Massachusetts Institute of Technology

DEEP LEARNING: Facial Expression Recognition

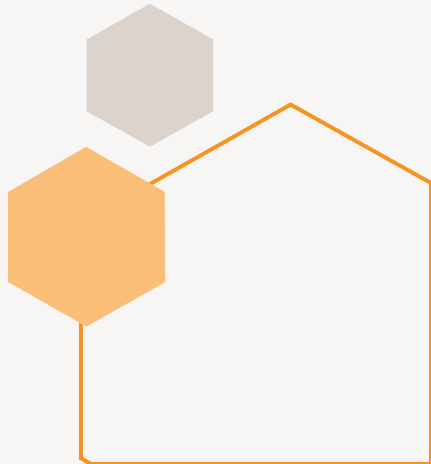
Samvel Sukiasian

MIT Applied Data Science Program: Leveraging AI for effective decision making, March 23’

Problem definition

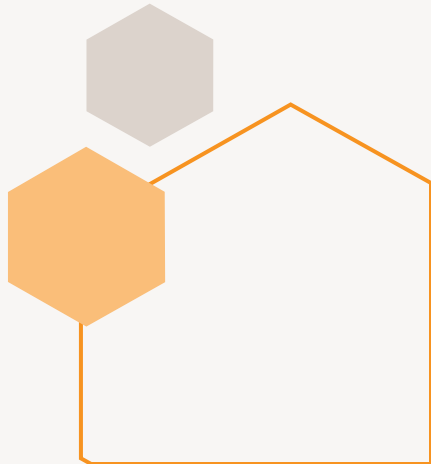
Recognize 1 of 4 facial expressions from an image:

- Happy
- Sad
- Surpised
- Neutral



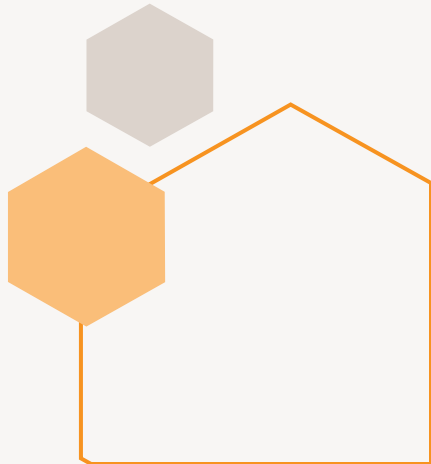
Models used

- 3 CNNs (including complex)
- VGG16
- Resnet
- EfficientNet

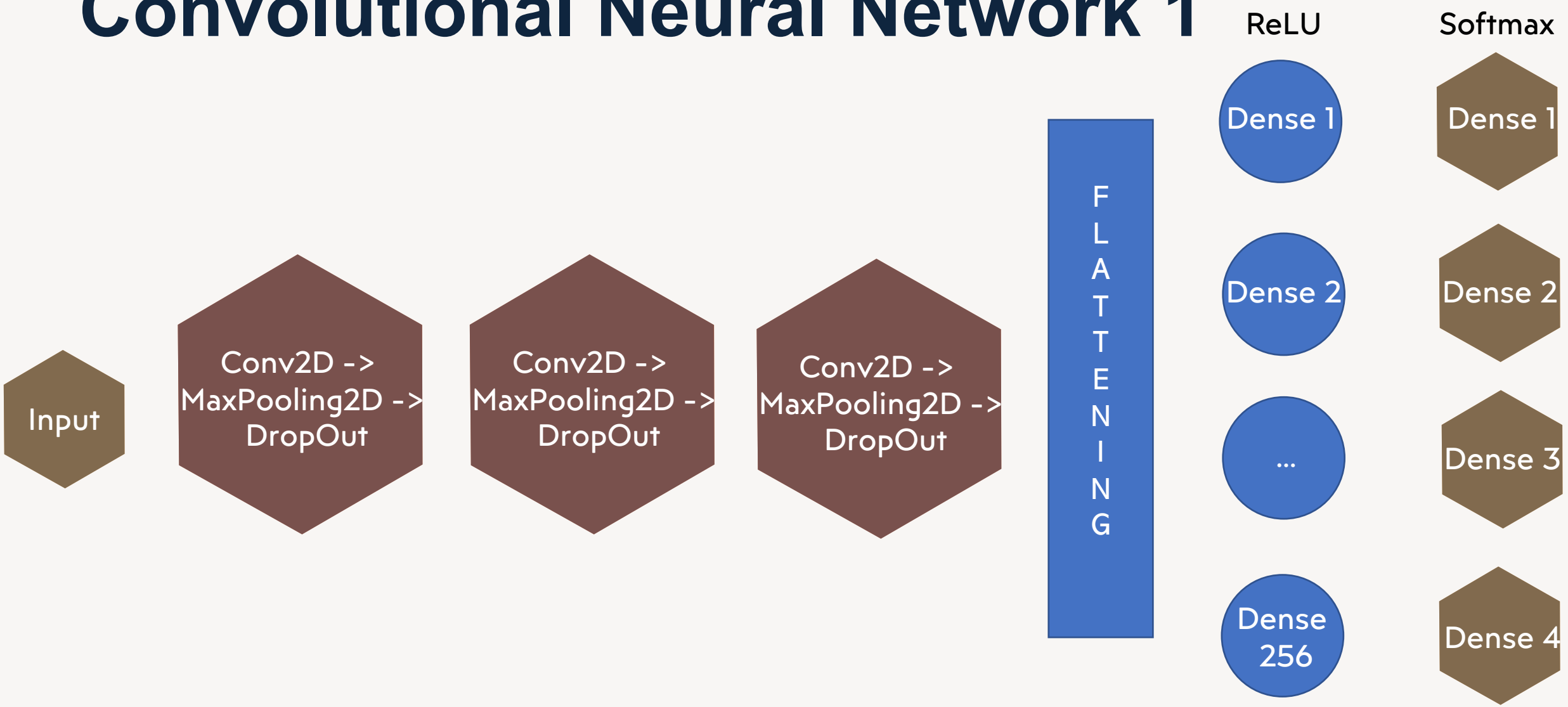


From Scratch Models

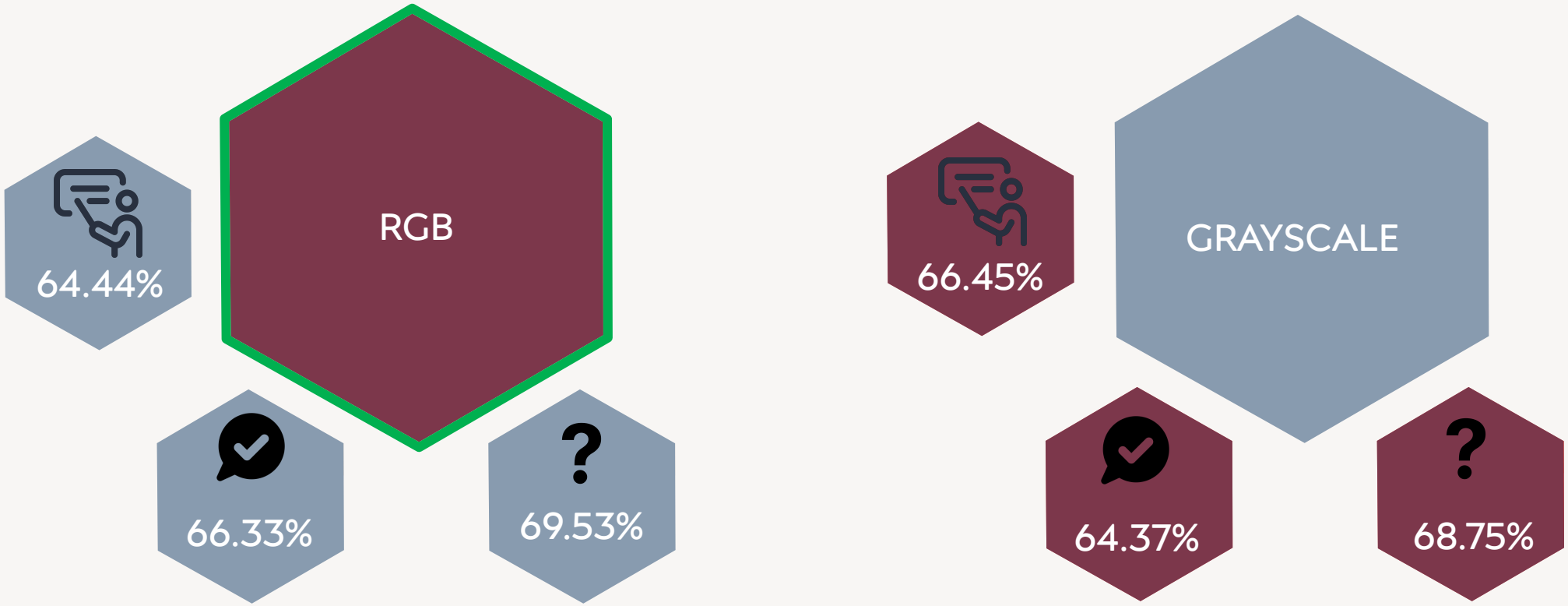
- Convolutional Neural Network 1
- Convolutional Neural Network 2
- Complex CNN



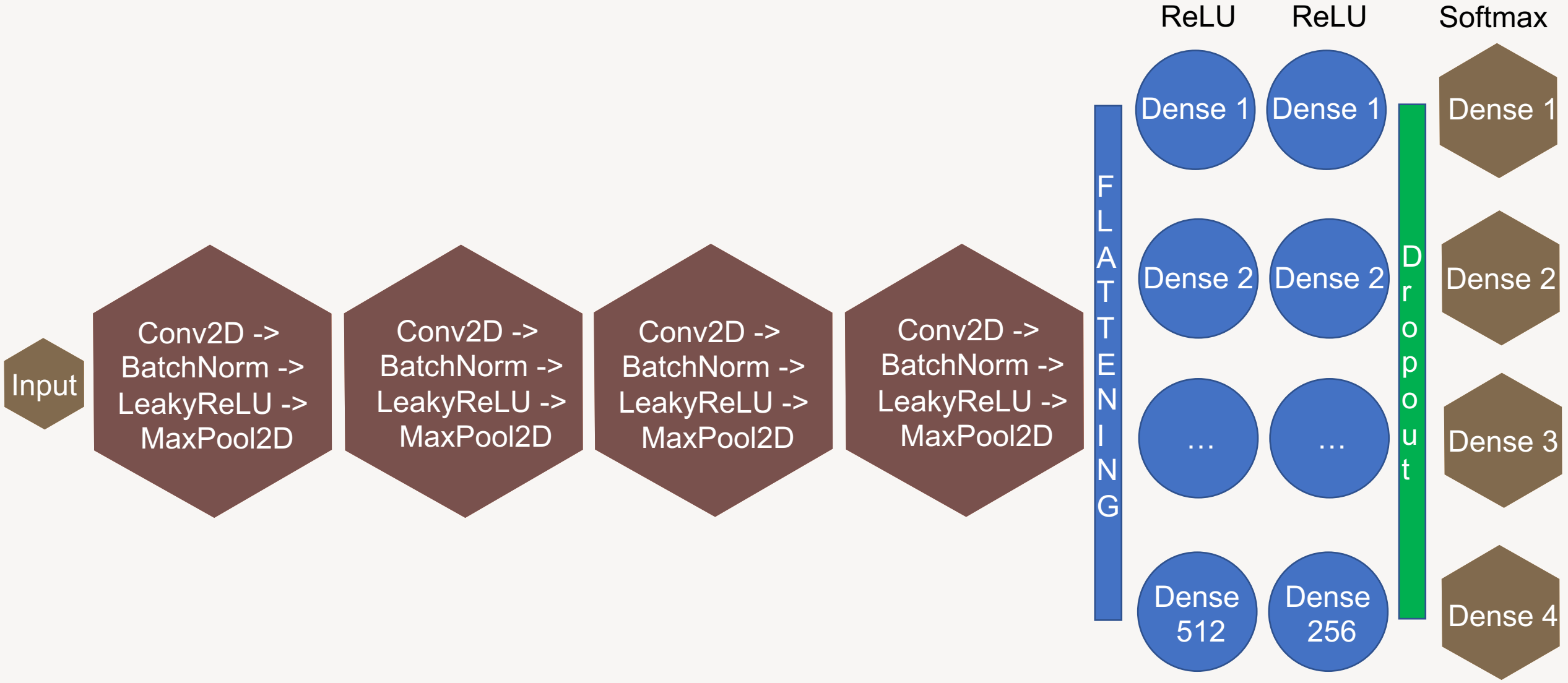
Convolutional Neural Network 1



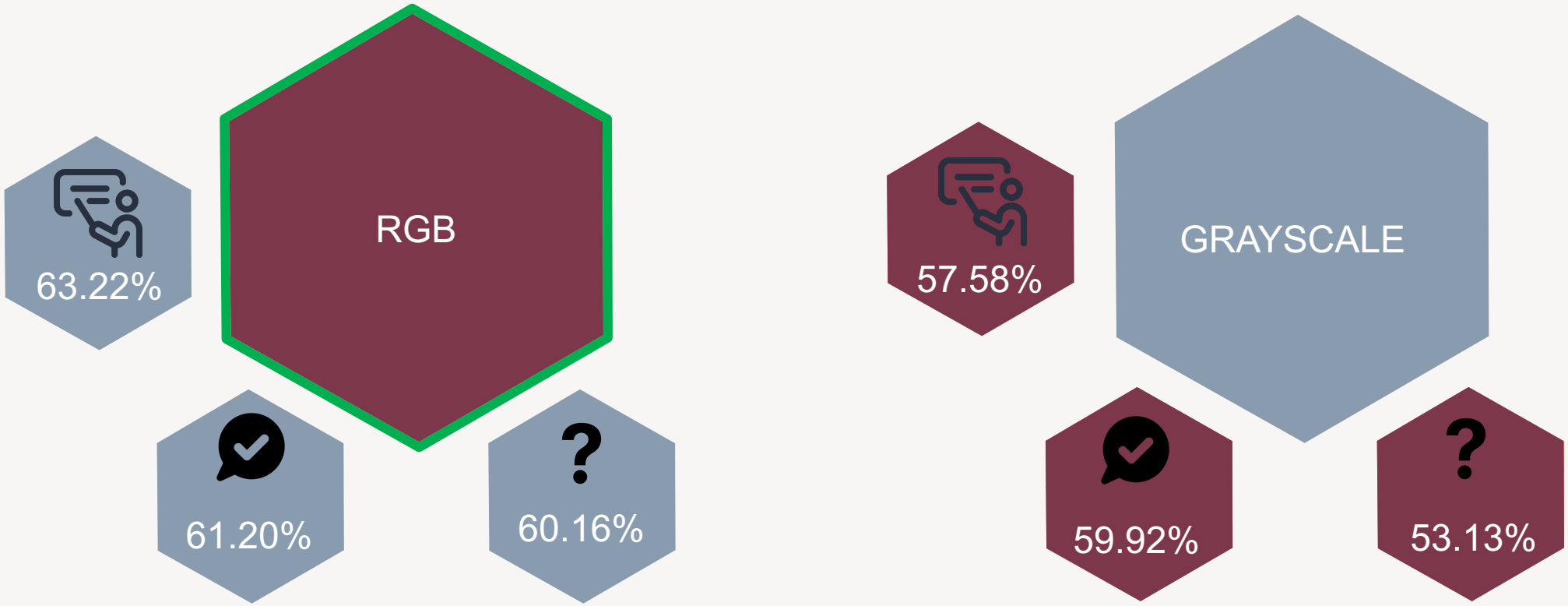
CNN 1: Accuracy results



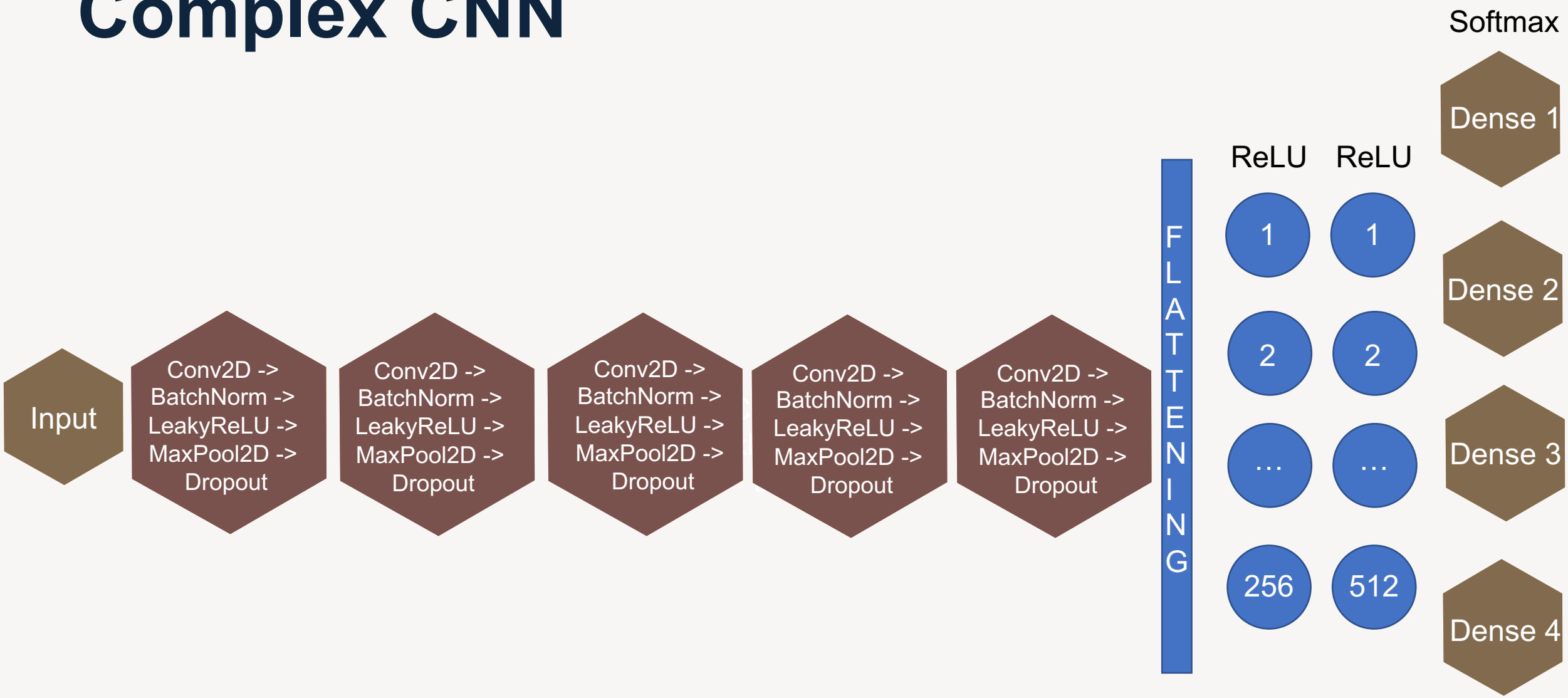
Convolutional Neural Network 2



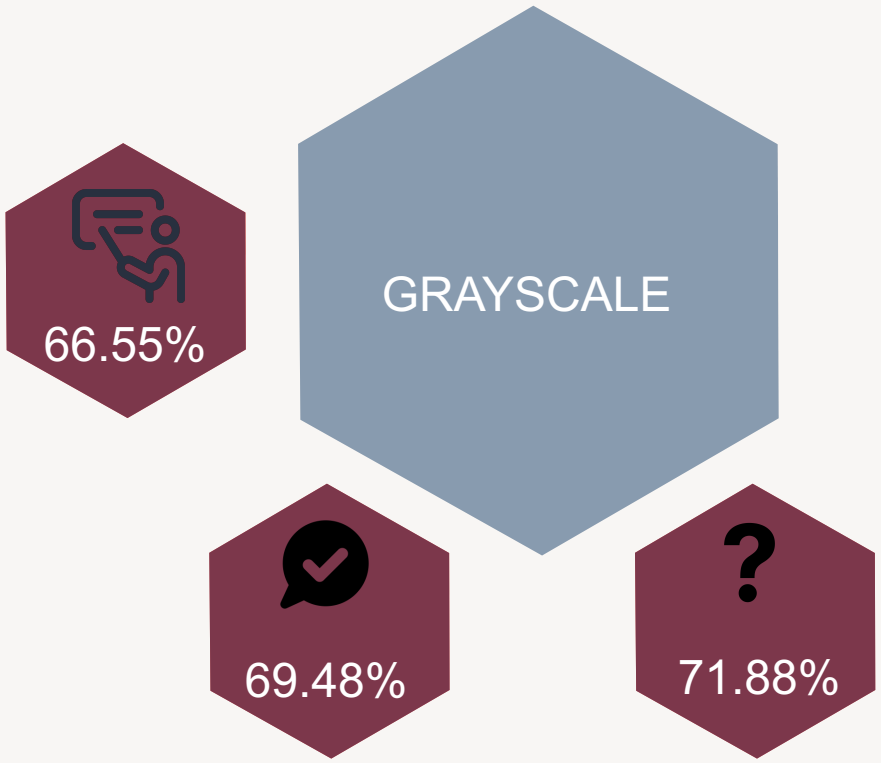
CNN 2: Accuracy results



Complex CNN

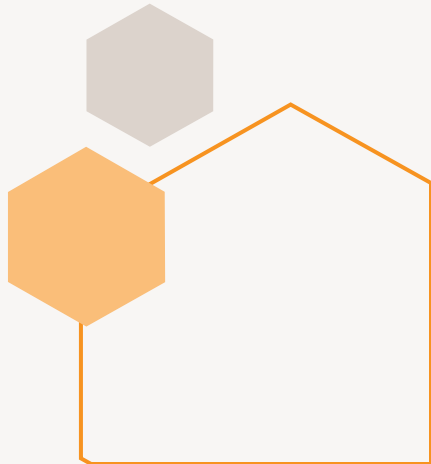

Softmax
Input
Conv2D -> BatchNorm -> LeakyReLU -> MaxPool2D -> Dropout
Conv2D -> BatchNorm -> LeakyReLU -> MaxPool2D -> Dropout
Conv2D -> BatchNorm -> LeakyReLU -> MaxPool2D -> Dropout
Conv2D -> BatchNorm -> LeakyReLU -> MaxPool2D -> Dropout
Conv2D -> BatchNorm -> LeakyReLU -> MaxPool2D -> Dropout
FLATTENING
ReLU
ReLU
1
1
2
2
...
...
256
512
Dense 1
Dense 2
Dense 3
Dense 4

Complex CNN: Accuracy results

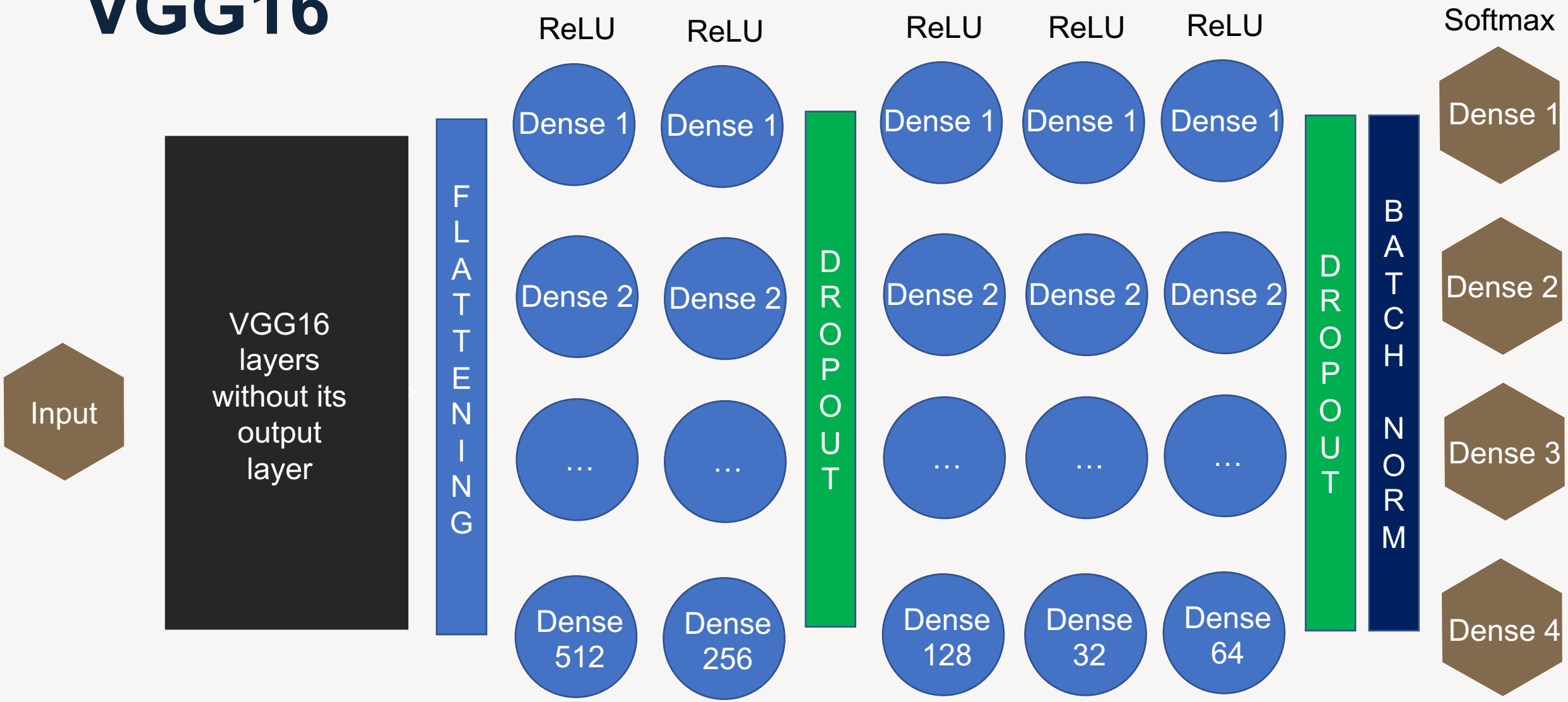


Transferable models

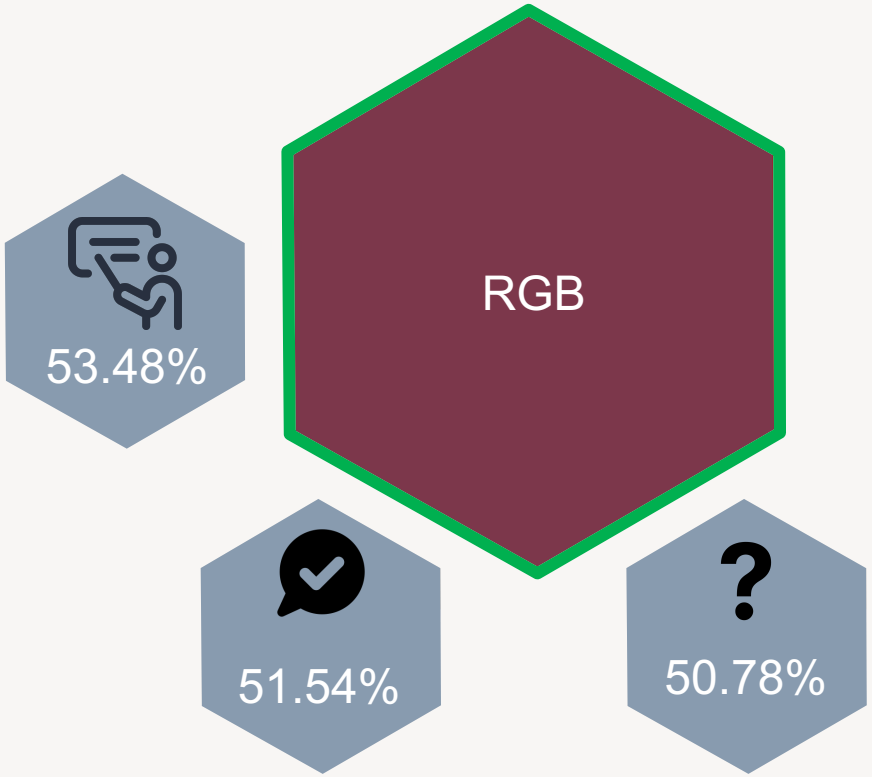
- VGG16
- ResNet
- EfficientNet



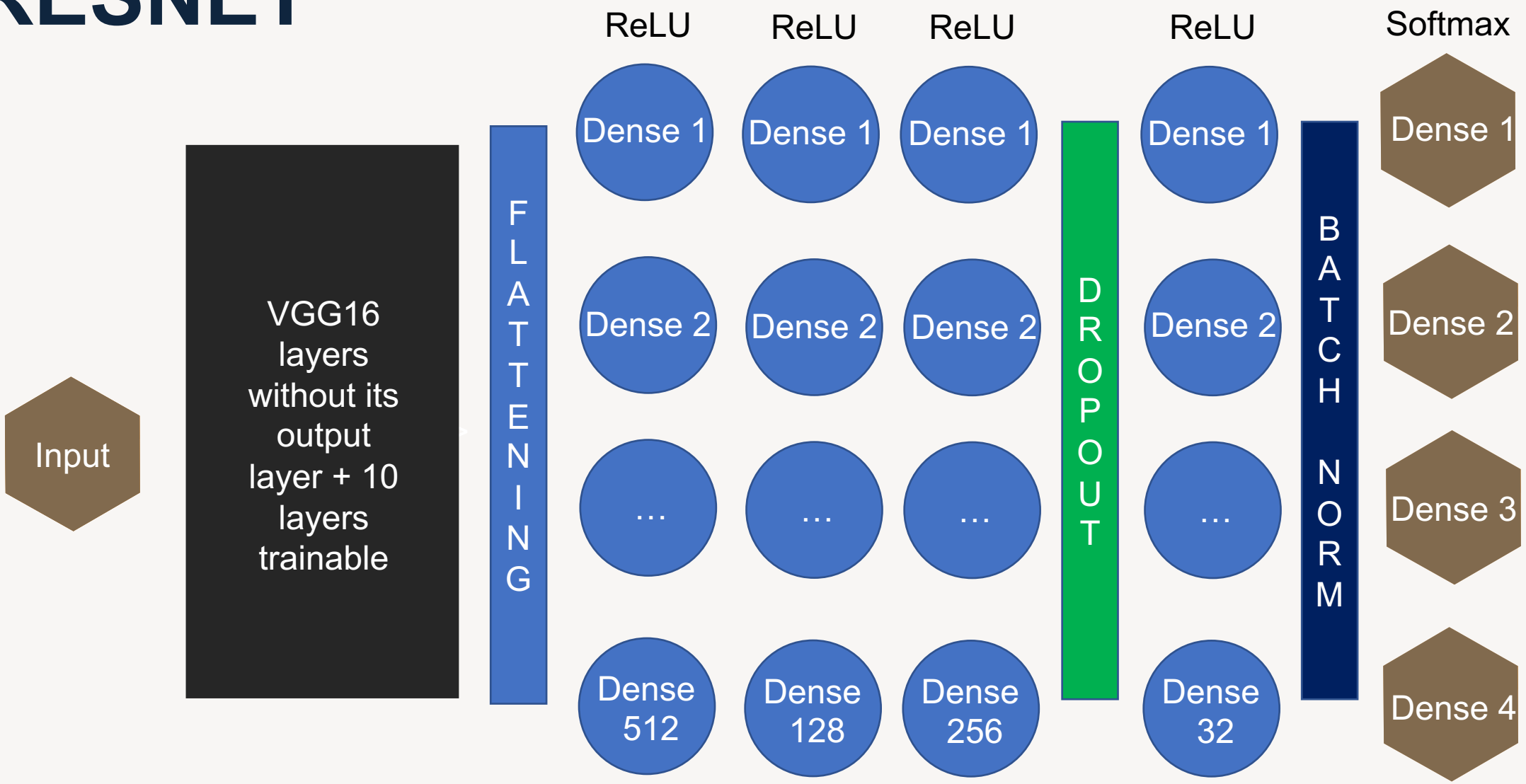
VGG16


Input
VGG16 layers without its output layer
FLATTENING
ReLU
Dense 1
Dense 2
…
Dense 512
ReLU
Dense 1
Dense 2
…
Dense 256
DROPOUT
ReLU
Dense 1
Dense 2
…
Dense 128
ReLU
Dense 1
Dense 2
…
Dense 32
ReLU
Dense 1
Dense 2
…
Dense 64
DROPOUT
BATCH NORM
Softmax
Dense 1
Dense 2
Dense 3
Dense 4

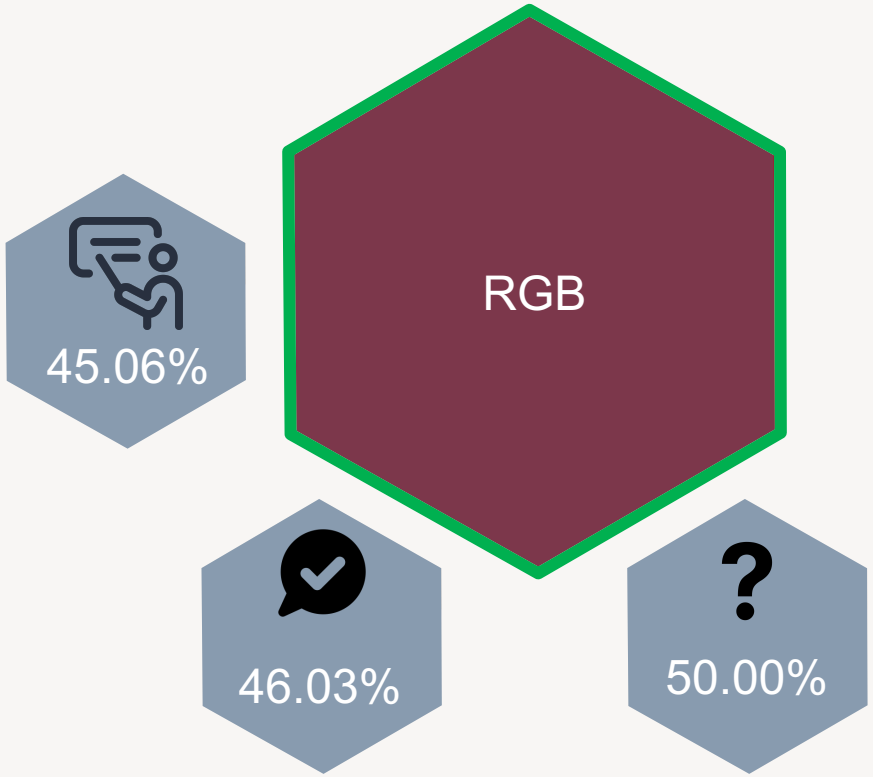
VGG16: Accuracy results



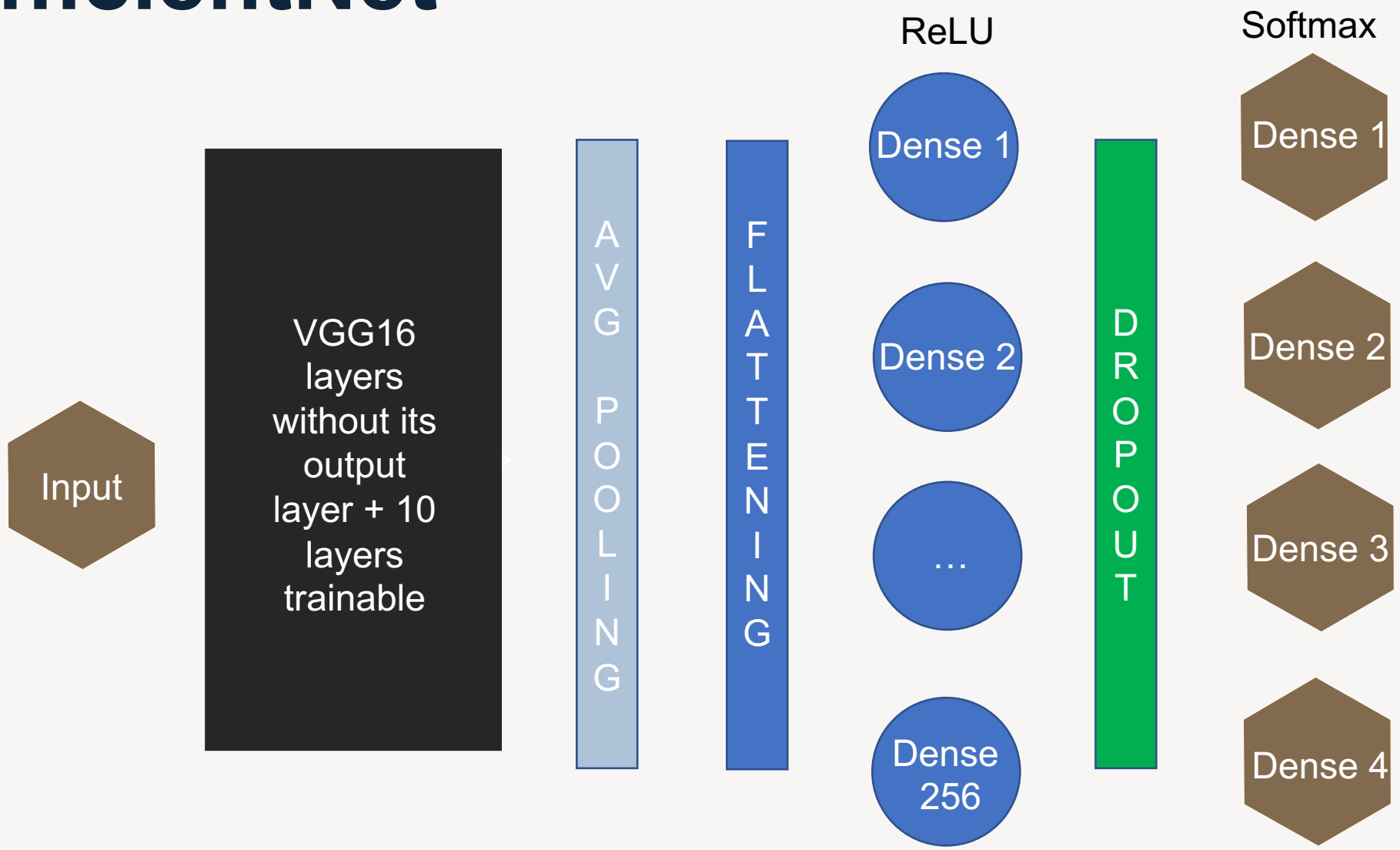
RESNET


ReLU
ReLU
ReLU
ReLU
Softmax
Input
VGG16 layers without its output layer + 10 layers trainable
FLATTENING
Dense 1
Dense 2
...
Dense 512
Dense 1
Dense 2
...
Dense 128
Dense 1
Dense 2
...
Dense 256
DROPOUT
Dense 1
Dense 2
...
Dense 32
BATCH NORM
Dense 1
Dense 2
Dense 3
Dense 4

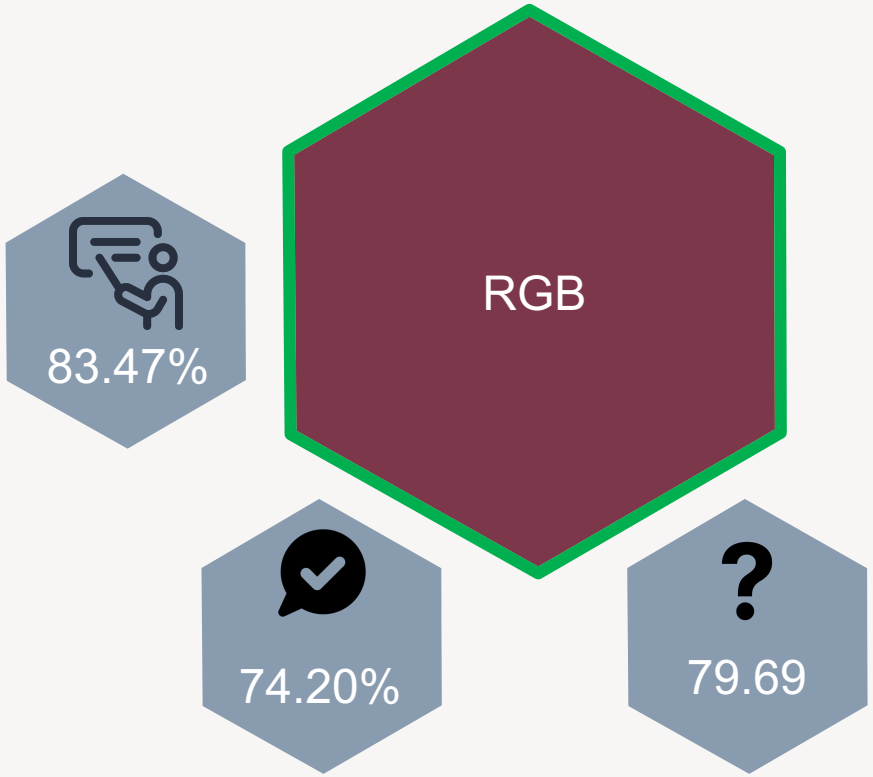
RESNET: Accuracy results



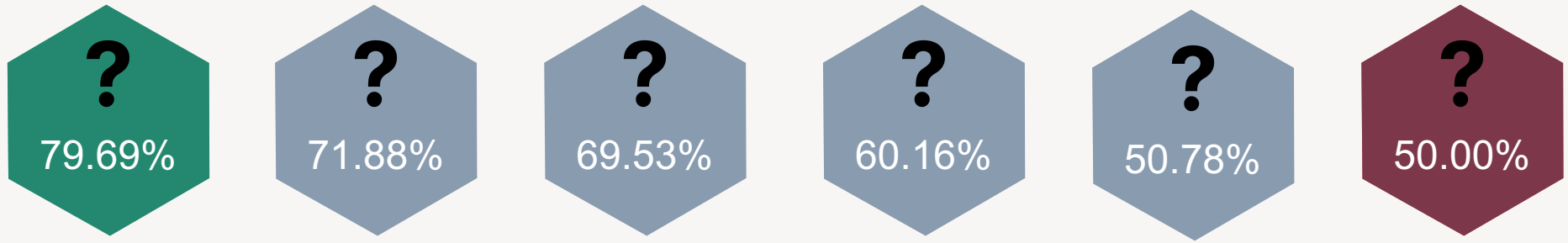
EfficientNet



EfficientNet: Accuracy results



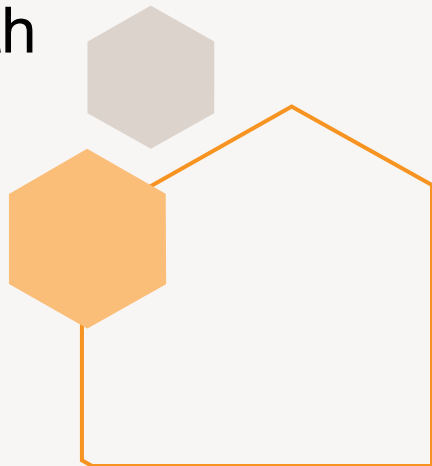
EfficientNet: Accuracy results



EfficientNet Complex CNN CNN 1 CNN 2 VGG ResNet

Recommendations

- Depending on the goal, to pay attention to high quality datasets for a better quality
- Access more powerful computational resources for better training
- Make an emphasize on Transfer Learning
- Use RGB instead of Grayscale
- ”Play” with different architectures and transfer layers, as well as with hyperparameters to improve performance



Thank you

Samvel Sukiasian

sukiasiansam@gmail.com