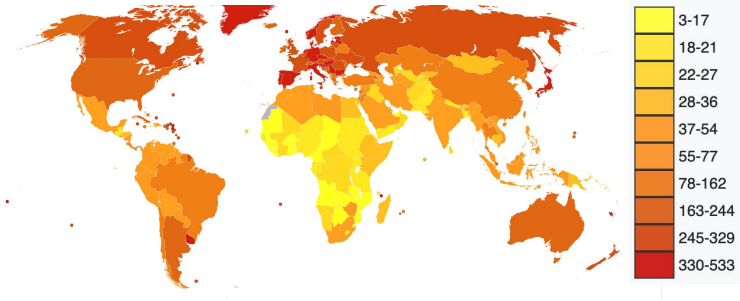
Classification of gastrointestinal lesions for resection

Project group 10

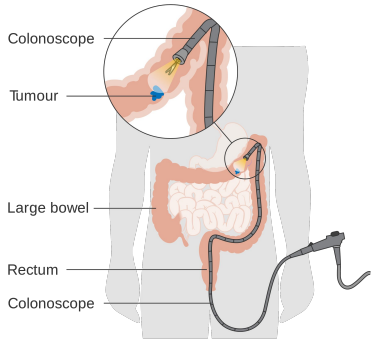
Athena Taymourtash
Bodnár Dániel
Alice Santilli

Objective – problem specification

- Colorectal cancer and the importance of the accuracy and the promptness of the diagnosis
  - Cause: Old age, lifestyle factors, genetic disorders
  - Diagnostic method: Tissue biopsy and colonoscopy



- Virtual Colonoscopy
  - tell surgeons in real time whether or not the lesion needs to be resected
  - discriminate the severity of individual lesions in patients with many polyps in real time



Examples of the available images

- videos, 25 frames/second, 768x576 pixels per frame
- each video length is 5 - 30 seconds
- 76 colonoscopy videos, both in NBI (narrow band imaging) and WL (white light)
- 15 serrated adenomas, 21 hyperplastic lesions and 40 adenomas
- Labels :
  Adenoma + serrated = resection
  hyperplastic = no resection

	Human Expert	Human Beginner
Acc.	76.54%	72.22%
Sen.	64.98%	54.50%
Spec.	81.35%	77.30%

Model Architecture

- By considering a video as a series of frames:

1. Loop over all extracted frames in the video file
2. For each frame, pass the frame through the CNN
3. Classify each frame *individually* and *independently* of each other
4. Choose the label with the largest corresponding probability

- For test data:
  - Computing the average of the predictions for all frames of a video and choosing the label with the largest corresponding probability

Data separation:

Training

(D:) › Gastro-Intestinal-Lesions › data › train

adenoma_01_WL
adenoma_02_WL
adenoma_03_WL
adenoma_04_WL
adenoma_05_WL
adenoma_06_WL
adenoma_07_WL
adenoma_08_WL
adenoma_09_WL
adenoma_13_WL
adenoma_14_WL
adenoma_15_WL
adenoma_16_WL
adenoma_17_WL
adenoma_18_WL
adenoma_19_WL
adenoma_20_WL
adenoma_22_WL
adenoma_24_WL
hyperplasic_08_WL
hyperplasic_09_WL
hyperplasic_13_WL
hyperplasic_14_WL
hyperplasic_15_WL
hyperplasic_16_WL
hyperplasic_18_WL
hyperplasic_19_WL
serrated_01_WL
serrated_02_WL
serrated_03_WL
serrated_04_WL
serrated_05_WL
serrated_06_WL
serrated_07_WL
serrated_08_WL
serrated_13_WL
serrated_14_WL
serrated_15_WL
adenoma_28_WL
adenoma_29_WL
adenoma_34_WL
adenoma_37_WL
hyperplasic_01_WL
hyperplasic_02_WL
hyperplasic_03_WL
hyperplasic_04_WL
hyperplasic_06_WL
hyperplasic_07_WL

Validation

(D:) › Gastro-Intestinal-Lesions › data › val

adenoma_10_WL
adenoma_11_WL
adenoma_12_WL
adenoma_25_WL
adenoma_30_WL
adenoma_31_WL
adenoma_39_WL
adenoma_40_WL
hyperplasic_10_WL
hyperplasic_11_WL
hyperplasic_12_WL
hyperplasic_20_WL
hyperplasic_21_WL
serrated_10_WL
serrated_12_WL

Test

(D:) › Gastro-Intestinal-Lesions › data › test


adenoma_21_WL

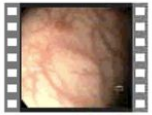

adenoma_23_WL


adenoma_26_WL


hyperplasic_05_WL


hyperplasic_17_WL


serrated_09_WL


serrated_11_WL

Preprocessing

Colonoscopy_Train.py

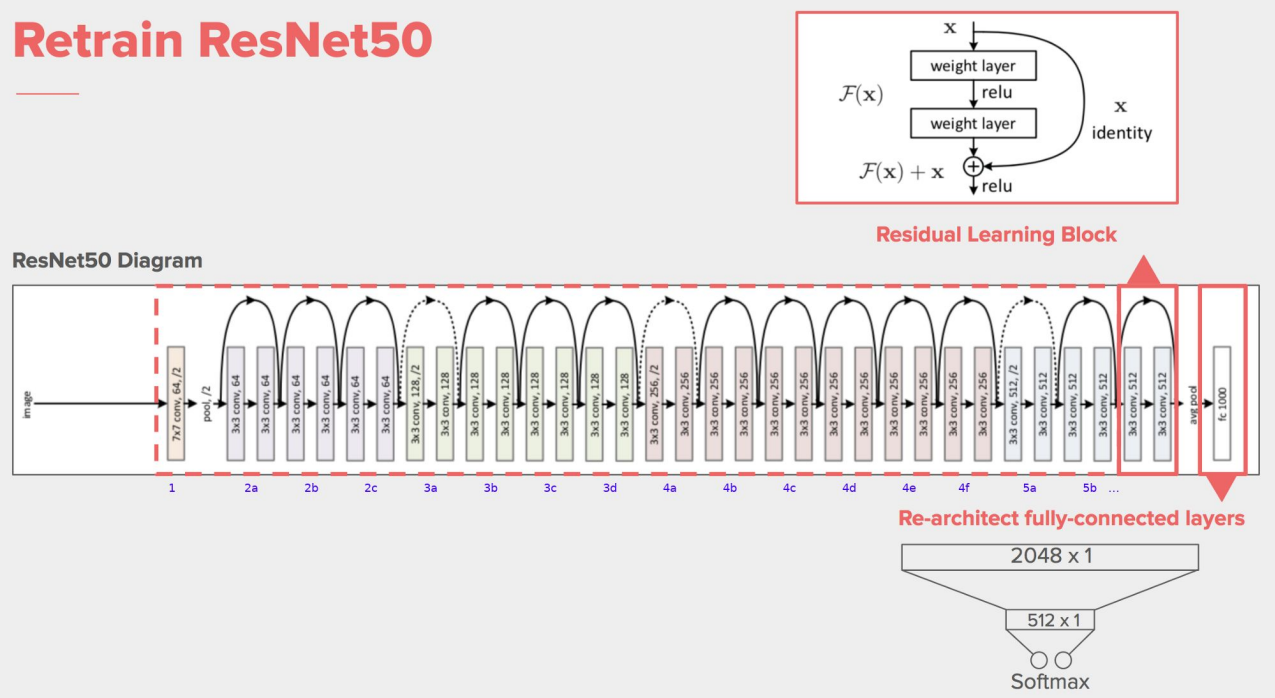
```
train_data, train_labels = DataLoader("train", datapath=args.dataset, step=10)
val_data, val_labels = DataLoader("val", datapath=args.dataset, step=10)
```

```
def DataLoader(what, datapath, step):

    STEP = step
    frames = []
    labels = []
    f_counter = 0

    if what == "train":
        print("loading training data......")
        directory = datapath + 'train/'

    if what == "val":
        print("loading validation data.....")
        directory = datapath + 'val/'
```

```
for filename in os.listdir(directory):

    video = cv2.VideoCapture(directory + filename)
    name = filename.split("_")[0] #get the class of the video
    if name=="adenoma" or name=="serrated":
        label = [0,1]
    elif name=='hyperplasic':
        label = [1,0] #one hot encoded labels
    else :
        print("issue with filename " + filename)

    while True:
        has_frame, frame = video.read()

        if not has_frame:
            break

        if f_counter % STEP == 0:
            frames.append(frame)
            labels.append(label)

        f_counter += 1
```

Training Model

- Transfer learning of ResNet50
- Froze all the base layers
- Fine tuned training to our dataset



Training Performance

- Short training time due to transfer learning
- 80% validation accuracy
- 97% training accuracy

Results

Predictions were made for the entire video

Test on 7 full videos :

5 needing resection

(*3 adenoma, 2 serrated)*

2 no resection

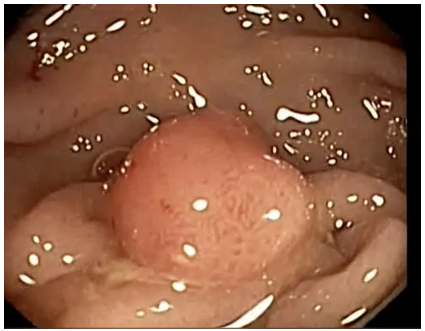
(hyperplasic polyp)

Type	True Class	Predicted Class	Confidence
Adenoma	1	1	100%
Adenoma	1	1	98%
Adenoma	1	1	99%
Serrated	1	1	97%
Serrated	1	1	51%
Hyperplasic	0	0	63%
Hyperplasic	0	0	56%

Future work

- Give surgeons information about which type of polyp they are viewing
  - This may help them make better clinical decisions
- Try different architectures or help extract more features from the images to aid the classification

Hyperplasic



Adenoma

Serrated

Conclusion

Colon Polyps and lesions can be successfully identified from colonoscopy videos.

This will give surgeons real time pathologic information about the polyps and make sure all important lesions are being resected.

Thank you