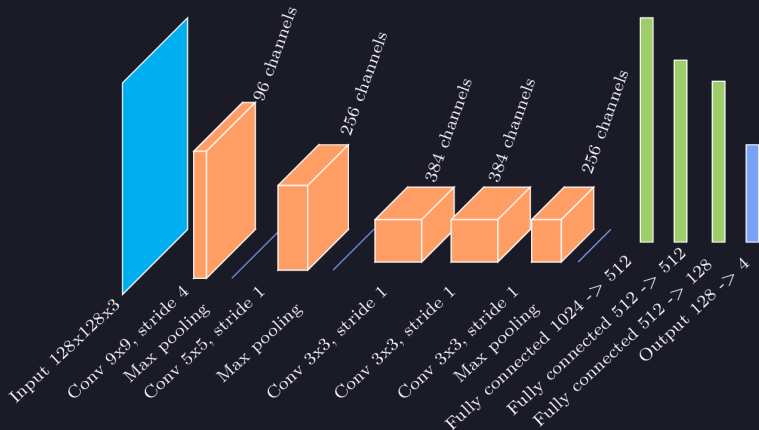
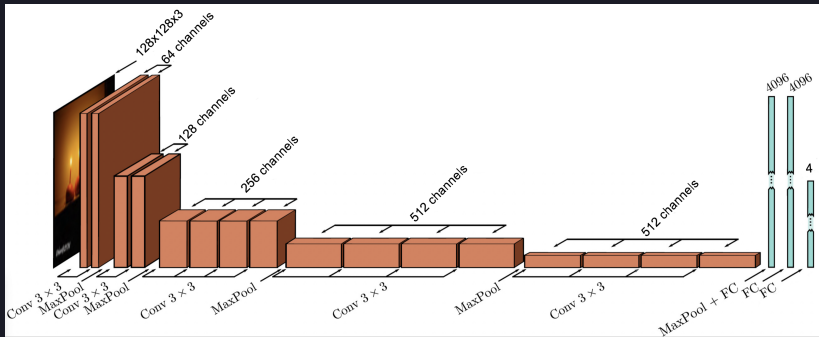


AlexNet



Number of parameters: 4589316

VGG



Number of parameters: 65070916

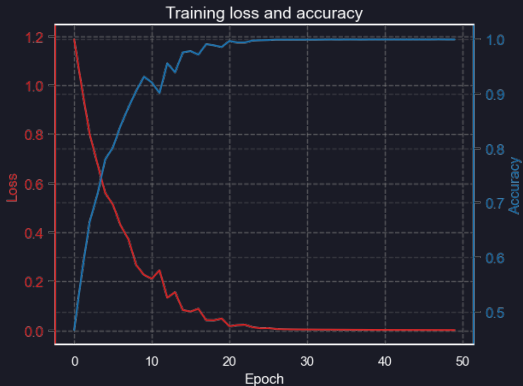
Dropout rate: 0.5

Setup Differences

Model	Data augmentation	Scheduler	Activation	L2 regularization
CustomCNN	Yes ✓	Yes ✓	<i>Mish</i>	Yes ✓
AlexNet	No ✗	Yes ✓	<i>ReLU</i>	Yes ✓
VGG16	No ✗	No ✗	<i>ReLU</i>	No ✗
VIT	Yes ✓	Yes ✓	<i>Mish</i>	Yes ✓

- All the other hyperparameters and settings are the same for all models (*batch size, optimizer, epochs, etc...*)
- Note that the **CustomCNN** is the one with less parameters (3,001,156) while **VGG16** is the one with more parameters(65,070,916)
- **VGG16** also has the highest dropout rate (0.5)

Training Loss and Accuracy for AlexNet



- Final training loss: $1.2 \cdot 10^{-3}$
- Final training accuracy: 99.9%

Confidence and Test Accuracy for AlexNet



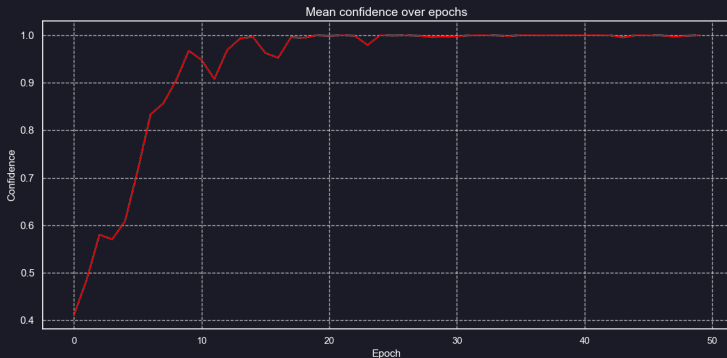
- Final training confidence: 99.9%
- Final test confidence: 96.5%
- Final test accuracy: 90%

Training Loss and Accuracy for VGG16



- Final training loss: $8.9 \cdot 10^{-6}$
- Final training accuracy: 99.9%

Confidence and Test Accuracy for VGG16



- Final training confidence: 100%
- Final test confidence: 98%
- Final test accuracy: 95%

Training Performance Comparison

Model	Loss	Accuracy	Confidence
CustomCNN	$1.4 \cdot 10^{-3}$	99%	100%
AlexNet	$1.2 \cdot 10^{-3}$	99%	99.9%
VGG16	$8.9 \cdot 10^{-6}$	99%	100%
VIT	0.27	90%	96.1%



Note that these are the values reached during the **last epoch**.

Focus on Accuracy



Test Performance Comparison

Model	Accuracy	Confidence
CustomCNN	99%	100%
AlexNet	90%	96.5%
VGG16	95%	98.0%
VIT	88%	93.3%



Note that these are the values reached after the **last epoch**.

Visualizing the first layer filters, CustomCNN



Visualizing the first layer filters, AlexNet



Visualizing the first layer filters, VGG16

