1. **Dataset**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dataset** | **Patients** | **Scans** | **Grade** | **Amount** |
| TCGA-GBM | 135 | Flair, T1, T1Gd and T2 | IV | 135 |
| TCGA-LGG | 106 | Flair, T1, T1Gd and T2 | II | 48 |
| III | 58 |
| LGG-1p19qDeletion | 159 | T1c and T2 | II | 104 |
| III | 55 |

All scans in three datasets are pre-operative.

|  |  |  |
| --- | --- | --- |
| **Grade II** | **Grade III** | **Grade IV** |
| 152 | 113 | 135 |

1. **Preprocessing**

2.1 **Skull-Stripping**: use same method to remove skull from each slice.

2.2 **Extract Tidy Brain**: remove unnecessary background.

2.3 **Scale Normalization**: resize all scans to same shape.

1. **Prepare for Training**

Generate training, validating and testing set.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dataset** | **Proportion** | **Grade II** | **Grade III** | **Grade IV** |
| Training | 60% | 91 | 68 | 81 |
| Validating | 20% | 31 | 23 | 27 |
| Testing | 20% | 30 | 22 | 27 |

Do **augmentation** on training set by horizontally flipping the volume.

1. **Train Sparse Autoencoder**

|  |  |  |
| --- | --- | --- |
|  | Minimize the **reconstruction loss** and **sparse penalty**. | Model structure and  Hyper-parameters:  learning rate,  number of epochs,  batch size,  dropout rate,  sparse coefficient,  regularization coefficients,  ...... |
|  | Minimize the **softmax cross entropy**. |

Use a table to record results of each training process.

1. **Evaluation**

5.1 Plot **learning curve** of loss and accuracy of each training process to evaluate parameter settings.

5.2 Compute **recall** and **precision** of three classes to describe model’s sensitivity.