neuromatch academy

The Differentiation of Parkinson's Disease Patients from Healthy Patients Through Training a Machine Learning Model on MRI Data

Luis Revilla, Hector Lorenzo Mebenga, Koffi Koudjonou, Vinicius Silvestre , Madhusudan Duwadi, Georg Zoidl

Group: OpenMinded

Pod TA: Shirin Vafaee
Project TAs: Yamil Vidal and Farrokh Karimi

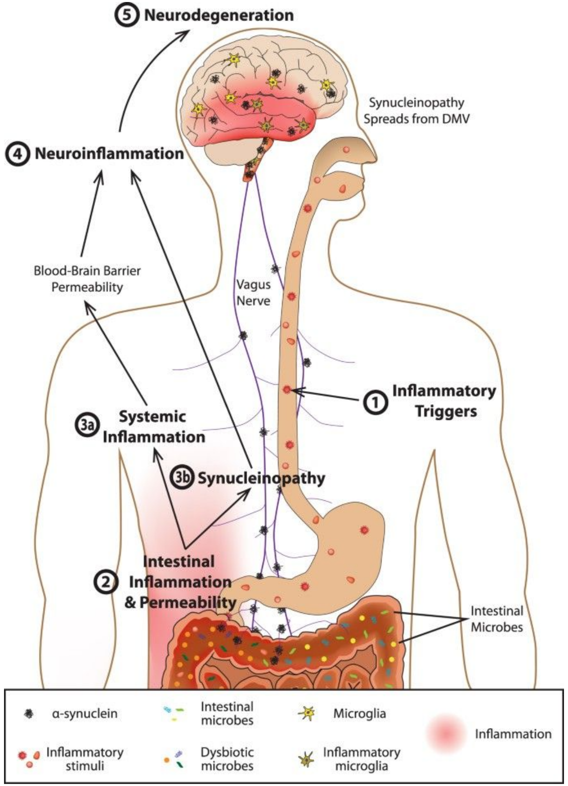
neuromatch academy

What is Parkinson's Disease?

- Neurodegenerative disorder **affecting dopaminergic neurons** in the brain especially the substantia nigra
- Parkinson's hinders affected person's abilities to do basic tasks due to:
  - Resting tremors, cognitive deficiencies, bradykinesia.

How is Parkinson's diagnosed?:

- Diagnosis is based on a person's **medical history** and a neurological examination, often involving **MRI and fMRI** of the brain.



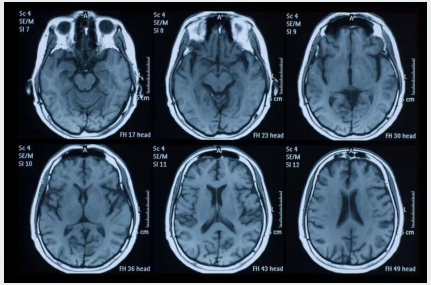
neuromatch academy

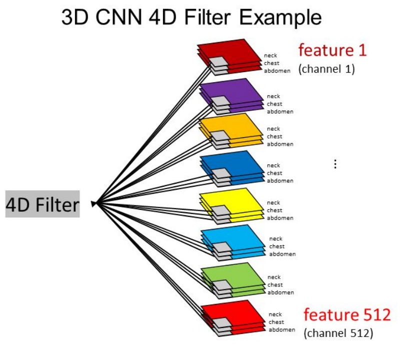
PARKINSON'S PROGRESSION MARKERS INITIATIVE

Play a Part in Parkinson's Research

Are we able to train a machine learning model to differentiate and detect Parkinson's disease based on MRI data?

- We used the Parkinson's Progression Markers Initiative (PPMI) dataset.
- In this project we trained a **2D and 3D convolutional neural network (CNN)** on MRI data.
- To train the model MRI data was taken from **healthy and Parkinson's disease patients** between the **ages of 50-75.**
- Slices of the brain were analyzed for differences in the two test groups.

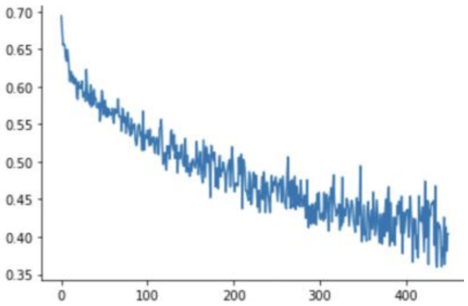




Test loss of 0.59 and test accuracy of 0.69 for 2D CNN

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 255, 255, 32)	416
batch_normalization (BatchNo	(None, 255, 255, 32)	128
max_pooling2d (MaxPooling2D)	(None, 127, 127, 32)	0
conv2d_1 (Conv2D)	(None, 126, 126, 32)	4128
batch_normalization_1 (Batch	(None, 126, 126, 32)	128
max_pooling2d_1 (MaxPooling2	(None, 63, 63, 32)	0
conv2d_2 (Conv2D)	(None, 62, 62, 64)	8256
batch_normalization_2 (Batch	(None, 62, 62, 64)	256
max_pooling2d_2 (MaxPooling2	(None, 31, 31, 64)	0
flatten (Flatten)	(None, 61504)	0
dense (Dense)	(None, 64)	3936320
dropout (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 2)	130

Total params: 3,949,762
Trainable params: 3,949,506
Non-trainable params: 256



Test loss: 0.5966757535934448 / Test Accuracy: 0.699999988079071

- We applied the 2D CNN.
- Total of **parameters tested:** Almost **4 million.**
- Faced some problems with the implementation of this model:
    - The number of parameters was a problem. Since we used a lot of images we've had to find ways to speed up the training without degrading the model performance.
    - The loss was bouncing between 0.6 and 0.7 so we tuned the hyperparameters of the model.

Test loss of 0.55 and test accuracy of 0.72 for 3D CNN

- We applied the 3D CNN to improve our 2D CNN test scores.
- Total **# of parameters** tested: **over 1.6 million.**
- We were able to have **higher accuracy** with **½ of the parameters** and **9x less epochs** than the 2D model.
- Faced some problems with the implementation of this model:
  - The **accuracy didn't decrease** at all in the beginning, but adding a dropout rate helped.
  - Formating of the data was difficult due to it being a .nii file. We used nilabel library to load it and reshape it in the right format

```
Epoch 49/50
8/8 [==============================] - 0s 43ms/step - loss: 0.4452 - accuracy: 0.7333
Epoch 50/50
8/8 [==============================] - 0s 43ms/step - loss: 0.4297 - accuracy: 0.7333
```

```
score = model.evaluate(np.array(x_test), y_test, batch_size=3)
print(f'Test loss: {score[0]} / Test Accuracy: {score[1]}')
```

```
9/9 [==============================] - 0s 11ms/step - loss: 0.5530 - accuracy: 0.7200
Test loss: 0.5530046224594116 / Test Accuracy: 0.7200000286102295
```

```
Model: "sequential_6"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv3d_12 (Conv3D)           (None, 62, 62, 14, 16)    1744
_________________________________________________________________
max_pooling3d_12 (MaxPooling (None, 31, 31, 7, 16)     0
_________________________________________________________________
conv3d_13 (Conv3D)           (None, 29, 29, 5, 32)     13856
_________________________________________________________________
max_pooling3d_13 (MaxPooling (None, 14, 14, 2, 32)     0
_________________________________________________________________
flatten_6 (Flatten)          (None, 12544)             0
_________________________________________________________________
dense_12 (Dense)             (None, 128)               1605760
_________________________________________________________________
dropout_6 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_13 (Dense)             (None, 1)                 129
=================================================================
Total params: 1,621,489
Trainable params: 1,621,489
Non-trainable params: 0
_________________________________________________________________
```

```
[53] nEpochs = 500 # Increase this value for better results (i.e., more training)
     batch size = 1  # Increasing this value might speed up fitting
```

Applying 3D-CNN weights on Early Onset PD

- We applied our model to 20-45 years old
- Testing with saved weights from 3D-CNN

```
# load model
savedModel = model.load_weights('/gdrive/MyDrive/PPMI/gfgModelWeights_2')
print('Model Loaded!')

loss, acc = model.evaluate(x_test, y_test, verbose=2)
print("Restored model, accuracy: {:5.2f}%".format(100 * acc))
```

```
Model Loaded!
1/1 - 0s - loss: 0.7183 - accuracy: 0.5789
Restored model, accuracy: 57.89%
```

Conclusion + Future directions

- Our maximum accuracy on tests was 72% in a 3D CNN.
- Our maximum accuracy on tests was 69% in a 2D CNN.
- Achieved a 57.9% accuracy on early onset PD

Future Directions:

- Use the trained model to improve accuracy of **predict early onset** of Parkinson's disease.
- Use more data to predict other neurodegenerative disorders such as **Alzheimer's** using **transfer learning.**

- Were we able to train the model?

Acknowledgement

- Shirin Vafaee for tutoring the course; Yamil Vidal, Farrokh Karimi for tutoring the project.
- Neuromatch Academy for organizing the summer program.

References

- https://colab.research.google.com/github/NeuromatchAcademy/course-content/blob/master/projects/fMRI/load_algonauts_videos.ipynb
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6732922/
- https://github.com/bsplku/3dcnn4fmri
- https://github.com/shailp52/Parkinson_Disease_Detection_fromMRI/blob/master/ParkinsonDetection.ipynb

Code

2D-CNN: https://colab.research.google.com/drive/1_oBy8XXl696Cv9N2EBGg_FR160aTT6UZ?usp=sharing
3D-CNN: https://colab.research.google.com/drive/1zDAEiVDeYGymmvXFeHnHlzzgrS1ESBkH?usp=sharing

Thank you!
Any questions?




Abstract

- The **application of machine learning** is becoming more prevalent in recent years in the **medical field**. Medical imaging techniques such as magnetic **resonance imaging (MRI)** and functional magnetic resonance imaging (fMRI) have benefited significantly from the application of machine learning due to machines' ability to **detect minor changes in the volume and shape** of the brain.
- Parkinson’s disease is a good model to apply machine learning to since patients affected by it have a significant loss of dopaminergic neurons in the brain.
- In the future, this model will be applied to younger patients to **detect the early onset of Parkinson’s disease** prior to the manifestations of clinical symptoms. Furthermore, using transfer learning we will apply the trained **CNN model to other neurodegenerative diseases** such as Alzheimer’s and dementia

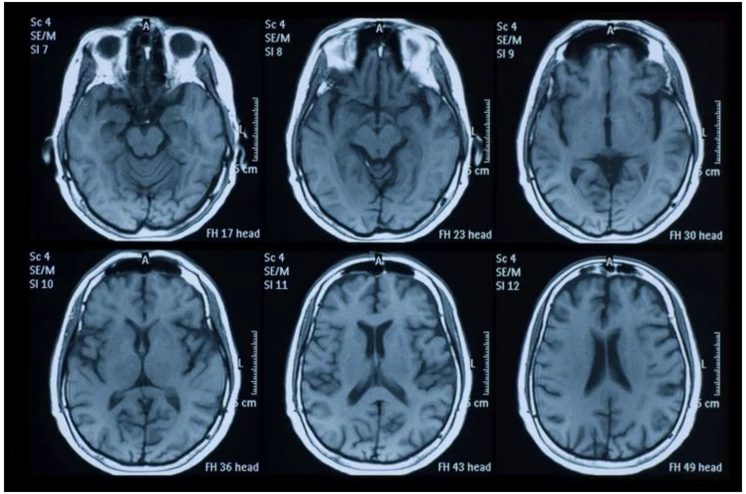


Fig 1. Resting state fMRI technique for imaging PD in early states. Researchers found that by using a certain threshold of connectivity, they could predict the onset of Parkinson’s in 11 out of 13 patients, or with 85% accuracy. Source: https://time.com/2860630/mri-scans-can-detect-early-onset-of-parkinsons-study-finds/

Dataset

Parkinson's Progression Markers Initiative (PPMI) dataset:

This database contains **MRI data of healthy and Parkinson's disease patients of all age groups**. PPMI is a landmark observational study to assess Parkinson's disease to speed therapeutic development. PPMI makes its data set and biorepository available to academia and industry.

- Population: Healthy and PD Patients from age 50-75.
- Sample: ## Slices of Control patients and ## Slices of PD patients

Fig 2. Parkinson's Progression Markers Initiative (PPMI) dataset. Source: https://www.ppmi-info.org/

neuromatch academy

Methods and Techniques

In this project, we train a **2D and 3D convolutional neural network (CNN)** on slices of the brain **(MRI images)** from the Parkinson's Progression Markers Initiative (PPMI) dataset.

The CNN will be used to detect and **differentiate between healthy and Parkinson's disease patients** with advanced Parkinson's disease between the ages of 50-75.

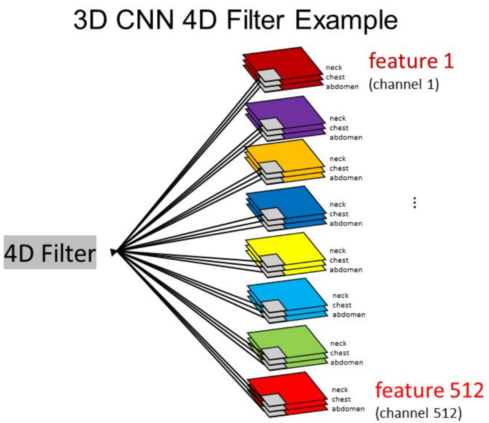


Fig 3. Pytorch 3D CNN Model design. Source: https://ichi.pro/es/diseno-de-cnn-personalizados-en-2d-y-3d-en-pytorch-124435025011963

Fig 4. Pytorch: Step by Step implementation 3D Convolution Neural Network: https://towardsdatascience.com/pytorch-step-by-step-implementation-3d-convolution-neural-network-8bf38c70e8b3