Magnet Loss - Experimentation on Clustering

Nithi Subbaian and Jinhan Zhang

METRIC LEARNING WITH ADAPTIVE DENSITY DISCRIMINATION

Oren Rippel
MIT, Facebook AI Research
rippel@math.mit.edu

Manohar Paluri
Facebook AI Research
mano@fb.com

Piotr Dollar
Facebook AI Research
pdollar@fb.com

Lubomir Bourdev
UC Berkeley
lubomir.bourdev@gmail.com

ABSTRACT

Distance metric learning (DML) approaches learn a transformation to a representation space where distance is in correspondence with a predefined notion of similarity. While such models offer a number of compelling benefits, it has been difficult for these to compete with modern classification algorithms in performance and even in feature extraction.

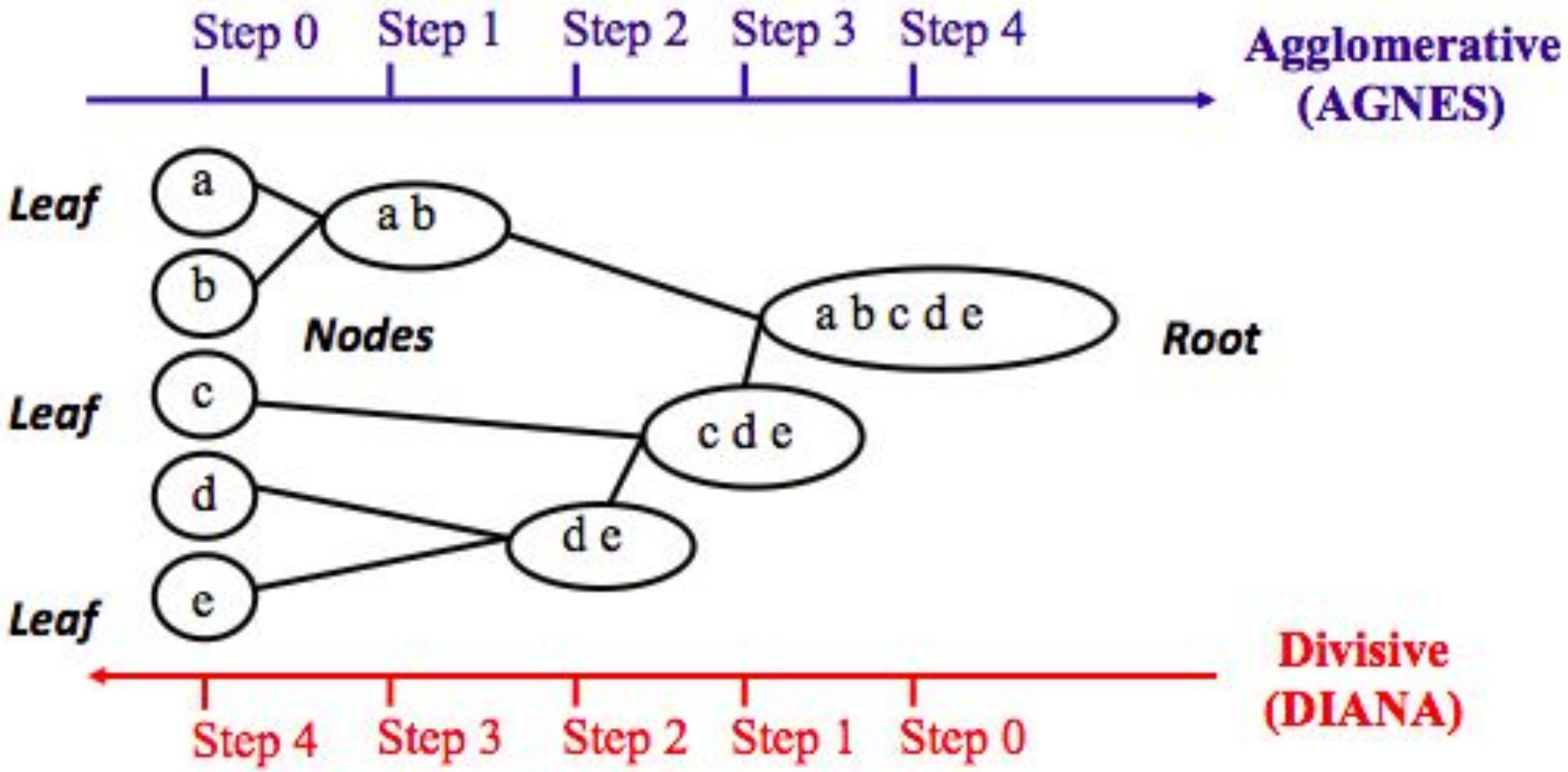
In this work, we propose a novel approach explicitly designed to address a number of subtle yet important issues which have stymied earlier DML algorithms. It maintains an explicit model of the distributions of the different classes in representation space. It then employs this knowledge to adaptively assess similarity, and achieve local discrimination by penalizing class distribution overlap.

We demonstrate the effectiveness of this idea on several tasks. Our approach achieves state-of-the-art classification results on a number of fine-grained visual recognition datasets, surpassing the standard softmax classifier and outperforming triplet loss by a relative margin of 30-40%. In terms of computational performance, it alleviates training inefficiencies in the traditional triplet loss, reaching the same error in 5-30 times fewer iterations. Beyond classification, we further validate the saliency of the learnt representations via their attribute concentration and hierarchy recovery properties, achieving 10-25% relative gains on the softmax classifier and 25-50% on triplet loss in these tasks.

Figure 1. On Page 2.

DML Improvement -

Use Agglomerative Clustering to adaptively determine the clusters number in Magnet Loss to improve intra class variation


Step 0
Step 1
Step 2
Step 3
Step 4
Agglomerative
(AGNES)
Leaf
a
b
a b
Nodes
a b c d e
Root
Leaf
c
c d e
d
d e
e
Leaf
Divisive
(DIANA)
Step 4
Step 3
Step 2
Step 1
Step 0

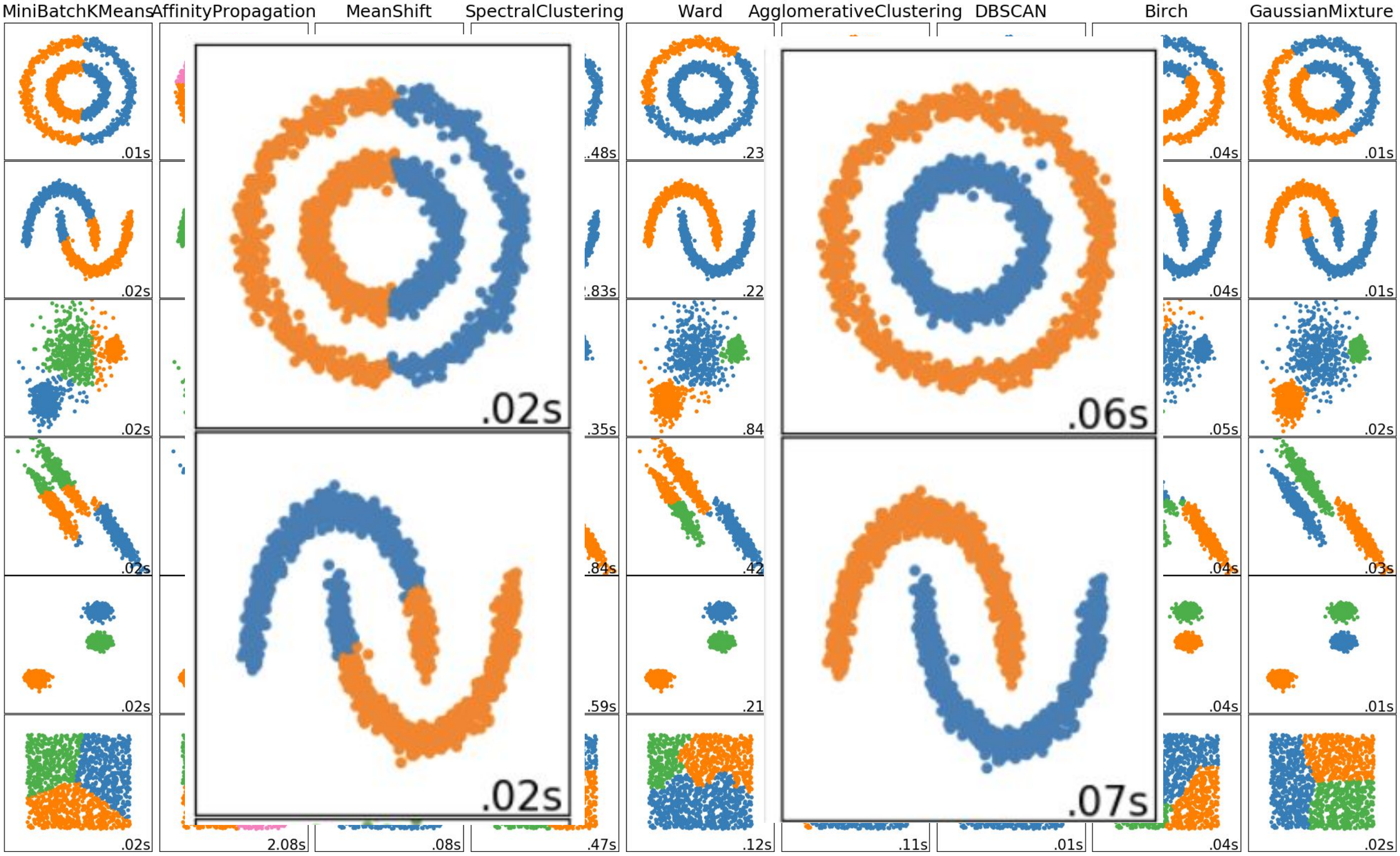
Improvements

Agglomerative Clustering Algorithm

VS K-means

possible to have varying number of clusters along the training and for different classes

Better at Spiral(elliptic) Clusters


MiniBatchKMeans
AffinityPropagation
MeanShift
SpectralClustering
Ward
AgglomerativeClustering
DBSCAN
Birch
GaussianMixture
.01s
.02s
.02s
.02s
.02s
.02s
2.08s
.08s
.48s
.35s
.84s
.59s
.47s
.23
.22
.84
.42
.21
.12s
.11s
.01s
.04s
.04s
.05s
.04s
.04s
.04s
.01s
.01s
.02s
.03s
.01s
.02s
.02s
.02s
.06s
.07s

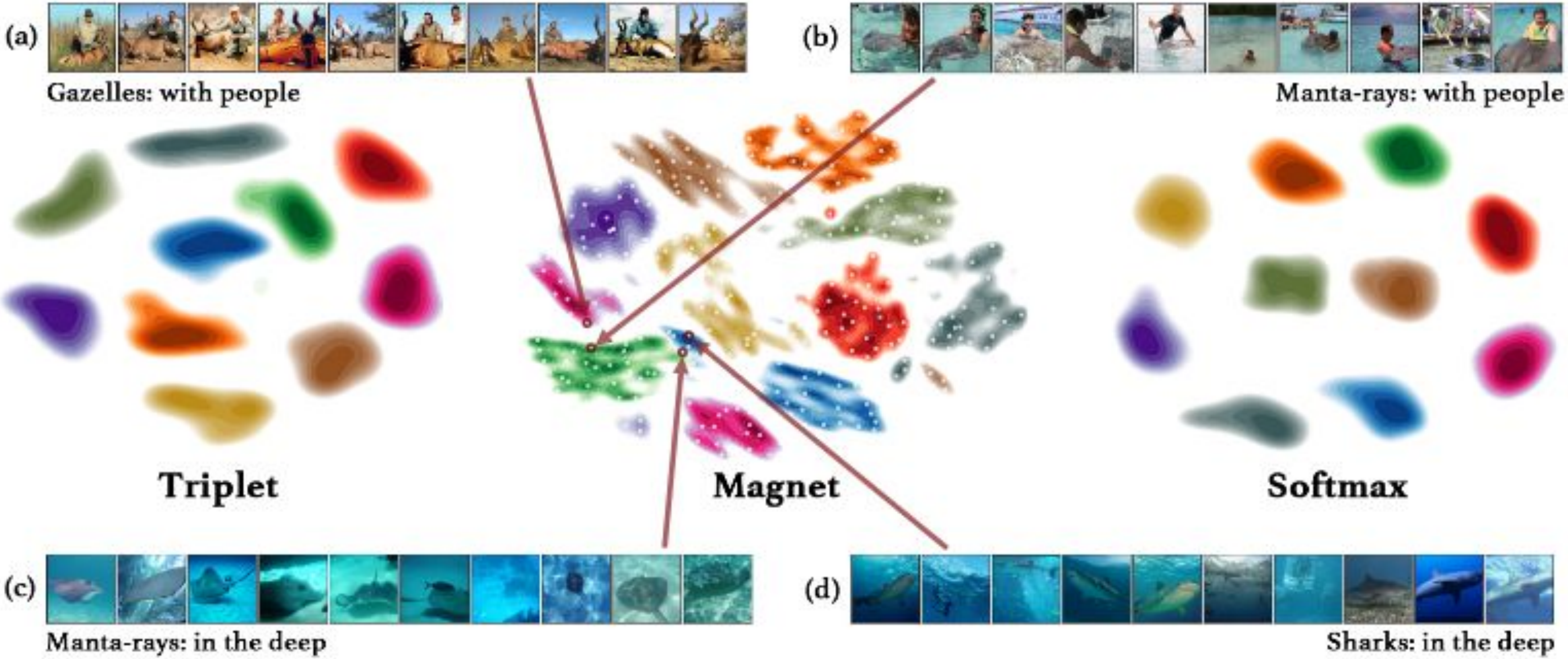
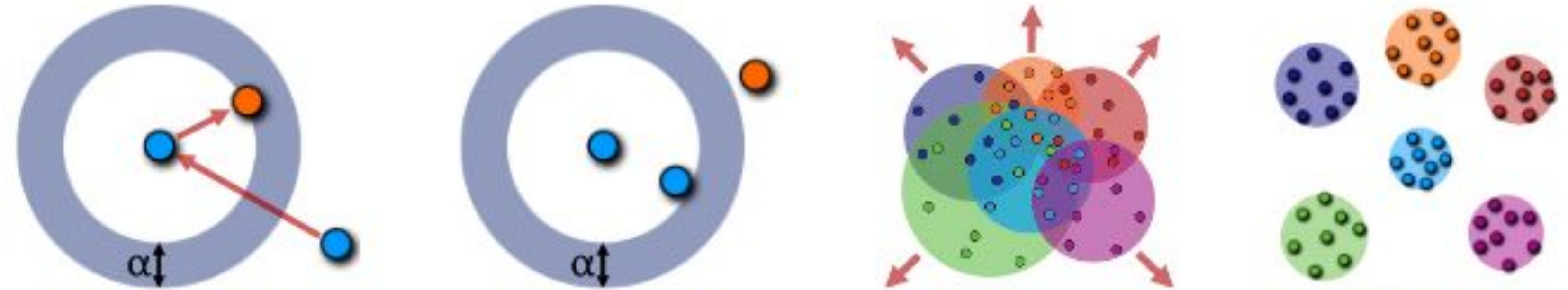
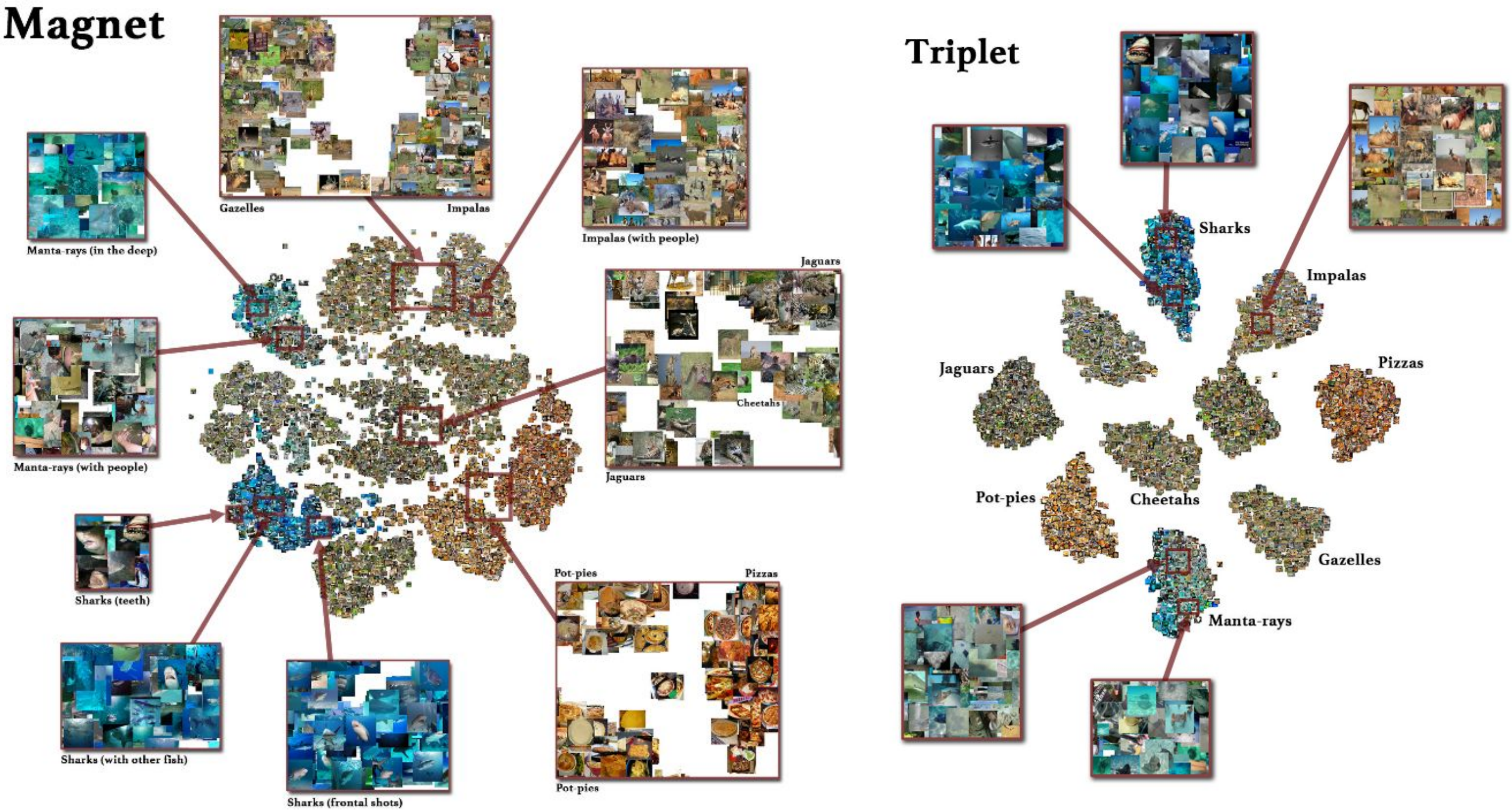


Figure 2. On Page 3.

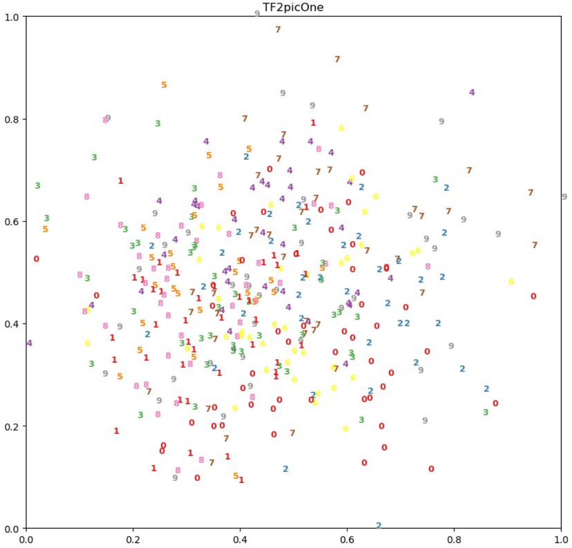


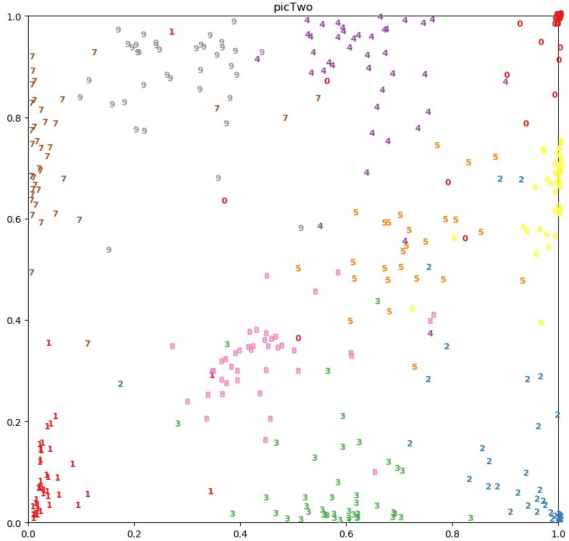
(a) Triplet: before. (b) Triplet: after. (c) Magnet: before. (d) Magnet: after.

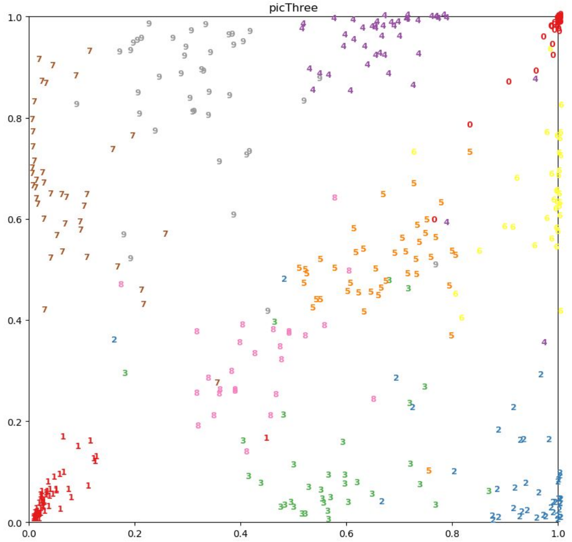
Figure 3. On Page 4.


Magnet
Gazelles
Impalas
Impalas (with people)
Manta-rays (in the deep)
Manta-rays (with people)
Jaguars
Cheetahs
Jaguars
Sharks (teeth)
Sharks (with other fish)
Sharks (frontal shots)
Pot-pies
Pizzas
Pot-pies
Triplet
Sharks
Impalas
Jaguars
Pizzas
Pot-pies
Cheetahs
Gazelles
Manta-rays

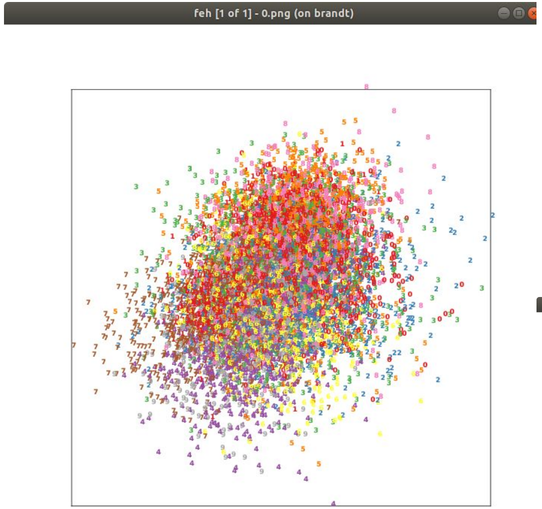
Magnet Loss - MNIST

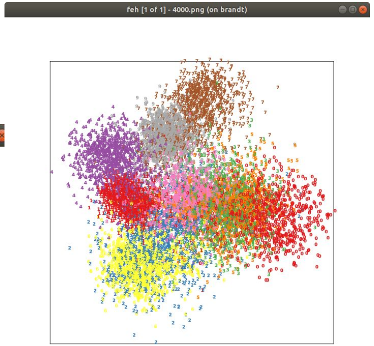


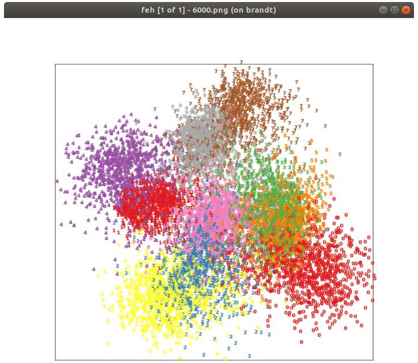


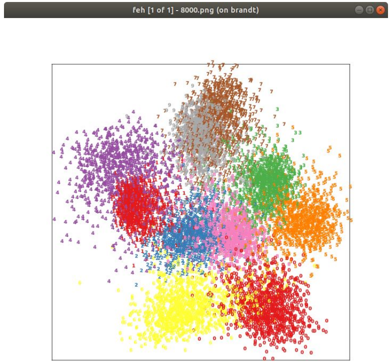


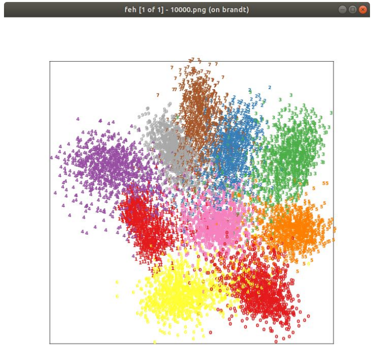
Magnet loss, k-means, fashion mnist

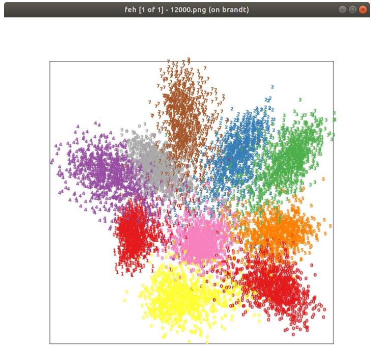


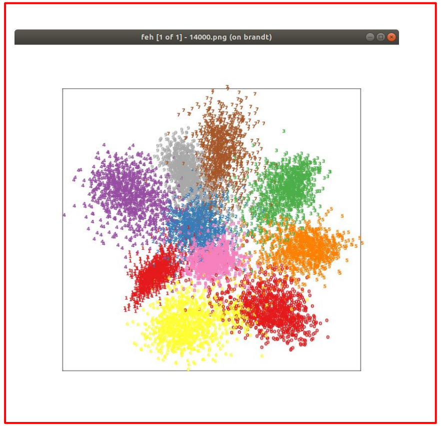




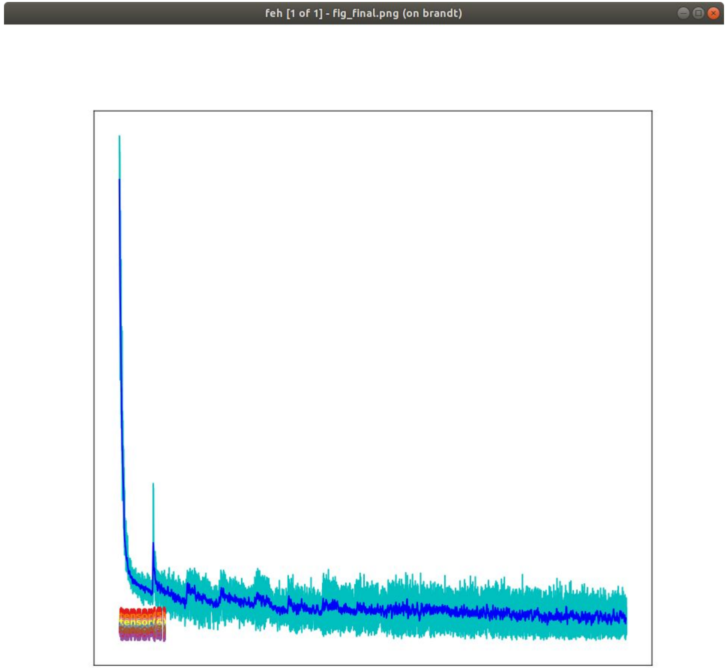




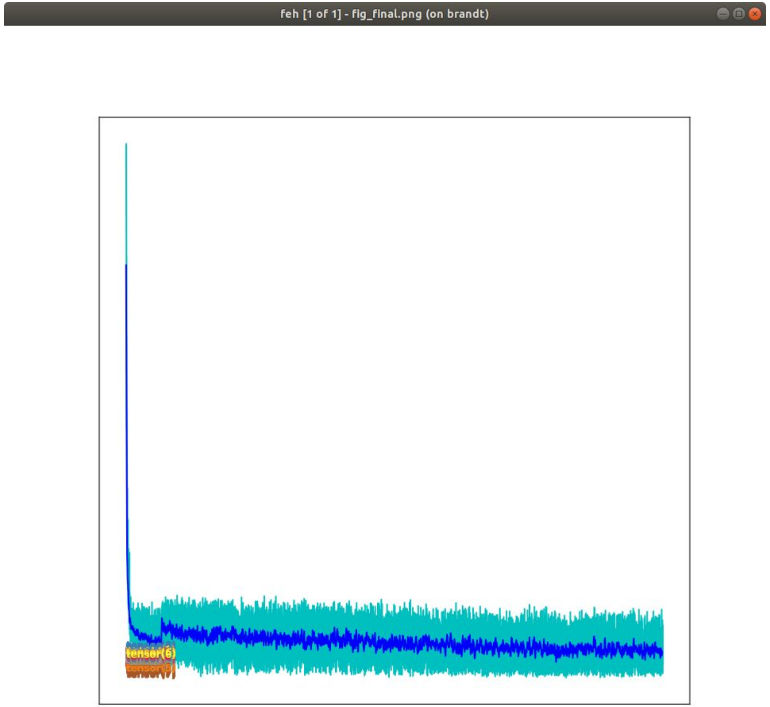




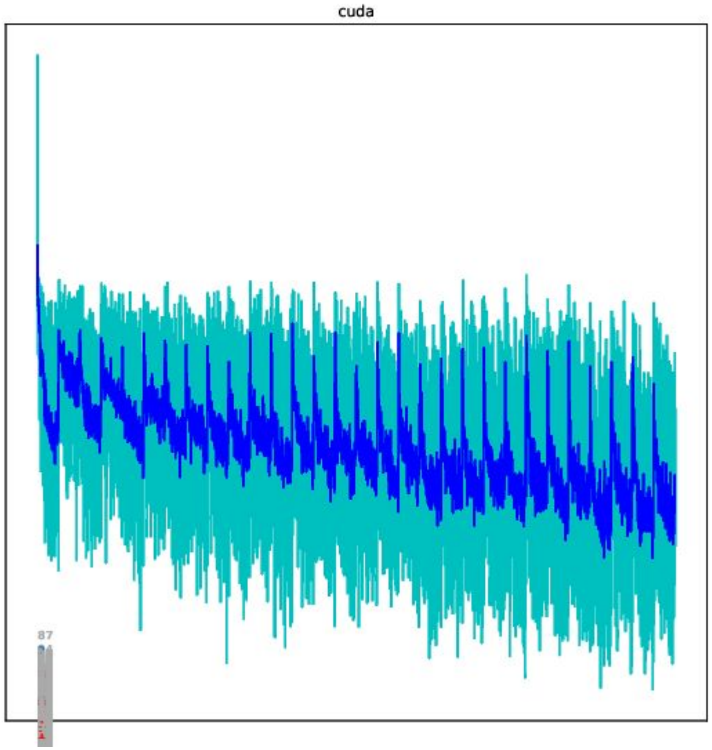
K means, mnist

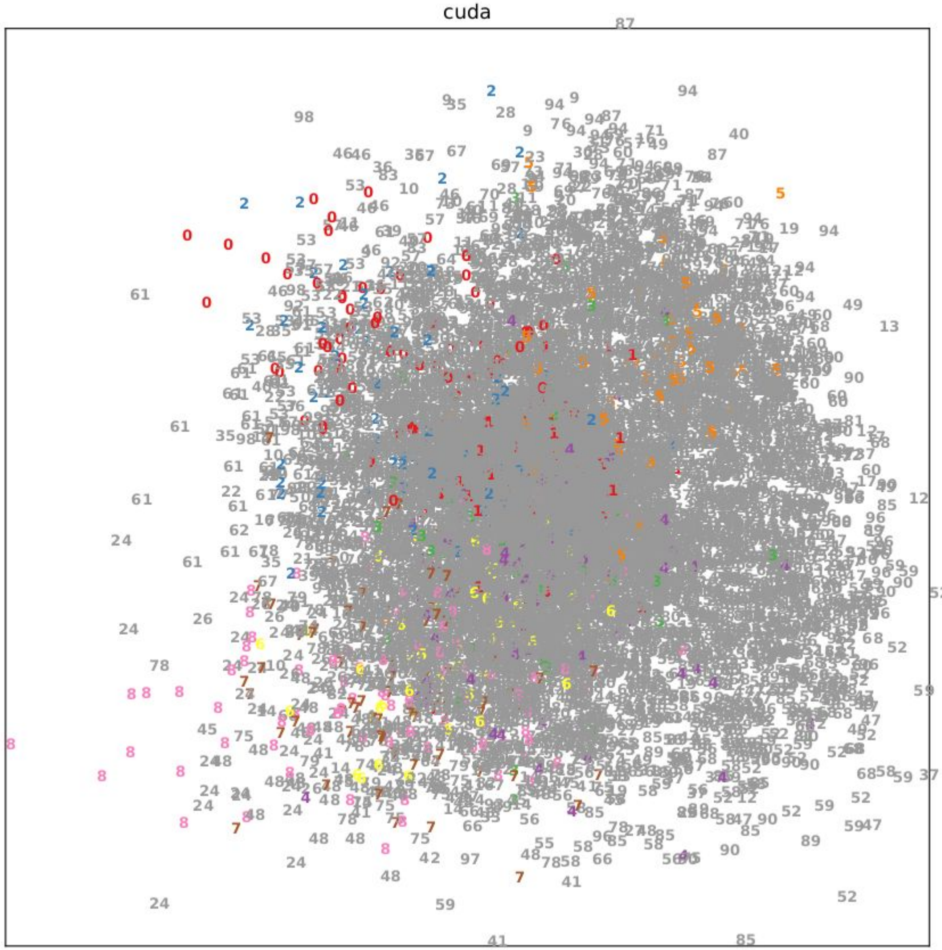


K means, fashion mnist

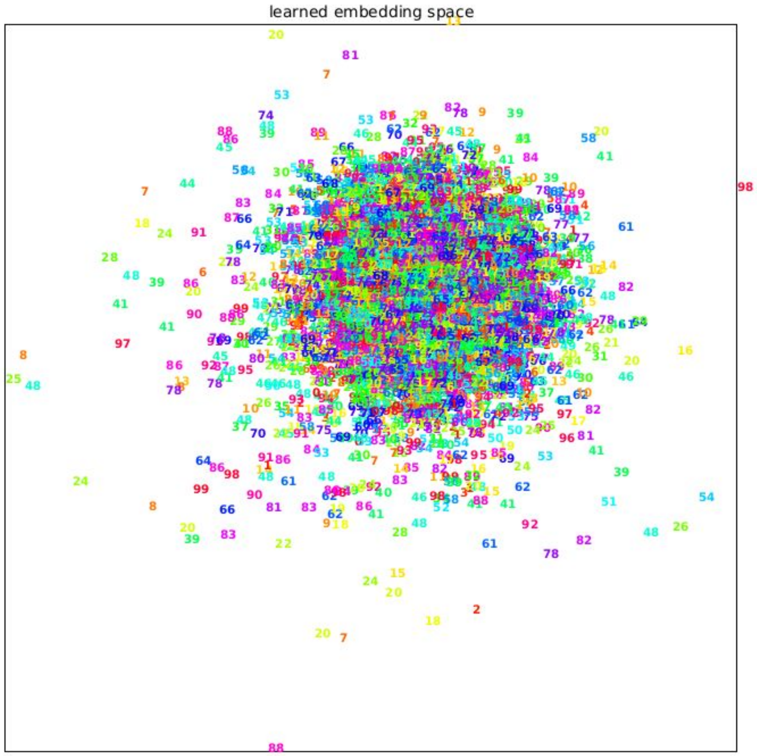


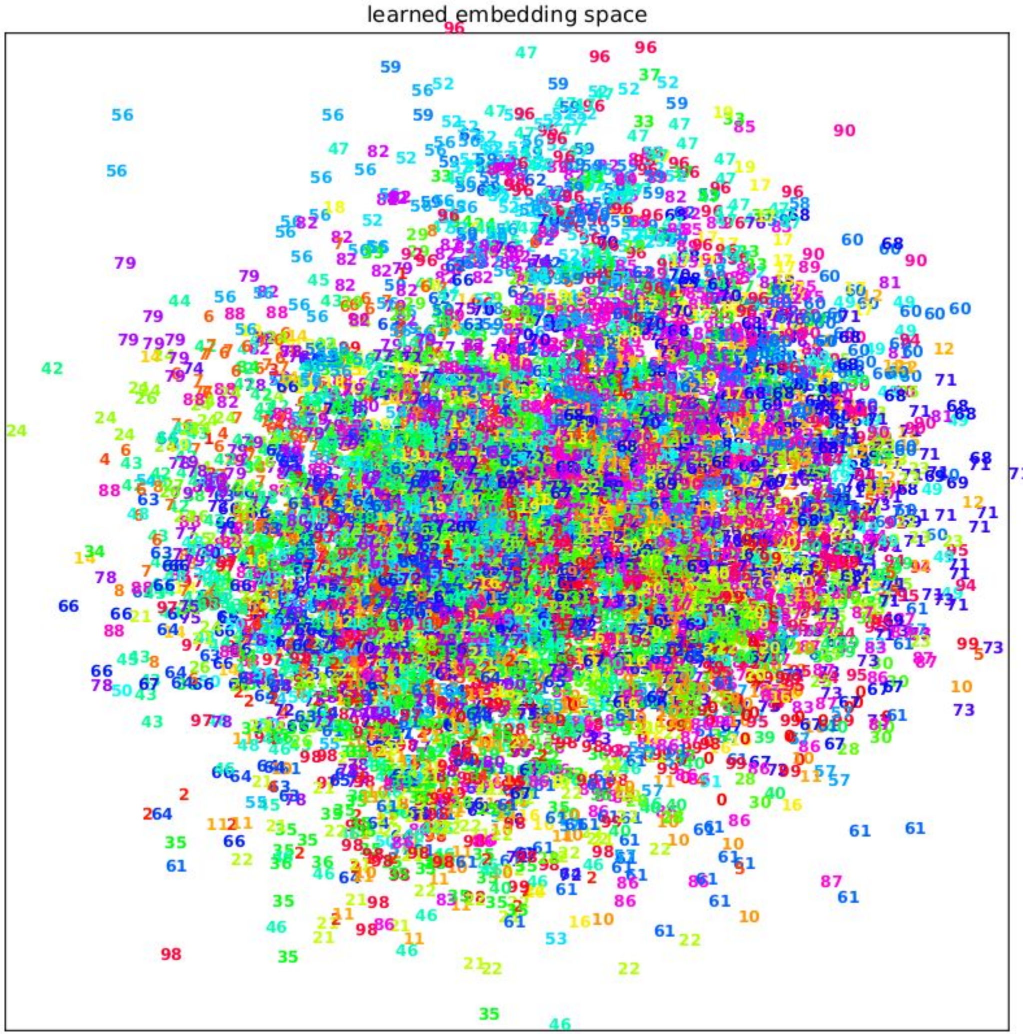
CIFAR-100, k means





CIFAR - 100, K means





Agglomerative Clustering, CIFAR 100