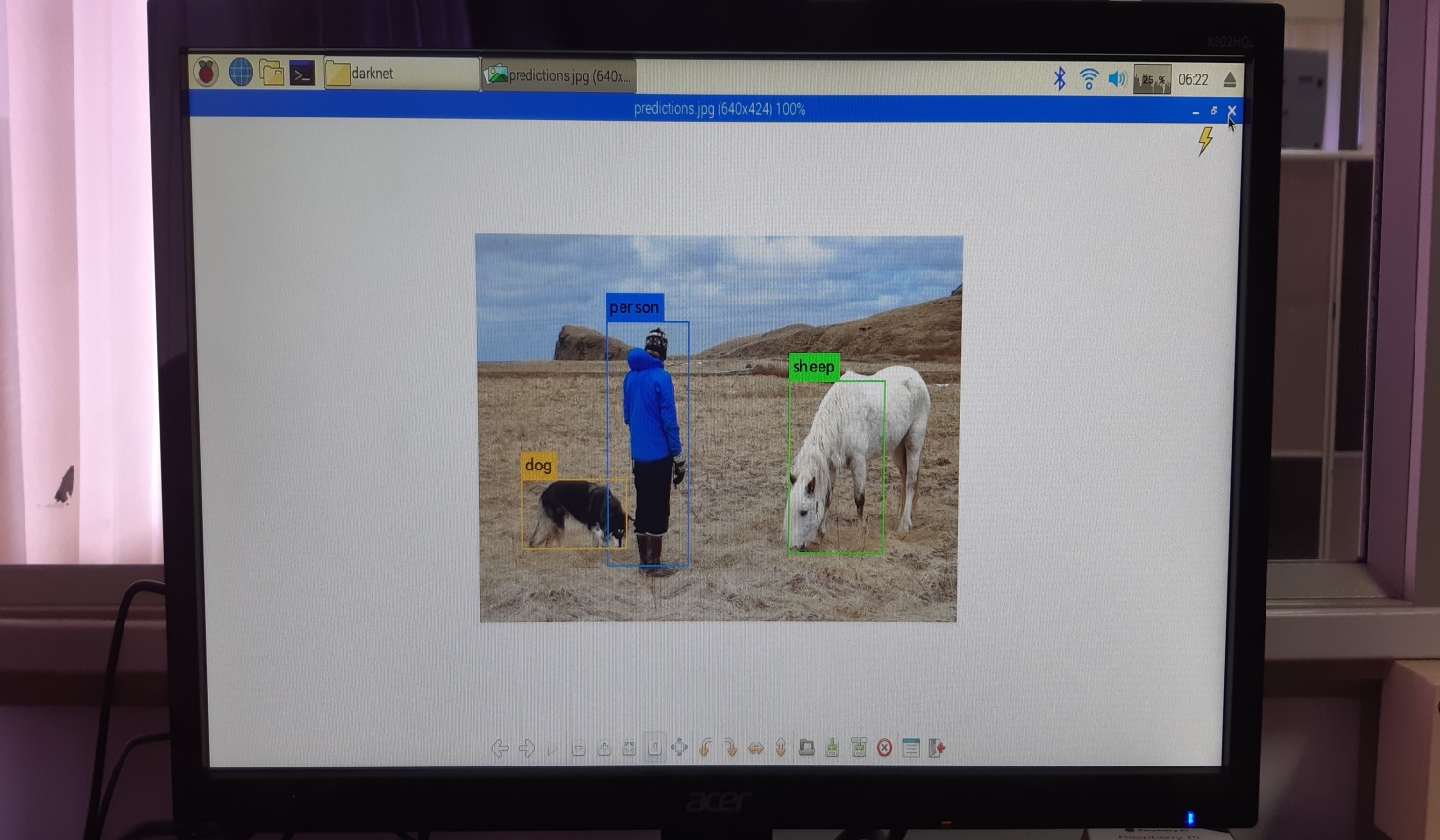
**FINAL RESULTS**

Prediction of images is done using the YOLO and the data is loaded into the data folder.

**Prediction 1:** Prediction on person and animals



layer filters size input output

0 conv 16 3 x 3 / 1 416 x 416 x 3 -> 416 x 416 x 16 0.150 BFLOPs

1 max 2 x 2 / 2 416 x 416 x 16 -> 208 x 208 x 16

2 conv 32 3 x 3 / 1 208 x 208 x 16 -> 208 x 208 x 32 0.399 BFLOPs

3 max 2 x 2 / 2 208 x 208 x 32 -> 104 x 104 x 32

4 conv 64 3 x 3 / 1 104 x 104 x 32 -> 104 x 104 x 64 0.399 BFLOPs

5 max 2 x 2 / 2 104 x 104 x 64 -> 52 x 52 x 64

6 conv 128 3 x 3 / 1 52 x 52 x 64 -> 52 x 52 x 128 0.399 BFLOPs

7 max 2 x 2 / 2 52 x 52 x 128 -> 26 x 26 x 128

8 conv 256 3 x 3 / 1 26 x 26 x 128 -> 26 x 26 x 256 0.399 BFLOPs

9 max 2 x 2 / 2 26 x 26 x 256 -> 13 x 13 x 256

10 conv 512 3 x 3 / 1 13 x 13 x 256 -> 13 x 13 x 512 0.399 BFLOPs

11 max 2 x 2 / 1 13 x 13 x 512 -> 13 x 13 x 512

12 conv 1024 3 x 3 / 1 13 x 13 x 512 -> 13 x 13 x1024 1.595 BFLOPs

13 conv 1024 3 x 3 / 1 13 x 13 x1024 -> 13 x 13 x1024 3.190 BFLOPs

14 conv 125 1 x 1 / 1 13 x 13 x1024 -> 13 x 13 x 125 0.043 BFLOPs

15 detection

mask\_scale: Using default '1.000000'

Loading weights from yolov2-tiny-voc.weights...Done!

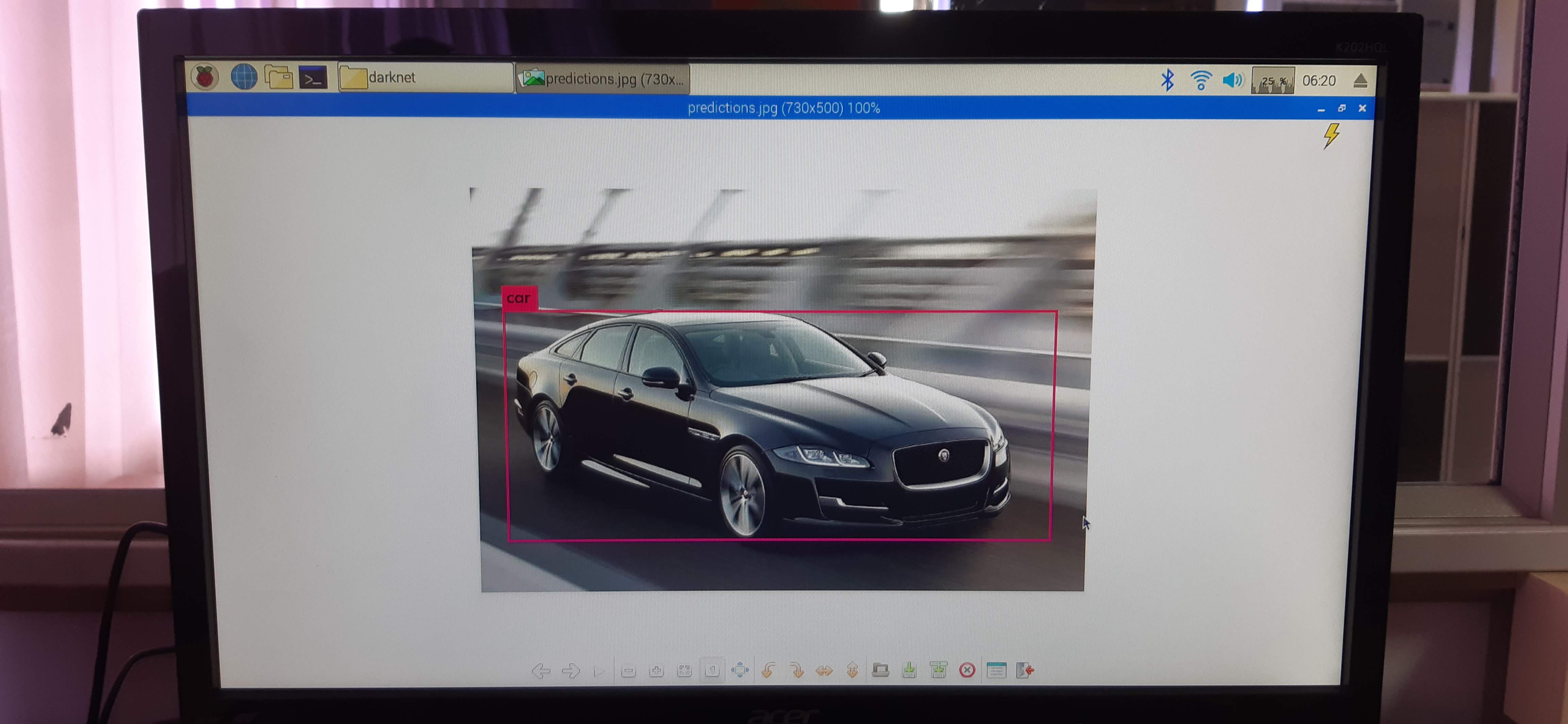
data/person.jpg: Predicted in 69.418173 seconds.

sheep: 60%

person: 73%

dog: 53%

**Prediction 2:** Predicting a car



layer filters size input output

0 conv 16 3 x 3 / 1 416 x 416 x 3 -> 416 x 416 x 16 0.150 BFLOPs

1 max 2 x 2 / 2 416 x 416 x 16 -> 208 x 208 x 16

2 conv 32 3 x 3 / 1 208 x 208 x 16 -> 208 x 208 x 32 0.399 BFLOPs

3 max 2 x 2 / 2 208 x 208 x 32 -> 104 x 104 x 32

4 conv 64 3 x 3 / 1 104 x 104 x 32 -> 104 x 104 x 64 0.399 BFLOPs

5 max 2 x 2 / 2 104 x 104 x 64 -> 52 x 52 x 64

6 conv 128 3 x 3 / 1 52 x 52 x 64 -> 52 x 52 x 128 0.399 BFLOPs

7 max 2 x 2 / 2 52 x 52 x 128 -> 26 x 26 x 128

8 conv 256 3 x 3 / 1 26 x 26 x 128 -> 26 x 26 x 256 0.399 BFLOPs

9 max 2 x 2 / 2 26 x 26 x 256 -> 13 x 13 x 256

10 conv 512 3 x 3 / 1 13 x 13 x 256 -> 13 x 13 x 512 0.399 BFLOPs

11 max 2 x 2 / 1 13 x 13 x 512 -> 13 x 13 x 512

12 conv 1024 3 x 3 / 1 13 x 13 x 512 -> 13 x 13 x1024 1.595 BFLOPs

13 conv 1024 3 x 3 / 1 13 x 13 x1024 -> 13 x 13 x1024 3.190 BFLOPs

14 conv 125 1 x 1 / 1 13 x 13 x1024 -> 13 x 13 x 125 0.043 BFLOPs

15 detection

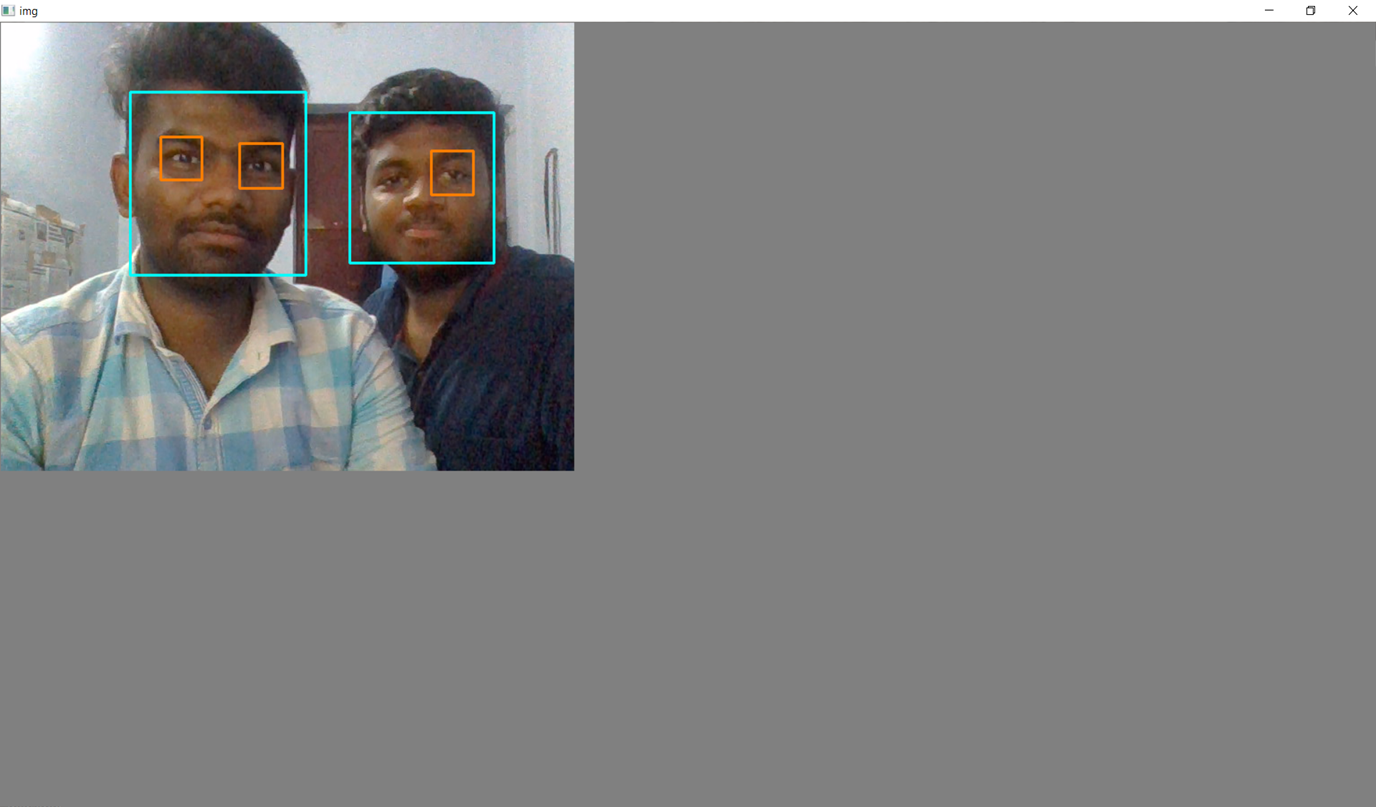
mask\_scale: Using default '1.000000'

Loading weights from yolov2-tiny-voc.weights...Done!

data/download.jpg: Predicted in 70.879140 seconds.

car: 96%

**Prediction 3:** Predicting human eyes



layer filters size input output

0 conv 16 3 x 3 / 1 416 x 416 x 3 -> 416 x 416 x 16 0.150 BFLOPs

1 max 2 x 2 / 2 416 x 416 x 16 -> 208 x 208 x 16

2 conv 32 3 x 3 / 1 208 x 208 x 16 -> 208 x 208 x 32 0.399 BFLOPs

3 max 2 x 2 / 2 208 x 208 x 32 -> 104 x 104 x 32

4 conv 64 3 x 3 / 1 104 x 104 x 32 -> 104 x 104 x 64 0.399 BFLOPs

5 max 2 x 2 / 2 104 x 104 x 64 -> 52 x 52 x 64

6 conv 128 3 x 3 / 1 52 x 52 x 64 -> 52 x 52 x 128 0.399 BFLOPs

7 max 2 x 2 / 2 52 x 52 x 128 -> 26 x 26 x 128

8 conv 256 3 x 3 / 1 26 x 26 x 128 -> 26 x 26 x 256 0.399 BFLOPs

9 max 2 x 2 / 2 26 x 26 x 256 -> 13 x 13 x 256

10 conv 512 3 x 3 / 1 13 x 13 x 256 -> 13 x 13 x 512 0.399 BFLOPs

11 max 2 x 2 / 1 13 x 13 x 512 -> 13 x 13 x 512

12 conv 1024 3 x 3 / 1 13 x 13 x 512 -> 13 x 13 x1024 1.595 BFLOPs

13 conv 1024 3 x 3 / 1 13 x 13 x1024 -> 13 x 13 x1024 3.190 BFLOPs

14 conv 125 1 x 1 / 1 13 x 13 x1024 -> 13 x 13 x 125 0.043 BFLOPs

15 detection

mask\_scale: Using default '1.000000'

Loading weights from yolov2-tiny-voc.weights...Done!

data/camera.jpg: Predicted in 68.389034 seconds.

person: 60%

person: 73%