**Methods to try**

1) <https://www.frontiersin.org/articles/10.3389/fnins.2018.00777/full>

To explore the multiple plane images in MRI, a 2.5D patch was formed by extracting three 32 x 32 patches from transverse, coronal, and sagittal plane centered at a same point.

Then, three patches were combined into a 2D RBG patch.

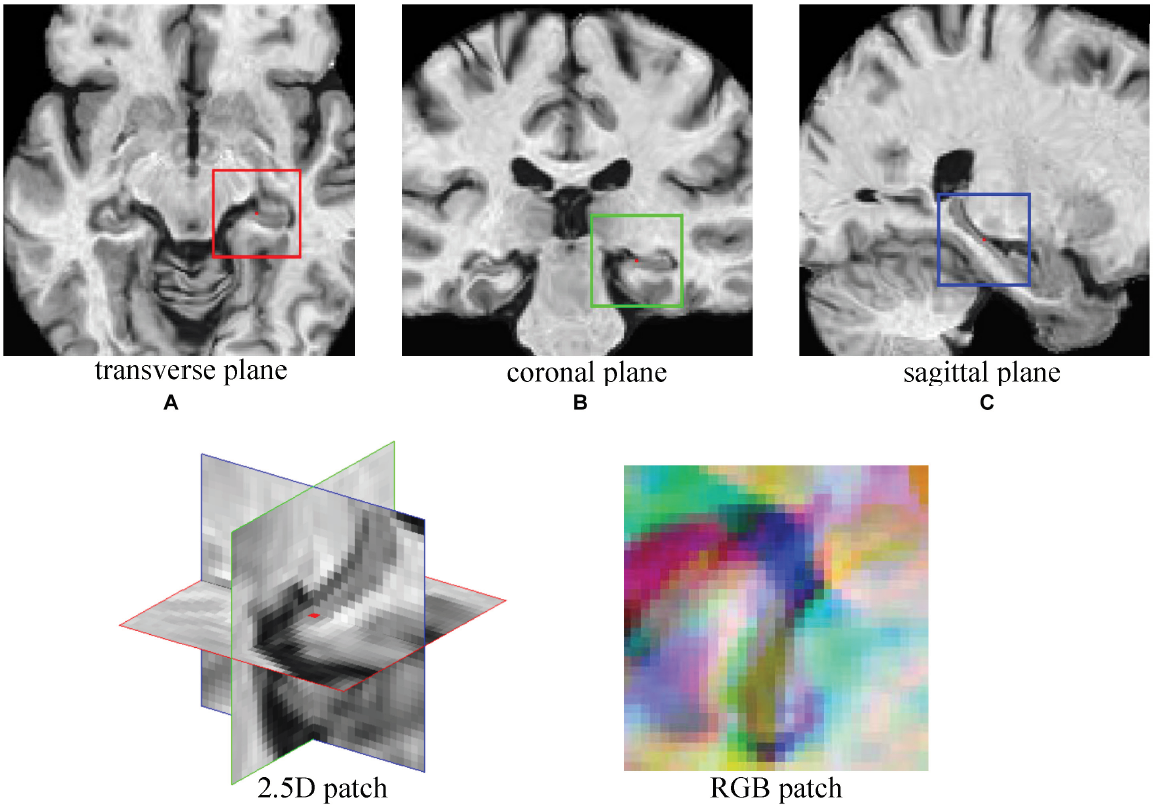
For a given voxel point, three patches of MRI are extracted from three planes and then concatenated into a three-channel cube.

Extracting multiple patches at different locations from MRI images. The choice of locations has three constraints,

* The patches must be originated in either left or right hippocampus region which have high correlation with AD;
* must be at least two voxels distance between each location;
* All locations were random chosen.

With these constraints, 151 patches were extracted from each image and the sampling positions were fixed during experiments.

The architecture of the CNN has an input of 32 x 32 RGB patch.



2) <https://web.stanford.edu/class/cs331b/2016/projects/glozman_liba.pdf>

As the expected input is 2D images, we create the following images from the volume data:

* ax key - axial middle slice in all three color channels.
* cor key - coronal middle slice in all three color channels.
* sag key - saggital middle slice in all three color channels.
* ax3 - mid-axial slice in the R channel, and 10 slices above and below the mid-axial slice in the G and B channels.
* cor3 - coronal middle slice in the R channel, and 10 slices before and behind the mid-coronal slice in the G and B channels.
* sag3 - mid-saggital slice in the R channel, and 10 slices before and behind the mid-saggital slice in the G and B channels.
* axcosag - mid-axial slice in the R channel, mid-coronal slice in the G channel and mid-saggital slice in the B channel.

**Dataset**

Data augmentation: mirror images.

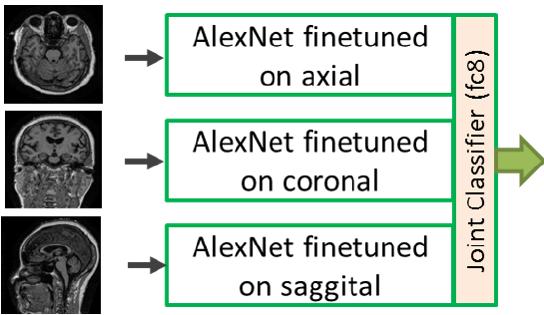
Data was randomly divided into training and testing sets on the subject level (rather than on the scan level, to avoid possible bias of having data of the same subject at different time points in both the training and testing sets).

**Training**

Learning rates of 0.001, 0.005, 0.01 and 0.015.

Dropout of 0.1 (keep rate of 0.9), mini-batch size of 30 images, and Adam optimizer

In order to combine information from different modalities or from different view angles (sagital/ coronal/axial), we implemented three-stream and two stream networks, which combine three or two fine-tuned AlexNets at the final classifier layer.



We train two or three separate networks to classify images of two or three different orientations - axial, coronal and sagittal correspondingly.

These networks are combined by concatenating their fine-tuned fc7 layers and then training the fc8 layer of the combined network (which now has 2 or 3 times more weights) on pairs or 3-tuples of images with corresponding views from volumes in the training set.

Two types of images are considered in these experiments:

* keyslices from each orientation - ax key, cor key, and sag key;
* aggregated slices from each orientation - ax3, cor3, sag3.

3) <https://www.researchgate.net/publication/319325795>

The Hippocampal ROI was selected from the 3D volume of a single projection (e.g. sagittal) of the brain.

Selecting voxels labelled as Hippocampal region in AAL and computing their bounding boxes, we get a sub-volume of the whole 3D projection.

The median slice inside this volume was chosen and its two closest neighbors were considered.

The input layer of the network constituted of 28x28x3 units receives data from three 28x28 central slices of the hippocampal region.

Fusion of the three projections networks (sagittal, coronal, and axial) which consists of a concatenation at the FC layer of three networks.

