UNSUPERVISED REPRESENTATION LEARNING FOR MEDICAL IMAGING

By Albert Garcia Sanchez

WHAT IS IT THAT MOVES ME?

- Data Science
- Artificial Intelligence
- **Machine Learning & Deep Learning**

MACHINE LEARNING

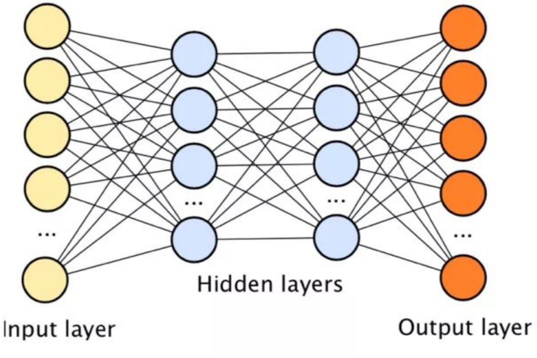
Enables the ability to develop techniques and models that learn from data

Artificial Neural Network is a biologically-inspired programming paradigm which has the ability to learn from observing data

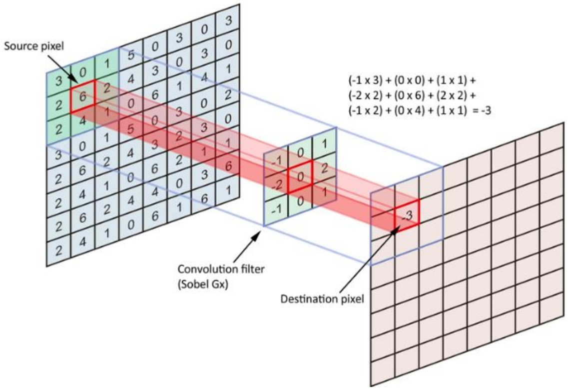
Broadly used due to the increase in data availability

DEEP LEARNING & CNNs

Deep Learning is a subfield of Machine Learning which has yielded great improvements



A Convolutional Neural Network (CNN) is a class of Deep Neural Network commonly applied to processing visual imagery



FINDING THE THESIS

Computer Vision based working with CNNs

The challenge of domain adaptation in a highly specialized image domain: gastrointestinal tract (GI)

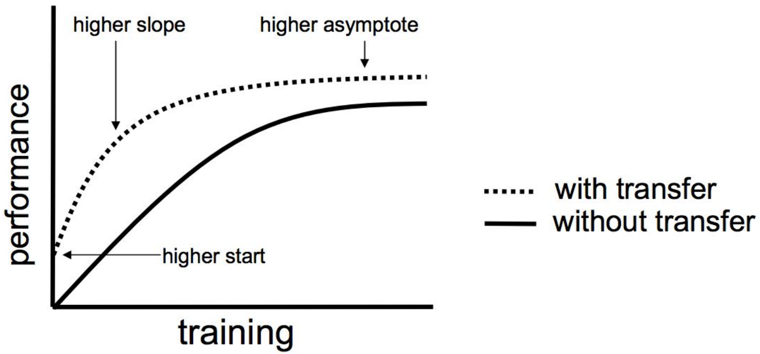
Human health evolution for the detection and prevention of polyps which develop into cancer

WIRELESS CAPSULE ENDOSCOPY (WCE)

A capsule provided with a camera which allows inner visualization of the entire GI tract with a minimal invasion for the patient

GOAL OF THE THESIS

Make Transfer Learning possible in the domain of WCE data through the implementation of a CNN



WHY THIS GOAL

ImageNet transfer learning underperforms when applied to specialized domains for obvious reasons

STEPS FOR A CNN IMPLEMENTATION

- Define the architecture
- Define the loss function
- Fetch and prepare the training and test data

RESNET50 ARCHITECTURE

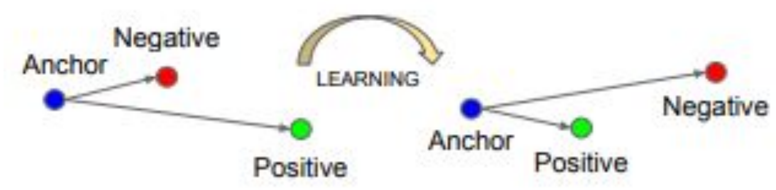
- State of the art CNN
- Widely used for its results in image recognition among other tasks
- Output layer produces a 2048 sized embedding
- Data Augmentation techniques implemented for comparison of the results

TRIPLET LOSS FUNCTION

The embedding of an image is a vector of *n* real values which represent the image

Given three image's embeddings A, P and N where A-P have the same identification and A-N have different ones

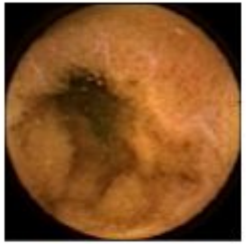
$$\mathcal{L}_{Triplet-Loss} = max(||A - P|| - ||A - N|| + \alpha,\ 0)$$

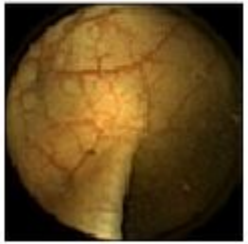


UNLABELED DATA FROM WCE

Two critical issues must be attended:

- Unlabeled images
- Undetermined number of classes



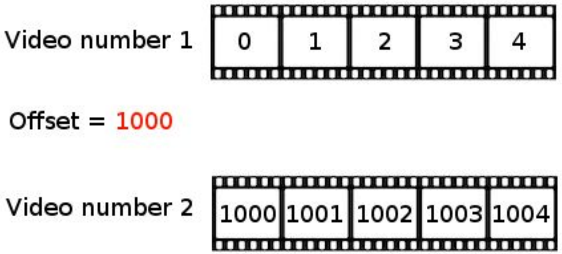



DATA PREPROCESSING

Labeling process:

- Supposing the images come from a continuous video then each image is labeled with its frame number

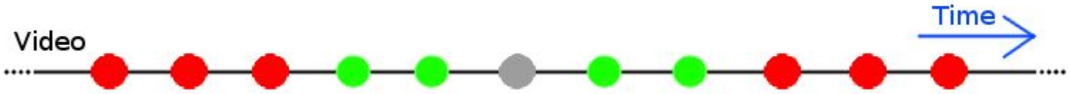
- With severals videos an offset is added



ADAPTING THE TRIPLET LOSS

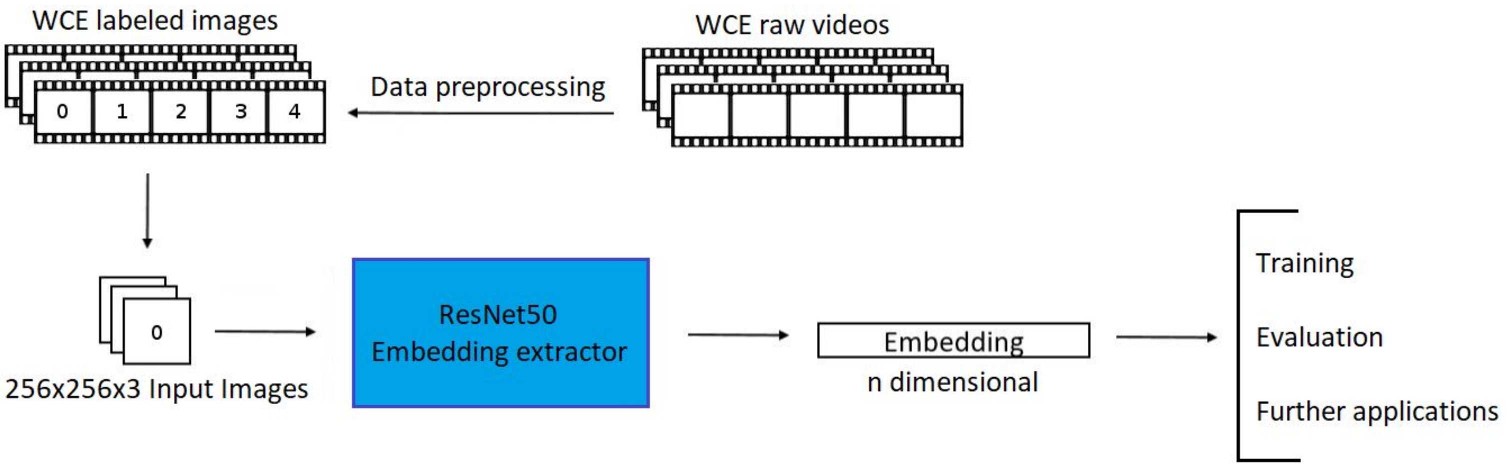
The identification of positives (P) and negatives (N) with respect to an anchor (A) must be modified

The positives for a given anchor are all the images whose label fall within the range (x-**ε**, x+**ε**) where x is the label of A and **ε** is a constant positive integer



$$|x - \text{label}| < \varepsilon$$

GENERAL PIPELINE PROPOSED



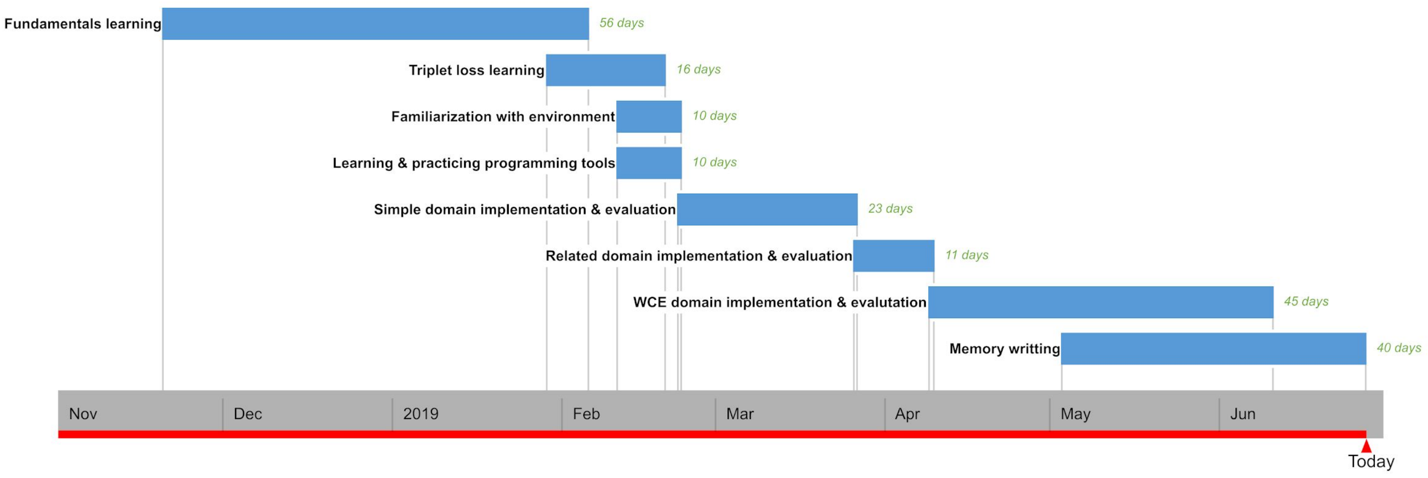
TECHNOLOGY USED

python™

colab

TensorFlow

TASKS & PLANNING



INCREASING IMPLEMENTATIONS

Formula 1 domain

First contact with the implementation of the triplet loss and a simple architecture.

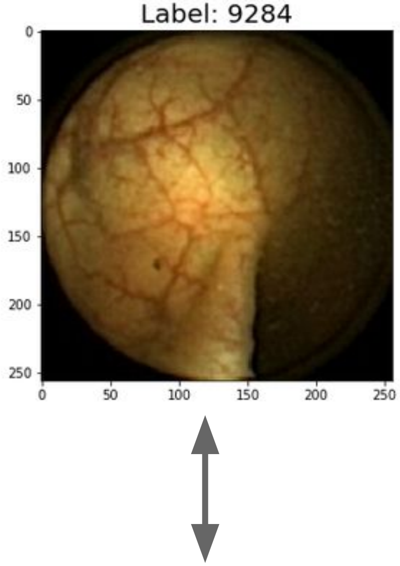
Colon domain

Transition implementation. Can be seen as a reduced domain within the WCE domain.

WCE domain

Final implementation with the ResNet50 architecture.

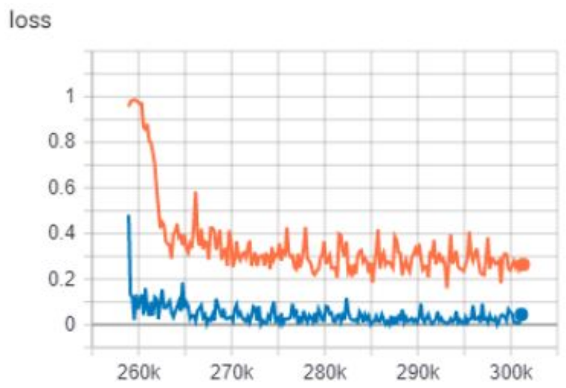
HOW TO EVALUATE THE RESULTS?



[0.278, 1.369, 5.136, -12.36, 7.336, ...]

TRIPLET LOSS LACK OF INSIGHTS

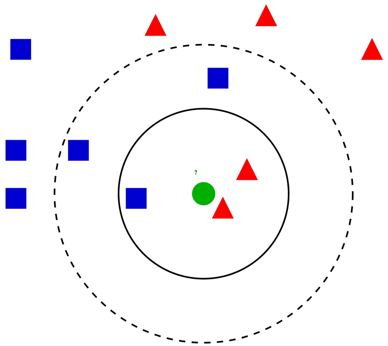
Triplet loss fails to provide information on whether the embeddings are improving in late stages of the training process



KNN BASED EVALUATIONS

KNN is an algorithm which computes the *K* nearest neighbours for a given sample within a set of samples

The sample set in this case is the set of all the image's embeddings contained in $\mathbb{R}^n$



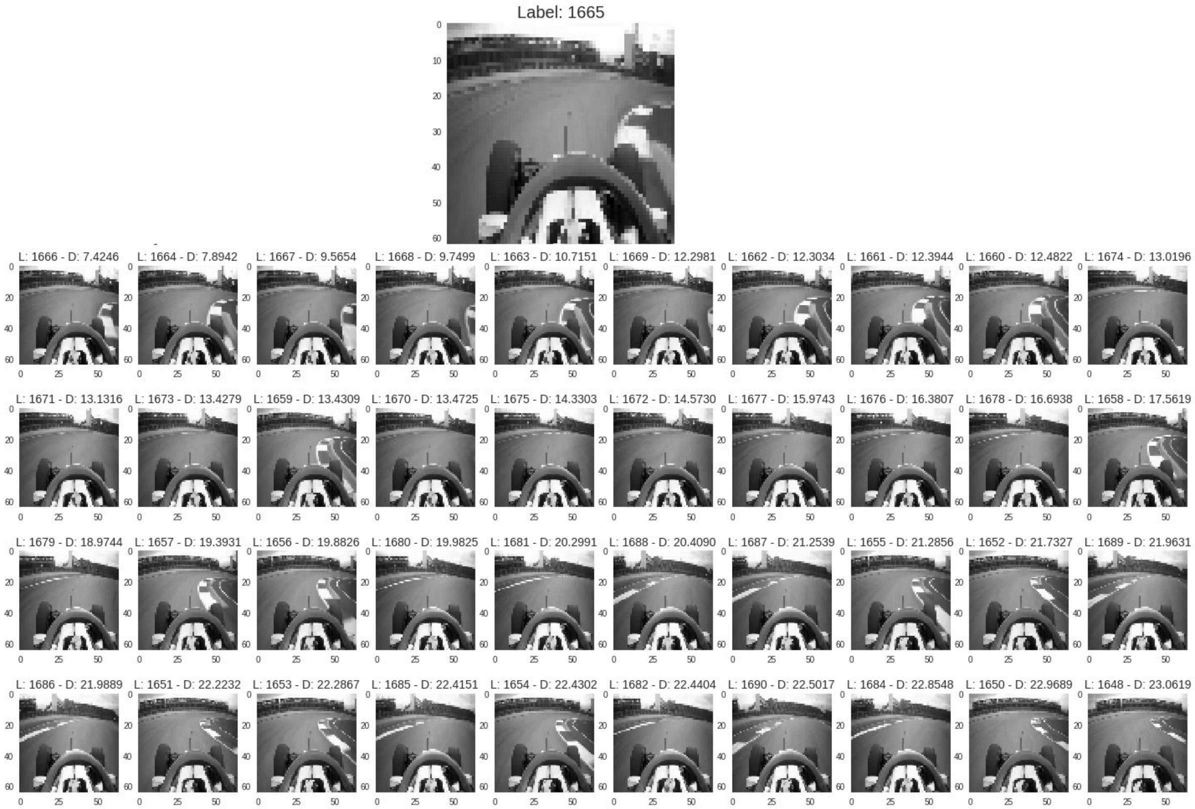
SIMPLE KNN SCORE (DEFINITION)

- For each anchor image, the KNN is computed with $K=2\varepsilon-2$ which is the maximum number of positives.
- Count the number of positives among the KNN
- Average this counter over all the images and compute the % ratio

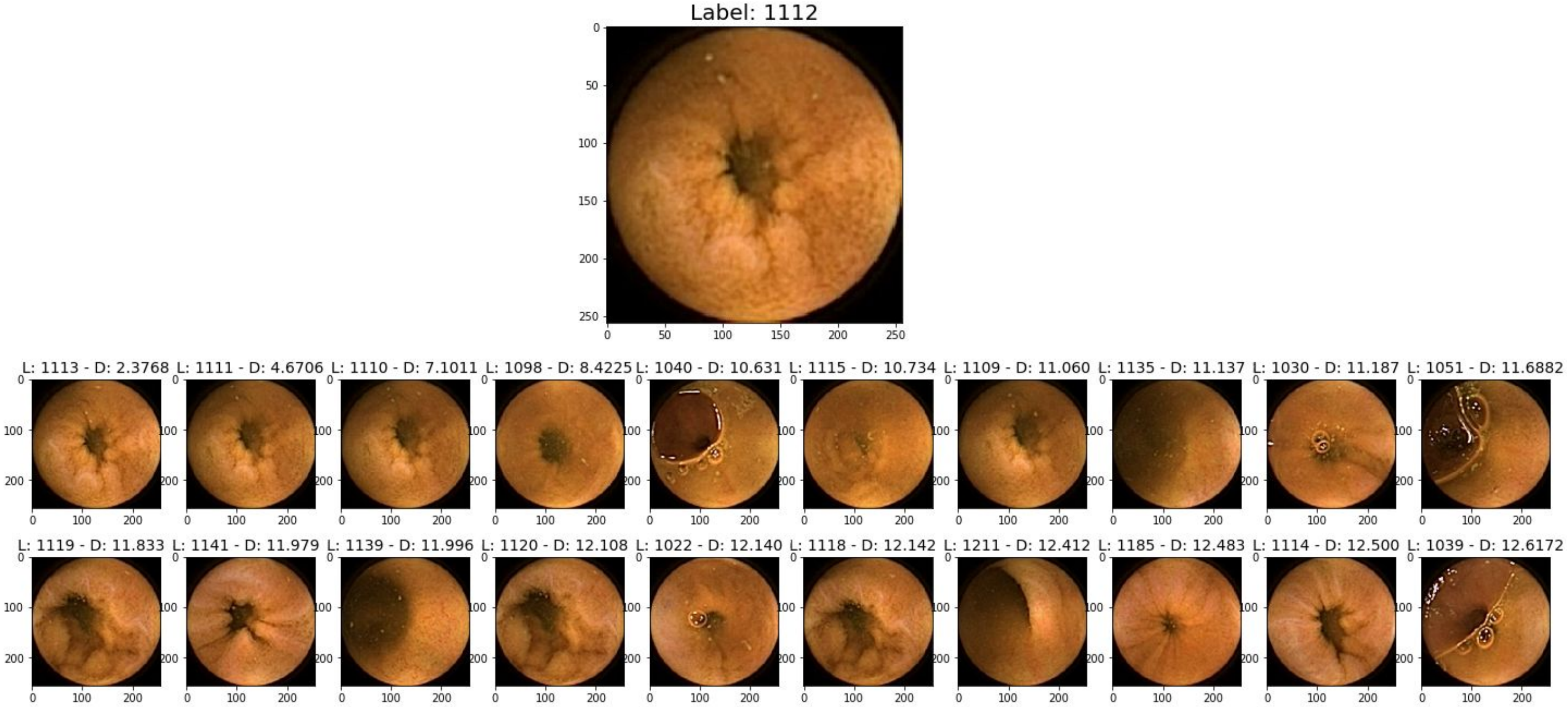
SIMPLE KNN SCORE (RESULTS)

	Training set	Test set
Formula 1 model	86.39%	69.96%
Colon model	22.86%	13.77%
WCE NDA model	36.36%	18.31%
WCE DA model	28.07%	16.19%

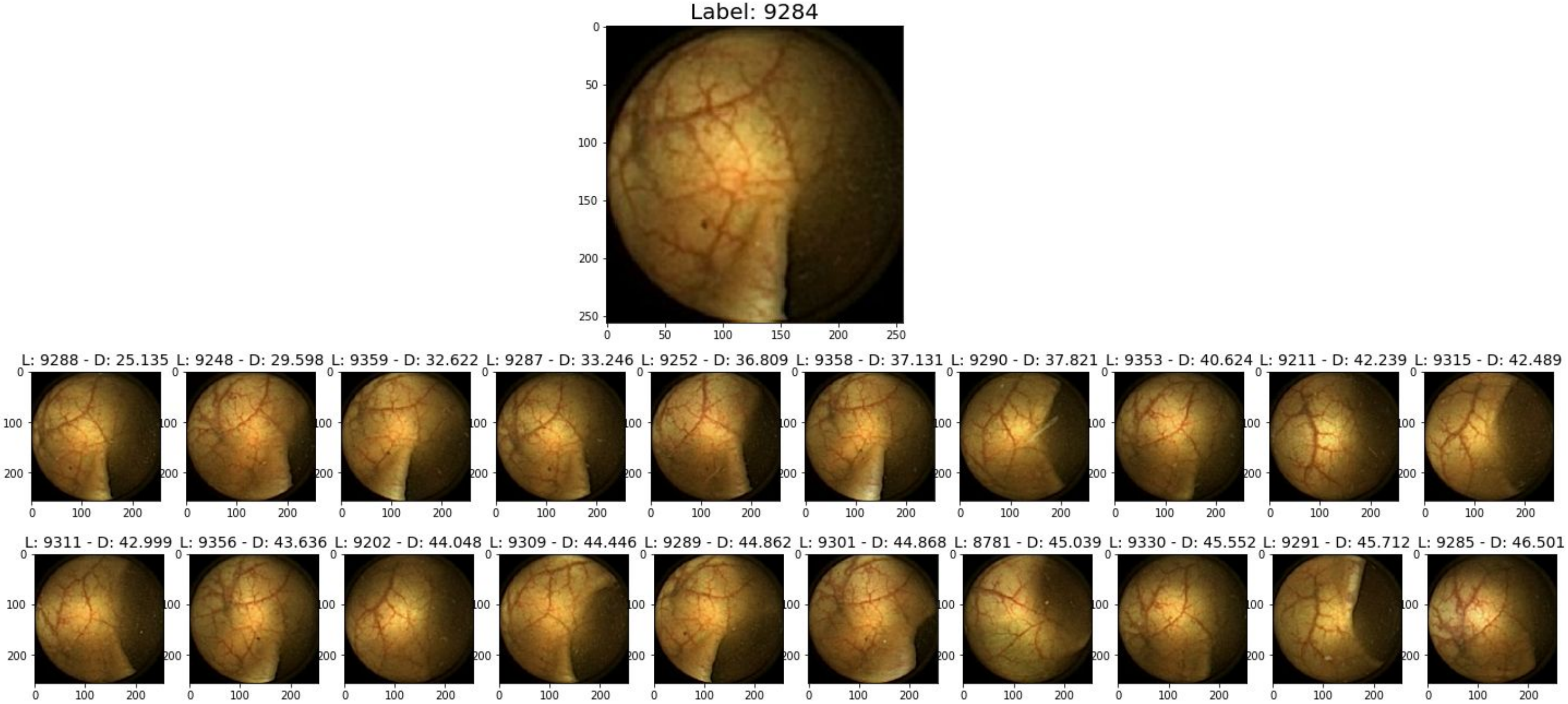
KNN QUALITATIVE RESULTS – FORMULA 1



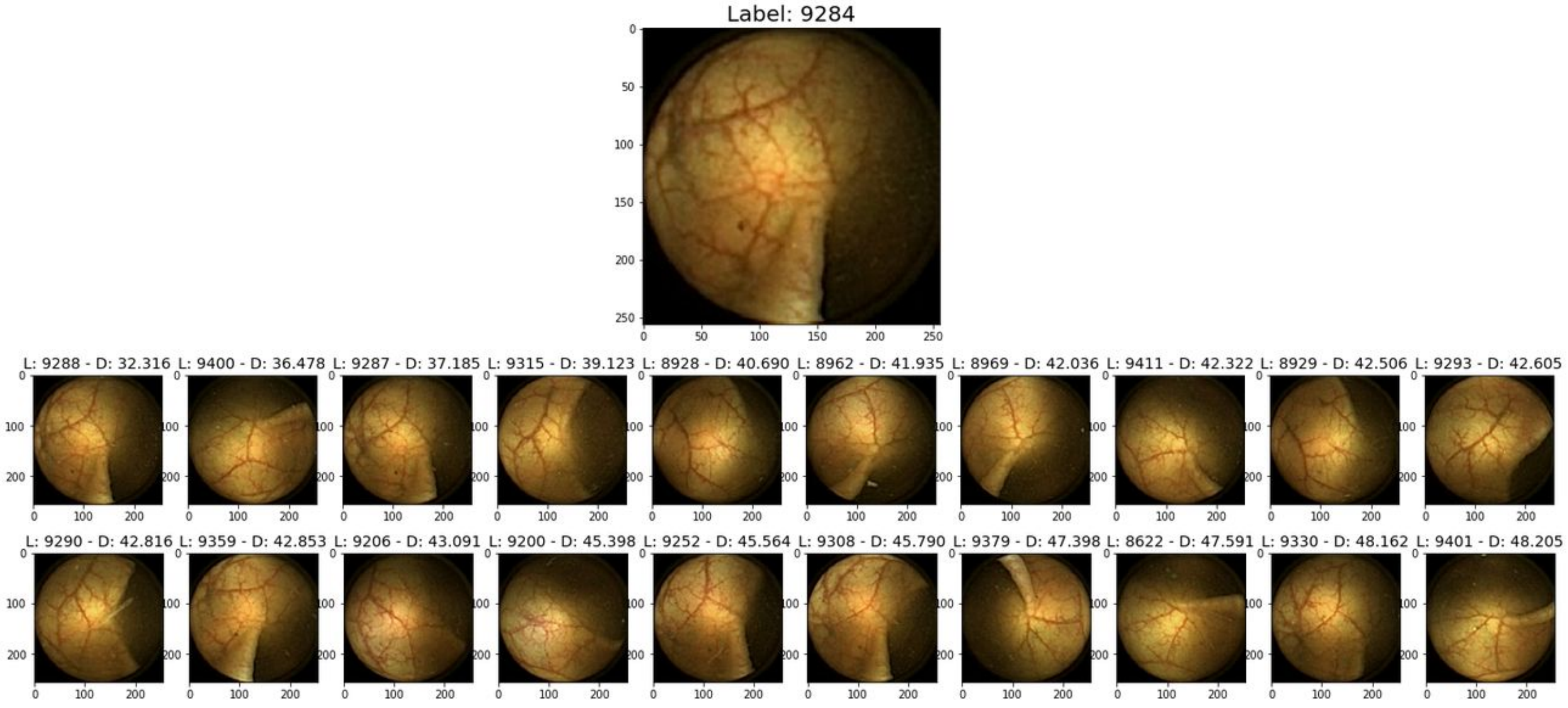
KNN QUALITATIVE RESULTS – COLON



KNN QUALITATIVE RESULTS – WCE NDA

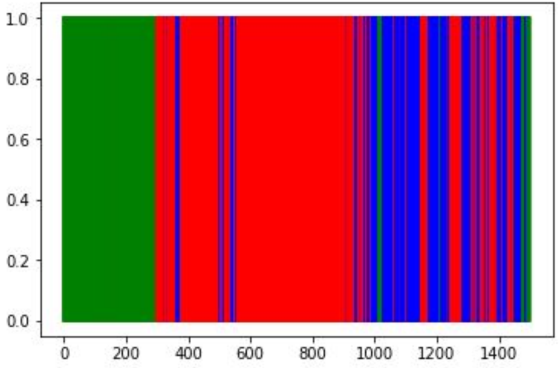


KNN QUALITATIVE RESULTS – WCE DA



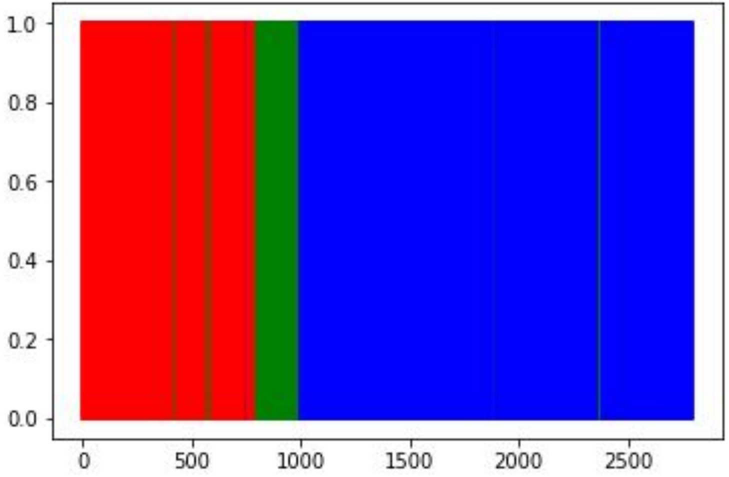
CLUSTERING EVALUATION – COLON

- Ideal clusters predefined
- Clustering embeddings with K-Means approach



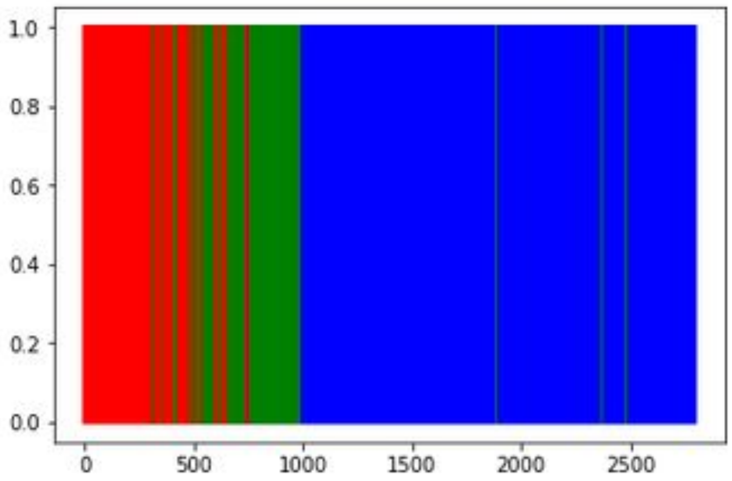
Cluster 1: 0-300 Cluster 2: 301-900 Cluster 3: 901 - 1500

CLUSTERING EVALUATION – WCE NDA



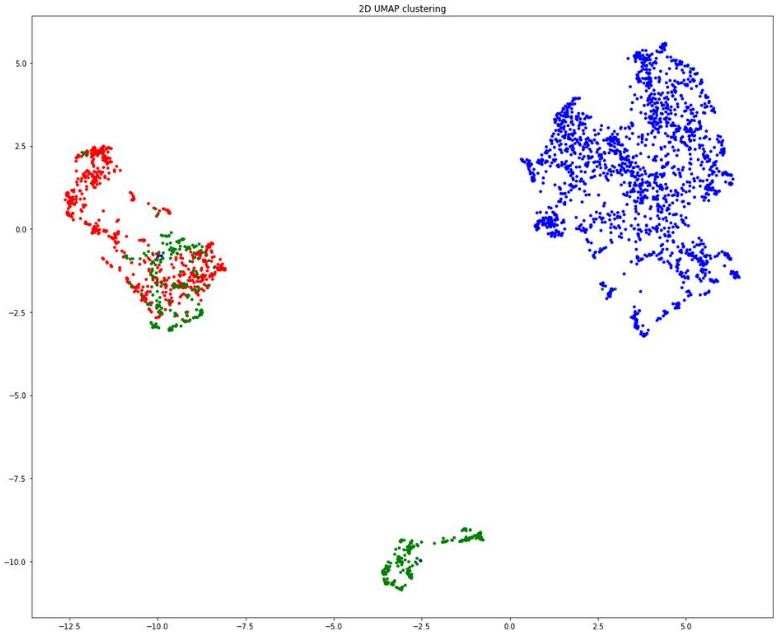
Cluster 1: 0-800 Cluster 2: 801 - 1000 Cluster 3: 1001 - 2796

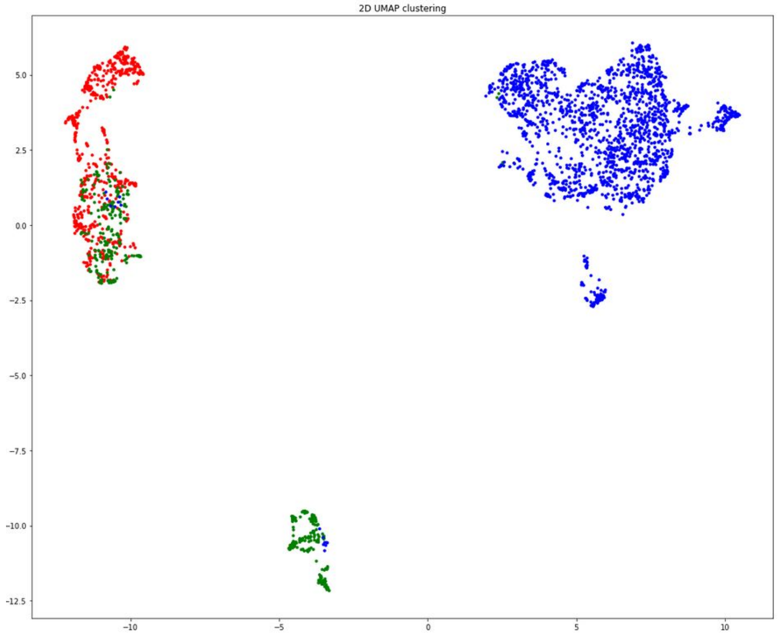
CLUSTERING EVALUATION – WCE DA



Cluster 1: 0-800   Cluster 2: 801 - 1000   Cluster 3: 1001 - 2796

EMBEDDINGS PROJECTION – NDA vs DA





CONCLUSIVE RESULTS

- Evaluations over a labeled WCE dataset

- Rank-1 evaluation consists in computing KNN with $k=1$ for each image of the dataset and checking whether the closest image has the same label.

- A classifier model is trained from the embeddings of a pretrained embedding extractor CNN

CLASSIFIER & RANK-1 RESULTS

	Training accuracy	Test accuracy	Rank-1 accuracy
Random init	68.55%	41.12%	65.86%
ImageNet init	25.64%	35.47%	63.17%
Triplet loss NDA	74.14%	67.68%	77.52%
Triplet loss DA	**80.20%**	**70%**	**78.15%**

Thank you for your time