**Name:** Antika Das (NU57PS)

**Internal Supervisor's Name**: Faragó Kinga Bettina and Andars Lorincz

**Title: *Investigation of similarities and dissimilarities of skin lesions (Melanoma/Non-Melanoma) Using Convolutional Neural Network*.**

**Motivation**:

Nevus and melanoma are a group of neoplasia. Although, nevi (plural of nevus) and melanomas are often treated as independent entities, there is evidence that a nevus can be a precursor for a melanoma. Melanoma is a type of skin cancer that develops when melanocytes (the cells that give the skin its tan or brown colour) start to grow out of control. So, our major goal is to build up a classifier to distinguish between **Melanoma** and **Non**-**melanoma** cases (our interest is to distinguish all other categories of skin problems from the most dangerous category, i.e., from melanoma).

In particular, the goal is to reduce the classifier's false negative rate (cases when the classifier labels it as non-melanoma, but the images are representing melanomas) to improve classifier’s overall performance.

**Dataset:**

ISIC 2019 (<https://challenge2019.isic-archive.com/>) and ISIC 2020 (<https://challenge2020.isic-archive.com/>) – Both are publicly available Skin Lesion Data Sets.

**Task List:-**

* DenseNet169 (<https://www.kaggle.com/pytorch/densenet169>) and other deep networks for pre-processing.
* Set up similarities and dissimilarities (pair) using [1], details:
  + Embedding into low dimensional space (for clustering and visualization)
  + human knowledge-based pairings based on common sense and on experts’ opinion
  + machine clustering using graph cut in 2D and 3D
* Use a TensorBoard-based (<https://www.tensorflow.org/tensorboard>) online, interactive interface (NIPGBoard) to project the results of the deep representation of the deep mesh into 3D space.
* Constrained pairwise loss for improvements
  + Considering methods to improve pairwise losses, e.g., triplet losses
* Evaluate the dataset:
  + Apply graph cutting methods such as Girvan-Newman community detection algorithm, Louvain community detection algorithm, Spectral clustering, Agglomerative Clustering, etc. to automate the search for clusters and look for false negative clusters that are of good quality (for clarification: good quality means that the items considered to be in the same cluster are mostly having the same information or label).
  + Extract the corresponding teacher and test data from the data set and improve the bad clusters (for clarification: bad cluster means that the items considered to be in the same cluster are mostly having different information or label)

**Three chosen subjects for final exam:**

|  |  |
| --- | --- |
| **Subject Code** | **Subject Name** |
| IPM-18AUTDMSSEG | Data Mining in Smart System |
| IPM-18AUTADLEG | Applied Deep Learning |
| IPM-18AUTIVPEG | Image and Video Processing |

**References:**

[1] Hsu, Y. C., & Kira, Z. (2015). Neural network-based clustering using pairwise constraints. arXiv preprint arXiv:1511.06321.