

Supplementary Material

July 6, 2024

1 Code and Data Repository

Link for the Dataset:

<https://github.com/mattgroh/fitzpatrick17k> for Fitzpatrick17k dataset

<https://ddi-dataset.github.io/> to access DDI dataset

Github Repository for the code can be accessed via : <https://github.com/aayushmanace/PatchAlign24>.

Follow the instructions in the Readme to reproduce the results.

General command to run the code: `python3 <file_name>.py <n_epochs> full <dataset_name>`

2 Hyperparameters Used

Below we provide the hyper parameters for our approach:

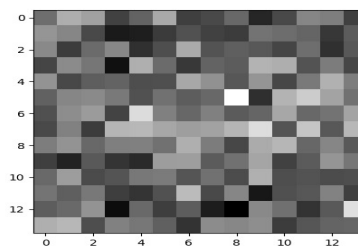
- alpha: 0.5
- beta: 1.0
- n_epochs: 20
- GOT, MGOT Threshold (β): 0.8
- Each experiment is repeated 5 times

3 Masking

The mask image [1b](#) for an image [1a](#) is generated by PatchAlign by taking the 14x14 patch embeddings of the ViT model showing the regions of the image where more importance is given for classification (lighter colour) and lesser importance is given (darker colour). Note that this is a random example.



(a)



(b)

Figure 1: (a) A Sample Image from the Fitzpatrick17k dataset (b) The Mask learnt by PatchAlign

4 Ablation Study

We have done an ablation study on entropy values with accuracy for $\beta = [0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]$ and as inferred from the graph, we found the best result for $\beta = 0.8$

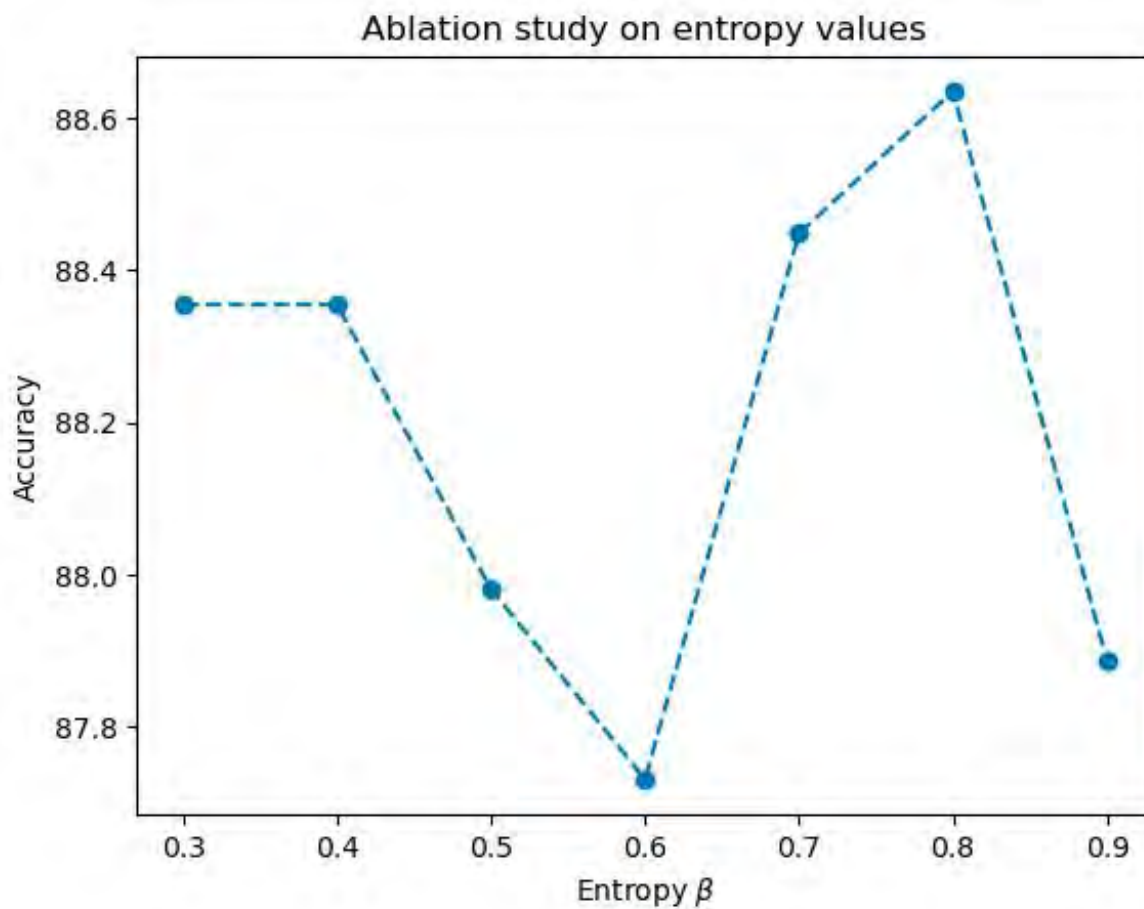


Figure 2