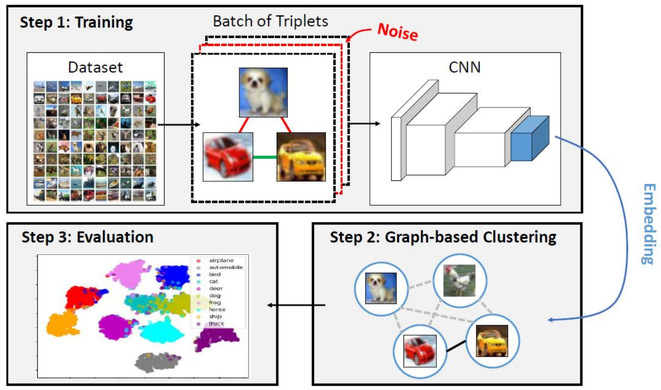
# Learning Embeddings for Image Clustering: An Empirical Study of Triplet Loss Approaches

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

[1]Fraunhofer Center Machine Learning – [2]Fraunhofer ITWM – [3]IMLA Offenburg University – [4]Data and Webscience Group, University of Mannheim

## Abstract

In this work, we evaluate two different image clustering objectives, k-means clustering and correlation clustering, in the context of Triplet Loss induced feature space embeddings. Specifically, we train a convolutional neural network to learn discriminative features by optimizing two popular versions of the Triplet Loss in order to study their clustering properties under the assumption of noisy labels.




## Study Setup

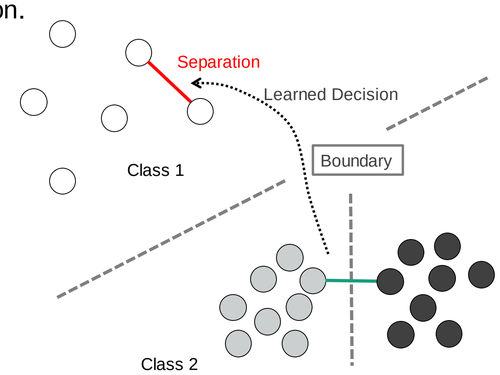
Our approach is based on the assumption that embedding features, learned from the regular Triplet Loss (1) can produce high variances in inter- and intra-cluster distances, because it only considers relative differences between the distances of positive and negative pairs. This objective is suitable for k-means clustering. Yet, the attempt to learn whether two data points should belong to the same or to a different class from their pairwise distances might fail, when the intra- and inter cluster samples are equally far away. This is shown in the following Figure where the correct decision boundaries are marked by green lines. In contrast, the red line, at the same Euclidean distance as the green lines, indicates a false separation of data. This motivates us to consider losses that preserve the distance equally between the positive pairs during the optimization.

<u>Triplet Losses</u>

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

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

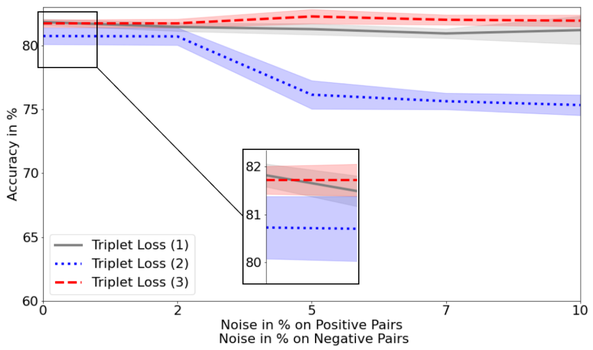
$$L_{triplet\_3} = \sum_{i=1}^{n} [\alpha - \|f(x_i^a) - f(x_j^n)\|^2]_+ + [\|f(x_i^a) - f(x_k^p)\|^2 - \beta]_+$$



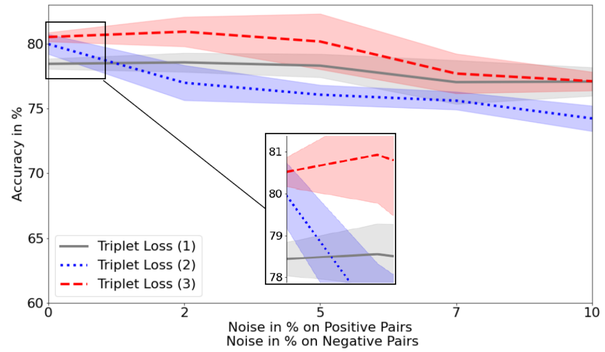

## Contribution

1) We conduct a thorough study of the clustering behavior of two clustering approaches, k-means and minimum cost multicuts, applied to learnt embedding spaces from three Triplet Loss formulations. 2) Our study reveals that, while the traditional Triplet Loss is well suited for k-means clustering, its performance drops under the looser assumptions made by minimum cost multicuts. 3) We propose a simplification of the Triplet Loss from (3), which allows to directly compute the probability of two data points for belonging to disjoint components and is robust against noise in both clustering scenarios.

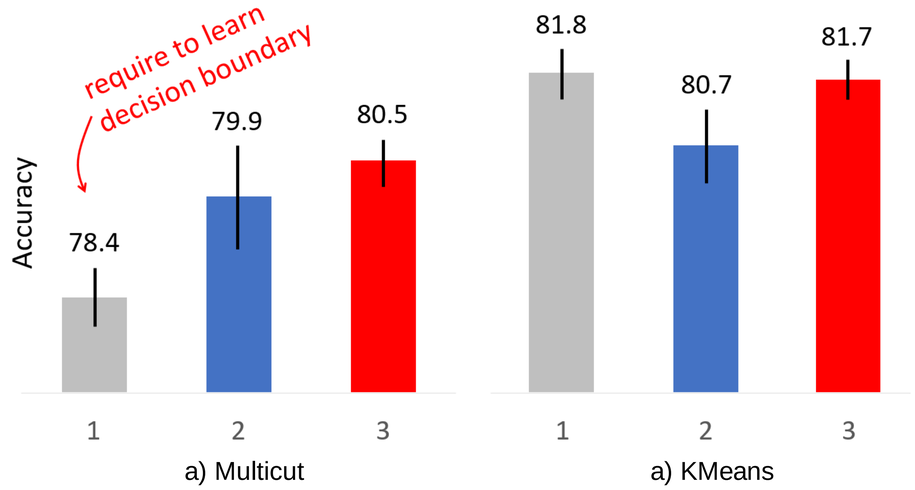
## Results




Clustering Performance with noisy data: KMeans




Clustering Performance with noisy data: Multicut




Clustering Performance: Multicut vs. KMeans