Learning Embeddings for Image Clustering: An Empirical Study of Triplet Loss Approaches

Kalun Ho[1, 2, 4], Janis Keuper[3], Franz-Josef Pfreundt[1, 2], and Margret Keuper[4]

[1]*Fraunhofer Center Machine Learning, Germany*
[2]*Competence Center High Performance Computing, Fraunhofer ITWM, Kaiserslautern, Germany*
[3]*Institute for Machine Learning and Analytics (IMLA), Offenburg University, Germany*
[4]*Data and Web Science Group, University of Mannheim, Germany*

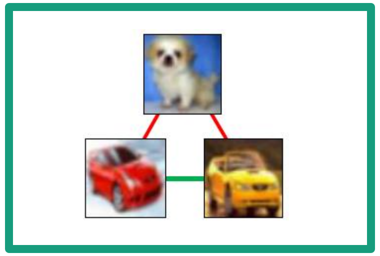
Fraunhofer

Our method has three contributions

Our Method

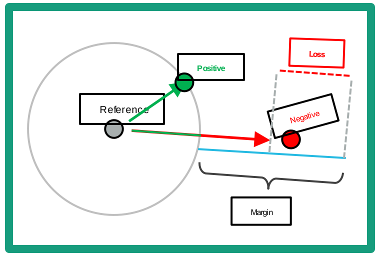
Contribution

1



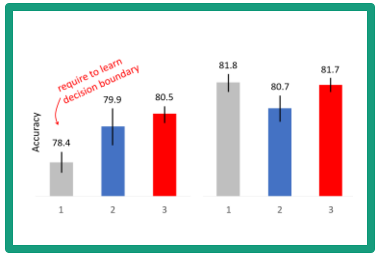
- Study on two clustering approaches applied on embeddings, learnt from three versions of Triplet Losses

2



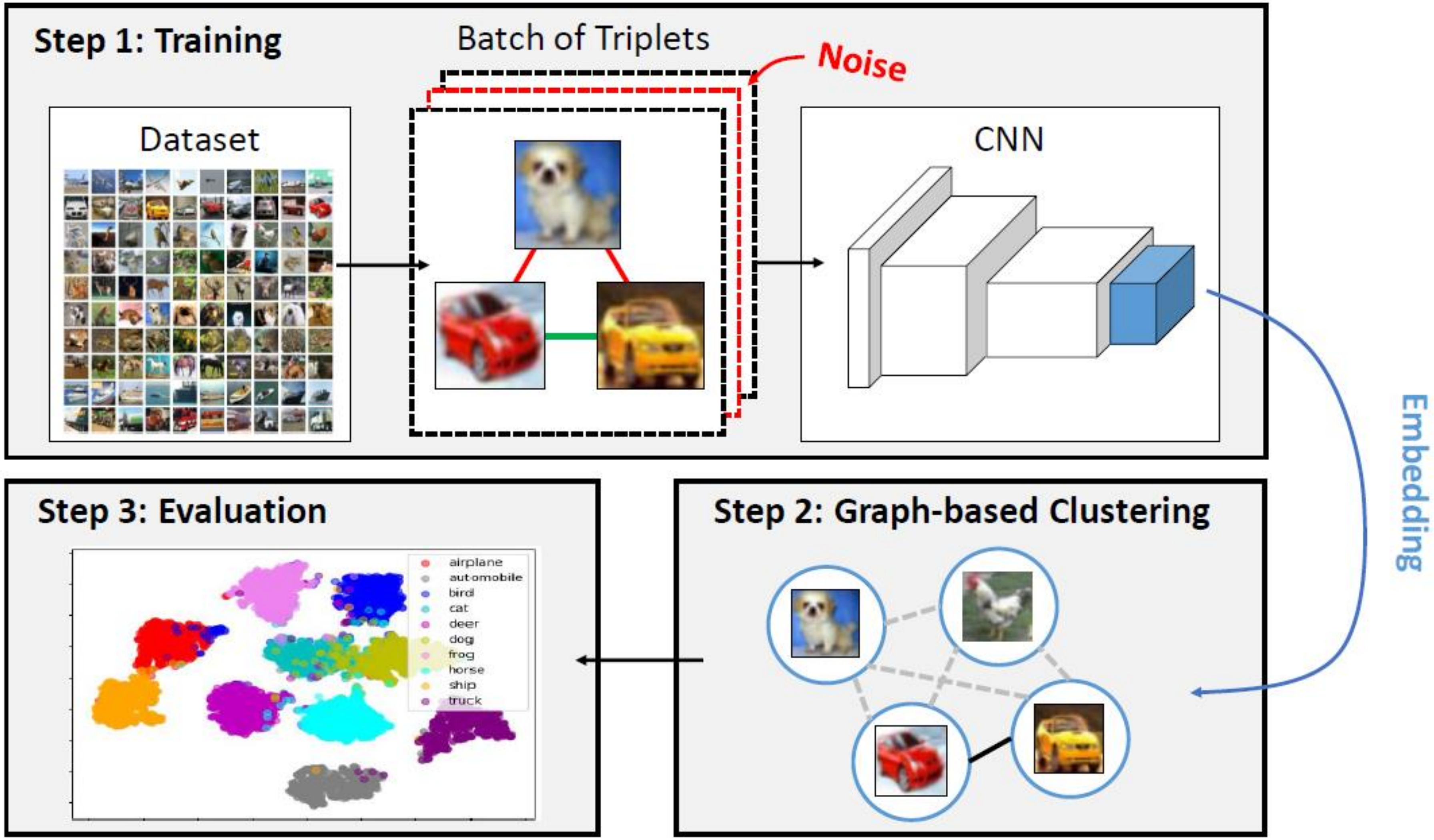
- Simplification of Triplet Loss, which allows to directly compute the probability of two data points for belonging to disjoint components

3



- Proposed Triplet Loss outperform previous versions on CIFAR-10 and is robust against noise

Overview


Step 1: Training
Batch of Triplets
Noise
Dataset
CNN
Embedding
Step 3: Evaluation
airplane
automobile
bird
cat
deer
dog
frog
horse
ship
truck
Step 2: Graph-based Clustering

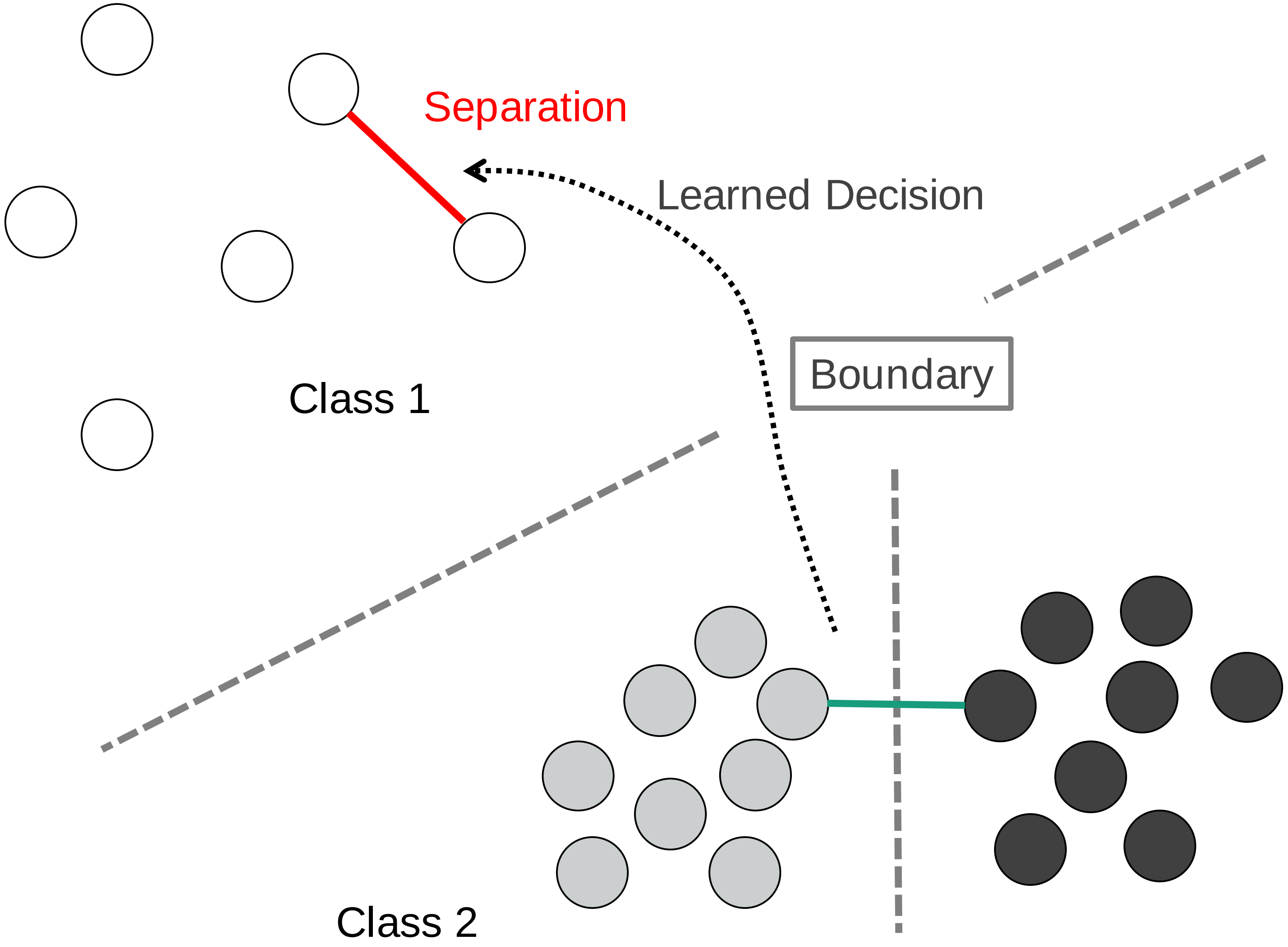
We investigated three versions of Triplet Loss

$$L_{triplet} = \sum_{i=1}^{n}[\|f(x_i^a) - f(x_i^p)\|^2 - \|f(x_i^a) - f(x_i^n\|^2) + \alpha]_+$$

$$L_{triplet\_2} = L_{triplet} + [\|f(x_i^a) - f(x_i^p)\|^2 - \beta]_+$$

$$L_{triplet\_3} = \sum_{i=1}^{n}[\alpha - \|f(x_i^a) - f(x_j^n)\|^2]_+ + [\|f(x_i^a) - f(x_k^p)\|^2 - \beta]_+$$

Different Intra-cluster Distance make it impossible to learn Threshold



We can derive the Threshold directly from training Parameters

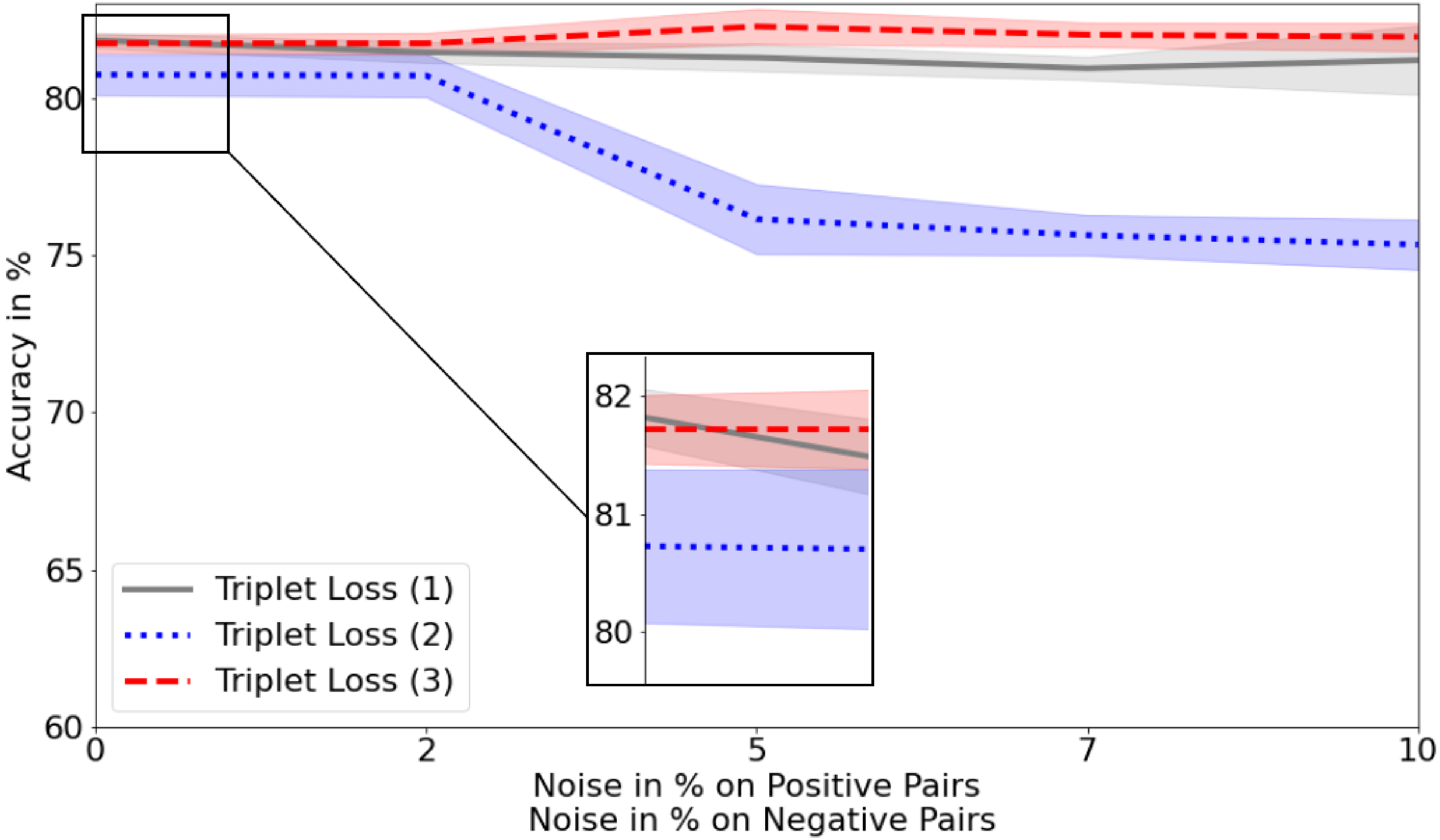
$$L_{triplet} = \sum_{i=1}^{n} [\|f(x_i^a) - f(x_i^p)\|^2 - \|f(x_i^a) - f(x_i^n\|^2) + \alpha]_+$$

$$L_{triplet\_2} = L_{triplet} + [\|f(x_i^a) - f(x_i^p)\|^2 - \beta]_+$$

$$L_{triplet\_3} = \sum_{i=1}^{n} [\alpha - \|f(x_i^a) - f(x_j^n)\|^2]_+ + [\|f(x_i^a) - f(x_k^p)\|^2 - \beta]_+$$

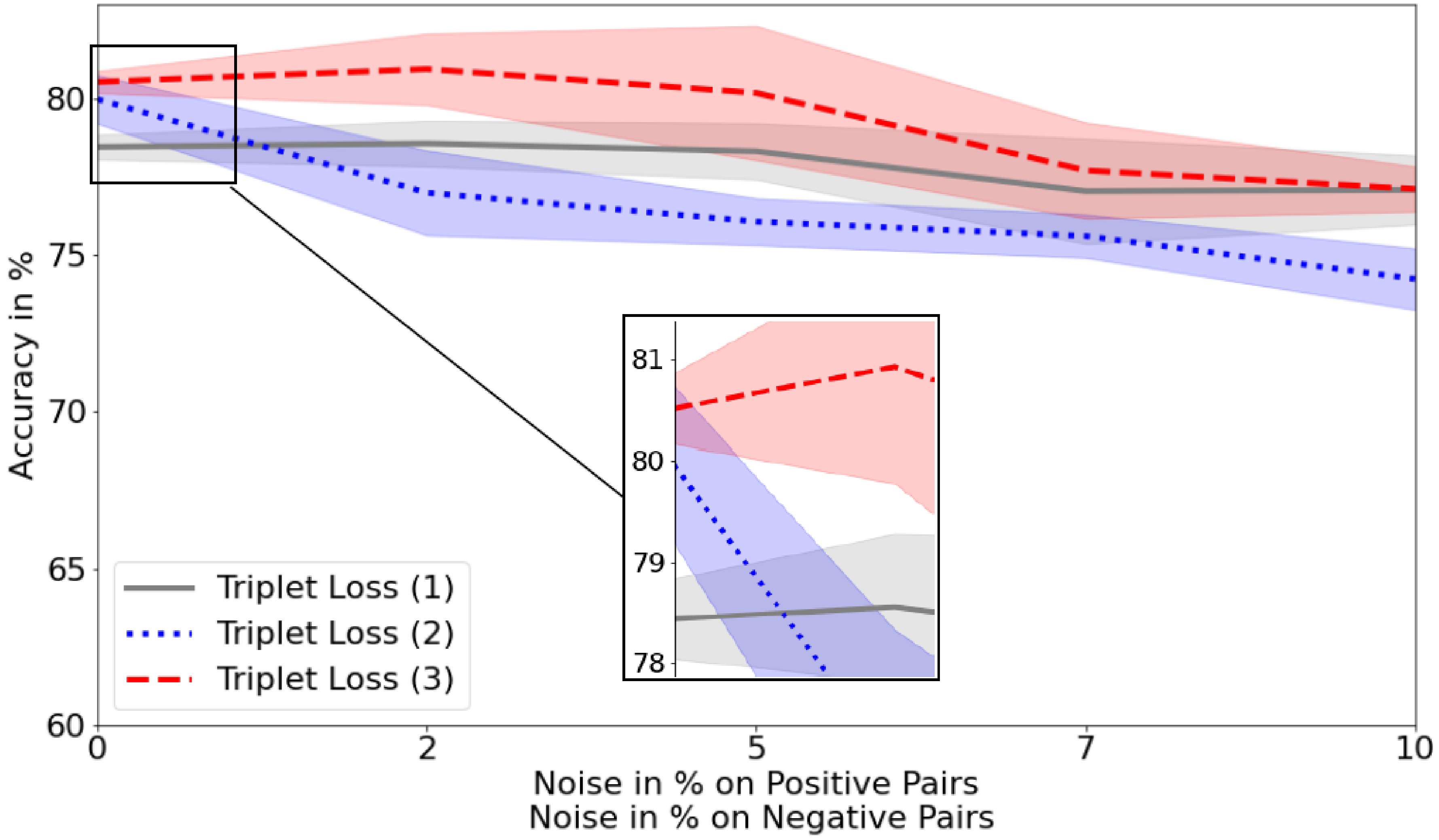
Threshold: $\tau = \sqrt{(\alpha + \beta)/2}$

Clustering Performance with noisy data: KMeans



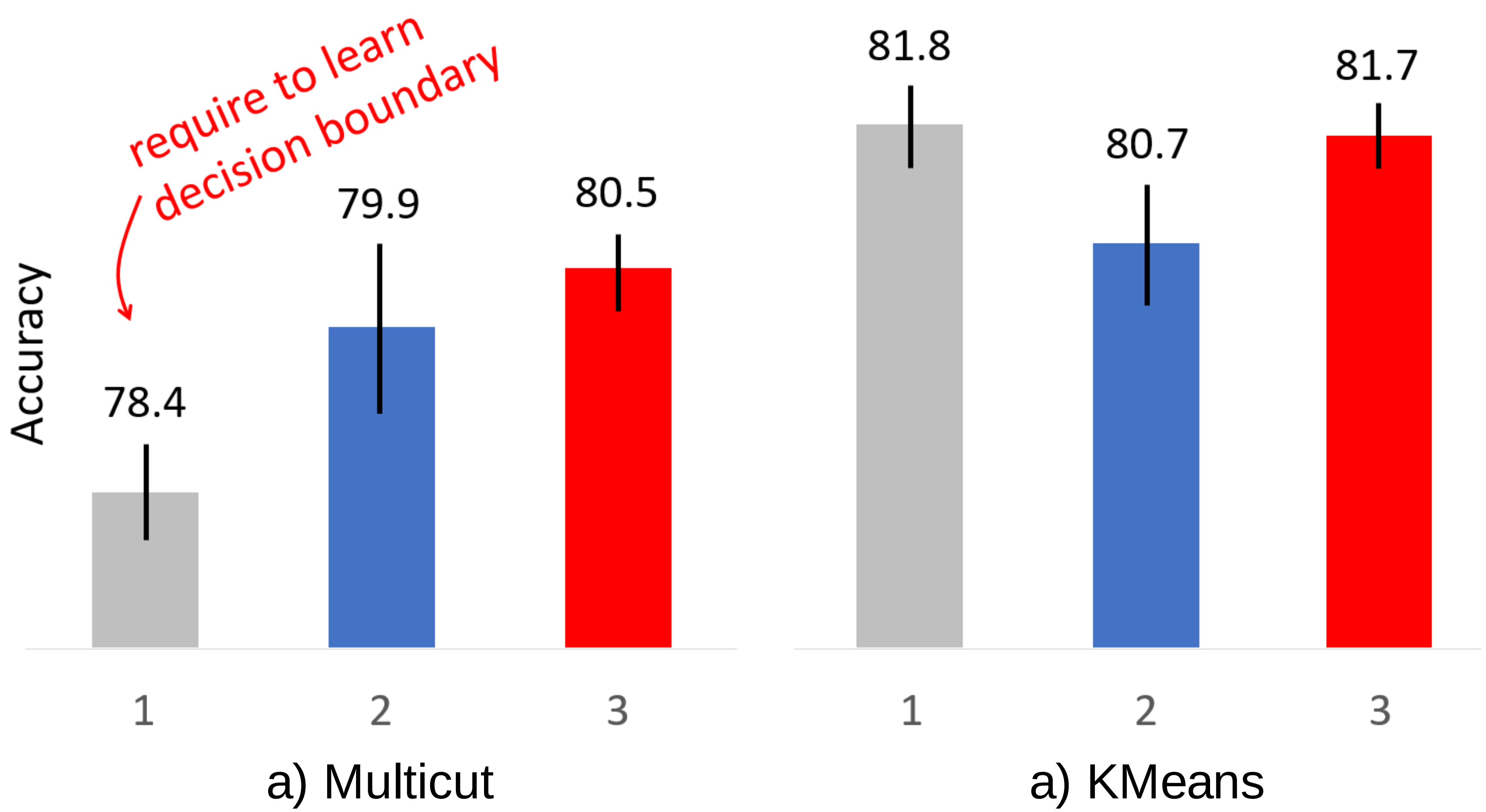
Fraunhofer

Clustering Performance with noisy data: Multicut



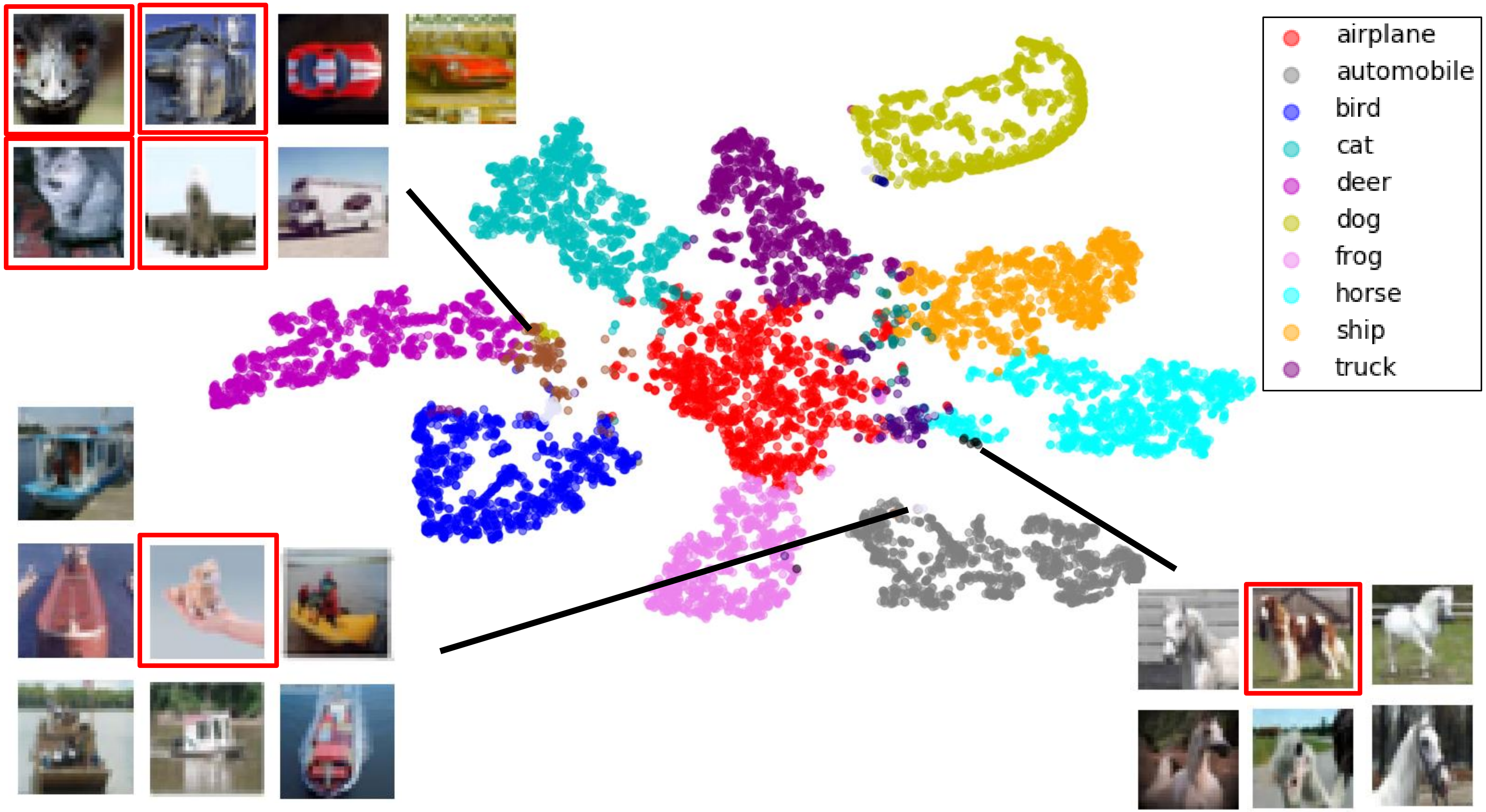
Fraunhofer

Clustering Performance: Multicut vs. KMeans



Fraunhofer

TSNE-Visualization: Multicut-Clustering on CIFAR10



Summary

- Study on two clustering approaches applied on embeddings, learnt from three versions of Triplet Losses

- Simplification of Triplet Loss, which allows to directly compute the probability of two data points for belonging to disjoint components

- Proposed Triplet Loss outperform previous versions on CIFAR-10 and is robust against noise