Classification of Brain Tumors Using Deep Learning Techniques

Tel Aviv University: DLMI (0553-5542)

July 2020

Marom Dadon Asaf Zorea Zrien

Program

1

Background: Brain tumors

2

Method: Image synthesis using a conditional GAN, Capsule classifier

3

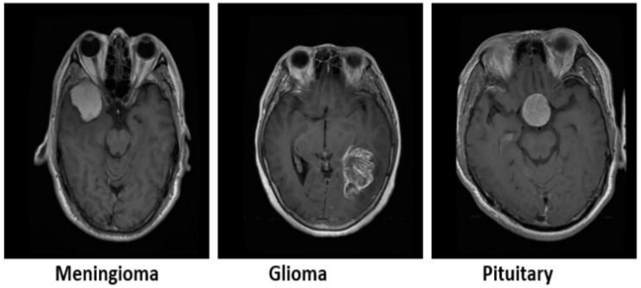
Experiments and **results**

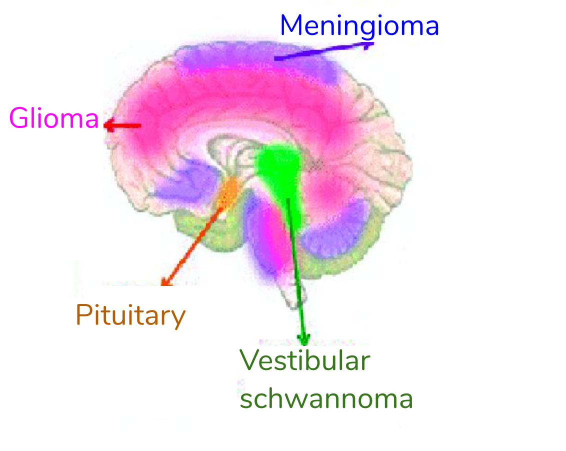
4

Conclusion and **suggestions** for future work

Background

Brain tumors can be divided into classes based on the affected area and shape, some are Glioma, Meningioma, and Pituitary.

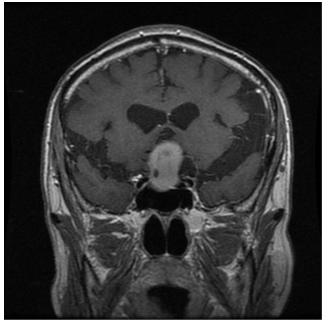




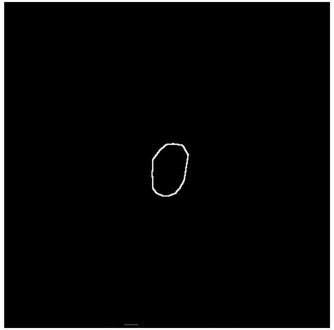
- Each of the classes has a certain level of malignancy, and can typically cause a different level of medical damage.
- Early detection of the disease will help in providing better treatments.

Method

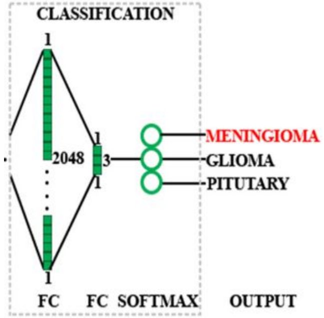
A few previous studies used machine learning models to segment and classify the tumors out of brain MRI images



2D max projection of brain

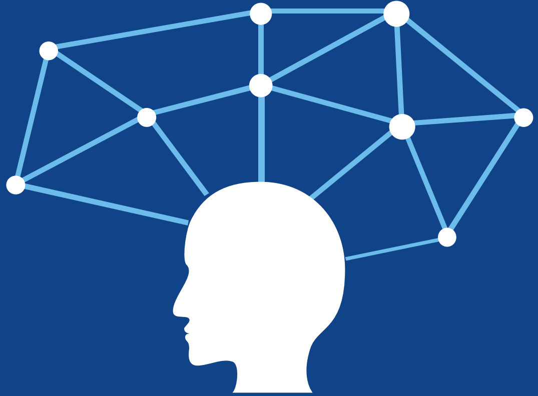


Tumor segmentation



Tumor classification

<u>Approach</u>: Generate synthetic images with a GAN, for data augmentation, to enlarge the dataset and enable to use **lighter** DL techniques for classification task



<u>Problem</u>: Shortage of data

Data

The dataset includes 3,064 brain MRI images of 233 patients, taken from different angles. Each image was given with a tumor mask and class label.

The generative models were trained on brain MRI images from the dataset. A mask of the skull and tumor was used as a condition to the GAN.

We trained three generative models, one for each kind of tumor, and generated 7,000 synthetic images for the classification model.

Overall, we used up to 80,000 augmented images to train the classification model.

Brain Tumor Dataset

(3,064 Images of Brain MRI Images)

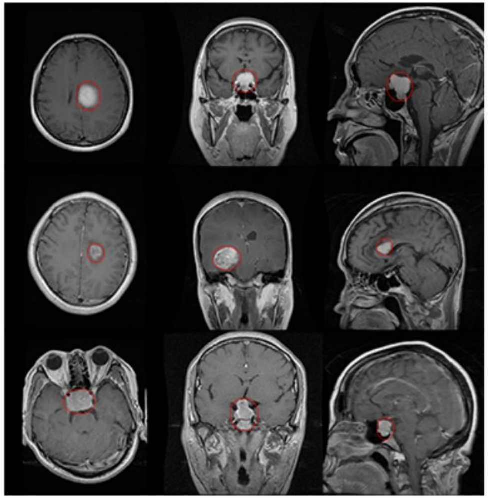
512

axial coronal sagittal

meningioma

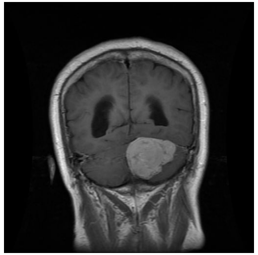
glioma

pituitary

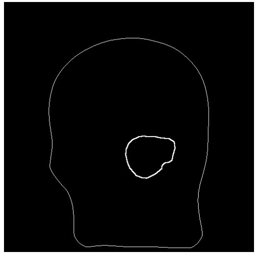


512

Brain MRI Image



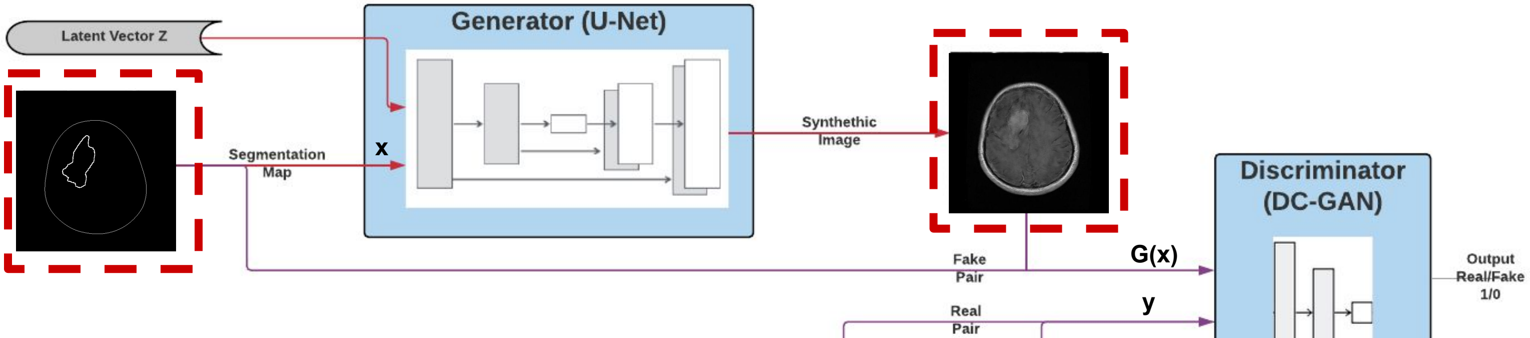
Mask extraction

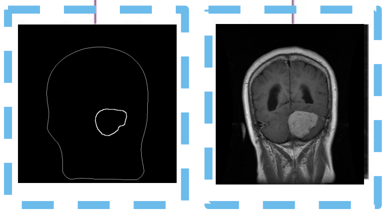


Conditional GAN [Mirza and Osindero, 2014]

The generator learns to generate a fake sample with a specific condition or characteristics, rather than an unknown noise vector

Reminder..
- ➔ Generator: generate fake images that can fool D
- ➔ Discriminator: classify fake pairs vs. real pairs





<u>Optimization Task</u>:

$$\min_G \max_D \mathbb{E}_{x,y}[\log D(G(x)) + \log(1 - D(y))]$$

Additional L1 generator loss for stable training and fast convergence

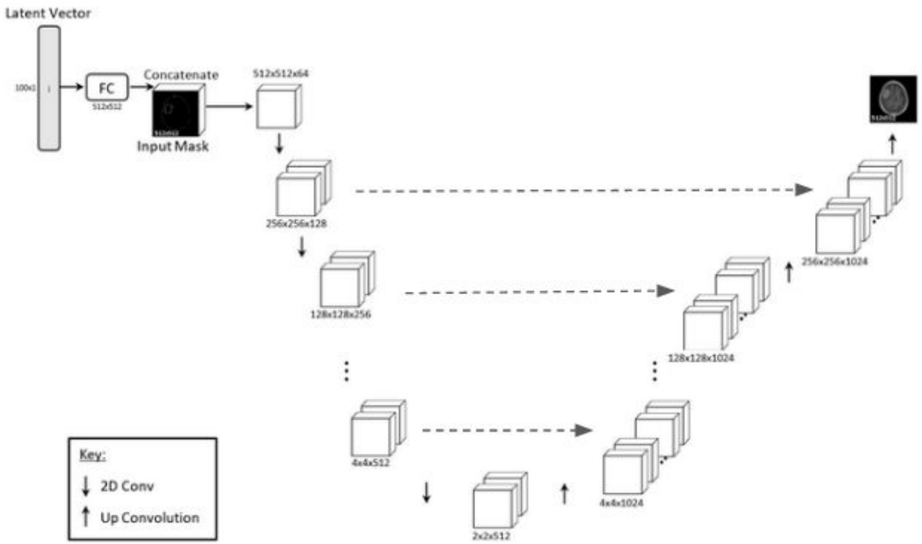
Convolutional GAN [Pix2Pix Isola et al, 2017]

➔ Built upon U-net encoder decoder network which encode both low and high level features with skip connections

Our contribute:

➔ Control the output image with a random latent vector z

➔ Reshape the label of the tumor (using rotate, shear and resize) to increase the output image variations



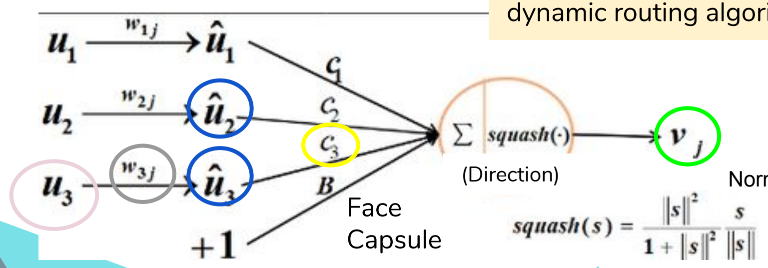
* Pix2pix network uses dropout instead of a latent vector

Capsule Network [Sabour et al., 2017]

Capsules are a ***vector*** specifying the ***features*** of the object and its ***likelihood***. These features can be any of the instantiation parameters like “pose” (position, size, orientation), hue, texture, etc.

Weight vector that encodes spatial relations between nose and face

C encode how v and u are related. Learned by the dynamic routing algorithm.

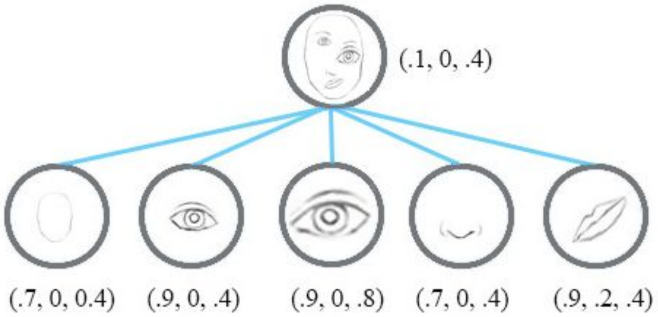


Vector output of low level capsule(nose)

Vectors that encode different features of face: noise, eye, etc..

High level features vector. Encodes probability and pose

How a CNN would classify this image?



Yes. Sub-sampling loses the spatial relations between higher-level parts such as a nose and a mouth.

Its better to aim for **equivariance** (instead of invariant): so that changes in viewpoint will lead to corresponding changes in neural activities (instead of weight).

Convolutional Capsule Network [LaLonde et al., 2018]

What is the con of capsule network?

➔ The <u>weight matrix</u> is memory intensive and limit the size of the images.

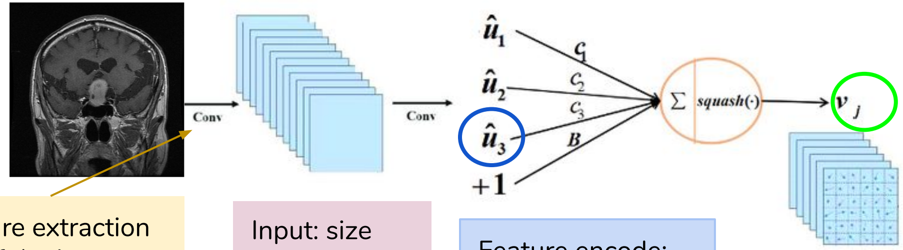
Replacing the weight matrix with <u>convolution</u> allows to reduce the computation and apply the method to larger images.

Additional Tricks:

★ Each spatial location (W,H) in the input capsule is routed to the corresponding location in the output capsules.

★ In the convolution step, the kernels Kw x KH is shared between the input capsules.

★ Squash function - normalize the output between [0,1].



Feature extraction out of the image

Input: size [B,I,Ci,Wi,Hi]

Feature encode: size [B,Wj,Hj,J,Cj]

Output is 4D vector. Each capsule is meant to represent different features or classes

** [B,I,Ci,Wi,Hi,Cj] = [Batch, Input Capsule, Channels, Weight,Height, Output Capsule]*

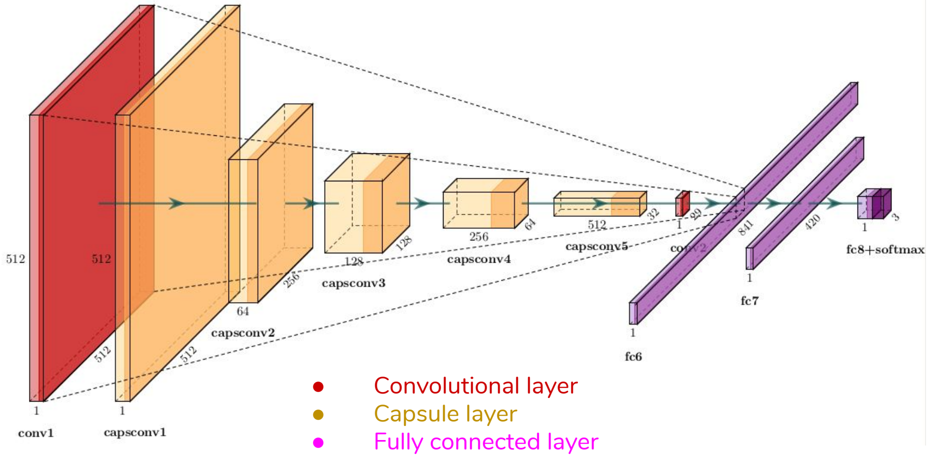
Classification model

Reminder..
ResNet is a CNN architecture which was designed to enable hundreds of convolutional layers, using skip connection to avoid vanishing gradient problem.

The classifier architecture was based on the discriminator of the generative model.

Key difference:

- ➔ We have a single image as input. The image passes a convolutional layer into a capsule unit.
- ➔ Each layer is split into blocks of convolutional capsules with a different set of channels.
- ➔ The output capsule passes a convolutional layer into fully connected layers for the classification



Capsule classifier architecture **(3.8M parameters)**

* We also tested a very deep state-of-the-art convolutional network - Resnet50 **(32M parameters).**

Evaluation

Goal: To test whether using the synthetic images could improve upon training with real images

Quantitative Analysis

Test the method by using classification task.

Train separate networks, on different datasets:

1. Real.
2. Synthetic images.

Performance was measured using:

Accuracy and F1 scores.

Qualitative Analysis

➔ Synthesised different images, for the same skull label, for each network.

➔ Linear interpolation of the latent space.

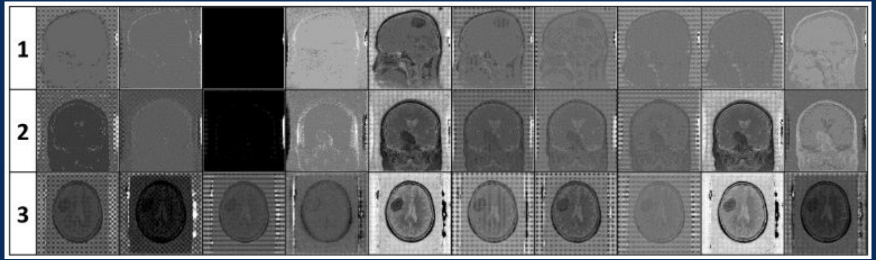
➔ Compare the last activation layers, produced from the same label.

GAN synthetic images

- The network haven't encoded information about the latent vector.
- We managed to produce diverse images by reshaping the tumor label.

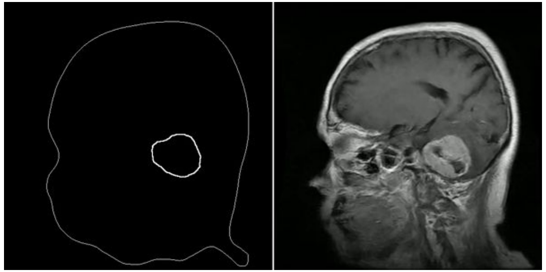
Last activation layers

- Some of the layers encode the skull pattern, while other encode the brain and tumor classes.

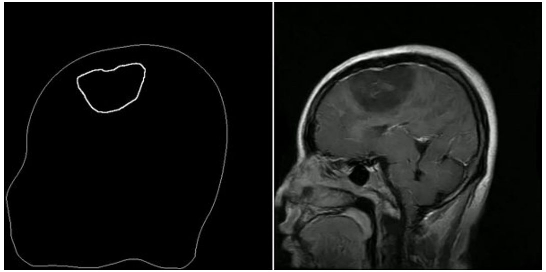


10 last activation layers out of 64

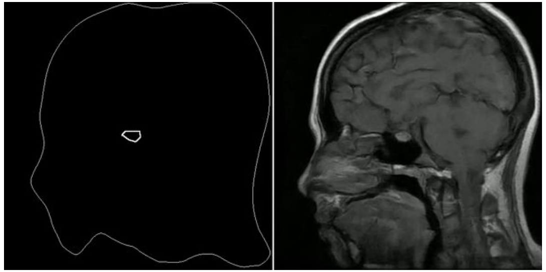
Meningioma



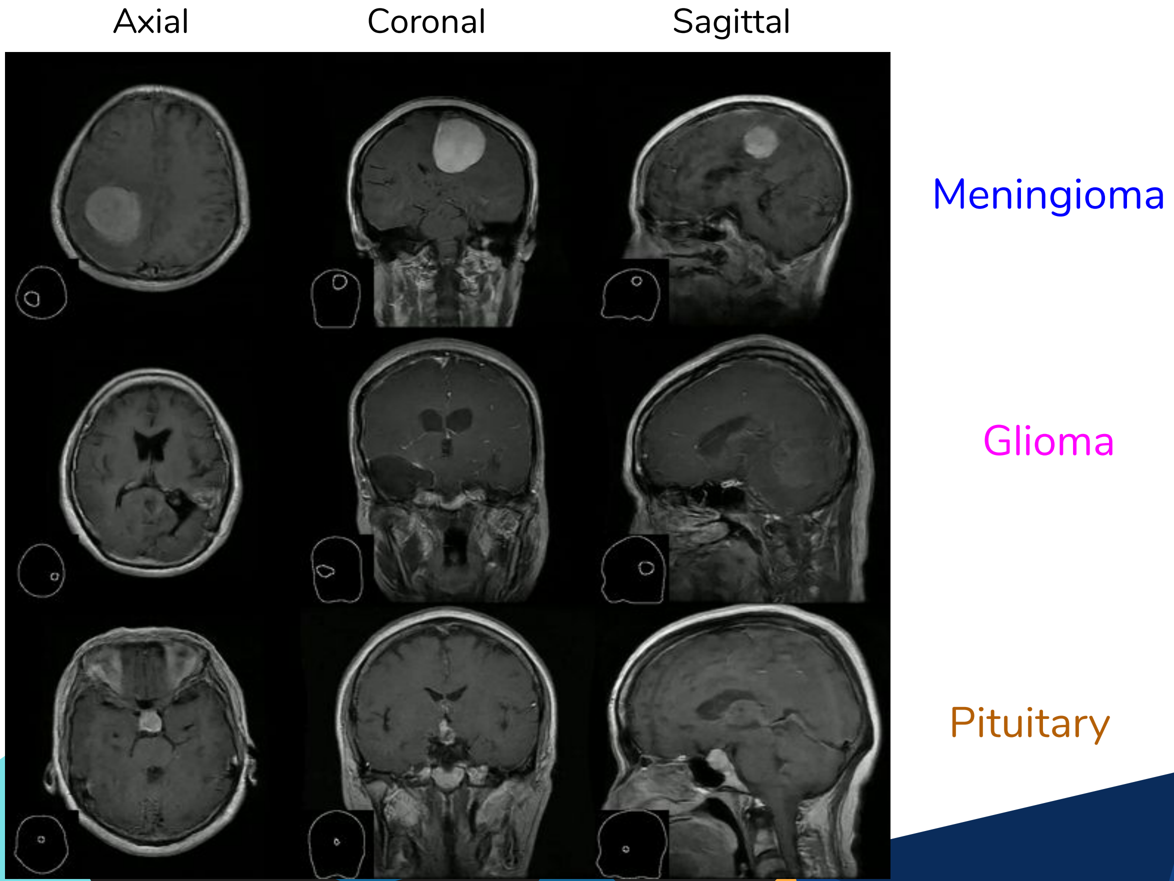
Glioma



Pituitary



Additional Synthesised Images..



Results - classification

Reminder..

➔ Dice/F1 calculates the harmonic mean of the recall and precision.

➔ Training on synthetic data performed relatively similarly to training on real data.

➔ Improve of performance when pretraining with synthetic images

➔ Train on real data got Resnet50 to perform relatively similarly to the pretrained capsule classifier.

Results Table

Classifier	Images	Pretrained	Accuracy	F1-score
Capsule	Real Data	No	0.85	0.86
Capsule	GAN	No	0.87	0.86
Capsule	Real Data	GAN	0.93	0.92
Resnet 50	Real Data	No	0.94	0.94

Conclusion

Performance

Pretraining with synthetic images lead the capsule model to comparable performance to the state of the art Resnet50 model, while using x9 fewer parameters.

Enlarge datasets

The model was able to synthesise diverse images and is useful when dealing with datasets with small amount of labeled data.

Suggestions for Future Work

CapsPix2Pix

Image synthesis could be much more variable using Capsule cGAN, because it's features could capture both tumor and noise classes.

Latent Vector

Test whether decrease the L1 coefficient in the cGAN loss will get the network to generate different and realistic images, for the same label.

Skull Reshaping

Increase the diversity of the brain in the synthesised images by also reshape the skulls.

Thanks!

Any questions?

You can find us at:
zoreasaf@gmail.com
marom.dadon@gmail.com