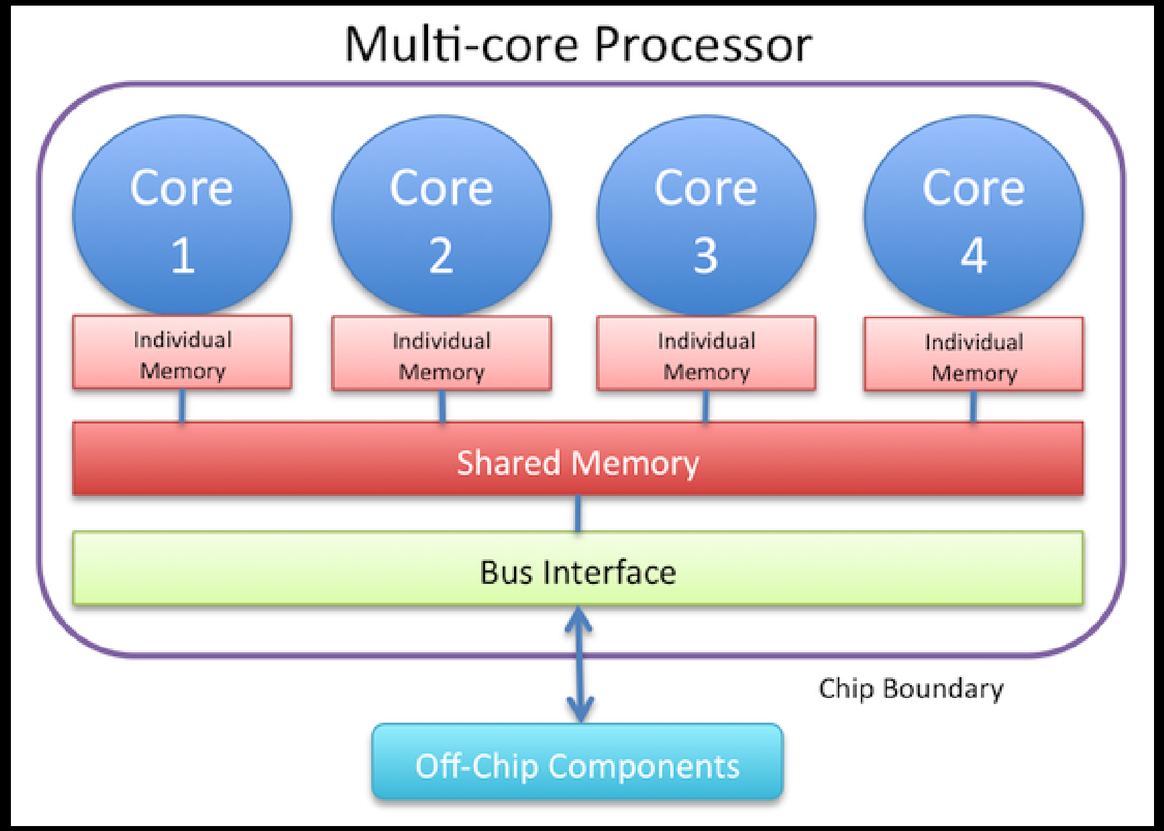
PARALLELISING ☹

CNNs

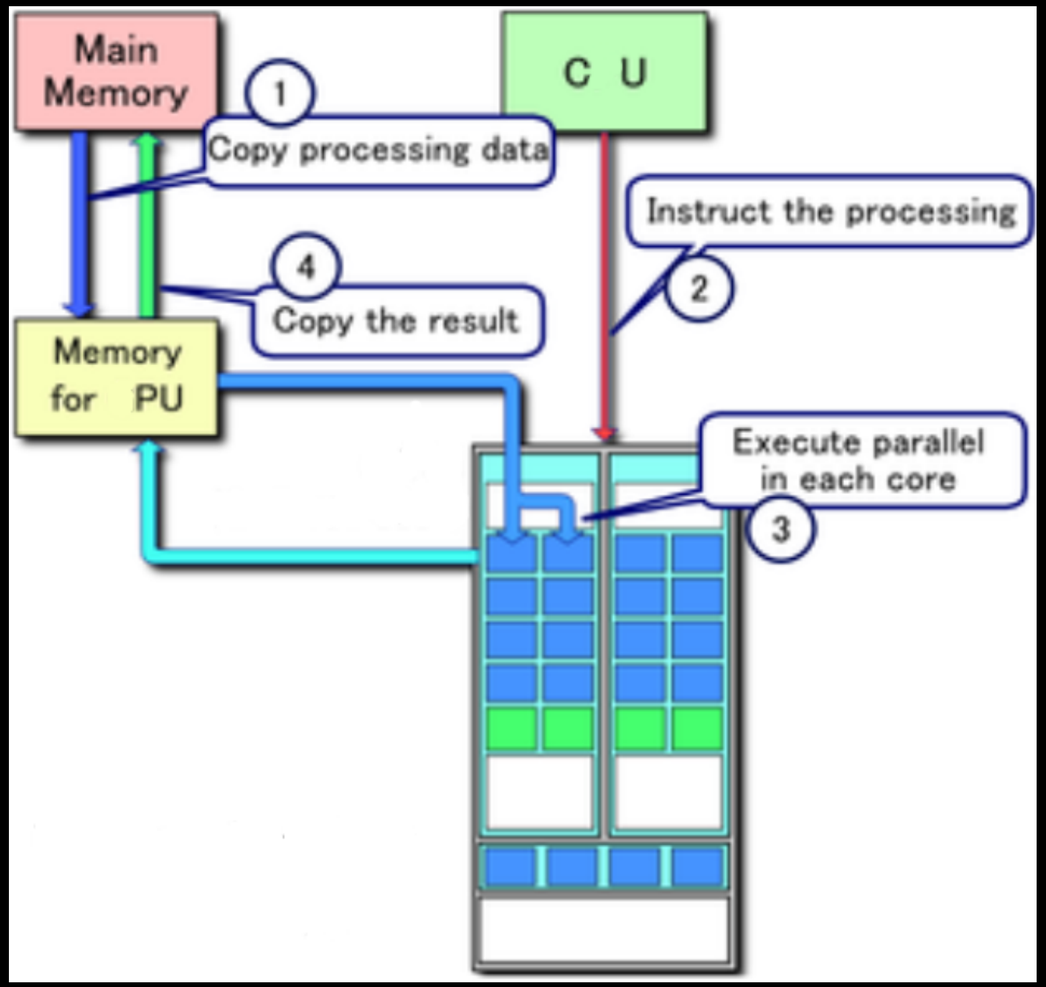
Achintya

Lohith

MULTI-CORE

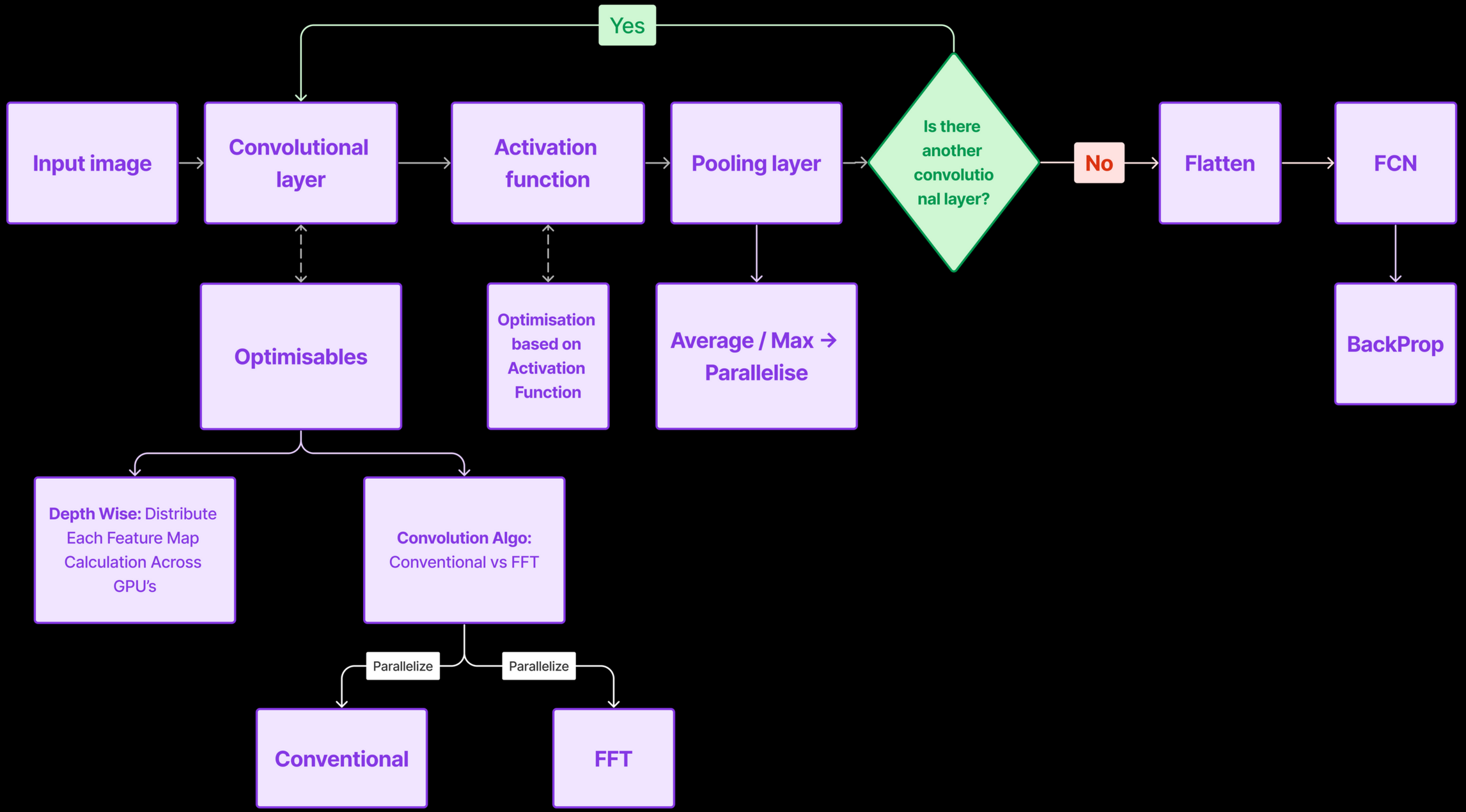


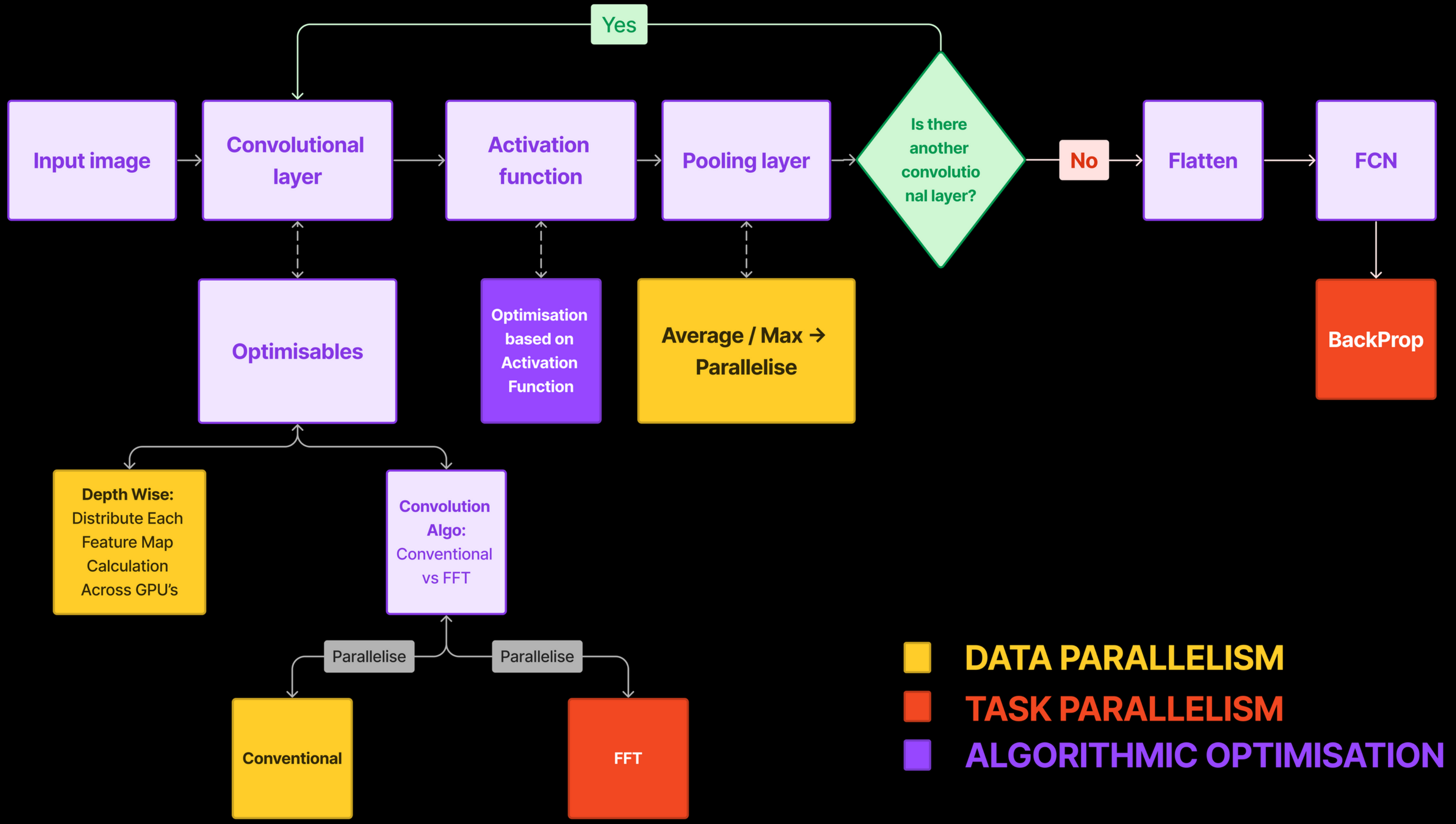
GPU

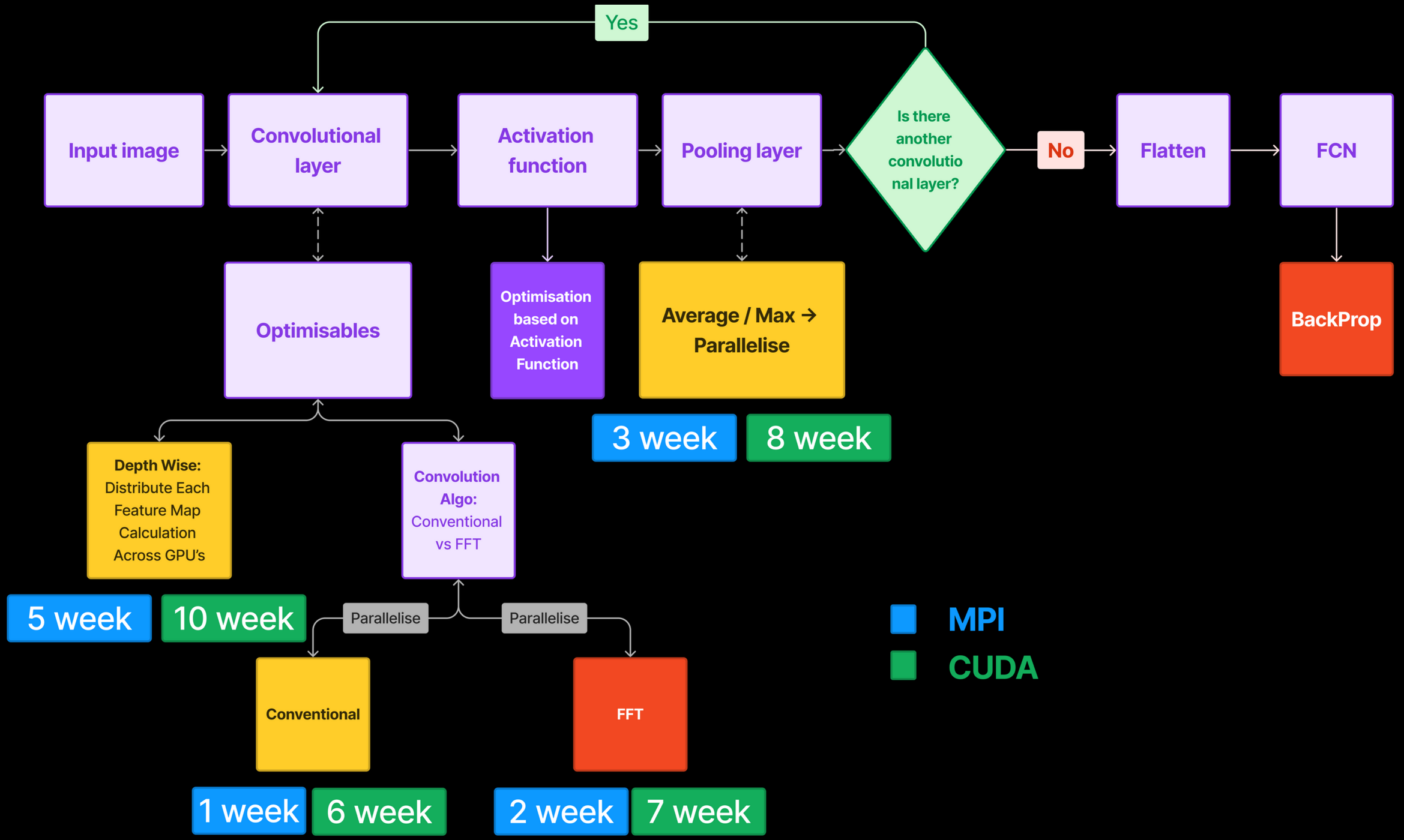


SIMD

BLOCK DIAGRAM




Yes
Input image
Convolutional layer
Activation function
Pooling layer
Is there another convolutional layer?
No
Flatten
FCN
BackProp
Optimisables
Optimisation based on Activation Function
Average / Max → Parallelise
Depth Wise: Distribute Each Feature Map Calculation Across GPU's
Convolution Algo: Conventional vs FFT
Parallelise
Parallelise
Conventional
FFT
DATA PARALLELISM
TASK PARALLELISM
ALGORITHMIC OPTIMISATION


Yes
Input image
Convolutional layer
Activation function
Pooling layer
Is there another convolutional layer?
No
Flatten
FCN
BackProp
Optimisables
Optimisation based on Activation Function
Average / Max → Parallelise
3 week
8 week
Depth Wise: Distribute Each Feature Map Calculation Across GPU's
Convolution Algo: Conventional vs FFT
5 week
10 week
Parallelise
Parallelise
Conventional
FFT
1 week
6 week
2 week
7 week
MPI
CUDA

CHALLENGES

CONSISTENCY ACROSS PLATFORMS

SERIAL NATURE OF NN'S

MEMORY MANAGEMENT

SCALABILITY

LYF