LUS Images classification with uncertainty detection and image similarity

Medical Imaging Diagnostic

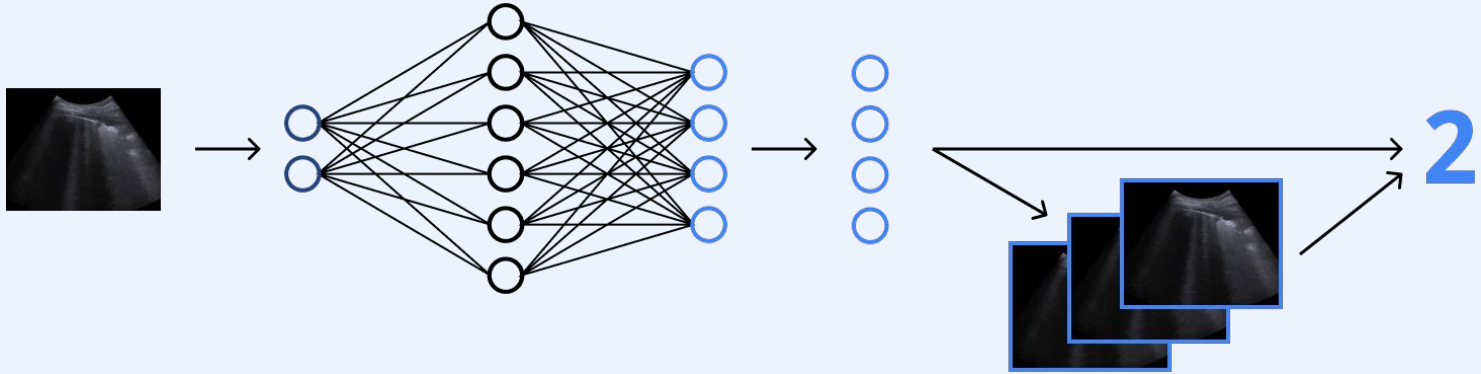
Davide Modolo - Università degli Studi di Trento

0

Quick overview

1 2 3

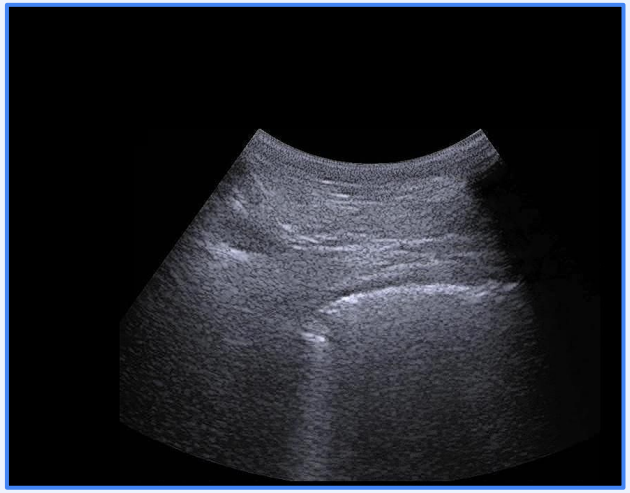
1. Frame into a classification model
2. Softmax output is analyzed by a classifier
3. If *false*, analyze the closer images score
4. Final score



1 Briefly on data

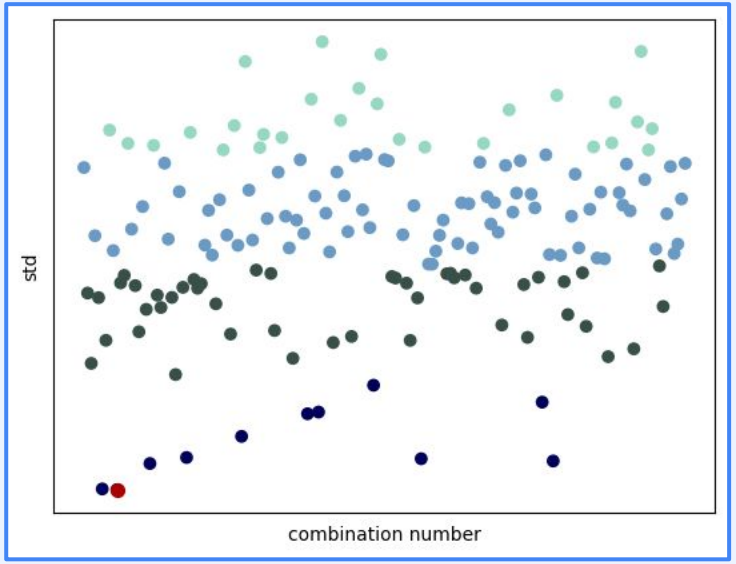
Lung Ultrasound Images

1. Images are scored from 0 to 3
2. This project works frame-wise
3. Augmentation was required
4. They seems in grayscale, but images are RGB



Dividing patients

To balance training, computed all possible 8-patients combinations stds between the number of frames per score -> took the smallest one



0	1	2	3
7888	7540	7189	7592

2

First classifier

3 4 5

Choosing the model

- ResNet18
- VGG16
- SqueezeNet
- CNN from scratch

Added layers to fine-tune, but easy overfitting. Working solution was to only add a layer to output the 4 classes

Augmentation

“Deep learning for classification and localization of covid-19 markers in point-of-care lung ultrasound”

E. Torri, R. Inchingolo, A. Smargiassi, G. Soldati, P. Rota, A. Passerini, R. J. G. van Sloun, E. Ricci, and L. Demi

- Affine transformations
- Multiplication with a constant
- Gaussian blurring
- Horizontal flipping

3

Confidence

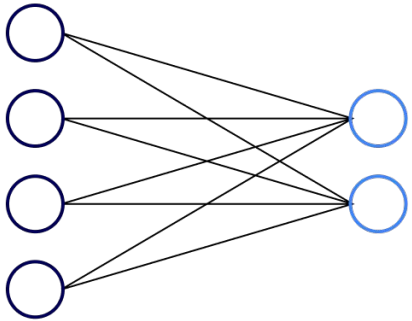
Threshold?

Independently from correct or wrong prediction and from the scores, the highest confidence in the softmax values were very similar

Understand the behaviour

Saving the softmax values and the prediction correctness from the fine-tuned ResNet18

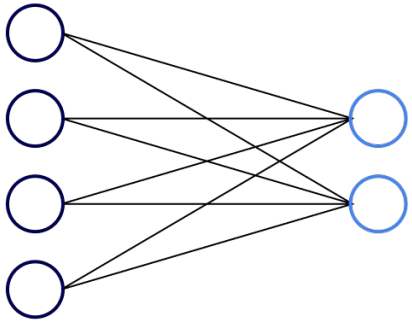
Neural Network



Sigmoid activation function

Basic model, any addition layer resulted in worse performance

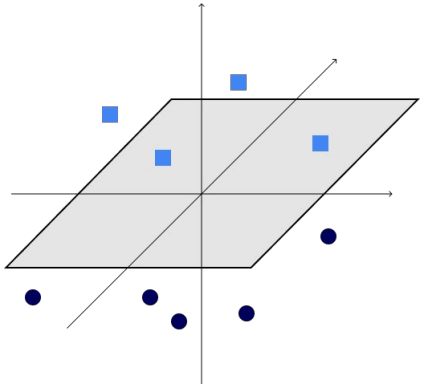
Neural Network



Sigmoid activation function

Basic model, any addition layer resulted in worse performance

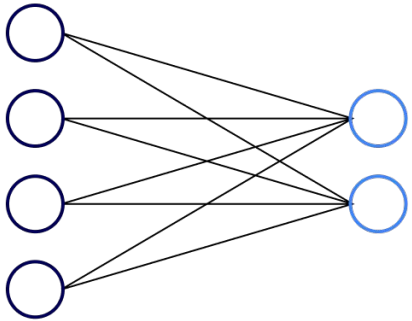
SVC



Support Vector Classifier

It required perfectly balanced dataset

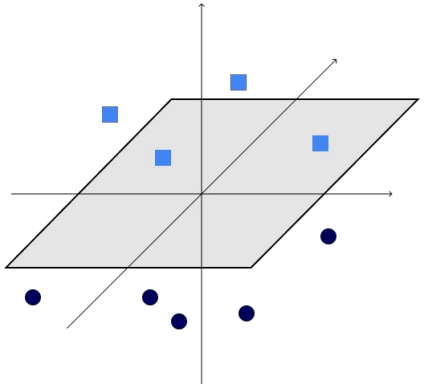
Neural Network



Sigmoid activation function

Basic model, any addition layer resulted in worse performance

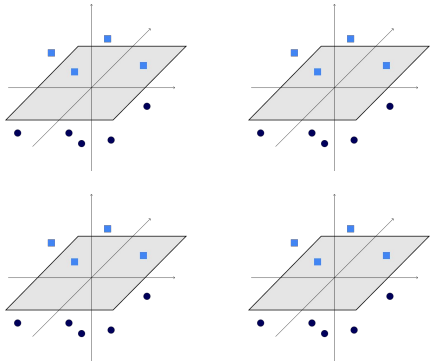
SVC



Support Vector Classifier

It required perfectly balanced dataset

4 SVCs



One for each class

The SVC used depended on the predicted score

4

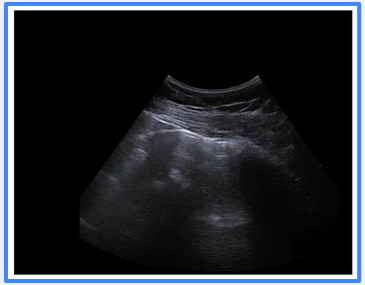
Similarity

Searching for similar images

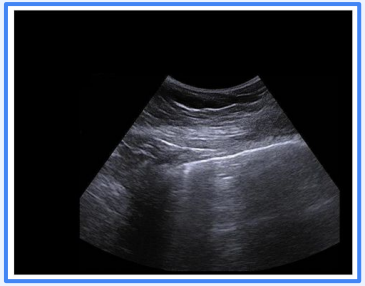
Only a portion of the available images from the train set has been used

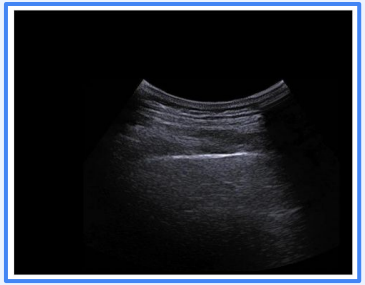
The same number of images for each score for each patient from the training set and again in the test set

QUERY



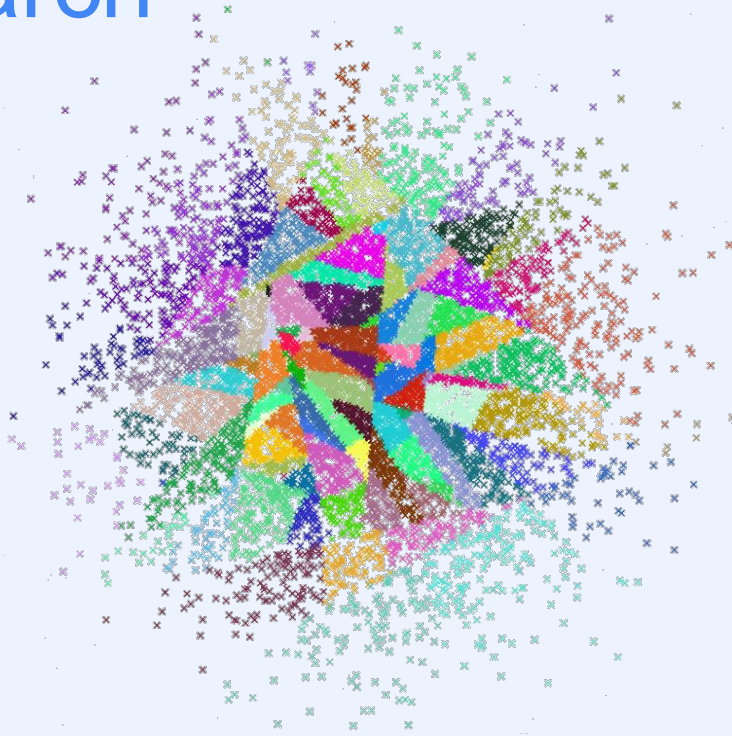
ANSWERS





Near Duplicate Image Search

Technique to find similar images using a nearest neighbour algorithm

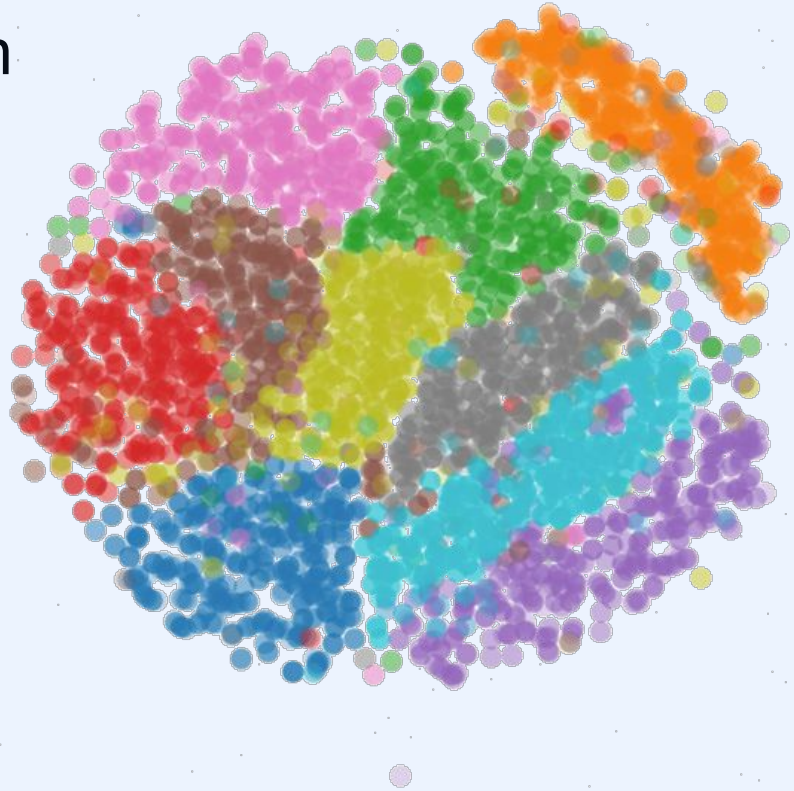


1. Convert each image into a high-dimensional feature vector
2. The feature vector is indexed using the Annoy library
3. Retrieval of the most similar images based on their Euclidean distance

t-SNE (t-Distributed Stochastic Neighbor Embedding)

Algorithm used for data visualization

It maps high-dimensional data points into a low-dimensional space, while preserving the pairwise similarities between the points



Raw Images

Flattened, worst results

Softmax

not based on visually similar images, best results

Embeddings

Taken from the fine-tuned ResNet18

Emb. alone

No t-SNE, cosine similarity

mean, *minimum*, *maximum*, *mode*, and *median* values were computed across the scores of 1 to 15 neighbours

5

Final results

Final Model

Many possible configurations tested

Behavior of each set of three modules was quite similar

Chose the three modules with best results on their own:

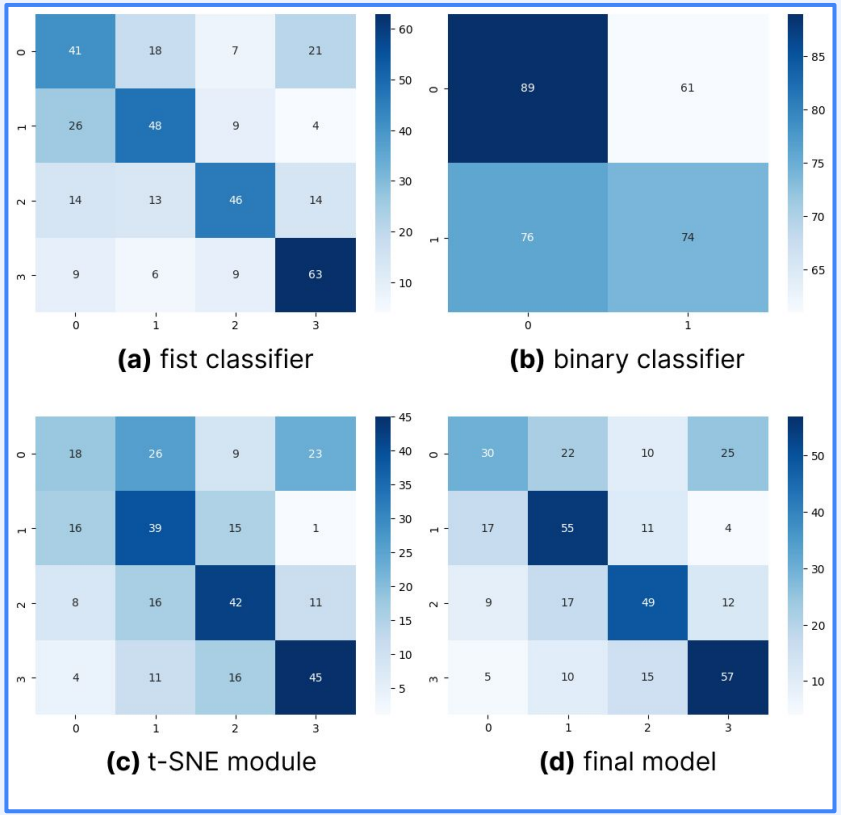
- **ResNet18**
- **Single SVC**
- **t-SNE softmax version**

Results

Final accuracy similar to the first classifier alone

It loses some accuracy in classes 0 and 3 to gain in 1 and 2

ResNet	SVC	t-SNE	Final
56.9%	54.33%	48%	54.89%



Binary Classifier Analysis

Right prediction on 3059/17512 images (the entire test set)

- getting almost always the correct result for the label 1
- less than 25% of the time with label 2
- and very few times with labels 0 and 3

t-SNE Analysis

It uses only a small subset of the training images

Maybe, predictions overlap with the wrong results from the with wrong results from binary

Conclusions

- There isn’t a single component responsible for the model’s lack of success
- Having multiple potential points of failure could have contributed to the final performance
- Overall, the proposed model needs further improvements and adjustments to achieve better or more valuable results

And that’s it, thank you

Medical Imaging Diagnostic

Davide Modolo - Università degli Studi di Trento