Parameters and Memory Consumption of CNNs

Jacob Idoko
Electrical and Software Engineering
Schulich School of Engineering

F2024

UNIVERSITY OF CALGARY

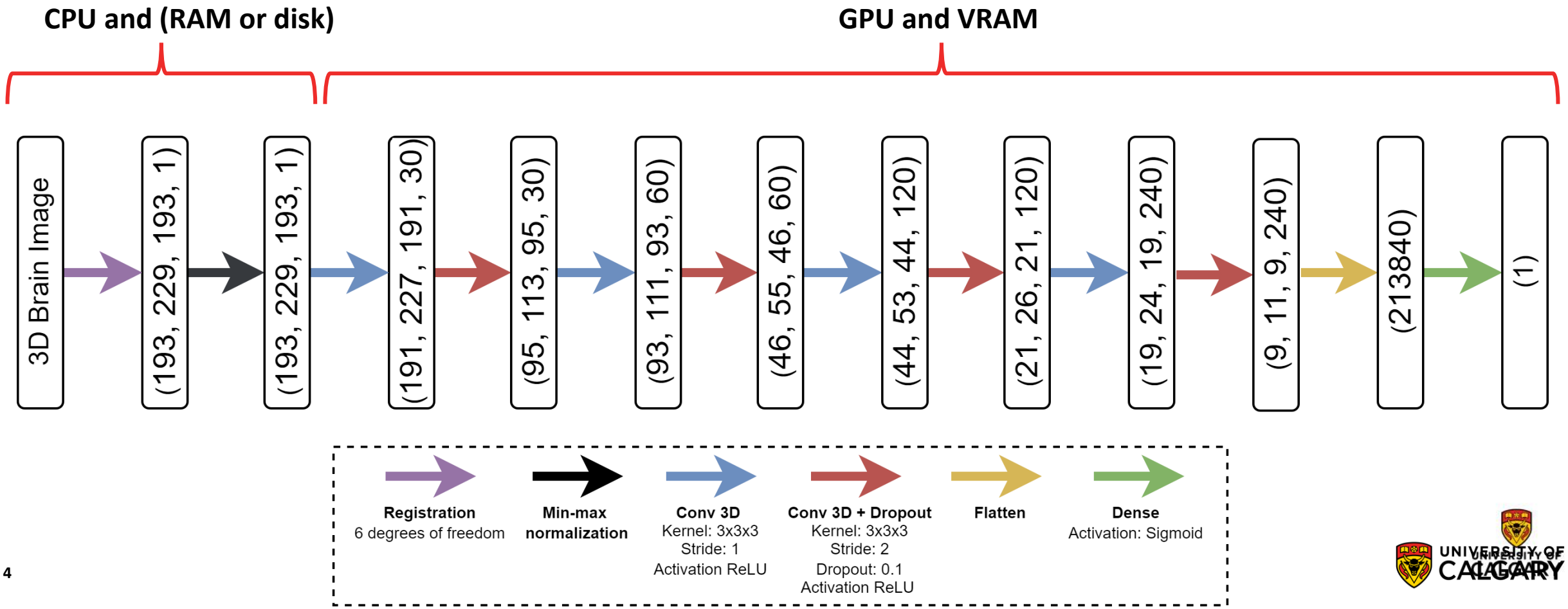
Outline

- Learning Goals
- Compute number of CNN parameters
- Compute GPU memory consumption
- Summary

Learning Goals

- Be able to compute number of parameters of a CNN
- Estimate GPU memory consumption during training/testing

3D CNN - Network Architecture



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**

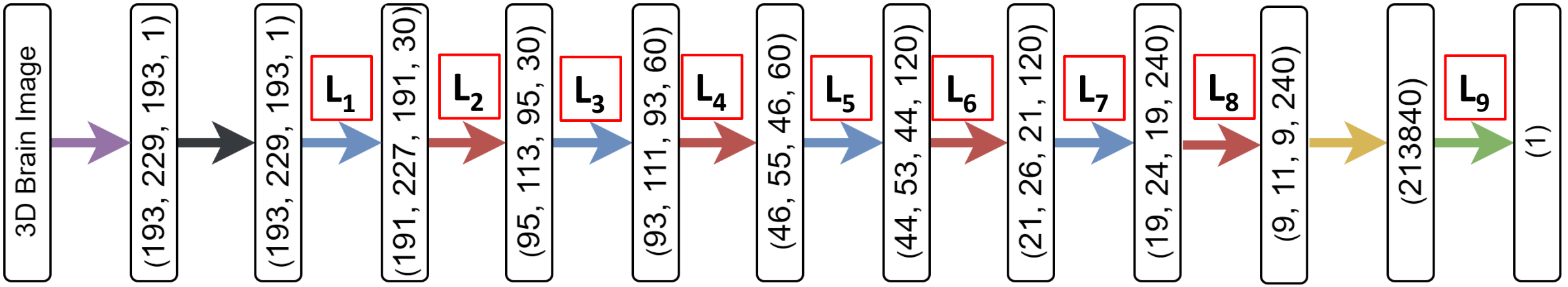
L_6 = (27*120 +1)*120 = **388,920**

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

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

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

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



Registration
6 degrees of freedom

Min-max normalization

Conv 3D
Kernel: 3x3x3
Stride: 1
Activation ReLU

Conv 3D + Dropout
Kernel: 3x3x3
Stride: 2
Dropout: 0.1
Activation ReLU

Flatten

Dense
Activation: Sigmoid

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**

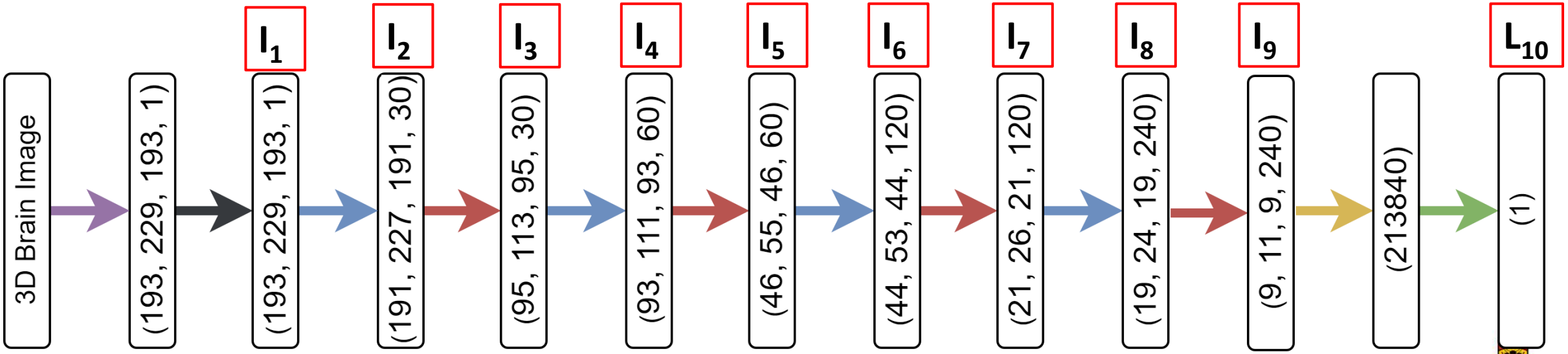
I_7 = 21*26*21*120*4 = **5.50 MB**

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**

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



Summary

- Understanding the number of parameters and GPU memory consumption is important:
  - It allows you to estimate if the hardware available is sufficient for training the desired model
  - It allows to identify layers with the most number of parameters and how to potentially alter them if facing problems, such as overfitting

Thank you!

UNIVERSITY OF CALGARY