VGG Neural Network

Thant Zin Bo, Aye Chan Myint

VGG

- Developed by Oxford's *Visual Geometry Group* (Simonyan & Zisserman, 2014) as "Very Deep Convolutional Networks for Large-Scale Image Recognition".

- Broke ImageNet records by proving depth + small filters outperform earlier wide, shallow CNNs.

- Core idea: stack many 3 × 3 conv layers instead of a few large ones—simple, uniform, easy to reproduce

- Two canonical variants: VGG-16 (16 weight layers) and VGG-19 (19 weight layers)

Architecture

- Input: fixed 224 × 224 RGB image.
- Convolution blocks
  - Repeated 3 × 3, stride 1 filters with ReLU after each.
  - Two to four conv layers per block, then 2 × 2 max-pool (stride 2) halves width & height.
- After 5 pooling blocks → flatten → three fully connected layers (4096-4096-1000) + soft-max.
- Parameter cost: ≈138 M (VGG-16) — memory-hungry but highly transferable

VGG-16 vs VGG-19 & Practical Considerations

Features	VGG16	VGG19
Convolution Layer	13	16
Fully Connected Layer	3	3
Trainable parameter	138,357,500	143,667,240 (Estimate)
Model size (MB)	527.79	548.1 (Estimate)

Impact, Applications & Limitations

- Backbone for transfer learning in:
    - Medical imaging (tumor detection, radiology)
    - Object & face recognition, biometric security
    - Satellite / environmental monitoring
    - Neural style transfer & Fast-RCNN object detection baseline.
- Strengths
    - Straightforward architecture = easy fine-tuning
    - Consistent 3×3 kernels generalize well across domains.
- Limitations
    - Huge parameter count ➔ high VRAM & inference latency.
    - Largely outperformed by newer, lighter models in accuracy : compute trade-off

HW3: Various CNN Model test Presentation

Results Case1: using "real" folder name

Input Images

⊘ bike   ⊘ bird   ⊘ butterfly   ⊘ cat   ⊘ chicken   ⊘ dog   ⊘ flower   ⊘ horse   ⊘ sheep   ⊘ woman

ResNet50	VGGNet16	InceptionV3	ConvNeXt	EfficientNet	ResNet50_prob	VGGNet16_prob	InceptionV3_prob	ConvNeXt_prob	EfficientNet_prob	label
moped	moped	stopwatch	moped	moped	0.80533	0.42419	1	0.82751	0.65037	bike
bee_eater	house_finch	web_site	bee_eater	bee_eater	0.52949	0.96203	1	0.92659	0.73385	bird
cabbage_butterfly	cabbage_butterfly	web_site	ringlet	ringlet	0.60616	0.77259	1	0.92665	0.75696	butterfly
tabby	tabby	web_site	Egyptian_cat	Egyptian_cat	0.65129	0.48065	0.99981	0.77957	0.45926	cat
hen	hen	pencil_sharper	hen	hen	0.99086	0.96126	0.85818	0.88911	0.69796	chicken
wire-haired_fox_terrier	German_shepherd	web_site	toy_terrier	kelpie	0.16791	0.14802	0.99998	0.47487	0.30531	dog
tray	pitcher	clog	hair_slide	lampshade	0.41643	0.14068	0.86002	0.29183	0.49116	flower
hartebeest	hartebeest	vase	sorrel	hartebeest	0.49483	0.87577	0.47337	0.2562	0.62734	horse
ram	ram	web_site	hog	ram	0.75918	0.72313	0.85689	0.76543	0.41133	sheep
sombrero	ice_lolly	saltshaker	tennis_ball	sombrero	0.64053	0.06982	0.99947	0.32151	0.14389	woman
Accuracy	Accuracy	Accuracy	Accuracy	Accuracy						
5.3	5	0	6	4.5						

Accuracy Score
1
0.8
0.5

Table 1: Model comparative result table for case1

Input Images

✓ bike ✓ bird ✓ butterfly ✓ cat ✓ chicken ✓ dog ✓ flower ✓ horse ✓ sheep ✓ woman

Predicted Label Images

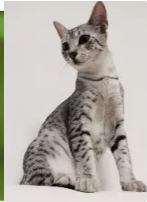Bee-eater | bird | Britannica.com

House Finch - eBird

Cabbage butterfly – Mort Bay

Ringlet | The Wildlife Trusts

Egyptian Cat

Egyptian Cats - Dagmara Mach

Tabby Free Stock Photo - Public

Wire Fox Terrier - Pictures, Information,

German Shepherd | Breed, Size, Coat &

Toy Fox Terrier - Pictures. Information

Wild Australian Kelpie

Sorrel Horse: Telling Sorrels Apart From Red

Hartebeest | Creatures of the
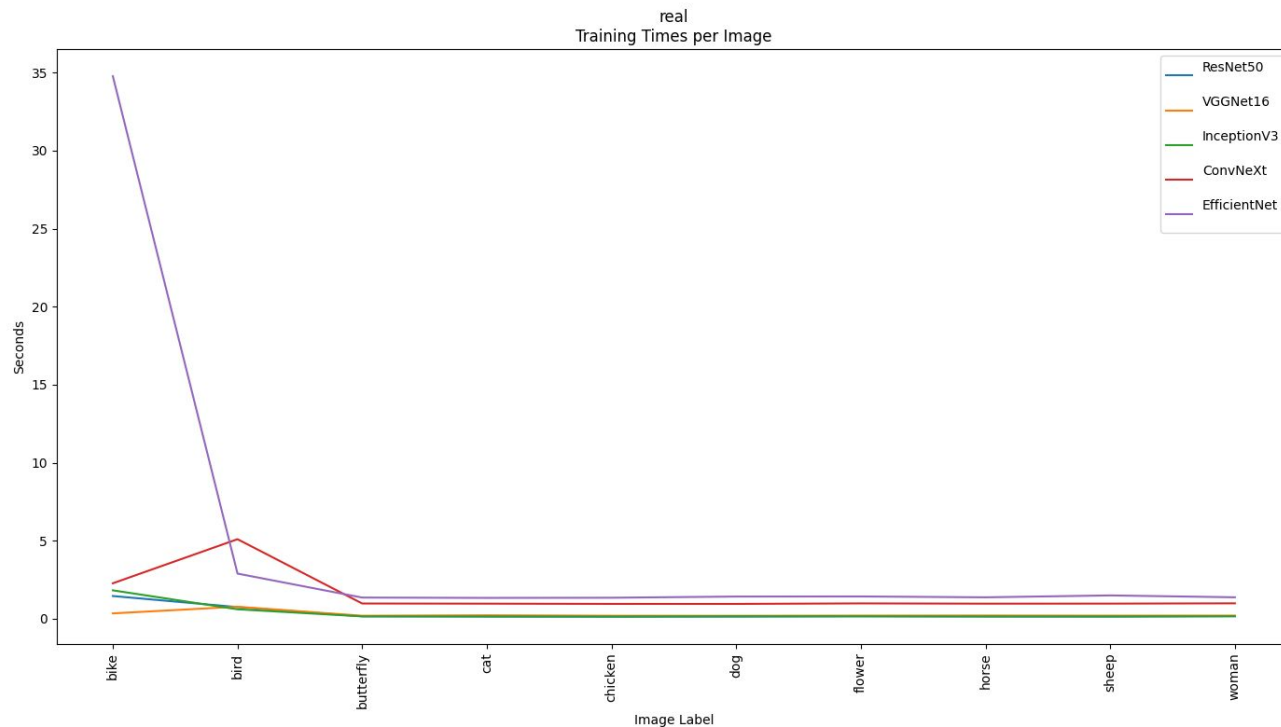
Feral Hog Management Program to

Sombrero Unisex Explorer

How to Start a Moped: Everything You Need to Know

Results Case1: using "real" folder name

real
Training Times per Image

Lowest inference time
•ResNet50
•InceptionV3
•VGGNet16

Results Case2: using "synthetic" folder name

Input Images

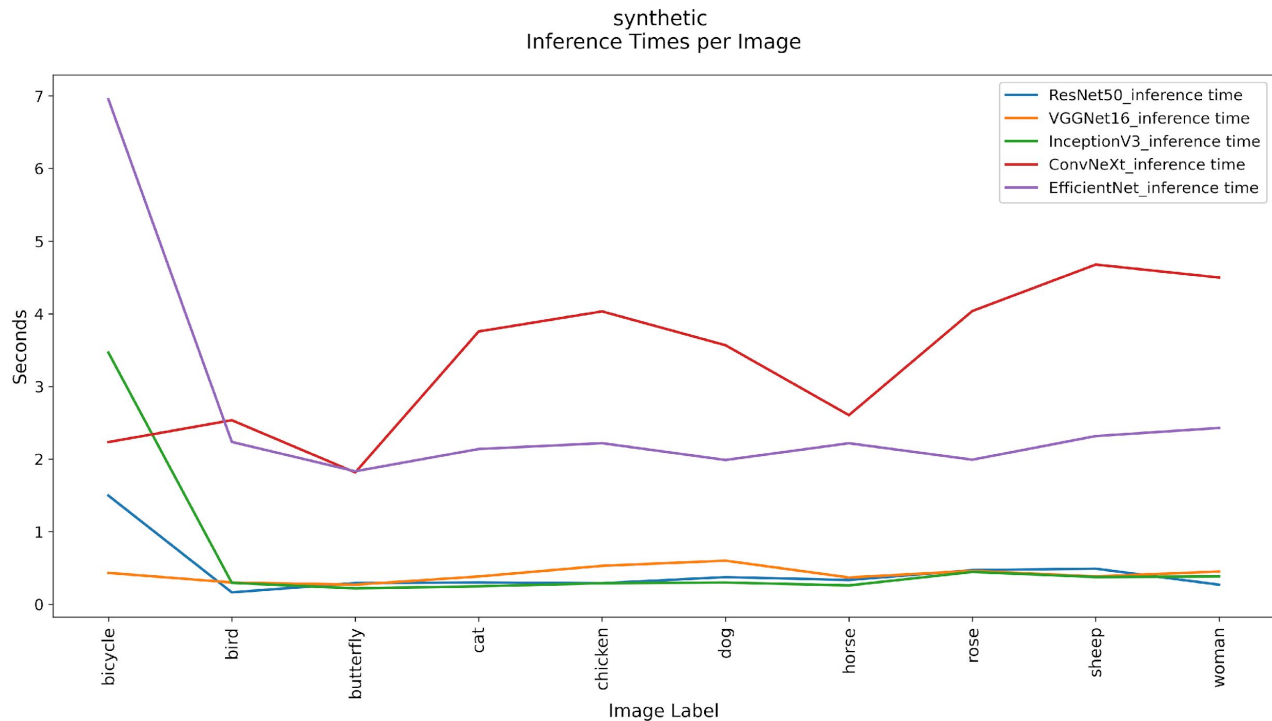⊘ bicycle ⊘ bird ⊘ butterfly ⊘ cat ⊘ chicken ⊘ dog ⊘ horse ⊘ rose ⊘ sheep ⊘ woman

ResNet50	VGGNet16	InceptionV3	ConvNeXt	EfficientNet	ResNet50_prob	VGGNet16_prob	InceptionV3_prob	ConvNeXt_prob	EfficientNet_prob	label
tricycle	tricycle	clog	bicycle-built-for-two	tricycle	0.84685	0.19208	0.59291	0.25444	0.81799	bicycle
pinwheel	jellyfish	flatworm	sulphur-crested_cockatoo	bee_eater	0.69915	0.17417	1	0.56962	0.35358	bird
pinwheel	bubble	flatworm	bubble	lacewing	0.49099	0.22225	0.89689	0.59244	0.64869	butterfly
laptop	tabby	web_site	tabby	Persian_cat	0.39527	0.2078	1	0.28785	0.49993	cat
hen	cock	stopwatch	hen	hen	0.50245	0.67628	0.58571	0.7616	0.6226	chicken
tennis_ball	golden_retriever	flatworm	tennis_ball	golden_retriever	0.91886	0.57009	0.92717	0.27182	0.60391	dog
alp	sorrel	flatworm	carousel	alp	0.37877	0.2142	0.98597	0.59683	0.38357	horse
vase	pot	web_site	necklace	coil	0.16037	0.3725	0.79715	0.24576	0.14973	rose
balloon	Arctic_fox	clog	Arctic_fox	Arctic_fox	0.09864	0.15921	0.99436	0.28605	0.14845	sheep
picket_fence	balloon	flatworm	picket_fence	picket_fence	0.99313	0.29855	0.99996	0.92295	0.77555	woman
Accuracy	Accuracy	Accuracy	Accuracy	Accuracy						
1.5	3.8	0	4.4	4						

Accuracy Score	
1	
0.8	
0.5	

Results Case2: using "synthetic" folder name

synthetic
Inference Times per Image

Lowest inference time
•ResNet50
•InceptionV3
•VGGNet16

Results Case3: using "test1" folder name

Input Images

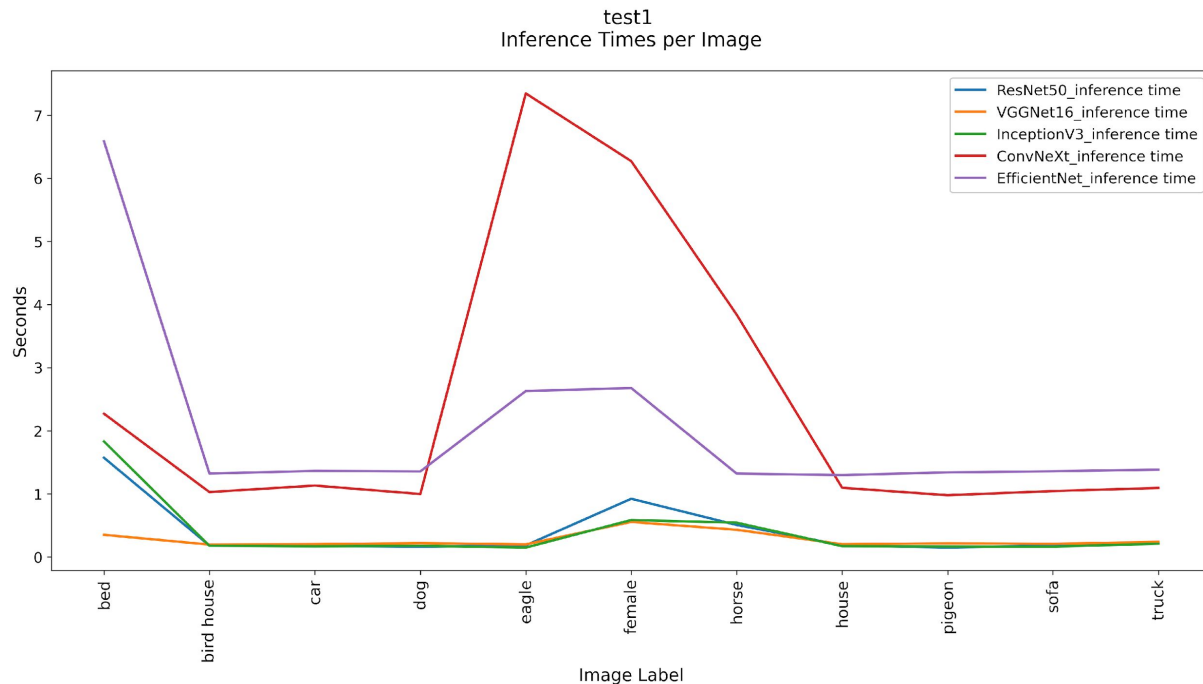| ✓ bed | ✓ bird house | ✓ car | ✓ dog | ✓ eagle | ✓ female | ✓ horse | ✓ house | ✓ pigeon | ✓ sofa | ✓ truck |

ResNet50	VGGNet16	InceptionV3	ConvNeXt	EfficientNet	ResNet50_prob	VGGNet16_prob	InceptionV3_prob	ConvNeXt_prob	EfficientNet_prob	label
home_theater	studio_couch	web_site	studio_couch	studio_couch	0.34067	0.26746	0.99992	0.25554	0.64738	bed
birdhouse	birdhouse	tennis_ball	birdhouse	birdhouse	0.99989	0.99863	0.99632	0.92486	0.81878	bird house
minivan	minivan	bow	beach_wagon	minivan	0.30592	0.16888	0.97962	0.81965	0.35422	car
whippet	Ibizan_hound	web_site	dingo	American_Staffordshire_terrier	0.55377	0.26041	1	0.3576	0.51846	dog
hornbill	bald_eagle	web_site	kite	macaw	0.30271	0.46598	1	0.25597	0.51434	eagle
trench_coat	trench_coat	stole	wig	wig	0.8986	0.18629	0.99968	0.45983	0.60992	female
sorrel	sorrel	web_site	gazelle	sorrel	0.86249	0.86109	0.99999	0.64218	0.63947	horse
picket_fence	patio	stopwatch	patio	patio	0.3221	0.16284	1	0.17431	0.2332	house
partridge	partridge	web_site	partridge	partridge	0.30066	0.25714	0.95404	0.37752	0.40252	pigeon
home_theater	studio_couch	pitcher	studio_couch	pedestal	0.5361	0.99577	0.99989	0.79505	0.44423	sofa
trailer_truck	garbage_truck	clog	garbage_truck	garbage_truck	0.72648	0.88189	0.99388	0.77973	0.58087	truck
Accuracy	Accuracy	Accuracy	Accuracy	Accuracy						
5.5	7	0	4	5.5						

Accuracy Score
1
0.8
0.5

Results Case3: using "test1" folder name

test1
Inference Times per Image

Lowest inference time
- ResNet50
- InceptionV3
- VGGNet16

Results Case4: using "test2" folder name

Input Images

✓ automobile ✓ bed ✓ bird house ✓ dog ✓ eagle ✓ female ✓ horse ✓ house ✓ pigeon ✓ sofa1 ✓ truck

ResNet50	VGGNet16	InceptionV3	ConvNeXt	EfficientNet	ResNet50_prob	VGGNet16_prob	InceptionV3_prob	ConvNeXt_prob	EfficientNet_prob	label
balloon	cab	stopwatch	solar_dish	lakeside	0.26035	0.25555	0.94278	0.67278	0.16456	automobile
studio_couch	studio_couch	flatworm	quilt	quilt	0.51293	0.81496	1	0.81613	0.75614	bed
birdhouse	birdhouse	web_site	birdhouse	birdhouse	1	1	1	0.76124	0.83851	bird house
kuvasz	Labrador_retriever	web_site	kuvasz	kuvasz	0.43705	0.4033	0.99177	0.29176	0.58353	dog
bald_eagle	bald_eagle	web_site	bald_eagle	bald_eagle	1	0.99968	1	0.88781	0.82653	eagle
umbrella	dumbbell	English_foxhound	comic_book	mask	0.06532	0.15229	0.76666	0.22708	0.10937	female
sorrel	sorrel	web_site	sorrel	Saluki	0.91842	0.99405	1	0.50848	0.35146	horse
window_screen	picket_fence	leatherback_turtle	patio	window_screen	0.1748	0.18918	0.7165	0.44986	0.53039	house
whippet	goose	flatworm	hen	partridge	0.11009	0.21103	1	0.14806	0.70049	pigeon
studio_couch	studio_couch	stopwatch	studio_couch	studio_couch	0.98675	0.7529	0.99865	0.93918	0.77844	sofa1
crane	crane	flatworm	garbage_truck	garbage_truck	0.44036	0.62959	1	0.77316	0.74995	truck
Accuracy	Accuracy	Accuracy	Accuracy	Accuracy						
5	6	0	6	6						

Accuracy Score
1
0.8
0.5

Results Case4: using "test2" folder name
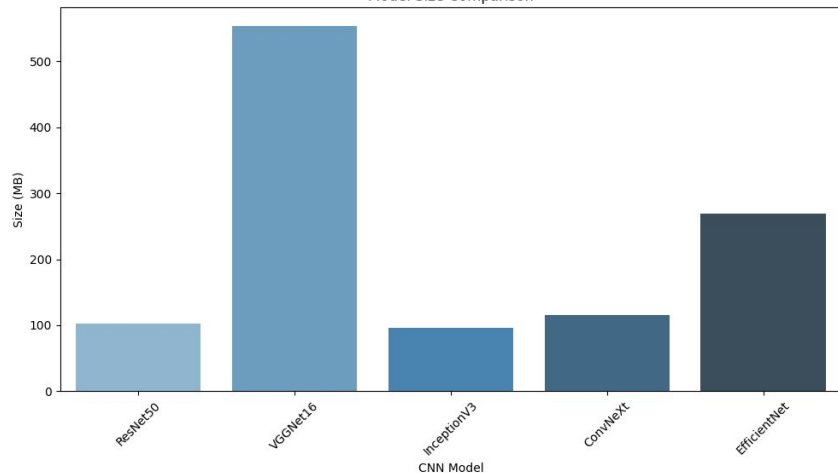

test2
Inference Times per Image
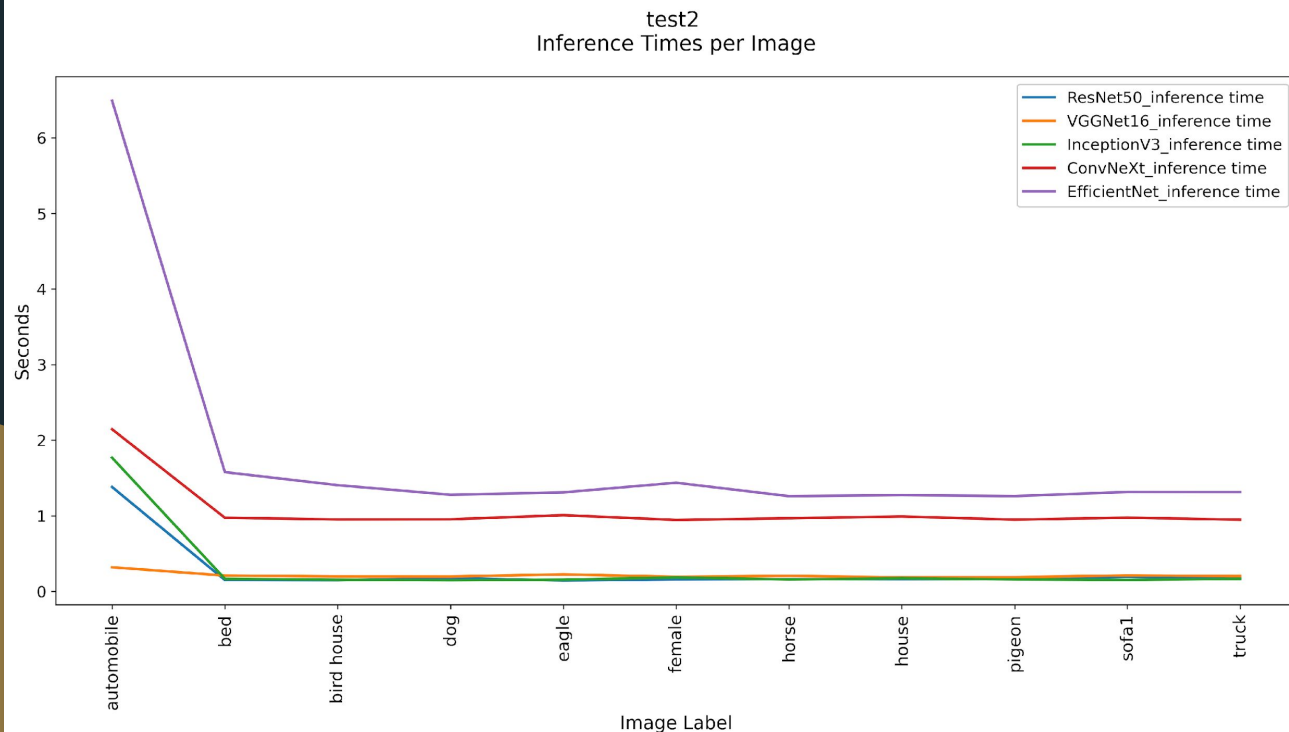
Lowest inference time
•ResNet50
•InceptionV3
•VGGNet16

Summary

Model	Accuracy (Out of 10) Real Folder Data	Accuracy (Out of 10) Synthetic Folder Data	Accuracy (Out of 10) Test1 Folder Data	Accuracy (Out of 10) Test2 Folder Data	Overall Average Accuracy Score (Out of 40)
ResNet50	5.3	1.5	5.5	5	17.3
VGGNet16	5	3.8	7	6	21.8
InceptionV3	0	0	0	0	0
ConvNeXt	6	4.4	4	6	20.4
EfficientNet	4.5	4	5.5	6	20

Model Size Comparison

Memory Usage Comparison
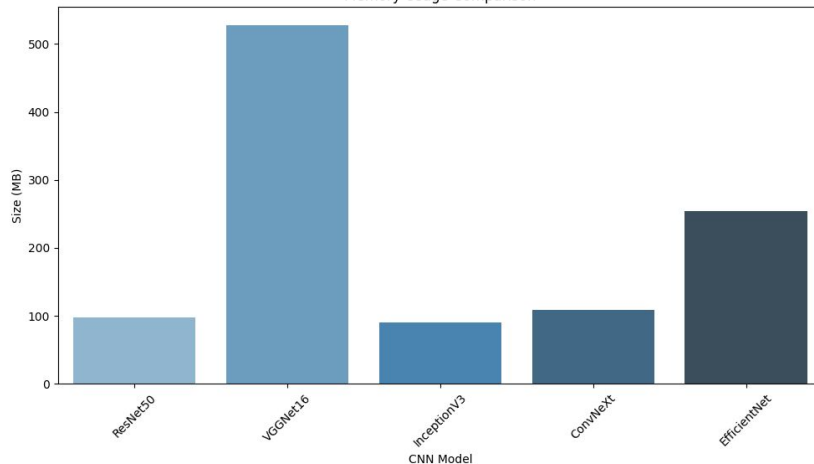
Summary

1. Top 3 Models by Prediction Accuracy

Rank	Model
1	VGGNet16
2	ConvNet
3	EfficientNet

These 3 models likely have more parameters or better architecture features that capture complex patterns.

2. Top 3 Models by Inference Speed

Rank	Model
1	ResNet50
2	InceptionV3
3	VGGNet16

Trade-off: While ResNet50 and InceptionV3 are faster, they sacrifice accuracy. VGGNet16 balances both speed and performance.

Summary

3. Key Metrics Comparison

Metric	Best Model	Worst Model
Accuracy	VGGNet16	InceptionV3
Speed	ResNet50	EfficientNet
Memory Usage and Model size	InceptionV3	VGGNet16

4. Observations
Prediction Quality: VGGNet16 consistently outperforms others, while inceptionV3 fails entirely (possibly due to trianing issues).
Speed Vs Accuracy: ResNet50 is fastest but moderately less accurate.

VGGNet16 offers the best model performance although it consumes large memory usage and model size.