Comparative Analysis of Pre-Trained CNNs for Classifying Equine Pain Faces

Dan Børge Jensen
Associate professor (Lektor)
IVH Department Meeting, 02/09/2024

KØBENHAVNS UNIVERSITET

Credit where it's due

Dan Børge Jensen
Associate professor

Natascha Jensen
Bachelor's student

Sofie Knudsen
Bachelor's student

Christina Larsen
Assistant professor

Main Aim

Identify the best pre-trained

Convolutional Neural Network

for extracting

Useful Image Features

for detecting

Equine Pain Face

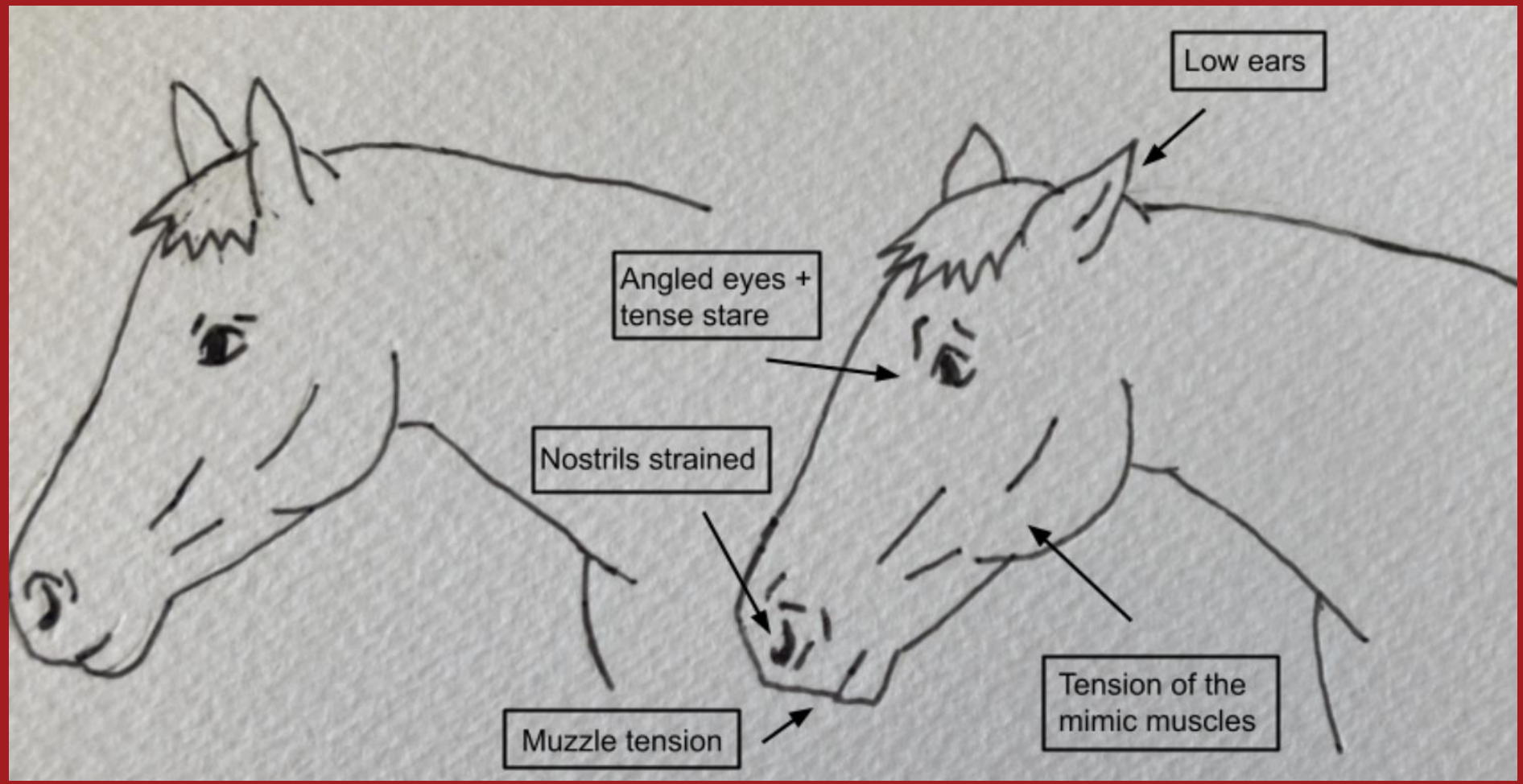
Motivation

1. Pain in horses is often misinterpreted as bad behaviour
2. Delayed identification of pain
   → delayed treatment
   → potentially euthanasia, could have been avoided
3. Automatic detection can help owners decide to call a veterinarian

Brief Data overview

Availible data: 22 videos; 11 with EPF, 11 without EPF; age: 16–25 years

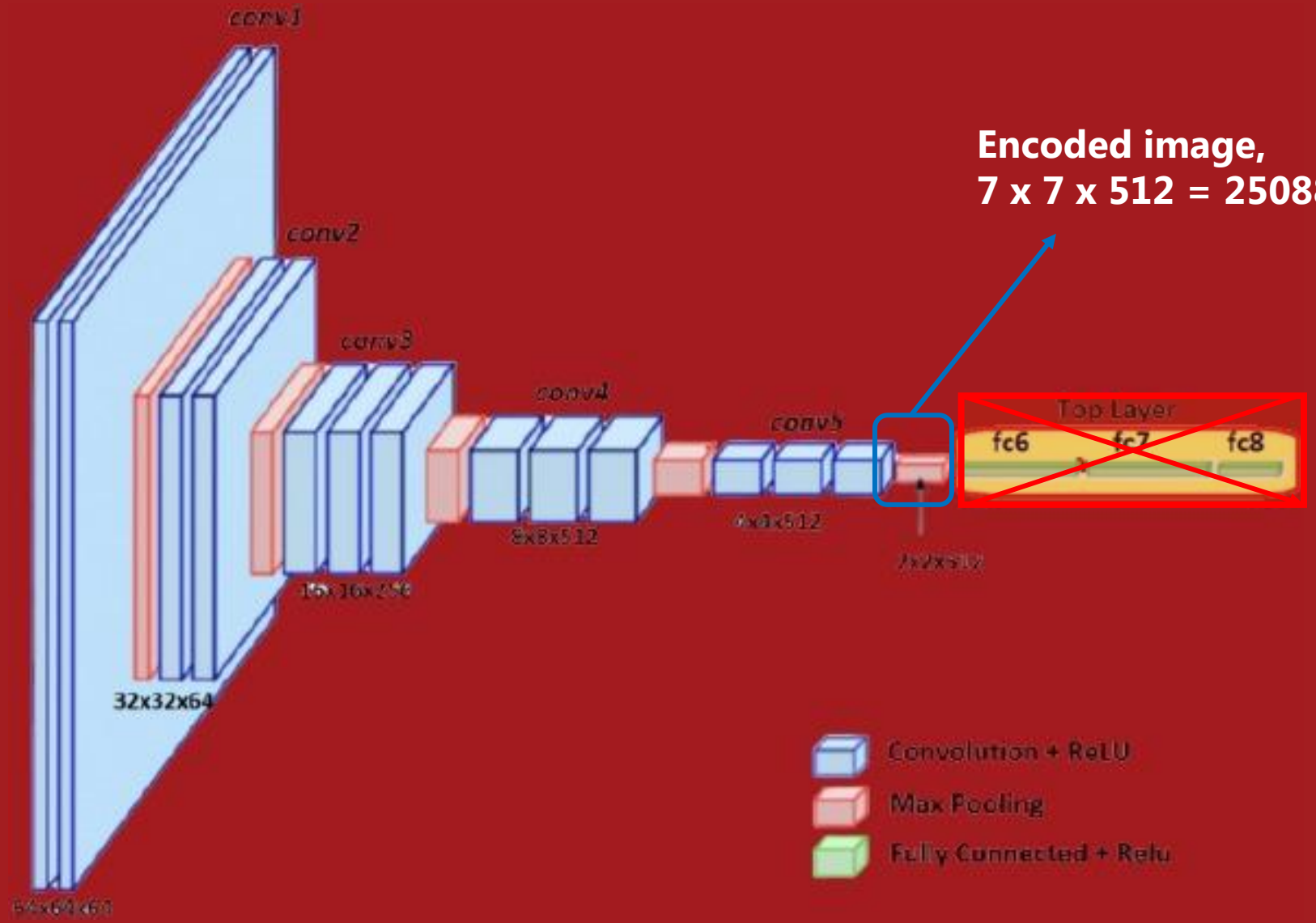
Ground truth EPF status annotated by Karina Glerup, ethologist



A horse without pain


colfix

Methods (1) – extracting features with a CNN



Encoded image,
7 x 7 x 512 = 25088 numerical values
(aka features)

Secondary model
FC-ANN

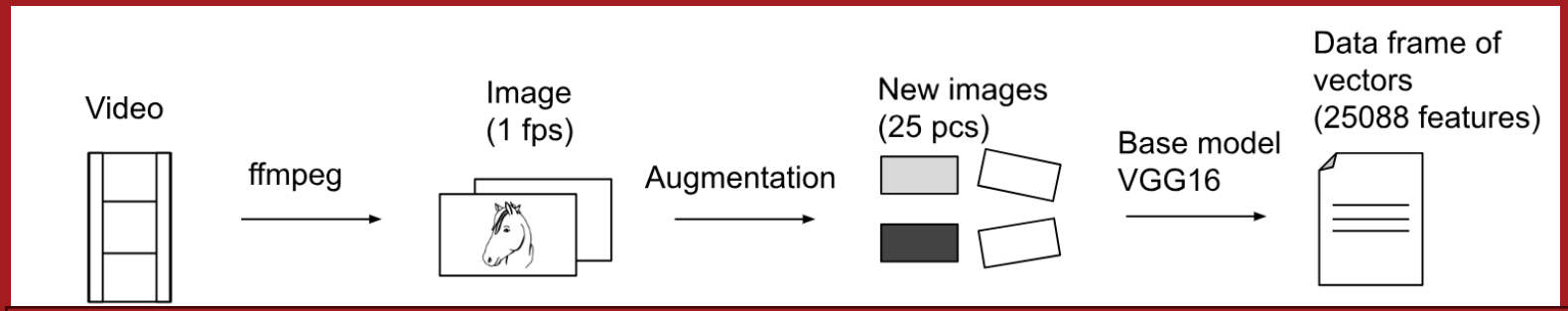
Binary classification

Methods (2) - Pipeline

Pretrained CNNs:

- VGG16
- MobileNetV2
- Xception
- ResNet101V2
- DenseNet201
- EfficientNetB7
- InceptionResNetV2

***All pre-trained for the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), image classification of 1000 classes**



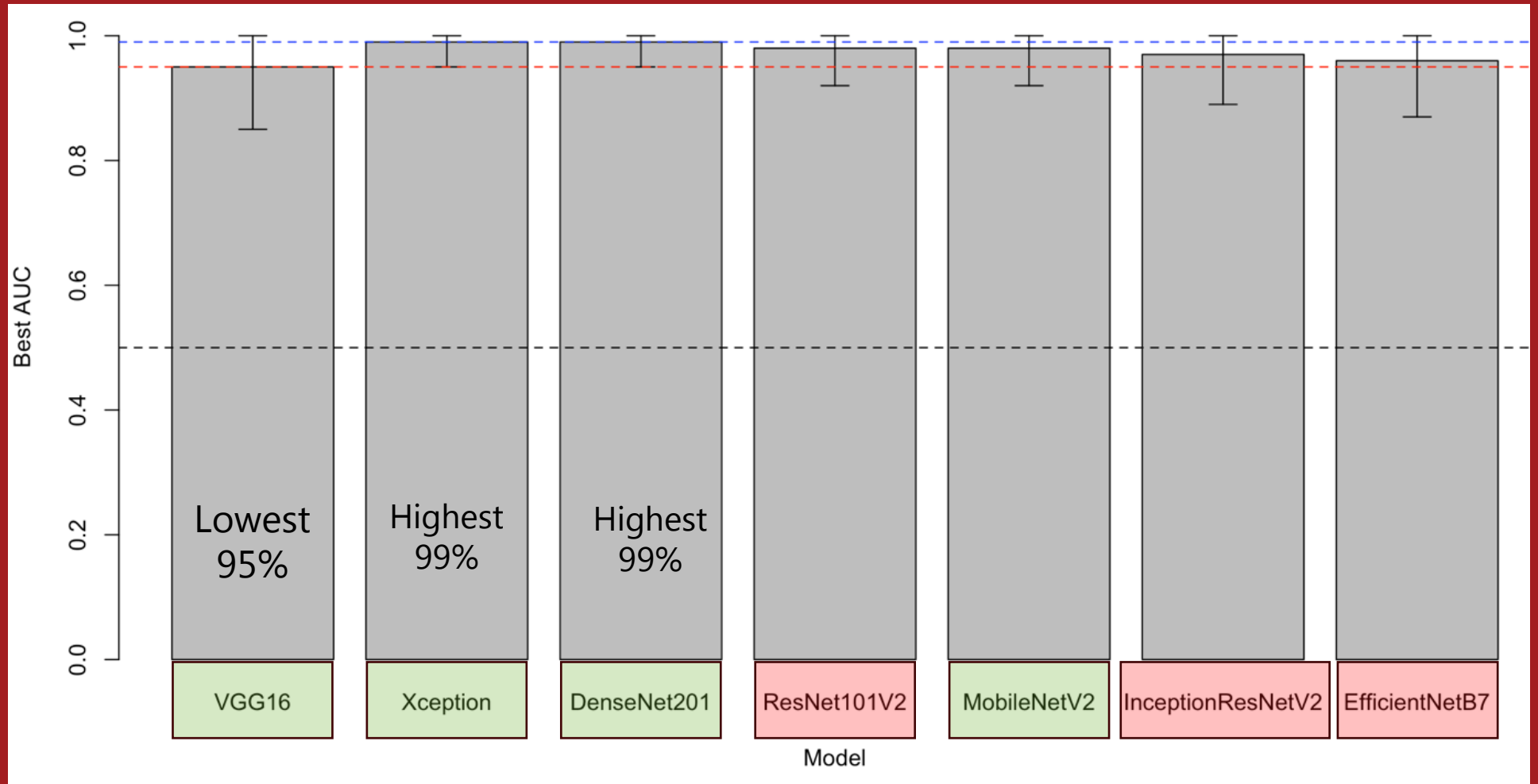
Results (1)
Comparison of pre-trained base-models

Worth further investigation!

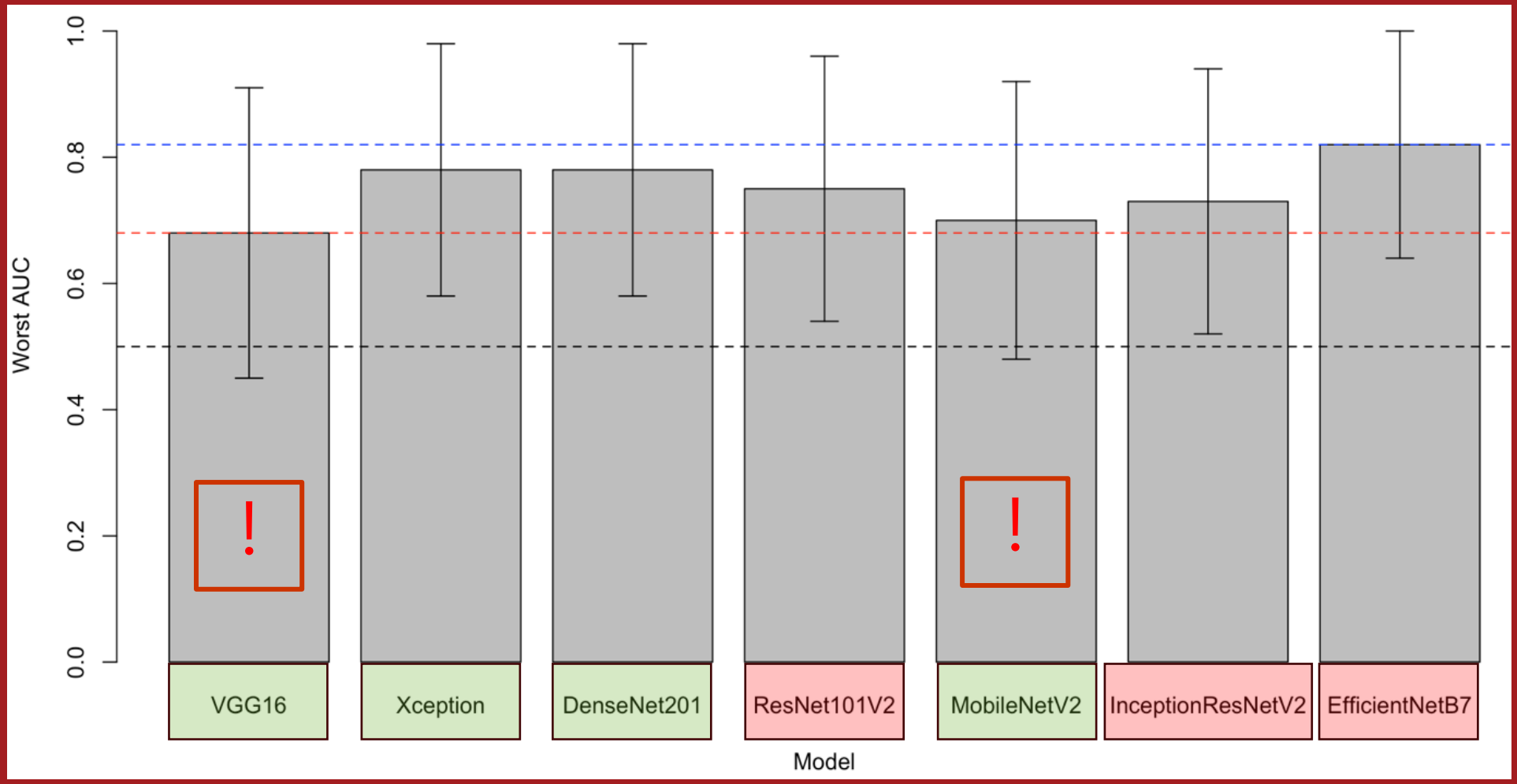
Base model	Base model inference speed	Optimization importance for secondary models	Best no. of input features for secondary model	Size (MB)
VGG16	Fast	Important	256	528
MobileNetV2			256	88
DenseNet201	Medium	Less important	128	80
Xception			2499	215
ResNet101V2	Slow		10782	171
InceptionResNetV2			64	14
EfficientNetB7			64	256

Too slow to be practical, not spectacular performance!

Results (2)
Best secondary models



Results (3)
Worst secondary models



Main conclusion

- No single pretrained model is “best” across the board
- DenseNet201 would be my recommendation

Base model	Base model inference speed	Optimization importance for secondary models	Best no. of input features for secondary model	Size (MB)
VGG16	Fast	Important	256	528
MobileNetV2			256	88
DenseNet201	Medium	Less important	128	80
Xception			2499	215
ResNet101V2	Slow		10782	171
InceptionResNetV2			64	14
EfficientNetB7			64	256

Extra slides

Automatic detection of Equine Pain Face
Results (1) – Encoding times

Base-models	Encoding time (total hours*)	Encoding time (sec. per frame*)	Est. encoding time (m) for 30 s video at 25 fps	Est. encoding time (s) for 30 s video at 1 fps
VGG16	1.01	0.24	3	6
MobileNetV2	1.39	0.32	4	12
DenseNet201	2.32	0.54	7	18
Xception	2.7	0.63	8	18
ResNet101V2	3.97	0.93	12	30
InceptionResNetV2	8.49	1.98	25	60
EfficientNetB7	11	2.57	32	78

*** 15400 frames in augmented data set**

Automatic detection of Equine Pain Face Methods (2) – Training & testing