Sex Prediction from T1-weighted Brain MRI

Traditional machine learning and deep learning perspectives

Roberto Souza

February 2023

UNIVERSITY OF CALGARY

Outline

- Introduction and Background
- Hypothesis and Objectives
- Materials and Methods
- Results
- Discussion and Next Steps

Introduction and Background

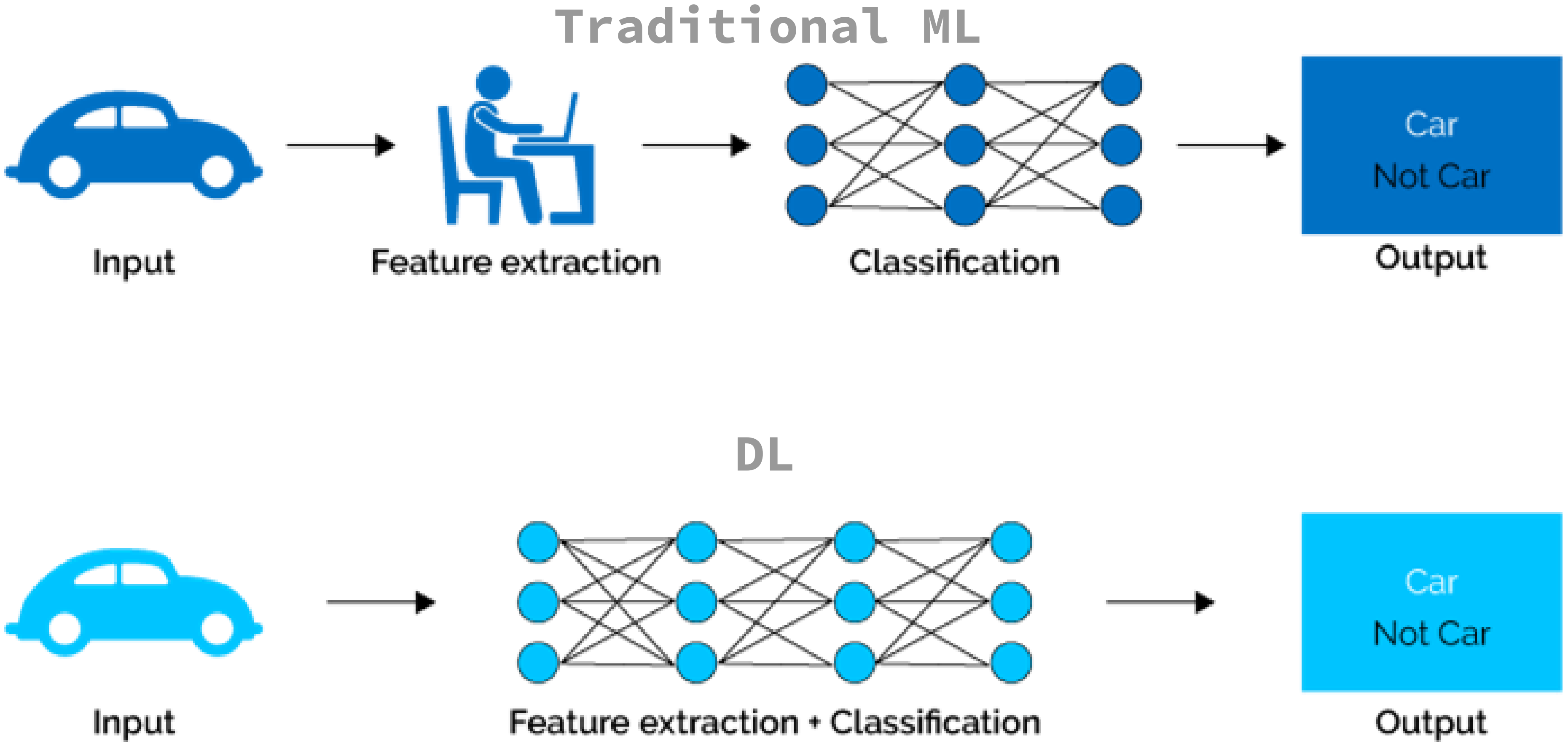
Introduction

- Many neurological diseases have different prevalence among women and men[1]
  - Alzheimer's -> women
  - Parkinson's - > men
- Understanding how the male and female brains differ from a traditional ML and DL perspectives is important and allows us to hypothesize whether sex plays an important role on the disease:
  - Ethics, fairness, and bias considerations
  - Inform your dataset construction if building a data-driven model
  - Decide whether to include confounding removal techniques, such as "sex" unlearning[2]

UNIVERSITY OF CALGARY

[1]Zielonka et al, Frontiers in Neurology, 2020 [2]Dinsdale et al, NeuroImage, 2021 ML = machine learning | DL = deep learning

Traditional ML versus DL

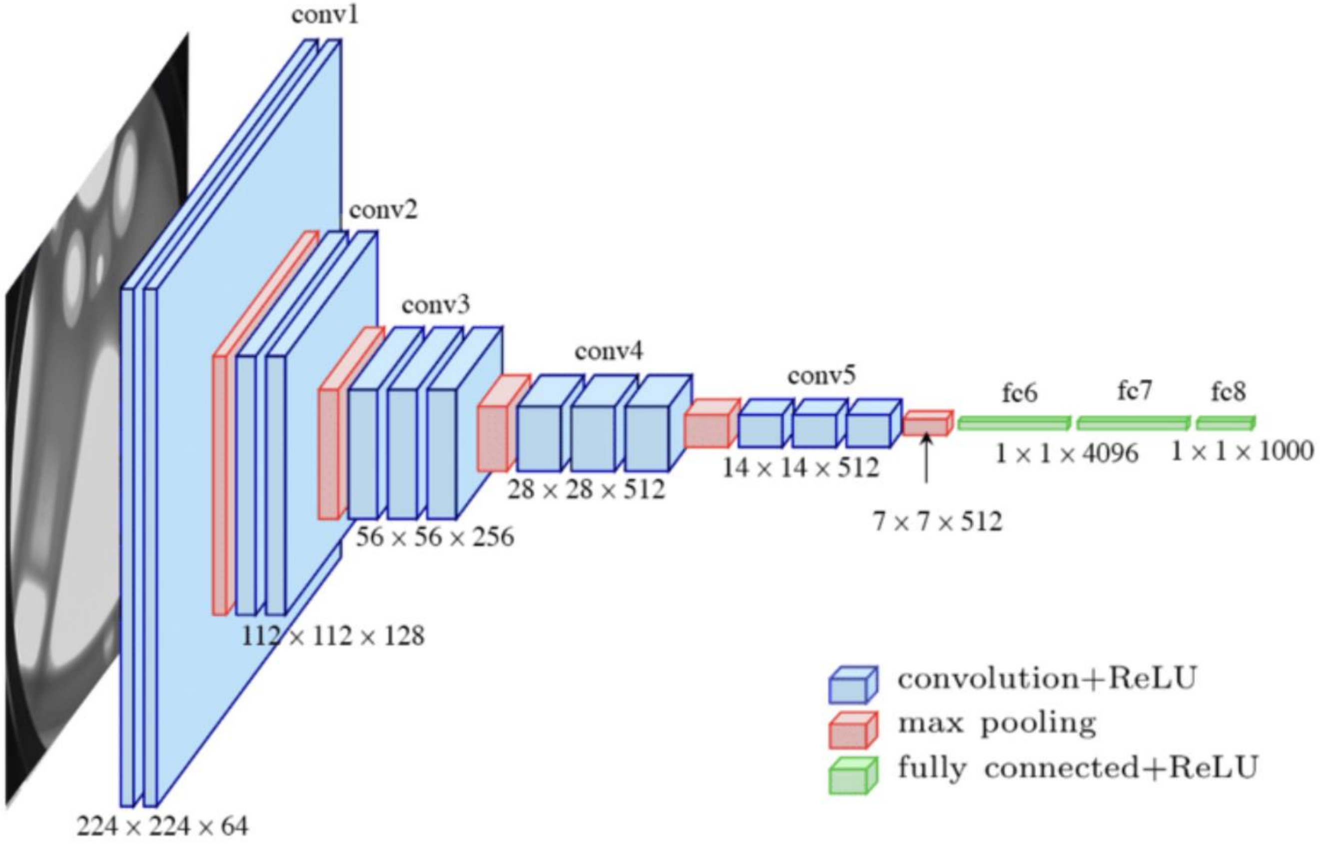


Trainable Parameters and GPU Memory

- Convolutional layer:
  - (w1*w2*w3*#channels + 1)*#filters
- Dense layer
  - (input_size + 1)*output_size
- Variables are often stored as float32 (i.e., 4 bytes)
  - Model parameters
  - Gradients
  - Intermediary results
  - Code overhead

VGG-16

- Won 2014 ImageNet
- Still often used in 2022
- Easy to code
- Drawbacks:
  - Many parameters
  - Computationally expensive



Hypothesis and Objectives

Hypothesis

We hypothesize that we can use ML obtain > 80% accuracy in distinguishing biological sex from presumed normal subjects using T1-weighted MRI

- Literature has obtained >90% accuracy but using other MRI sequences (e.g., DTI)

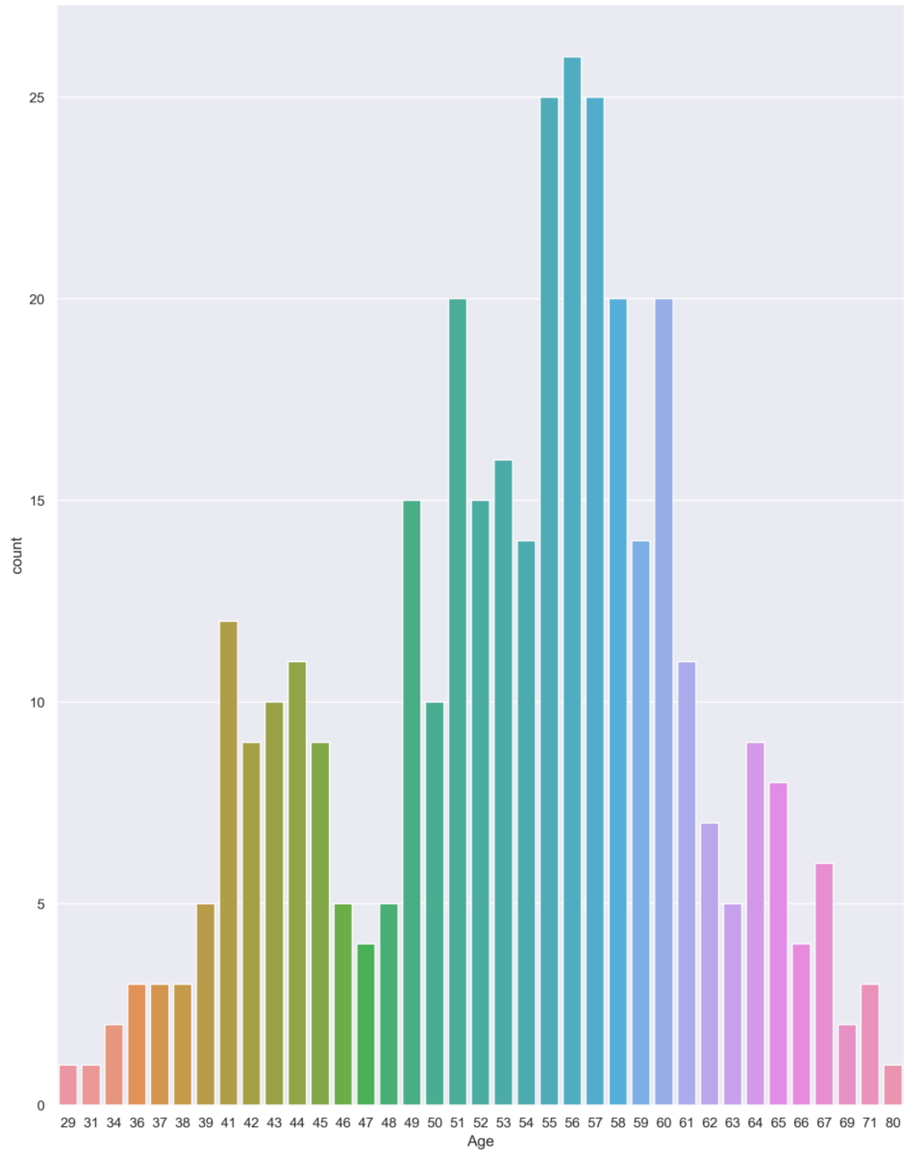
Objectives

1. Develop traditional ML models of sex classification using volumetric features
2. Perform feature importance analysis
3. Develop DL models of sex classification
4. Perform visual explanation of the model using GRAD-CAM

Materials and Methods

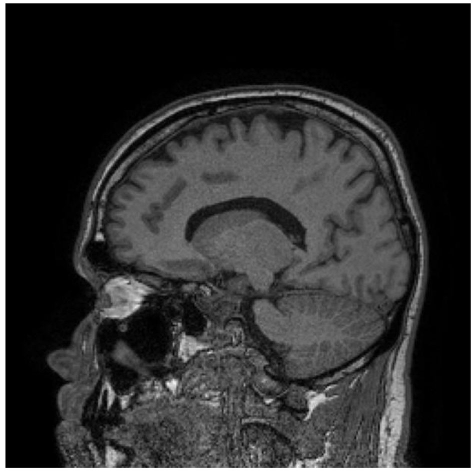
Calgary-Campinas Dataset

Vendor	Field	Age	Gender	Datasets
Siemens	1.5 T	53.9 ± 7.3	30M/30F	60
	3 T	56.6 ± 6.9	30M/30F	60
Philips	1.5 T	52.8 ± 9.6	26M/33F	59
	3 T	50.0 ± 9.3	30M/30F	60
GE	1.5 T	53.9 ± 5.8	30M/30F	60
	3 T	53.6 ± 5.7	30M/30F	60
All	1.5 and 3 T	53.5 ± 7.8	176M/183F	359

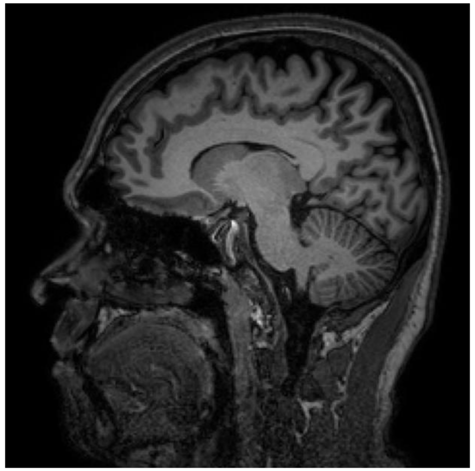


Calgary-Campinas Dataset

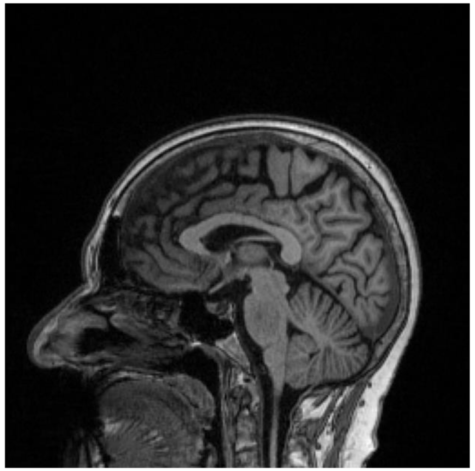
philips_15



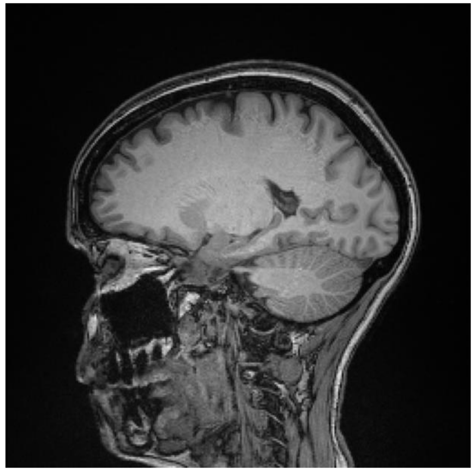
philips_3



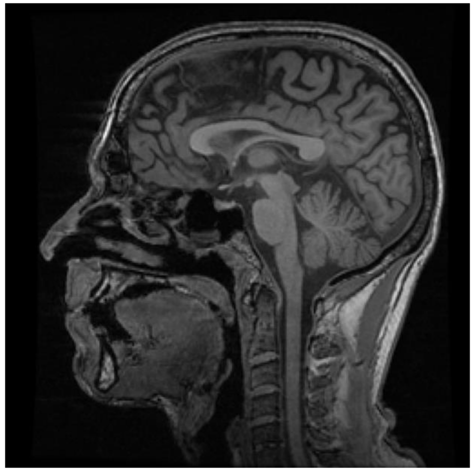
siemens_15



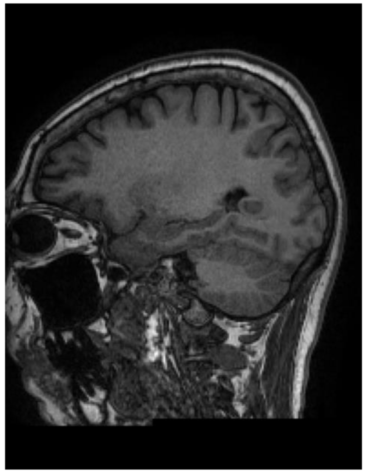
siemens_3



ge_3



ge_15



Traditional ML Analysis

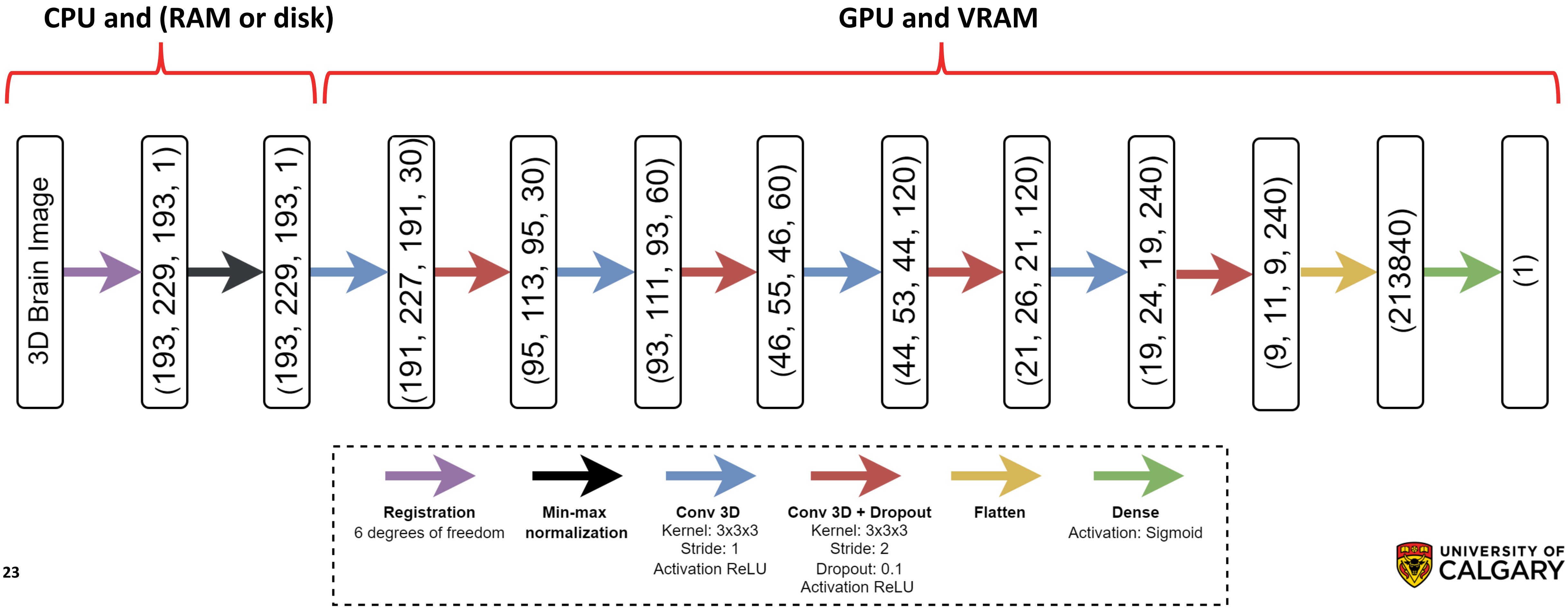
- Volumetrics features extracted from FreeSurfer + age
- Models:
  - Random Forest
  - XGBOOST
- Grid-search for hyperparameter tuning
- Stratified 10-fold cross-validation
  - Sex
  - Scanner vendor
  - Magnetic field

DL Analysis

- 3D Convolutional Neural Network (CNN)
- Stratified 10-fold cross-validation
  - Sex
  - Scanner vendor
  - Magnetic field
- Network inputs
  - 6-dof registration to MNI space + skull-stripping + min-max normalization
  - No data augmentation

UNIVERSITY OF CALGARY

3D CNN - Network Architecture



UNIVERSITY OF CALGARY

Number of Model Parameters

L_1 = (27*1 +1)*30 = **840**

L_2 = (27*30 +1)*30 = **24,330**

L_3 = (27*30 +1)*60 = **48,660**

L_4 = (27*60 +1)*60 = **97,260**

L_5 = (27*60 +1)*120 = **194,520**

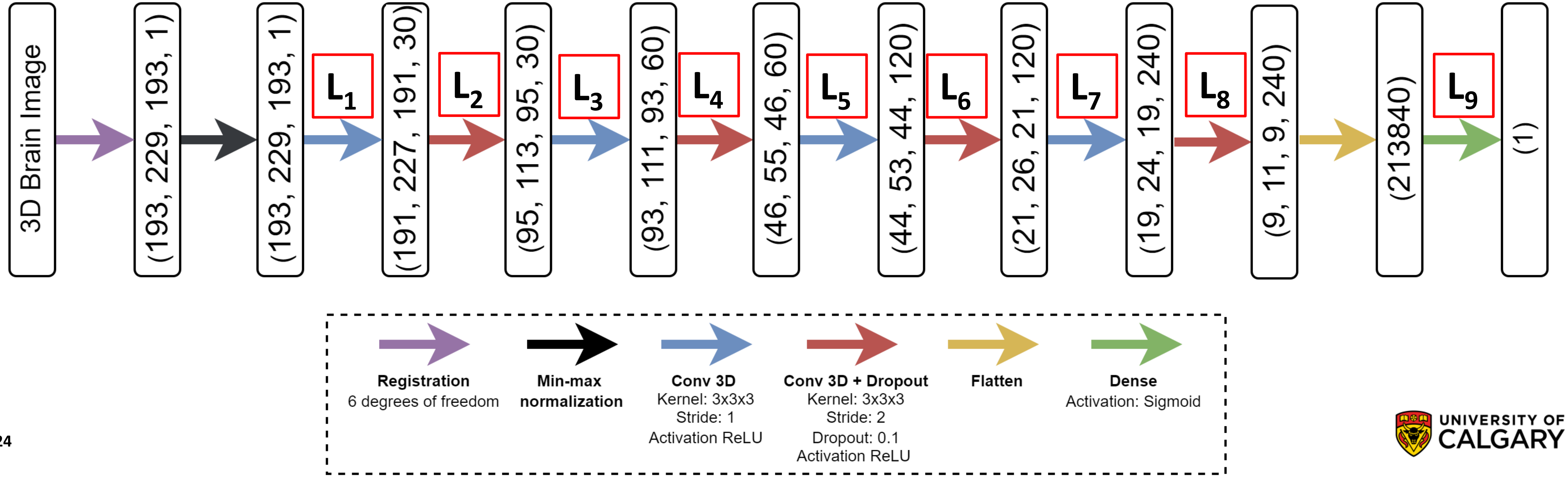
L_6 = (27*120 +1)*120 = **388,920**

L_7 = (27*120 +1)*240 = **777,840**

L_8 = (27*240 +1)*240 = **1,555,440**

L_9 = (213840 + 1)*1 = **213,841**

$$\boxed{\sum_{i=1}^{9} L_i = 3,301,651}$$



UNIVERSITY OF CALGARY

GPU Memory Consumption

Params =3,301,651*4 **= 13.21 MB**

Grads =3,301,651*4 **= 13.21 MB**

I_1 = 193*229*193*1*4 = **34.12 MB**

I_2 = 191*227*191*30*4 = **993.74 MB**

I_3 = 95*113*95*30*4 = **122.38 MB**

I_4 = 93*111*93*60*4 = **230.4 MB**

I_5 = 46*55*46*60*4 = **51.50 MB**

I_6 = 44*53*44*120*4 **= 49.25 MB**

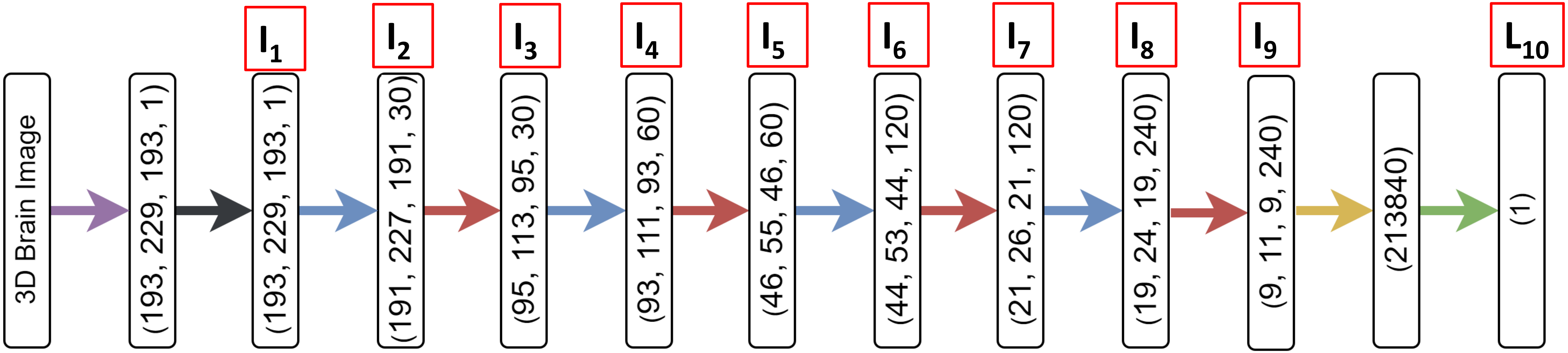
I_7 = 21*26*21*120*4 = **5.50 MB**

I_8 = 19*24*19*240*4 = **8.32 MB**

I_9 = 9*11*9*240*4 **= 0.86 MB**

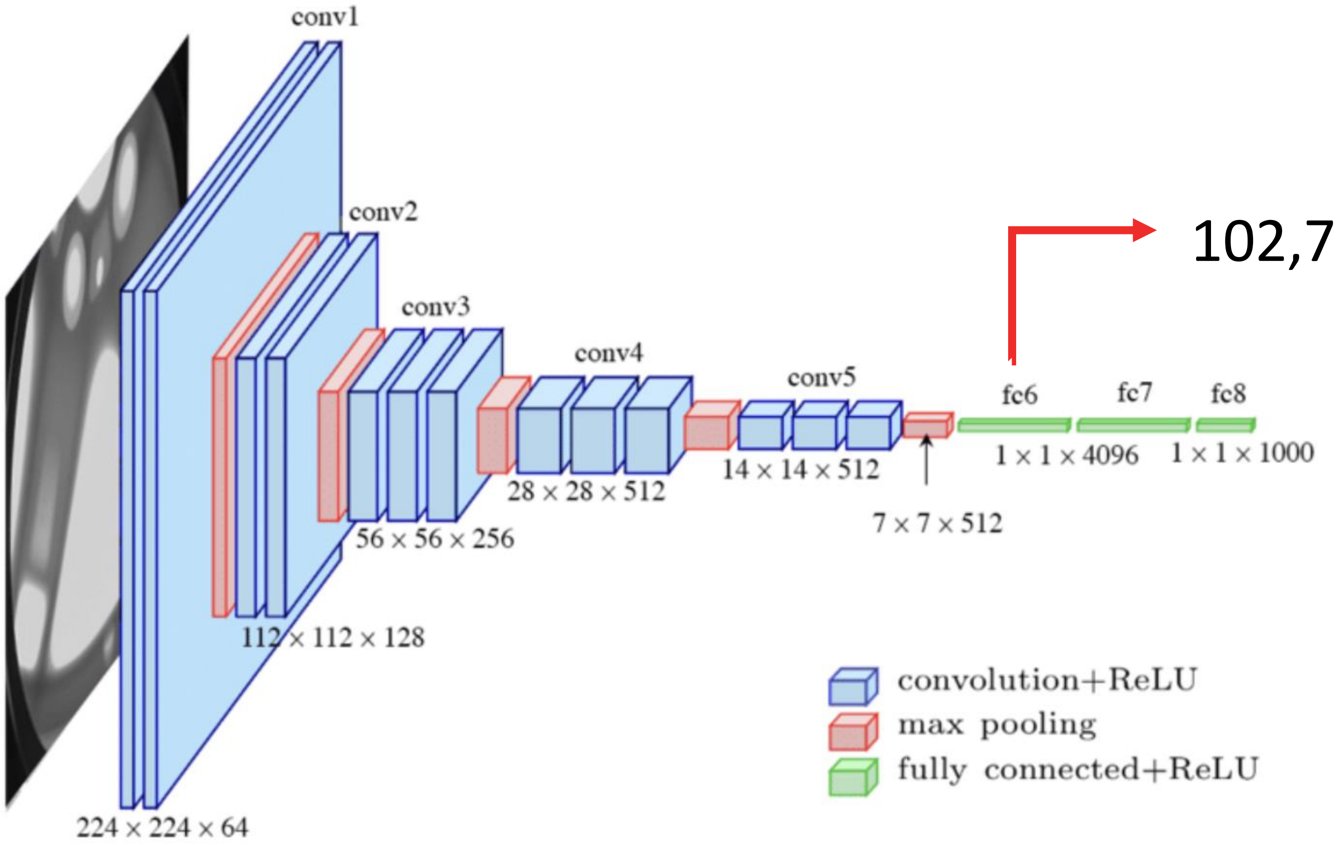
I_{10} = 1*4 = **4e-6 MB**

$$\boxed{\boldsymbol{Batch\ mem = Params + Grads + bs \times \left(I_1 + 2 \times \sum_{i=2}^{10} I_i\right)}}$$



UNIVERSITY OF CALGARY

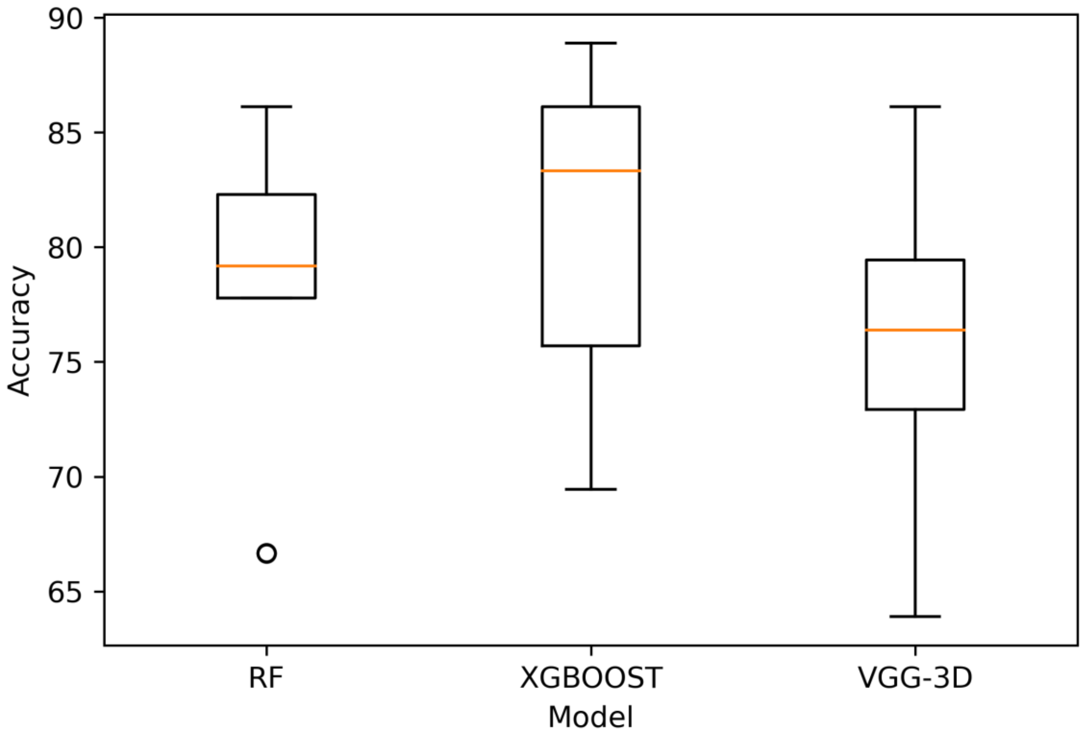
VGG-16 Revisited

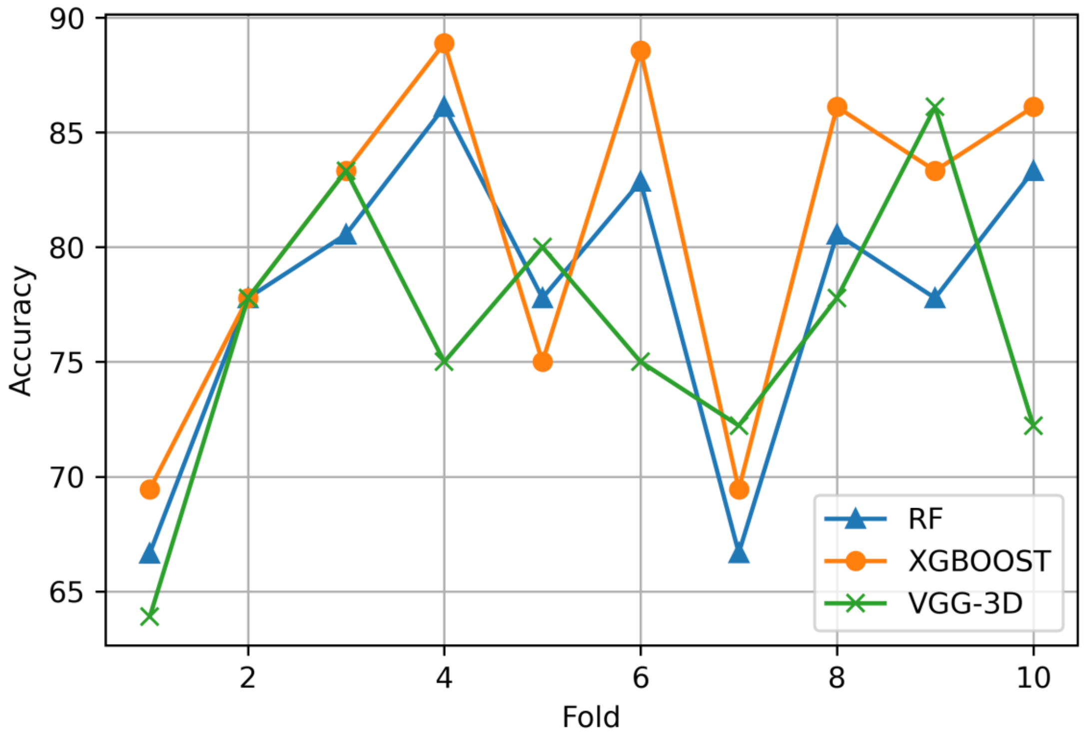


- Total number of parameters 138,357,544

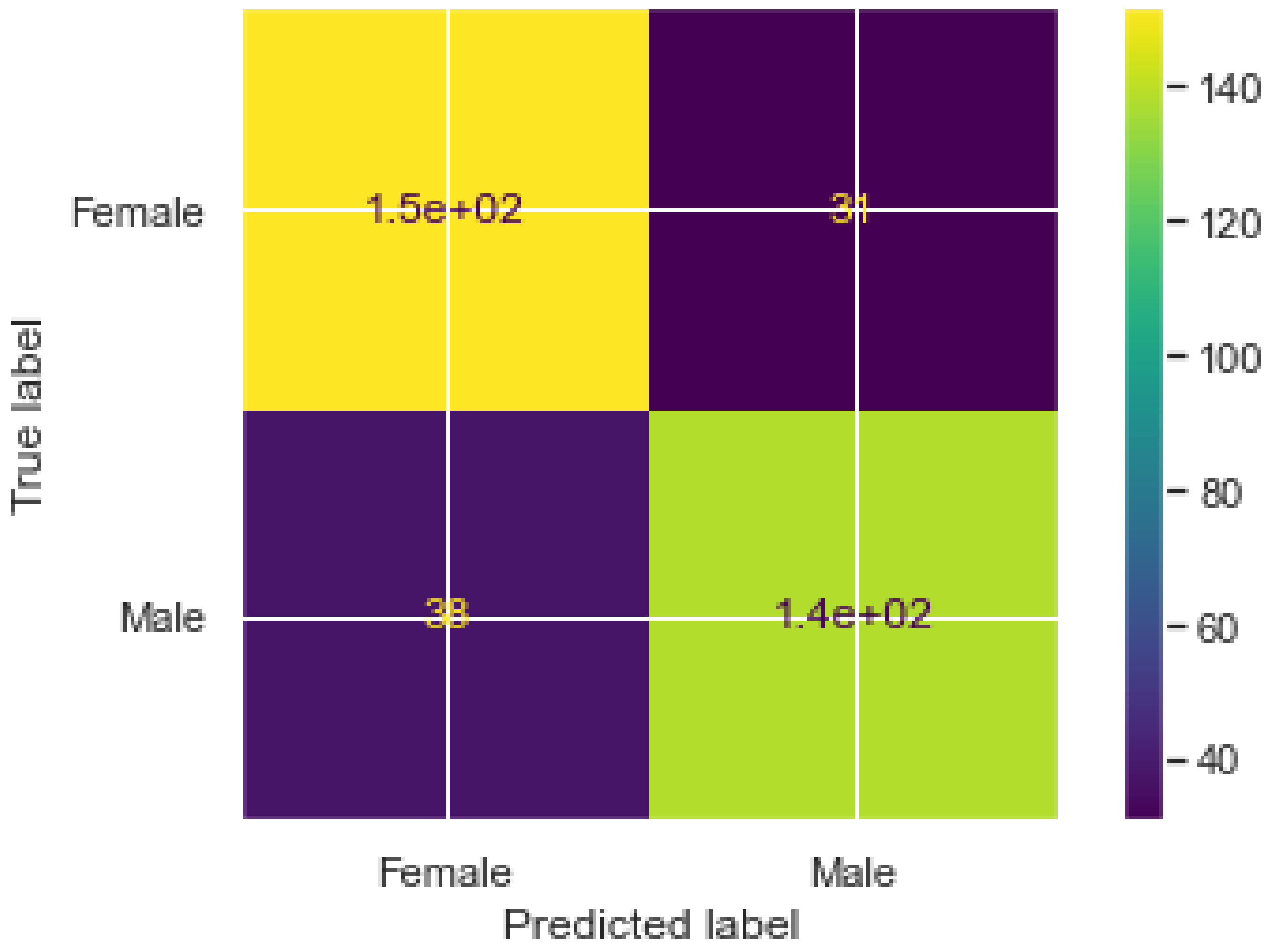
Results

Accuracy Results

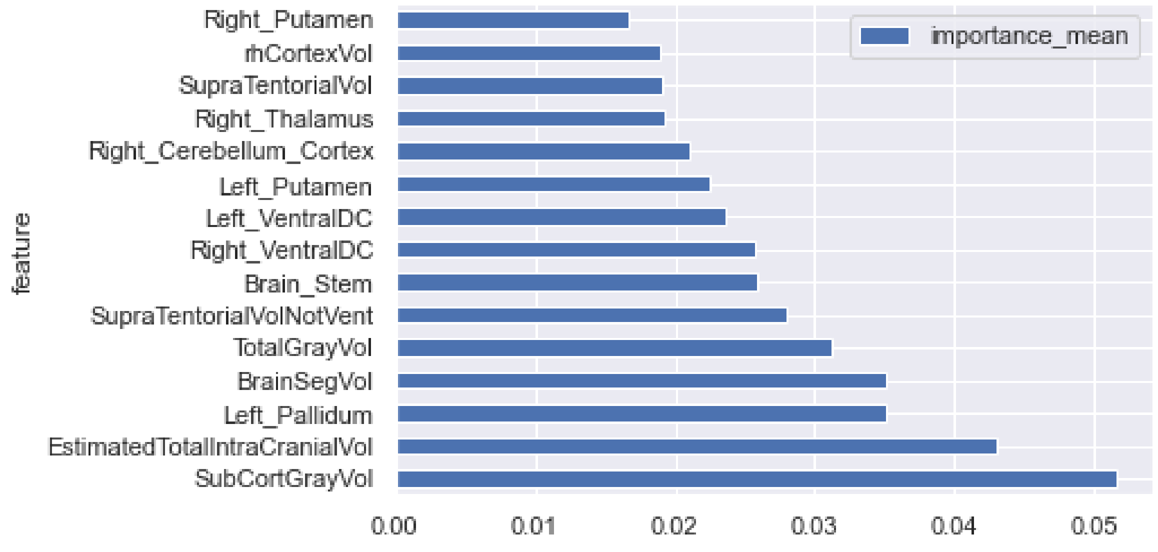




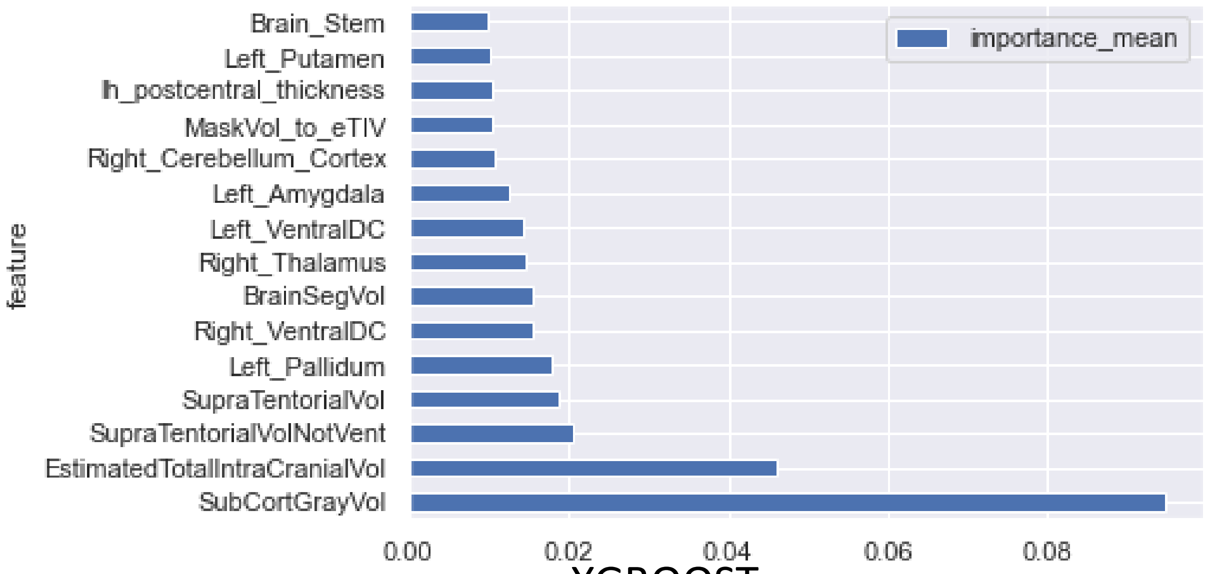
Confusion Matrix



Feature Importance Analysis (Top-15 Features)

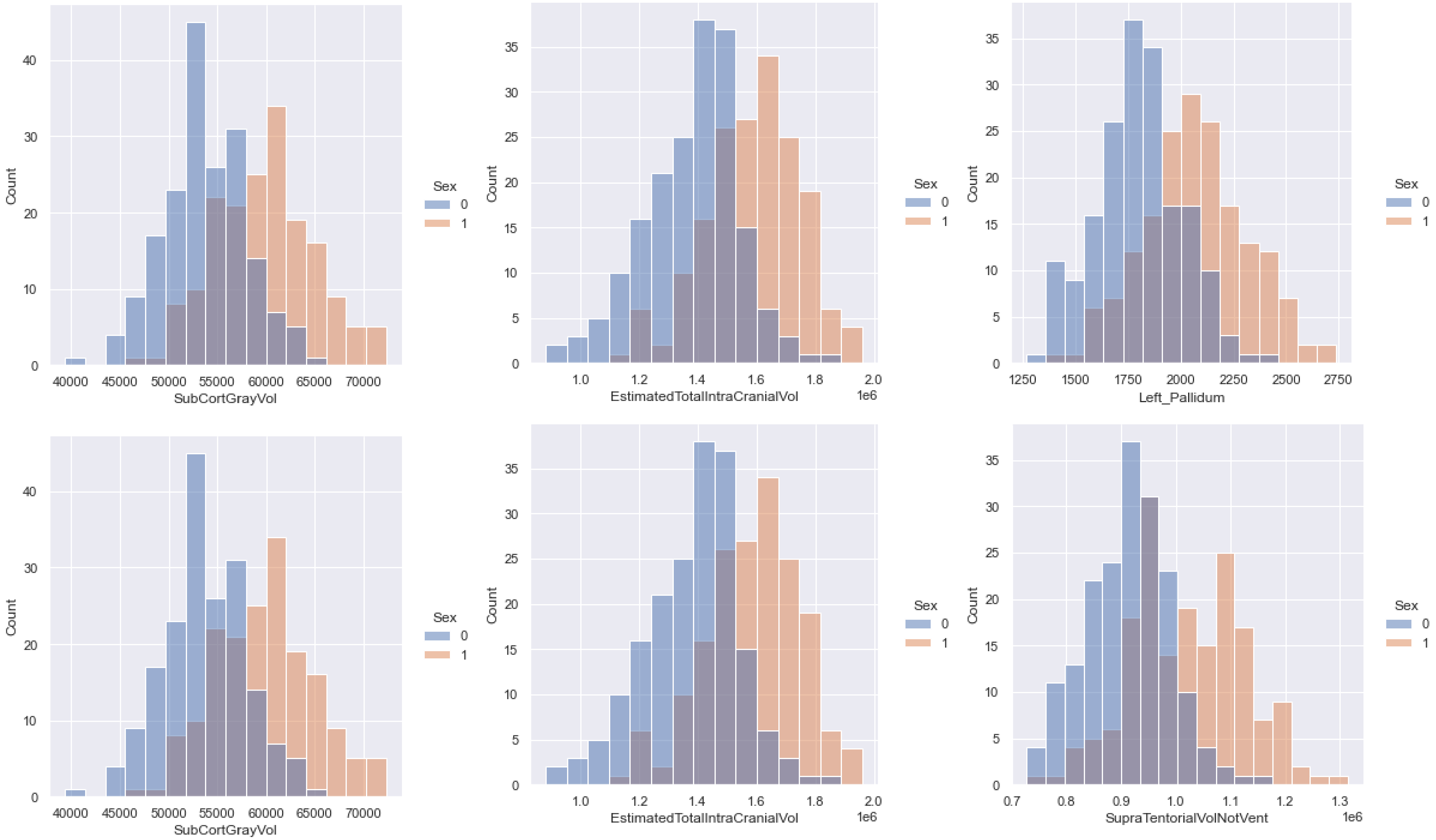


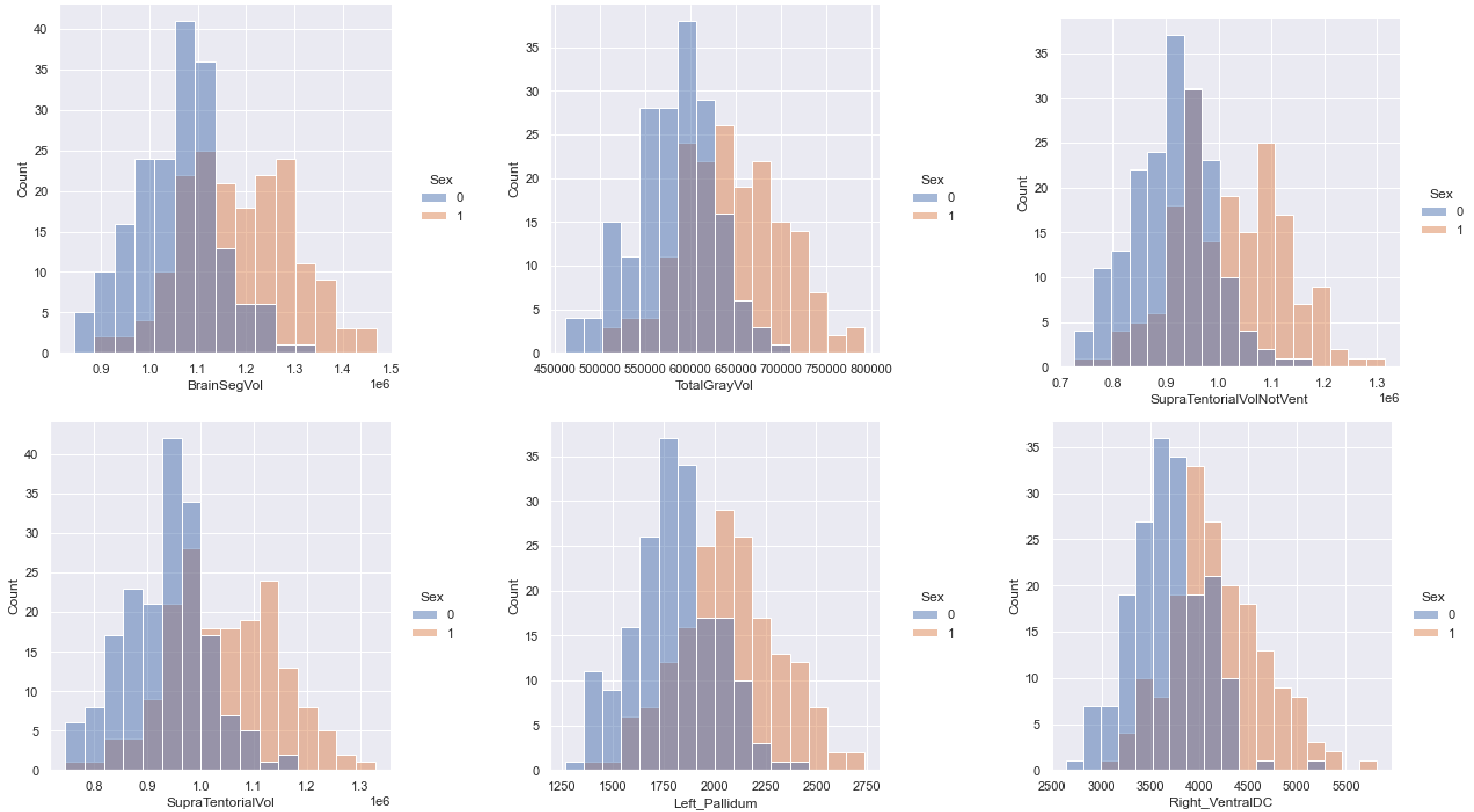
Random Forest



XGBOOST

UNIVERSITY OF CALGARY

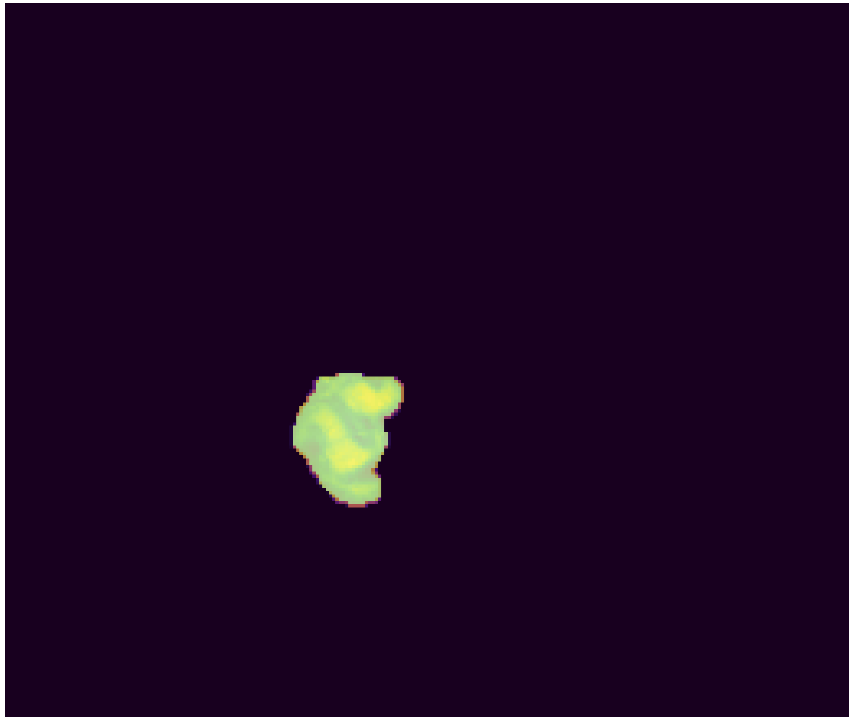

Count
SubCortGrayVol
40000 45000 50000 55000 60000 65000 70000
Sex
0
1
Count
EstimatedTotalIntraCranialVol
1.0 1.2 1.4 1.6 1.8 2.0
1e6
Sex
0
1
Count
Left_Pallidum
1250 1500 1750 2000 2250 2500 2750
Sex
0
1
Count
SubCortGrayVol
40000 45000 50000 55000 60000 65000 70000
Sex
0
1
Count
EstimatedTotalIntraCranialVol
1.0 1.2 1.4 1.6 1.8 2.0
1e6
Sex
0
1
Count
SupraTentorialVolNotVent
0.7 0.8 0.9 1.0 1.1 1.2 1.3
1e6
Sex
0
1



Count
BrainSegVol
1e6
Sex
0
1
0 5 10 15 20 25 30 35 40
0.9 1.0 1.1 1.2 1.3 1.4 1.5
TotalGrayVol
0 5 10 15 20 25 30 35
450000 500000 550000 600000 650000 700000 750000 800000
SupraTentorialVolNotVent
0.7 0.8 0.9 1.0 1.1 1.2 1.3
SupraTentorialVol
0.8 0.9 1.0 1.1 1.2 1.3
Left_Pallidum
1250 1500 1750 2000 2250 2500 2750
Right_VentralDC
2500 3000 3500 4000 4500 5000 5500

GRAD-CAM Visualization

Male



Female

inferno palette

Grad-CAM Visualization

Discussion and Next Steps

Discussion

- We were able to obtain > 80% accuracy as hypothesized
- Traditional ML outperformed DL, perhaps due to the small dataset size for training?
- Models seems to have learned different things
  - Model combination can be used to improve results?
- Better hyperparameter tuning and feature selection could be used to improve traditional ML results
- More data, data augmentation or a more modern CNN architecture could be used to improve DL results
- GRAD-CAM results are difficult to interpret

Next Steps

- Further improve the models – perhaps add ADNI dtata?
- Build atlas using GRAD-CAM visualizations highlighting differences between male and female brain
- Assess the utility of the information learned to inform the development of classification models of neurological diseases that are known to affect men and women differently

Thank you!