CNN Model Performance Analysis

Comparison of 5 Pretrained CNNs on 20 Test Images

CNN Model Comparison Summary Table

Model	Top-1 Accuracy (%)	Top-2 Accuracy (%)	Top-3 Accuracy (%)	Avg Time (s)	Avg Mem (MB)	Model Size (Params)
ResNet50	9.09	0.0	0.0	0.49	97.8	25,636,712
VGGNet16	9.09	0.0	0.0	0.32	527.79	138,357,544
InceptionV3	0.0	0.0	0.0	0.37	90.99	23,851,784
ConvNeXt	9.09	0.0	0.0	0.80	109.06	28,589,128
EfficientNet	0.0	4.55	0.0	1.63	254.28	66,658,687

Prediction Accuracy

- **Top-1 Accuracy**: ResNet50, VGGNet16, and ConvNeXt each scored **9.09%**, predicting 2 out of 20 images

- **Top-2 Accuracy**: Only EfficientNet achieved **4.55%** (1 image had correct label in second rank).

- **Top-3 Accuracy**: All models failed to place correct labels in top-3.

Observation: *Most models didn't perform very well, which could be because the pretrained models weren't specifically trained on images like mine. Also, since I only tested on 20 images, it's a small sample, so even one or two correct predictions make a big difference in the accuracy numbers.*

Inference Time per Image (Efficiency)

- **Fastest Models**: VGGNet16 (~0.32s) and InceptionV3 (~0.37s)

- **Slowest Model**: EfficientNet (~1.63s), likely due to its deeper architecture and complexity.

Observation: Speed really matters in real-time use, and models like ResNet50 are a better fit when quick results are needed.

Model Size & Memory Usage

- **Lightest Model (Memory)**: InceptionV3 (~91MB avg usage)
- **Heaviest Model**: VGGNet16 (~528MB), despite modest accuracy.
- **EfficientNet:** used ~254MB, suggesting a balance between complexity and inference.

Observation: VGGNet uses a lot of memory and feels a bit outdated compared to newer, more efficient models.

Final Conclusion

- ✓ **Balanced Choice**: **ResNet50** offers a good balance between accuracy and speed with moderate memory usage.
- ✓ **Avoid**: **InceptionV3** and **EfficientNet** in their current form due to poor Top-1 accuracy.
- ✓ **Next Steps**: Fine-tuning or transfer learning is essential for better real-world accuracy on our dataset.