

Facial Expression Recognition.

Convolutional
Networks
& XAI

Maria Aalto
Eetu Oinonen
Pavel Degterev
Suphitcha Pitkänen

Our goals.

Build facial-emotion classifier model.

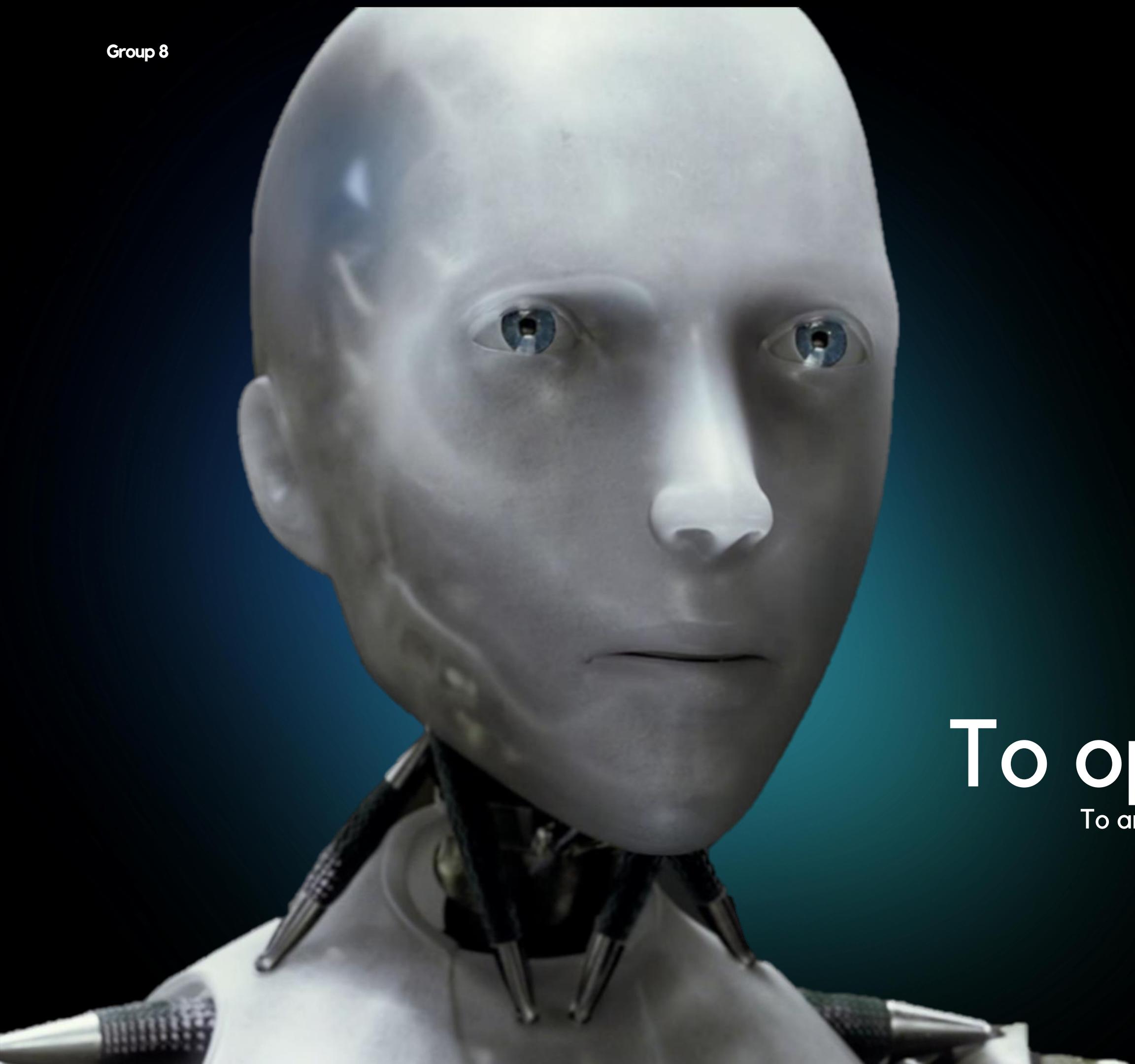
Data sets that we've used for that:

AffectNet : 30.6k (96*96) images – hand picked high q dataset

FER-2013 : 35.9k (48*48) images – noisy one

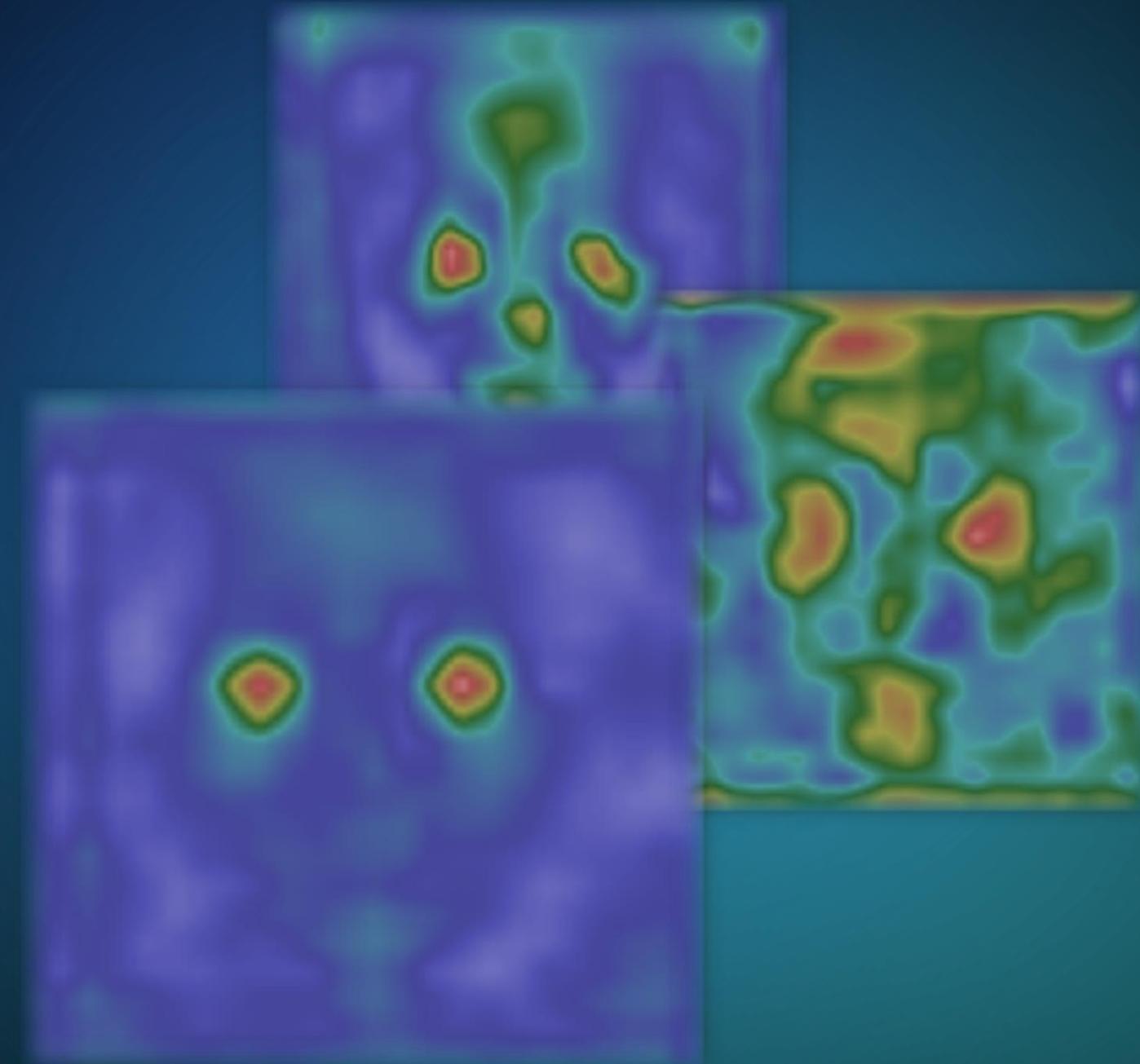
CK+ : 0.9k (48*48) laboratory made images





To open the Black Box

To understand why models failed on certain images, we used
XAI (*Explainable AI*), specifically Grad-CAM heatmaps.



To open the Black Box

To understand why models failed on certain images, we used **XAI** (*Explainable AI*), specifically Grad-CAM heatmaps.

We build the model

Our initial results was:

(pre-fine-tuned): 0.375

Final test accuracy: 0.602

Trainable params: 14,522,119



Trained on different data

6 rows ▾ 6 rows × 3 cols

	Test_AffectNet	Test_FER2013	Test_CKPlus
Double_48_data	58.0%	61.5%	81.9%
Double_96_data	70.9%	65.1%	88.3%
Model_AffectNet	69.3%	44.0%	66.1%
Model_FER2013	29.0%	65.3%	80.3%
Initial_model	28.1%	60.2%	80.4%
Model_CKPlus	26.4%	29.7%	95.8%

Compared the results



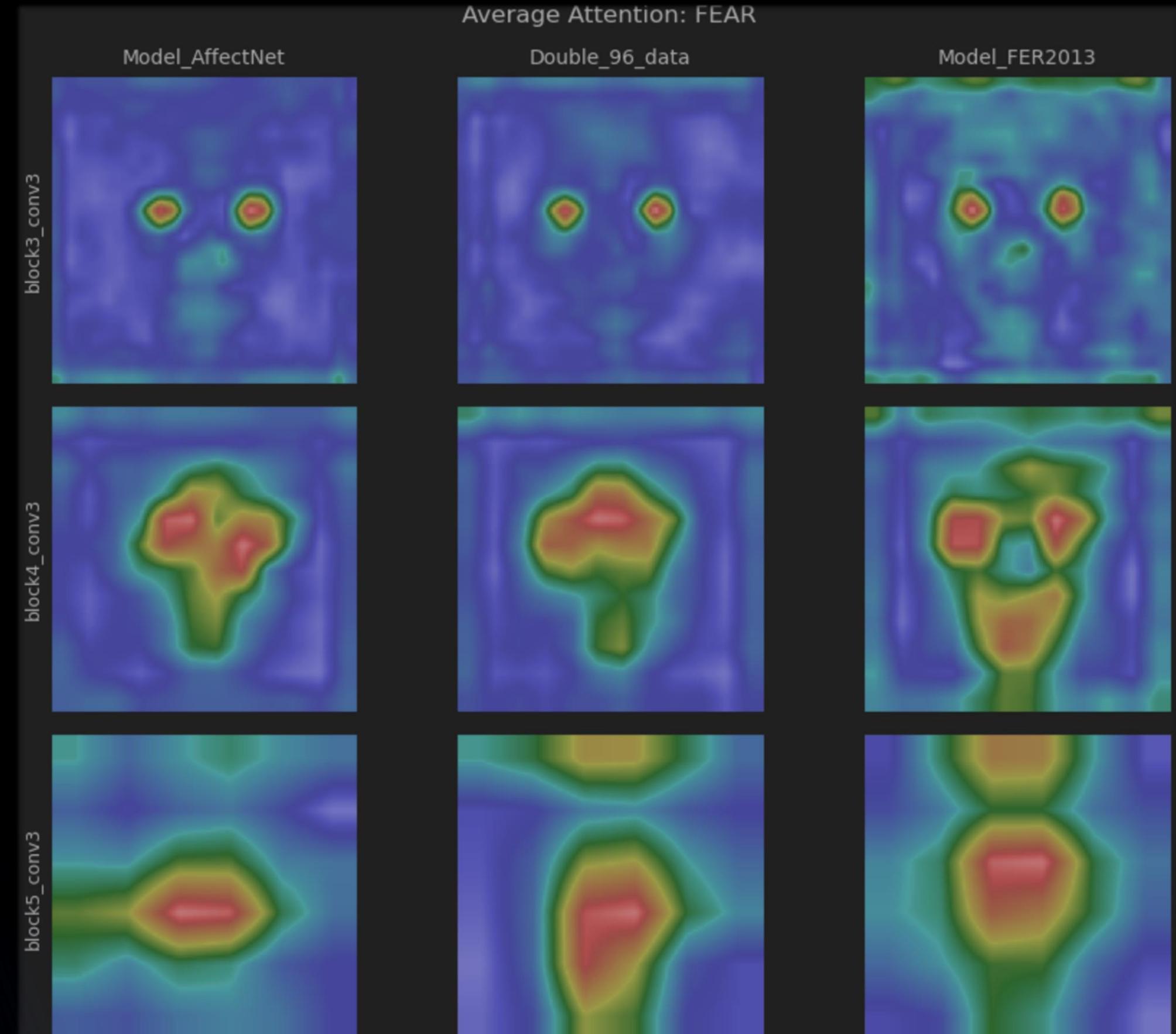
What did we find out?

What did we find out?

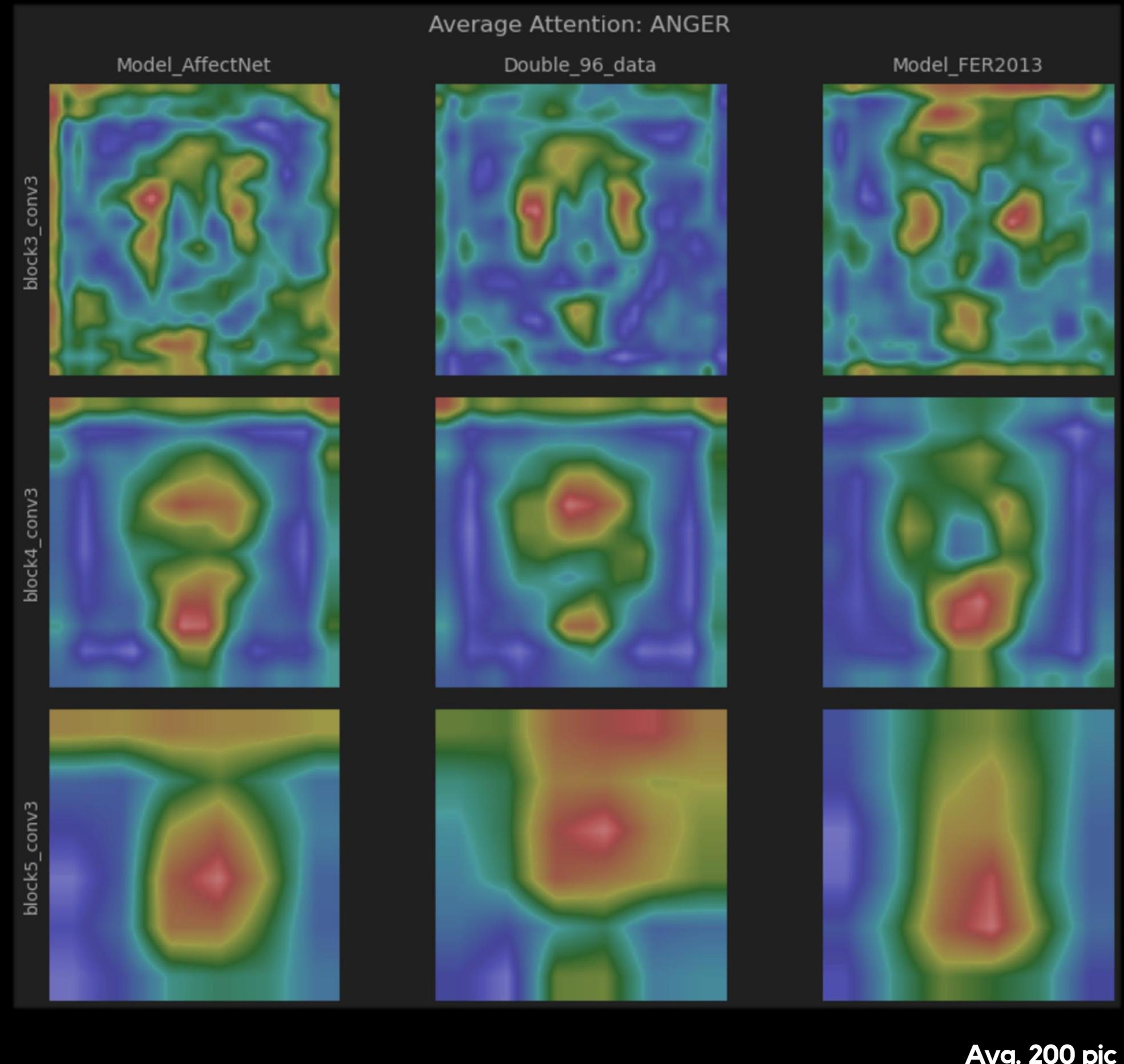
	Test_AffectNet	Test_Fer2013	Test_CKplus
All_96_data	71.1 %	65.8 %	88.5 %
Double_96_data	70.9 %	65.1 %	88.3 %
Model_AffectNet	69.3 %	44.0 %	66.1 %
Model_Fer2013	29.0 %	65.3 %	80.3 %
Initial_model	28.1 %	60.2 %	80.4 %
Model_CKplus	26.4 %	29.7 %	92.8 %

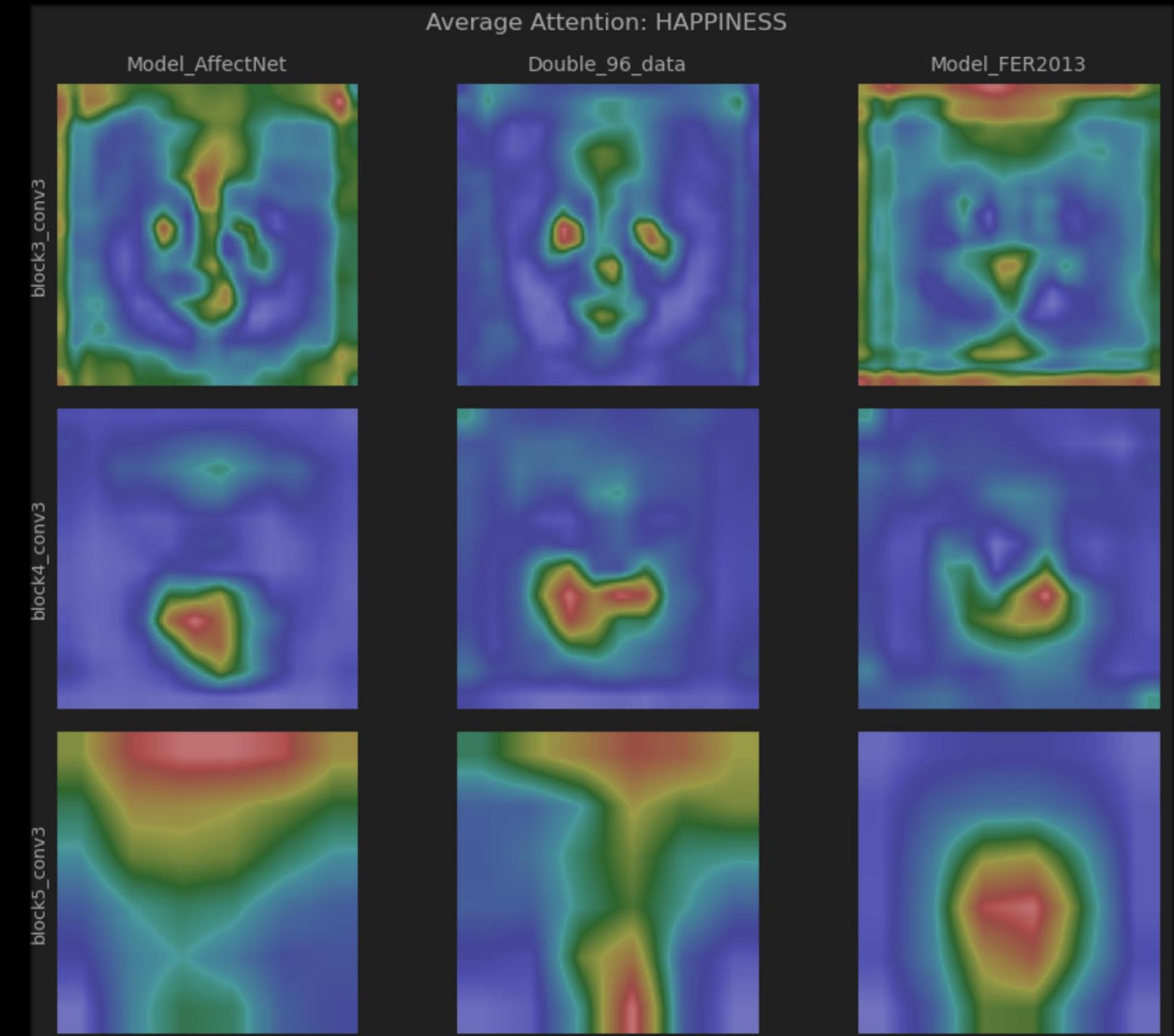
Where does it look,
and did it get smarter?

Indeed



Avg. 200 pic



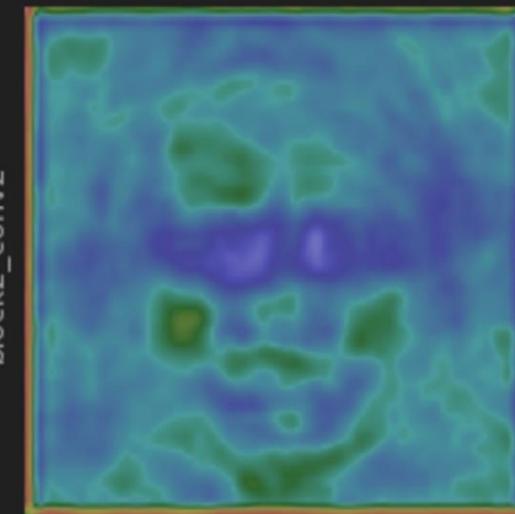


Avg. 200 pic

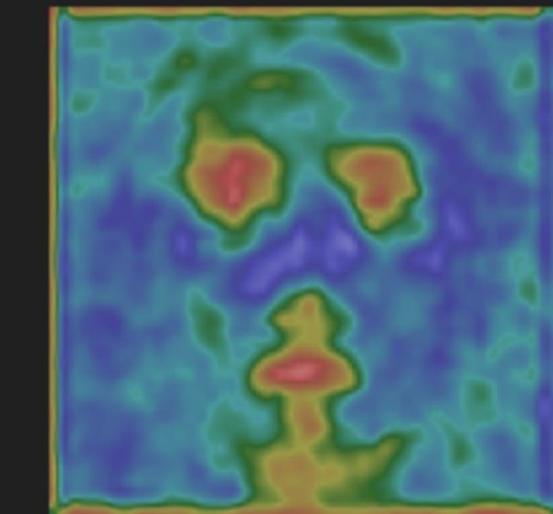
Group 8

Average Attention: SURPRISE

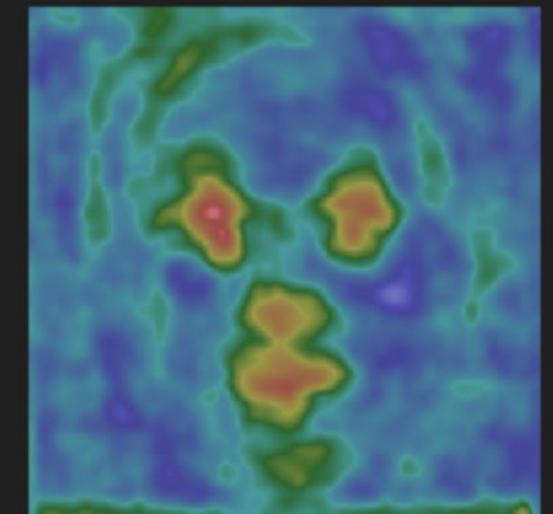
Model_AffectNet



Double_96_data

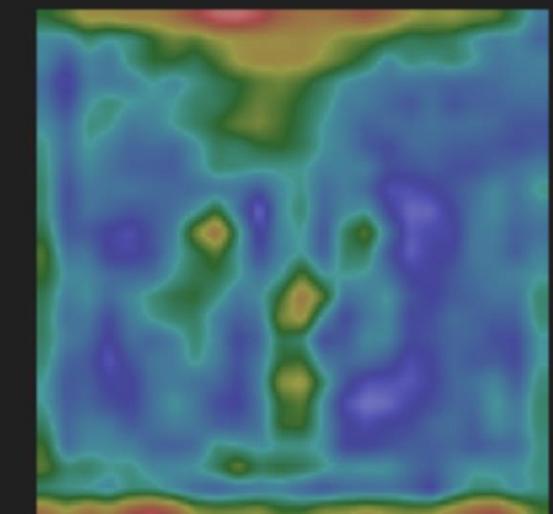
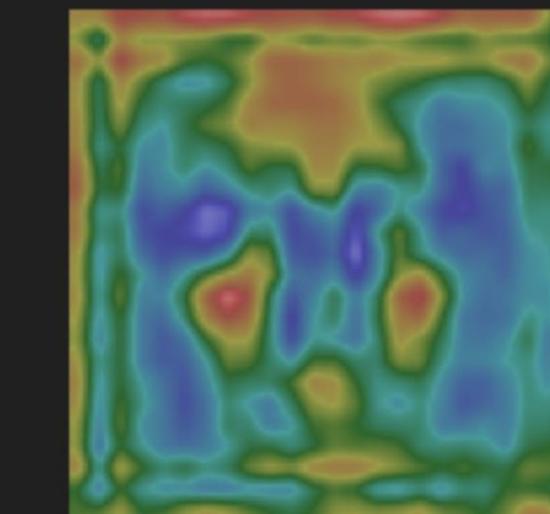
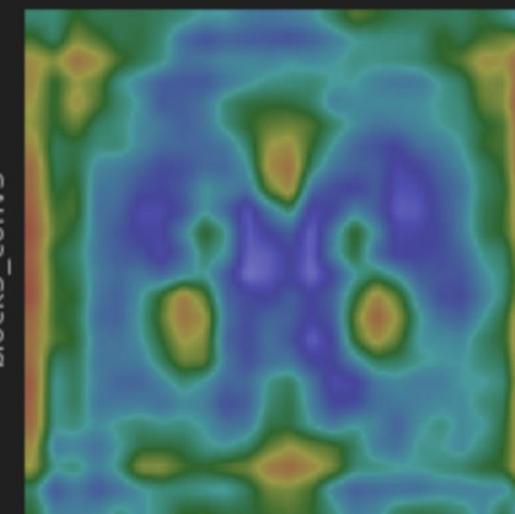


Model_FER2013



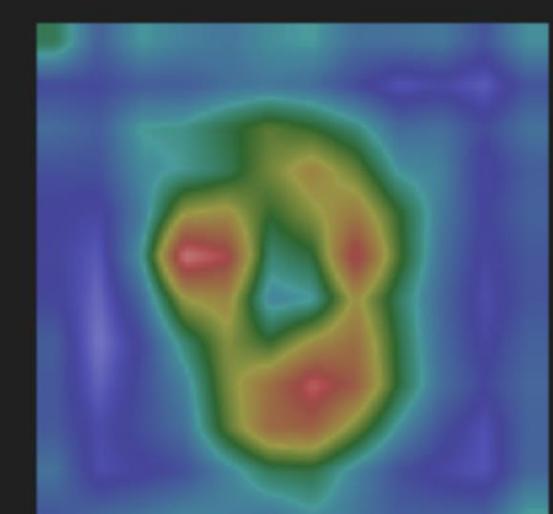
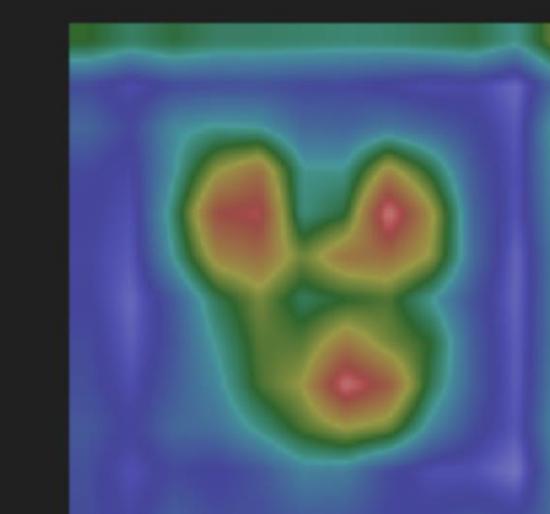
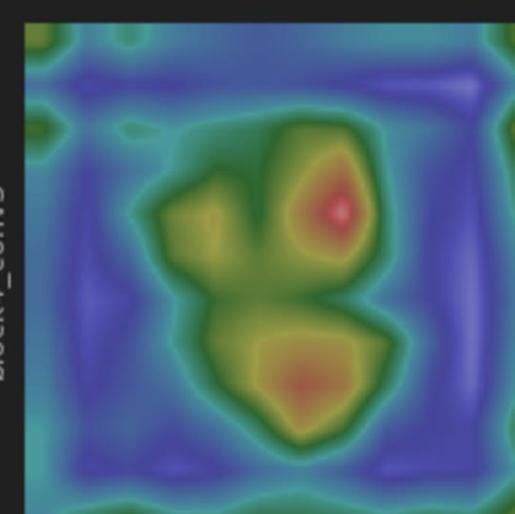
block2_conv2

Model_AffectNet



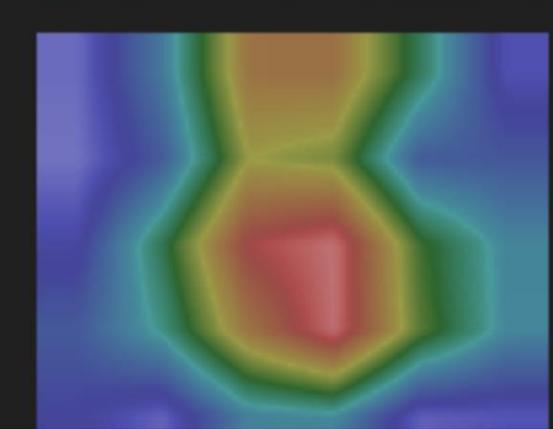
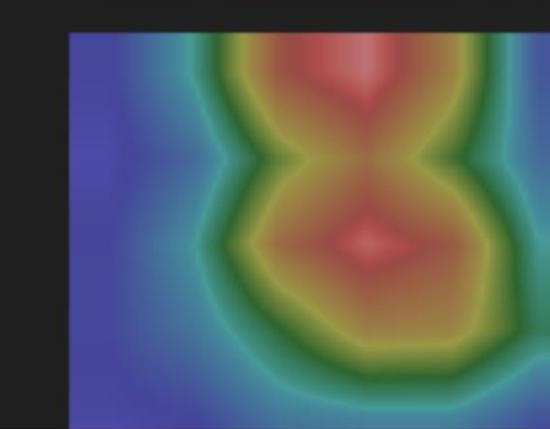
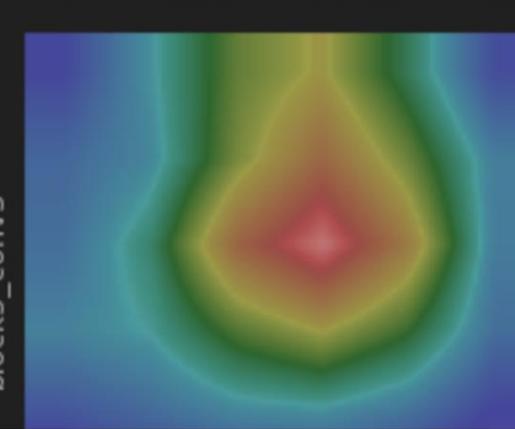
block3_conv3

Model_AffectNet



block4_conv3

Model_AffectNet



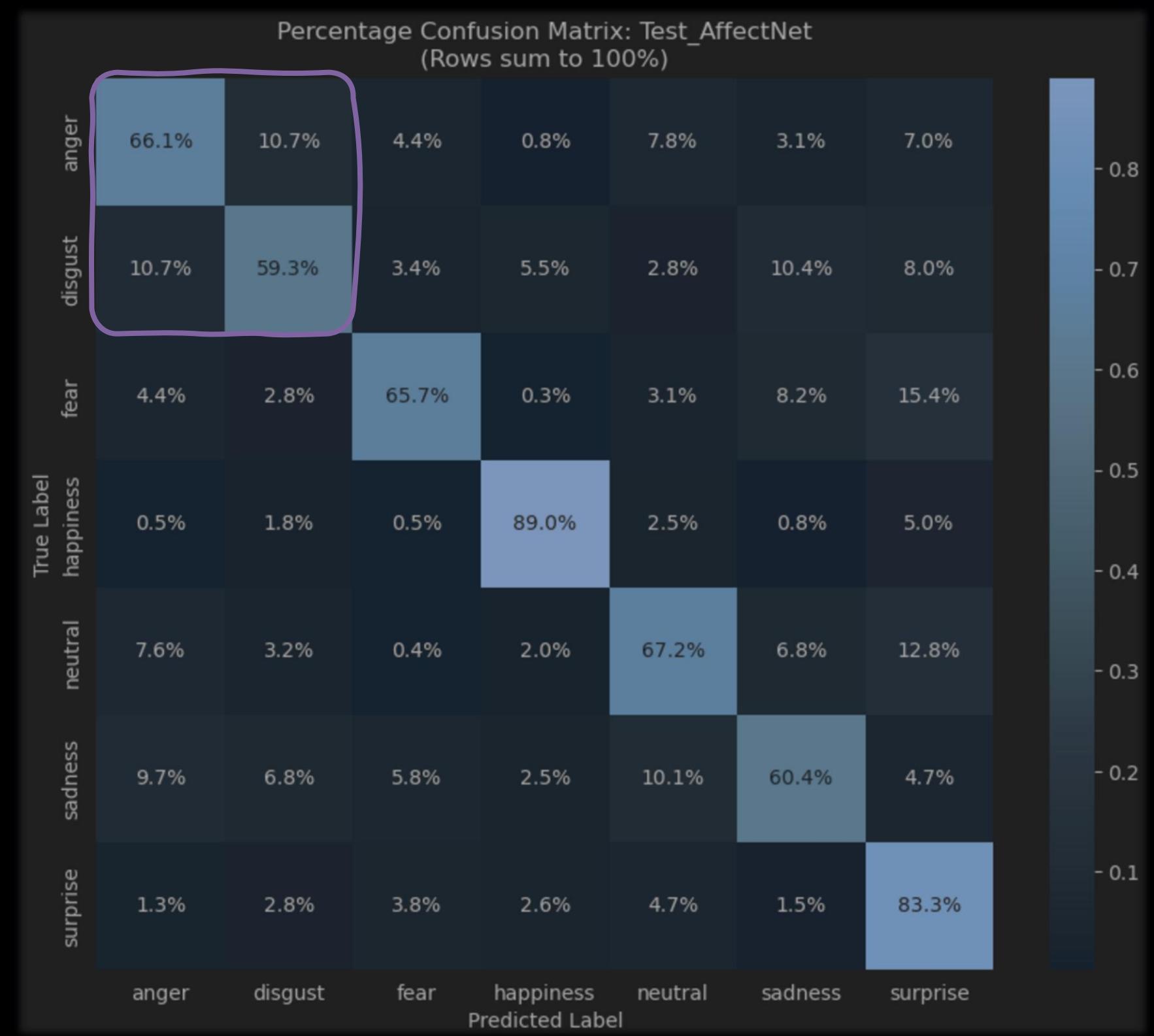
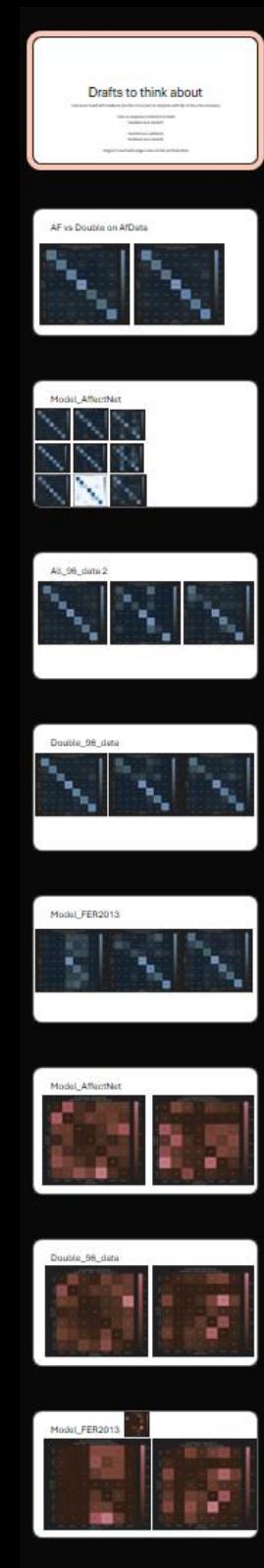
block5_conv3

Why not 100% accuracy?

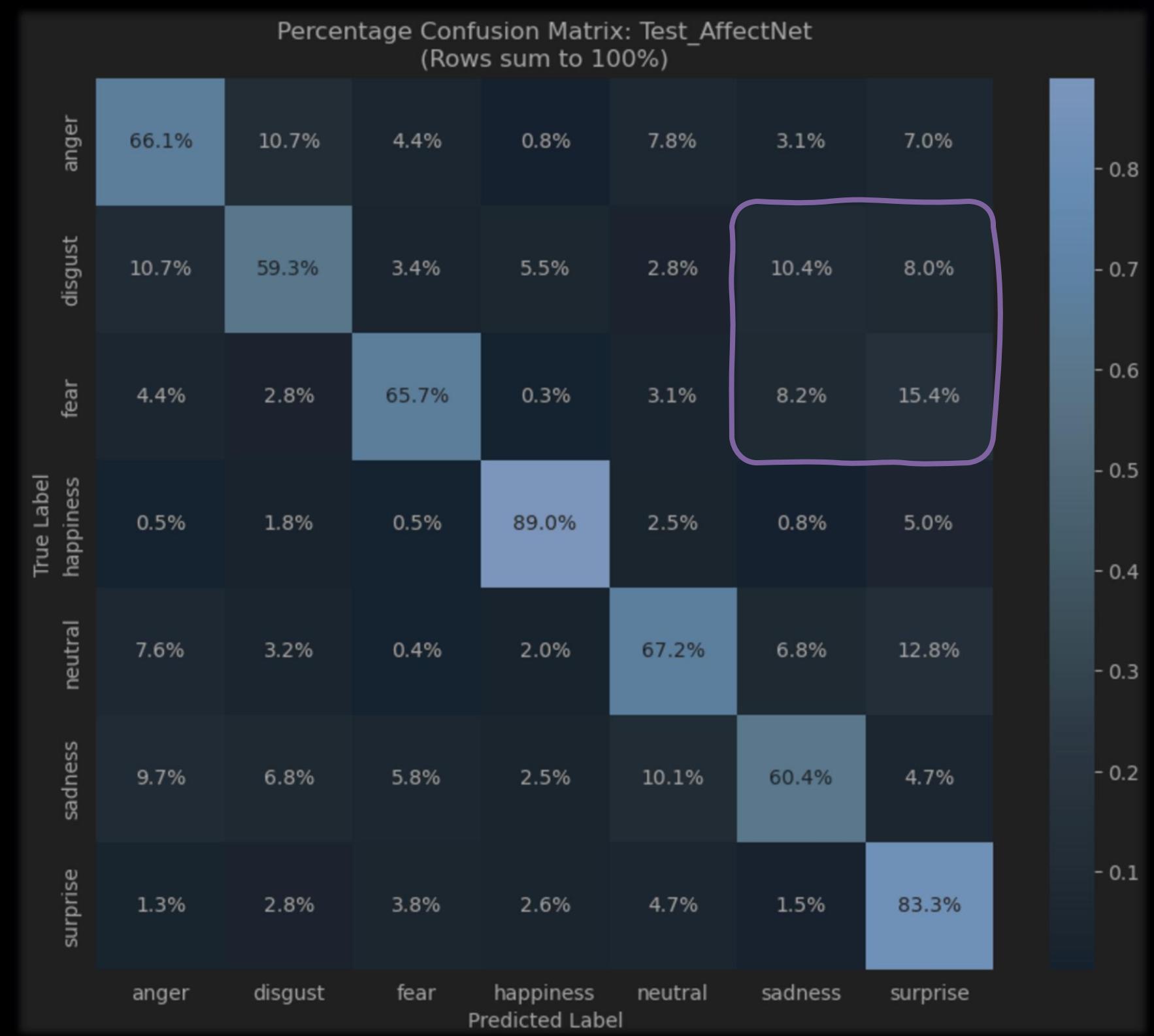
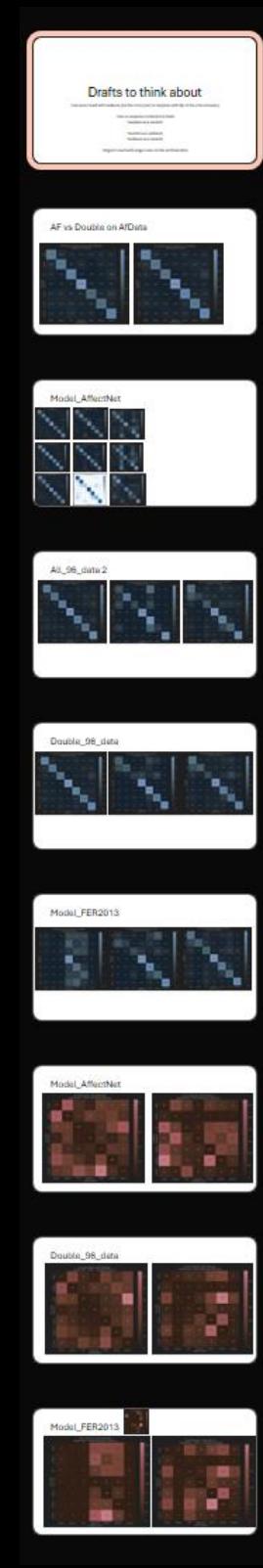
or something at least close to it



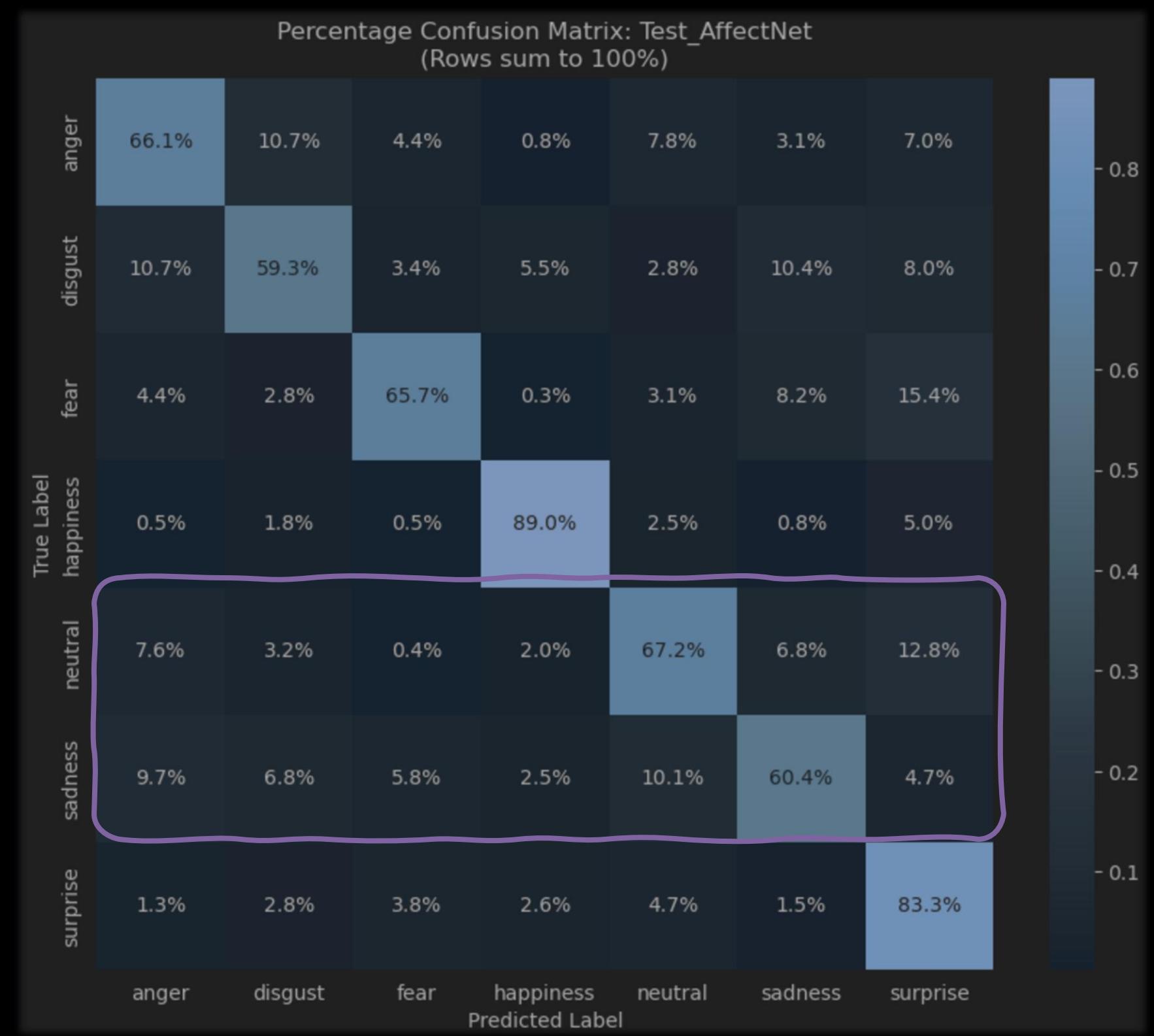
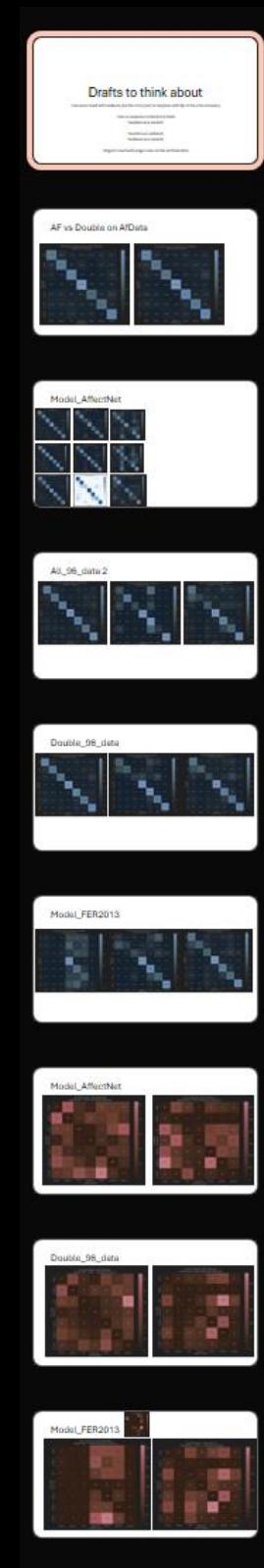
Why not 100%



Why not 100%

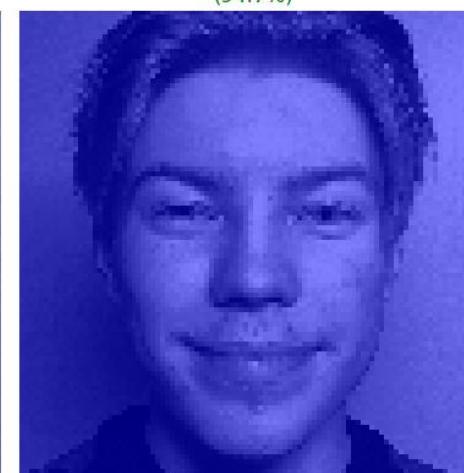
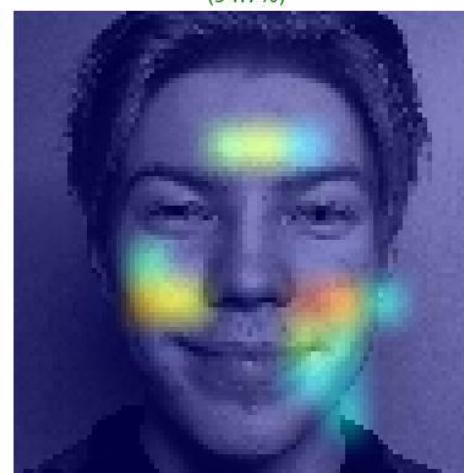
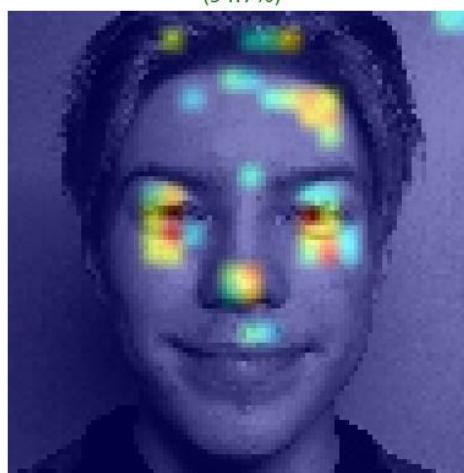
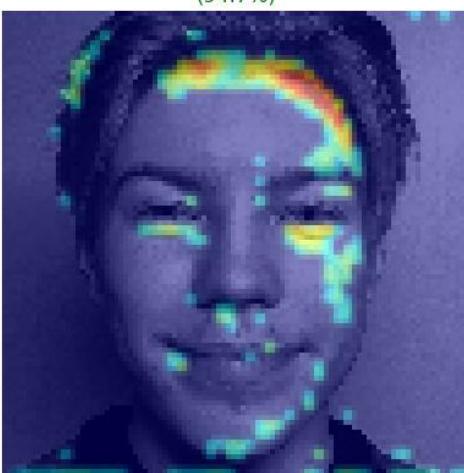


Why not 100%

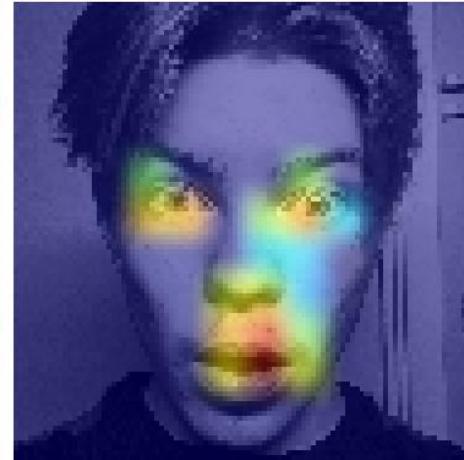
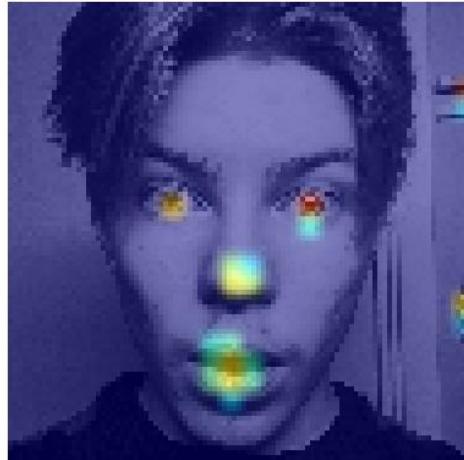
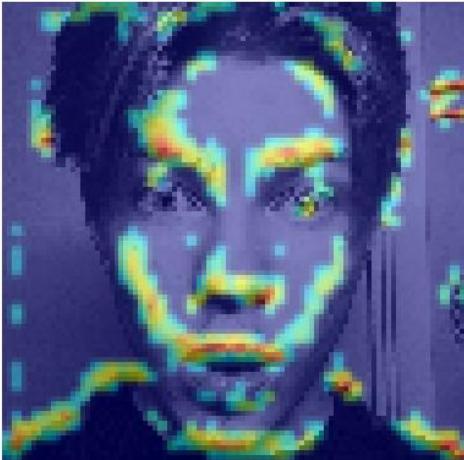
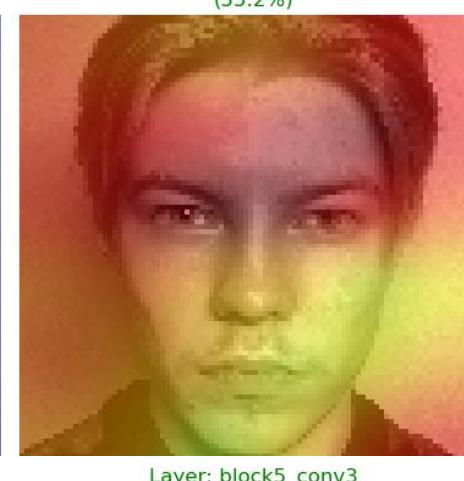
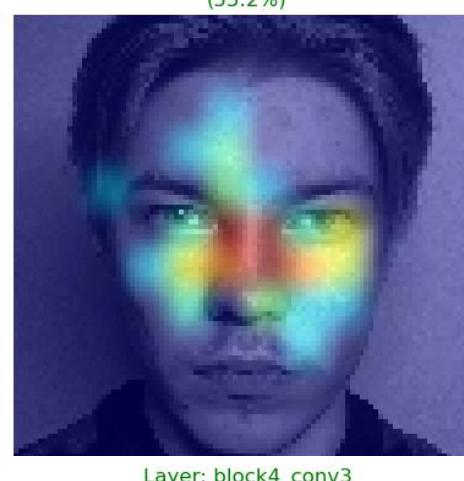
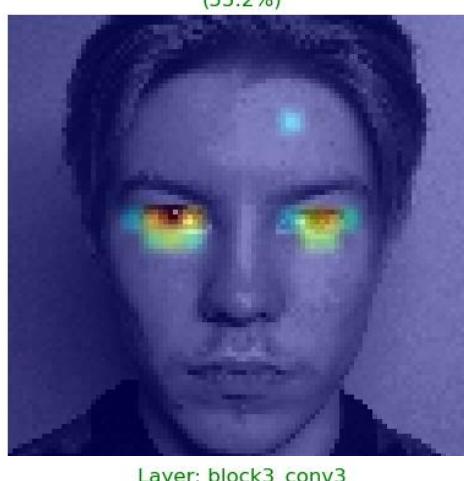
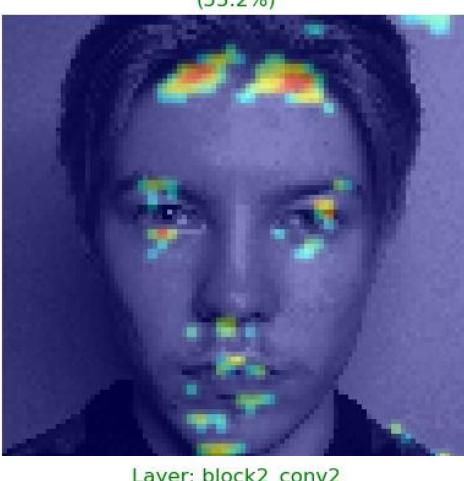
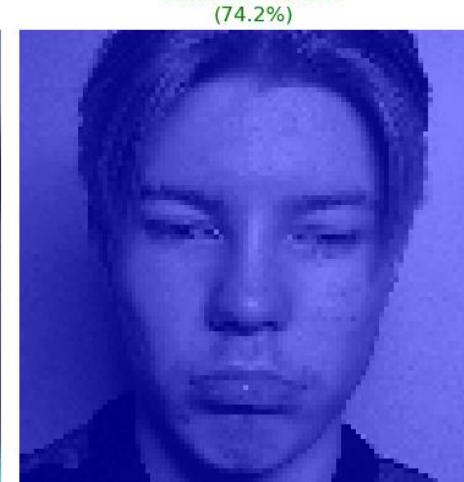
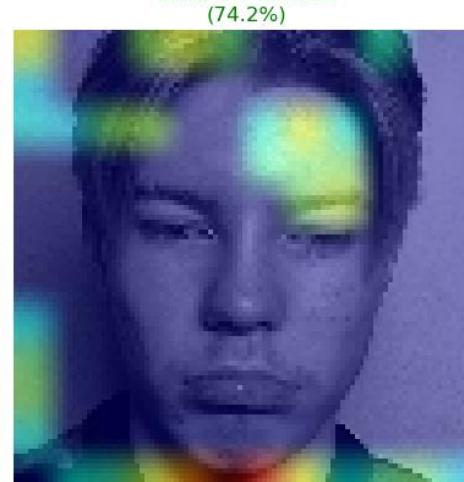
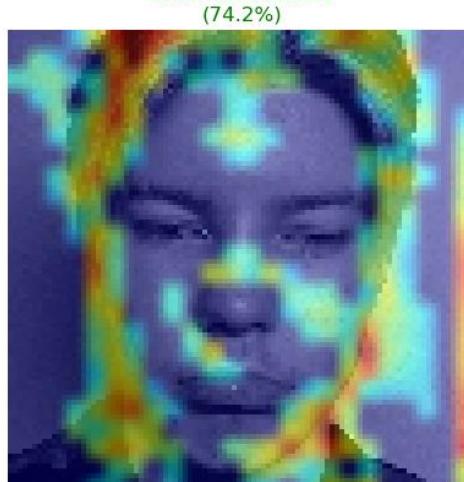
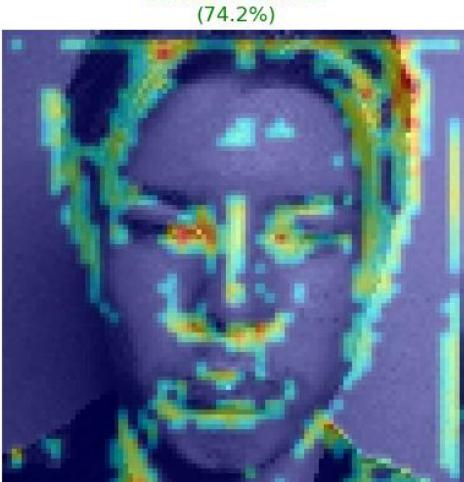


wetup pictures

Group 8



12/12/2025



Ethical questions

Does emotion sell?



A neutral face can become
hunger, grief, or desire depending
on the context

<https://www.youtube.com/watch?v=ZkfihE6JZJw>

– Kuleshov Effect