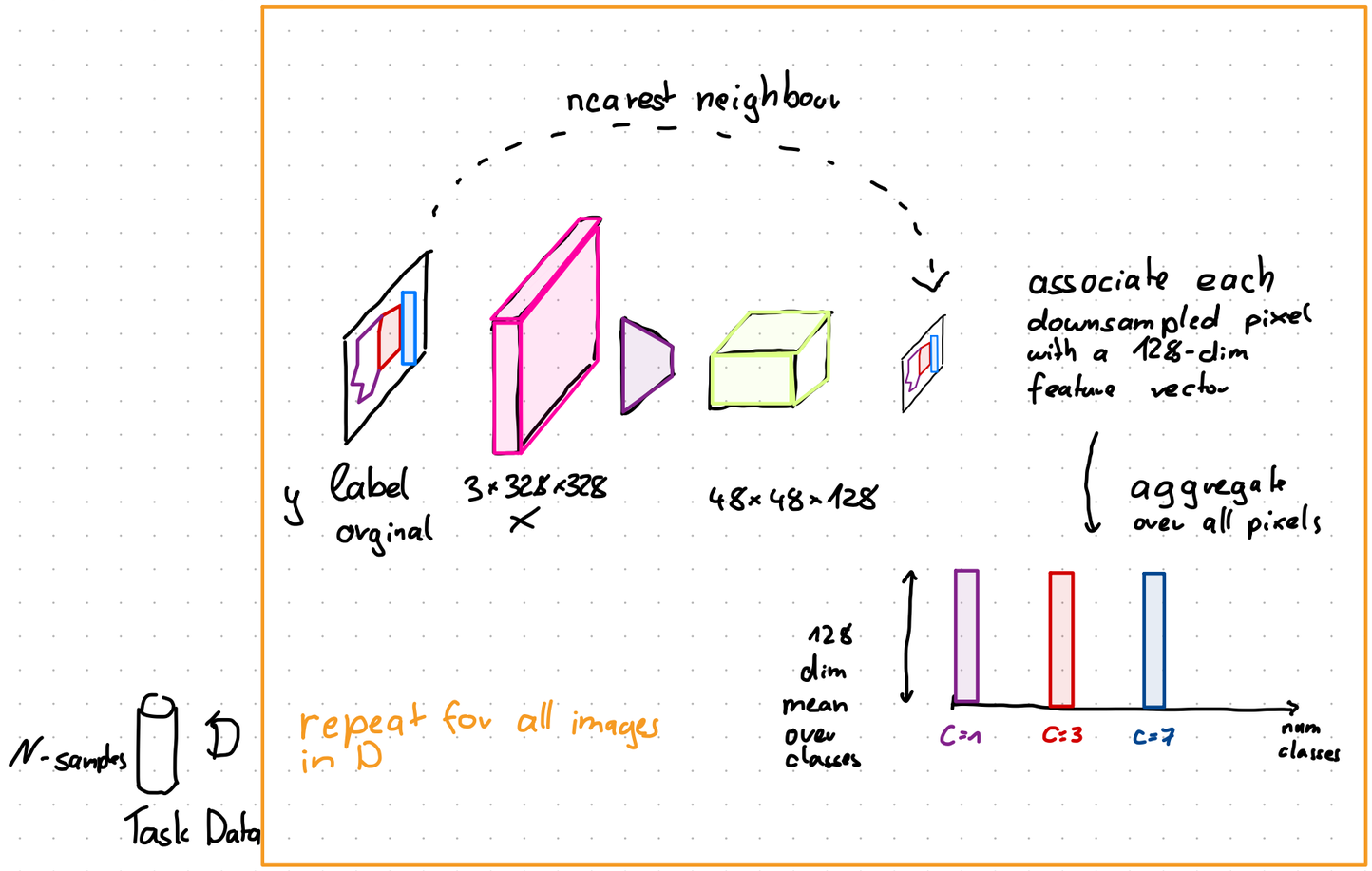

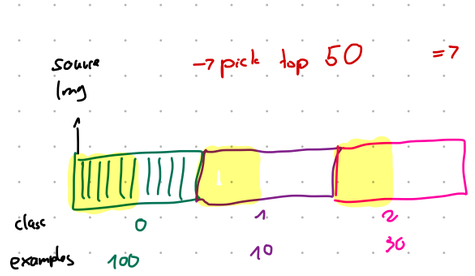

class balanced sampling Top-K Best

0 -> pick 10

1 -> .. 1

2 -> 3

Assumptions

1: Fast R-CNN latent features are connected to pixel wise label at same pose givenn CNN-archi!

2: The feature mean over an Image is meaningful (somehow clustered together even pixelwise)

$N \times NC \times 128$ dataset embedding

-> calculate classwise centroids $NC \times 128$

-> find classwise Top-K Images with highest distance to centroid

$K \times NC$ Image candidates for buffer

↳ put K' Images from Image candidates into buffer

3. Distance metric from centroid of class reflects uncertainty or indicates that a sample is on decision boundary

Aggregation

- Should we associate a latent feature to a single class or maybe all possible ones in the label
- Associate them according to soft-max output

Distance Metric

- cosine-similarity -> do we need length invariance
- some norm

Ideas:

- incooperate the network predicted label only use correctly classified when aggregating
- measure distance to other clusters