MVision Internship

Objective

Test if self-supervised approaches can help improve the results obtained from supervised segmentation models

- Some self supervised approaches -
    - Siamese Twins
    - BYOL
    - Barlow Twins
- This research focuses on testing barlow twins

Initial Plan

- Test transformer based segmentation models to produce a baseline
  - Medical transformers
- Use the same model configuration to pretrain on unlabelled data using the barlow twins approach
- Test if the pretrained model used as a backbone helps to improve the performance

Update in plan

- Transformer model used in medical transformer was too big to be run in barlow twins using the computation resources we had, so started to experiment with small models like resnet50 and resnet18
- It is not easy to attribute the success of a segmentation model on the pretrained encoder, as the decoder layer could have helped to learn the features.
- Hence, we decided to create a multi label image classification problem to identify if the ROI is present in the image or not.

How to make the data suitable for classification

- If the slice had the ROI, then its class is 1
- If the slice does not have the ROI, but other slices of that 3D image have the ROI, then its class is -1
- If the ROI is none for that image then its class is -1
- Else its class is 0

Challenge 1: Not able to support the batchsize barlow twins was trained with

- The model does not converge until the batchsize is 256, so I used a small model like resnet18, with a very low learning rate to be able to fit the model in memory
- The final model uses a learning rate weights of 0.002 and the learning rate biases of 0.000048 after trial and error.

Challenge 2: Setup the loss function when training multi label classification, when each label has 3 (not 2 labels).

- Classes are 0,1, -1
- We don't care about the samples where the class is 0 as we don’t know if the ROI is present or not.
- Get weights which is 0 if the class is 0
- Convert the class for -1 to 0.
- Multiply the final loss with weights so that records wit ROI class 0 does not impact the loss.

Challenge 3: Biased models

The above methodology led to completely biased model that were not able to classify labels 0 and 1 -

- Then we used the true positives, false positives, true negatives and false negatives to identify which labels were not getting identified correctly and realised that the label 1 classes was too over sampled.
- This further made me realize that there was data loading error due to which the data wasn’t being loaded properly.
- I fixed the data loading problem which led to better results

Challenge 4: The barlows twins model was not doing better than a random model

- This could be due to some transforms making the performance of the overall model worse
- To test this hypothesis, we tried to disentangle the transforms and run them individually over a common base model.
- As hypothesized some transforms were not performing as well as the others.

Experiments

- Freeze the resnet backbone and get results of classification for -
  - Random weights
  - Imagenet pretrained weights
  - Barlow twins pretrained weights (trained from scratch)
  - Barlow twins pretrained weights (using imagenet pretrained)
  - Barlow twins pretrained weights (using imagenet pretrained) with individual transforms for 50 epochs
    - Blur
    - Color jitter
    - Sharpen
    - Cutout
    - Elastic transformation
    - Crop
- Finetune the resnet backbone
  - Random initialization
  - Imagenet pretrained weights
  - Barlow twins pretrained weights (trained from scratch)

Results of the experiments

- First experiment was to understand if a barlow twins trained model can improve the performance of a model which is randomly initialized, where we freeze the resnet layers
- With the initial transforms the barlow twins doe not improve the performance over randomly initialised weights.
- Initial transforms were - Cropping, blurring, sharpening, random brightness contrast

	structure	random	barlows scratch initial transforms freeze
0	Brain	0.719068	0.639075
1	Bone_Mandible	0.522015	0.447632
2	SpinalCanal	0.561902	0.376504
3	Glnd_Lacrimal_L	0.488618	0.488618
4	Lung_R	0.786583	0.702765
5	Glnd_Submand_R	0.477321	0.477321
6	Glnd_Lacrimal_R	0.488324	0.488324
7	Cochlea_L	0.494660	0.494660
8	OpticNrv_cnv_R	0.492643	0.492643
9	Lens_R	0.493509	0.493509
10	SpinalCord	0.522654	0.366833
11	Parotid_R	0.472943	0.462152
12	Glnd_Submand_L	0.478424	0.478424
13	Brainstem	0.478645	0.460979
14	OpticNrv_cnv_L	0.492527	0.492527
15	Cochlea_R	0.495005	0.495005
16	Eye_R	0.483451	0.483451
17	Lens_L	0.493625	0.493625
18	Lung_L	0.790141	0.697944
19	Eye_L	0.483571	0.483571
20	Parotid_L	0.471582	0.462282

Results of the experiments

Second experiment we ran was to see if barlow twins model trained with pre-trained imagenet weights could perform better. We can clearly see that this significantly improved the performance for all the ROIs

	structure	random	barlows scratch initial transforms freeze	barlows pre trained imgnet freeze
0	Brain	0.719068	0.639075	0.829654
1	Bone_Mandible	0.522015	0.447632	0.757724
2	SpinalCanal	0.561902	0.376504	0.720642
3	Glnd_Lacrimal_L	0.488618	0.488618	0.550690
4	Lung_R	0.786583	0.702765	0.865998
5	Glnd_Submand_R	0.477321	0.477321	0.708313
6	Glnd_Lacrimal_R	0.488324	0.488324	0.574866
7	Cochlea_L	0.494660	0.494660	0.494660
8	OpticNrv_cnv_R	0.492643	0.492643	0.521833
9	Lens_R	0.493509	0.493509	0.525646
10	SpinalCord	0.522654	0.366833	0.693499
11	Parotid_R	0.472943	0.462152	0.731060
12	Glnd_Submand_L	0.478424	0.478424	0.716745
13	Brainstem	0.478645	0.460979	0.722195
14	OpticNrv_cnv_L	0.492527	0.492527	0.551998
15	Cochlea_R	0.495005	0.495005	0.506297
16	Eye_R	0.483451	0.483451	0.607314
17	Lens_L	0.493625	0.493625	0.534601
18	Lung_L	0.790141	0.697944	0.866130
19	Eye_L	0.483571	0.483571	0.616248
20	Parotid_L	0.471582	0.462282	0.733489

Results of the experiments

Third experiment we ran was to try individual transforms on top of the barlows pre trained imagenet model to see which transformations improve the performance the most.

	structure	barlows pre trained imgnet freeze	barlows pre trained imgnet elastic freeze	barlows pre trained imgnet crop freeze	barlows pre trained imgnet blur freeze	barlows pre trained imgnet cutout freeze	barlows pre trained imgnet color jitter freeze	barlows pre trained imgnet sharpen freeze	imgnet freeze
0	Brain	0.829654	0.894654	0.875919	0.860216	0.838084	0.838230	0.827304	0.896997
1	Bone_Mandible	0.757724	0.859947	0.822382	0.792158	0.764504	0.756379	0.736001	0.849415
2	SpinalCanal	0.720642	0.861675	0.843611	0.778711	0.757408	0.747412	0.736828	0.839836
3	Glnd_Lacrimal_L	0.550690	0.736652	0.621827	0.591008	0.568279	0.575222	0.614604	0.721879
4	Lung_R	0.865998	0.893062	0.878605	0.874981	0.866055	0.874405	0.868646	0.924239
5	Glnd_Submand_R	0.708313	0.790101	0.760919	0.732874	0.696444	0.688284	0.692636	0.775675
6	Glnd_Lacrimal_R	0.574866	0.721381	0.638580	0.617021	0.614066	0.610398	0.583002	0.724869
7	Cochlea_L	0.494660	0.624895	0.514543	0.494660	0.514802	0.505076	0.515334	0.647065
8	OpticNrv_cnv_R	0.521833	0.681163	0.535727	0.594056	0.566016	0.590659	0.587043	0.639219
9	Lens_R	0.525646	0.665040	0.536285	0.557221	0.518241	0.524407	0.541120	0.594850
10	SpinalCord	0.693499	0.840034	0.825308	0.770120	0.744105	0.735529	0.714287	0.826309
11	Parotid_R	0.731060	0.826743	0.793418	0.761239	0.721893	0.719452	0.701844	0.812012
12	Glnd_Submand_L	0.716745	0.799018	0.749253	0.724631	0.690002	0.693539	0.710918	0.790306
13	Brainstem	0.722195	0.850619	0.797157	0.792906	0.734233	0.739052	0.706473	0.831099
14	OpticNrv_cnv_L	0.551998	0.709343	0.568219	0.590003	0.568282	0.567216	0.625543	0.653826
15	Cochlea_R	0.506297	0.533871	0.505584	0.494947	0.506297	0.506297	0.538183	0.642640
16	Eye_R	0.607314	0.795340	0.691965	0.687952	0.663988	0.675654	0.679612	0.757235
17	Lens_L	0.534601	0.643353	0.510058	0.534998	0.565900	0.524407	0.549672	0.602427
18	Lung_L	0.866130	0.893596	0.874903	0.870223	0.871862	0.875091	0.871236	0.925947
19	Eye_L	0.616248	0.802958	0.694182	0.675629	0.636582	0.673763	0.658600	0.748210
20	Parotid_L	0.733489	0.822348	0.790727	0.764270	0.715113	0.718250	0.711027	0.820480

Results of the experiments

- We can observe in the third experiment that the transformations improve performance over the base model in this order - elastic transformation, crop, blur, cutout, color jitter and sharpen.
- Elastic transformation was able to reach imagenet and surpass it performance in several tasks as well.
- Sharpen leads to a decrease in performance for some ROIs over the base model and hence was discarded in the final model

Results of the experiments

Final model was trained with the following transformations - elastic transformation, crop, blur, cutout, color jitter. This is shown to improve performance over random initialization for some ROIs like Brain, Bone mandible, Spinal cord, Spinal canal, Lung L, Lung R, Brainstem

	structure	random	barlows scratch initial transforms freeze	barlows scratch transforms other than sharpen freeze
0	Brain	0.719068	0.639075	0.749298
1	Bone_Mandible	0.522015	0.447632	0.570391
2	SpinalCanal	0.561902	0.376504	0.606247
3	Glnd_Lacrimal_L	0.488618	0.488618	0.488618
4	Lung_R	0.786583	0.702765	0.799989
5	Glnd_Submand_R	0.477321	0.477321	0.477321
6	Glnd_Lacrimal_R	0.488324	0.488324	0.488324
7	Cochlea_L	0.494660	0.494660	0.494660
8	OpticNrv_cnv_R	0.492643	0.492643	0.492643
9	Lens_R	0.493509	0.493509	0.493509
10	SpinalCord	0.522654	0.366833	0.582068
11	Parotid_R	0.472943	0.462152	0.471262
12	Glnd_Submand_L	0.478424	0.478424	0.478424
13	Brainstem	0.478645	0.460979	0.489704
14	OpticNrv_cnv_L	0.492527	0.492527	0.492527
15	Cochlea_R	0.495005	0.495005	0.495005
16	Eye_R	0.483451	0.483451	0.483451
17	Lens_L	0.493625	0.493625	0.493625
18	Lung_L	0.790141	0.697944	0.794700
19	Eye_L	0.483571	0.483571	0.483571
20	Parotid_L	0.471582	0.462282	0.479301

Results of the experiments

We also tested the final model without freezing the layers and it improved the performance over random initialization though it is still not as good as imgnet pretrained model.

	structure	baseline	barlows scratch initial transform	barlows scratch final transform	barlows pre trained imgnet
0	Brain	0.985061	0.979202	0.986091	0.997672
1	Bone_Mandible	0.966569	0.969487	0.968554	0.995386
2	SpinalCanal	0.991332	0.984994	0.991092	0.999064
3	Glnd_Lacrimal_L	0.905114	0.902252	0.922145	0.981849
4	Lung_R	0.985563	0.981184	0.986158	0.996933
5	Glnd_Submand_R	0.940586	0.938606	0.954674	0.985882
6	Glnd_Lacrimal_R	0.899035	0.900128	0.908538	0.975011
7	Cochlea_L	0.808092	0.815408	0.815244	0.968061
8	OpticNrv_cnv_R	0.858003	0.857067	0.867614	0.958283
9	Lens_R	0.829786	0.823032	0.863821	0.925918
10	SpinalCord	0.990067	0.985206	0.990760	0.997924
11	Parotid_R	0.964117	0.955827	0.955458	0.986655
12	Glnd_Submand_L	0.947922	0.938025	0.944551	0.987578
13	Brainstem	0.954158	0.957301	0.965816	0.992756
14	OpticNrv_cnv_L	0.866323	0.853859	0.862499	0.960600
15	Cochlea_R	0.839928	0.816214	0.829891	0.961153
16	Eye_R	0.952199	0.942546	0.953290	0.981615
17	Lens_L	0.838060	0.823033	0.873218	0.948194
18	Lung_L	0.987485	0.981430	0.987784	0.998323
19	Eye_L	0.952893	0.940534	0.956021	0.982506
20	Parotid_L	0.954975	0.961141	0.959365	0.987073

Links to project code

Barlow twins : https://github.com/sborar/barlowtwins

Resnet classification:
https://github.com/sborar/resnet_multilabel_image_classification