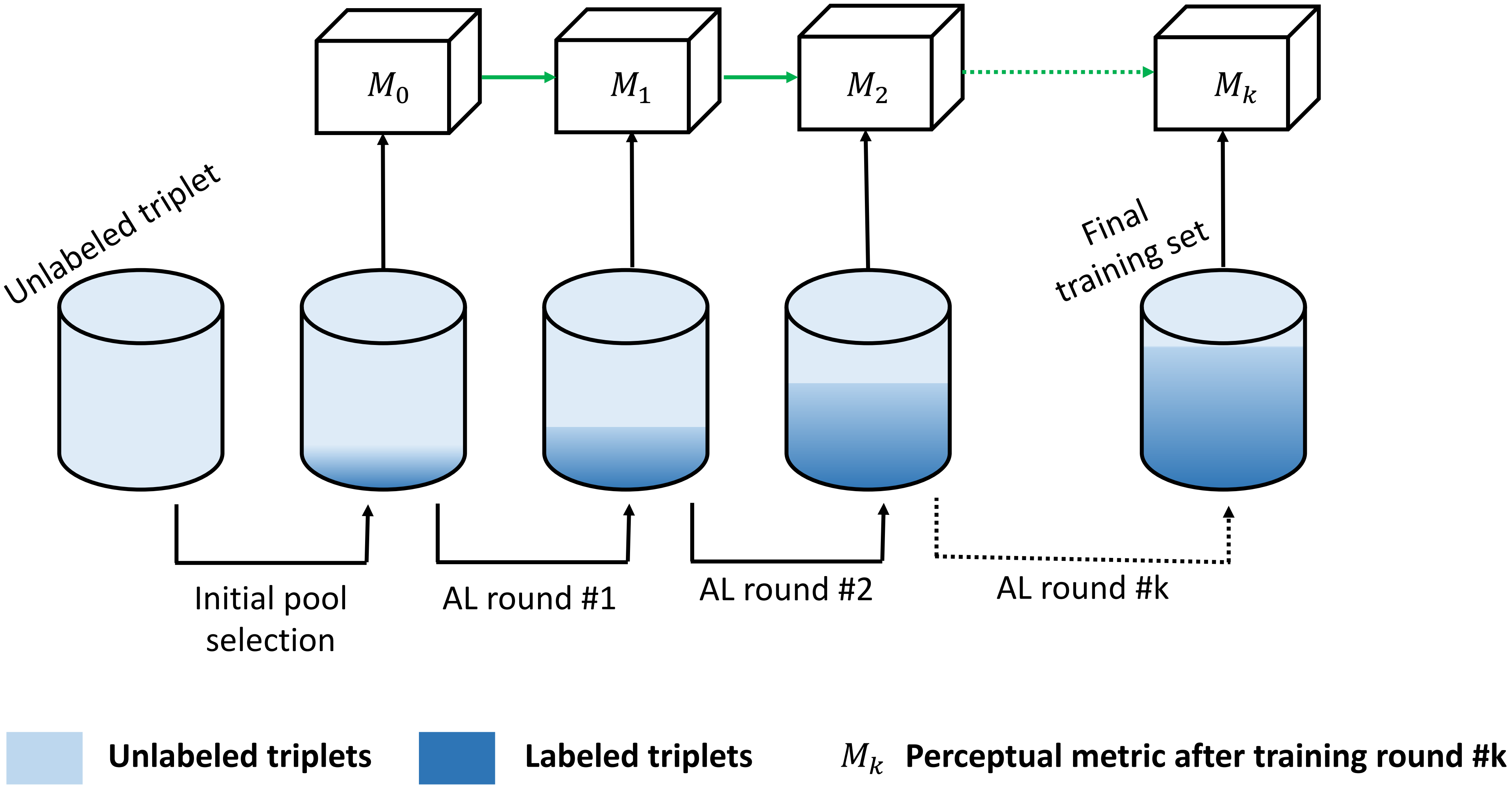


Order $-O(n^3)$

Reduce dependence on human guidance

Google Image

Batch-Mode Active Metric Learning Framework



Smart Labeling – All triplets are not equally important.

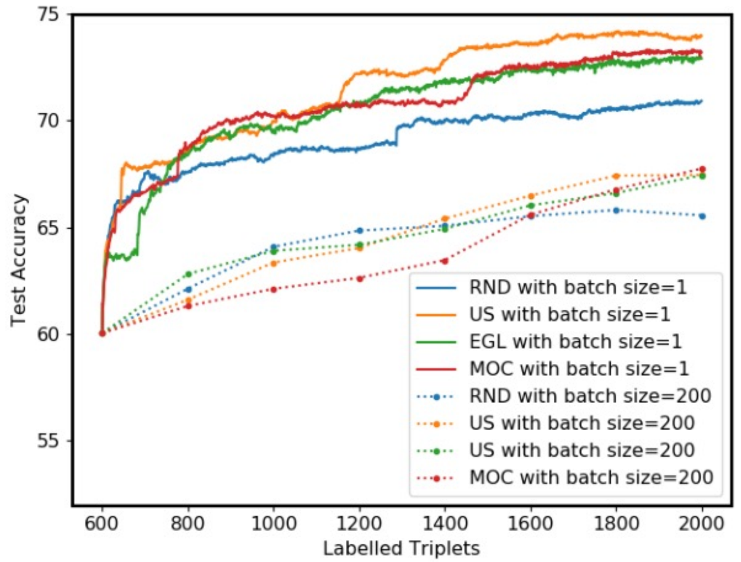
Identify informative triplet

Uncertainty sampling: Identify subset S of triplets about which the current model (ϕ) is highly uncertain in predicting its order: $d_{\phi(x_i,x_j)} \approx d_\phi(x_i, x_k)$

$$S^* = \underset{\{S \subset U\}}{\operatorname{argmax}} H(S)$$

$$= \sum_{\{t \in S\}} -p_t log p_t - (1 - p_t)\log(1 - p_t)$$

$$p_t = \frac{\mu + d_\phi^2(x_i,x_k)}{2\mu + d_\phi^2(x_i,x_k) + d_\phi^2(x_i,x_j)} \; ; \qquad \mu > 0$$



Active learning strategies are effective when subset size is one

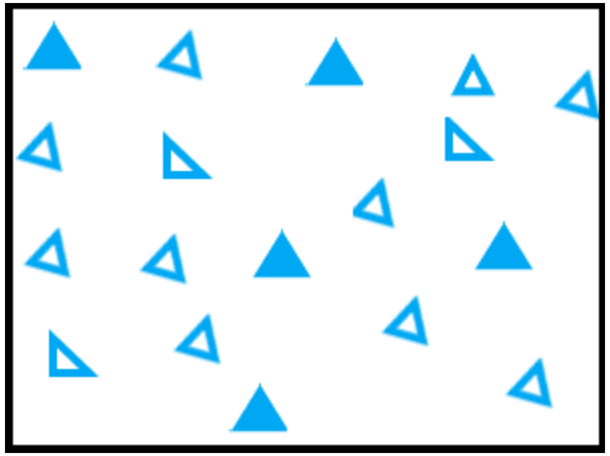
Smart Labeling – Diversity is necessary

Integrate informativeness and diversity

1. Select an overcomplete set of informative triplets

$$S = argmax_{S \subset U, |S|=k} H(S); \quad k > b$$

2. Pick b diverse triplets using farthest point sampling (FPS)



FPS

$$S^* = argmax_{\{t_i, t_j\} \subset S} \rho_\phi(t_i, t_j)$$

$$for\ n = 3, \dots, b\ do$$

$$S^* \leftarrow S^* \cup \{argmax_{t \in S \setminus S^*} argmin_{t' \in S^*} \rho_\phi(t, t')\}$$

$$end$$

How to define $\rho_\phi(t_i, t_j)$?

Smart Labeling – Diversity is necessary

Integrate informativeness and diversity

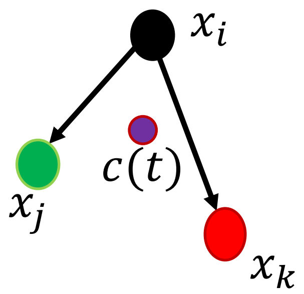
$$\rho_{\phi}(t, t') = \underbrace{H(t) \times H(t')}_{\text{Informativeness}} \times \underbrace{d(t, t')}_{\text{Diversity}}$$

Gradient distance: Each triplet is represented by expected gradient of loss function with respect to last layer of model parameters: $g(t)$

$$d(t, t') = 1 - \langle \frac{g(t)}{|g(t)|}, \frac{g(t')}{|g(t')|} \rangle$$

Centroidal distance: Each triplet is represented by centroid of the embedding of three objects

$$d(t, t') = \|c(t) - c(t')\|; \;\; c(t) = \frac{1}{3}(\phi(x_i) + \phi(x_j) + \phi(x_k))$$



Smart Labeling – Diversity is necessary

Integrate informativeness and diversity

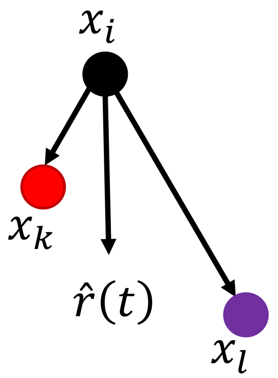
Euclidean distance: Triplet is represented by concatenated object embedding

$$\phi(x_i, x_j, x_k) = \phi(x_i) \oplus \phi(x_j) \oplus \phi(x_k)$$

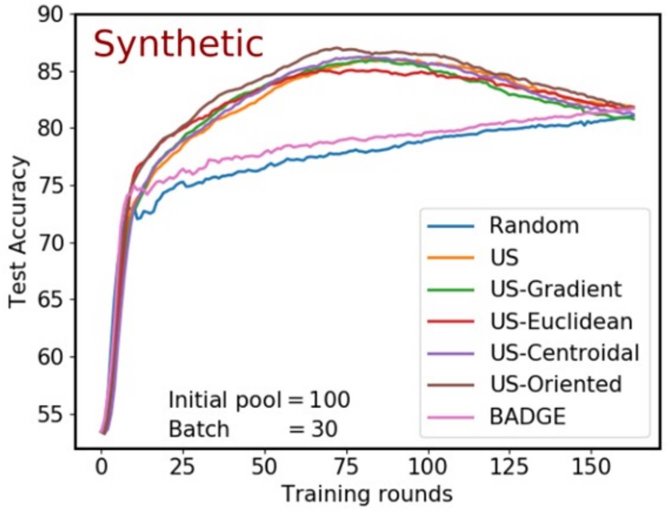
$$d(t, t') = \sum_{y \in \{ijk, ikj\}} \frac{1}{2} \| \phi(t^y) - \phi(t') \|$$

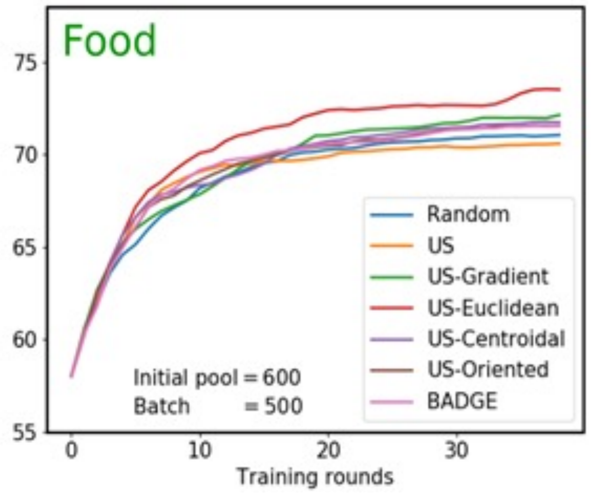
Oriented distance: Distance between the anchors of two triplets + cosine distance between resultant vectors

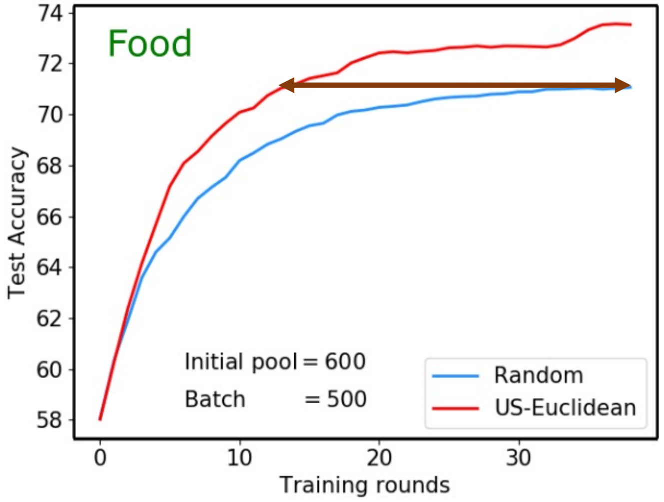
$$d(t, t') = d_\phi(x_i, x_i') + (1 - \langle \hat{r}(t), \hat{r}(t') \rangle)$$



Results on Different Datasets







- All variants of decorrelated AL performs better than random
- Our method achieves higher performance gain over random and the US method with larger subset sizes or initial pool

Less than half (39%) as many labeled triplets needed by our method

Source code: https://priyadarshini-k.com/publications/