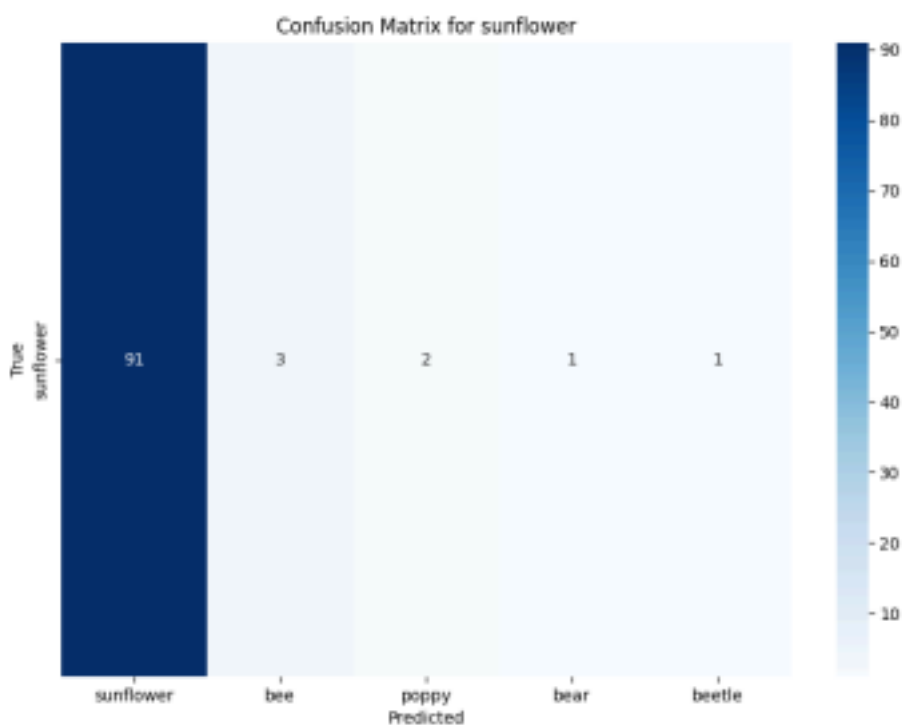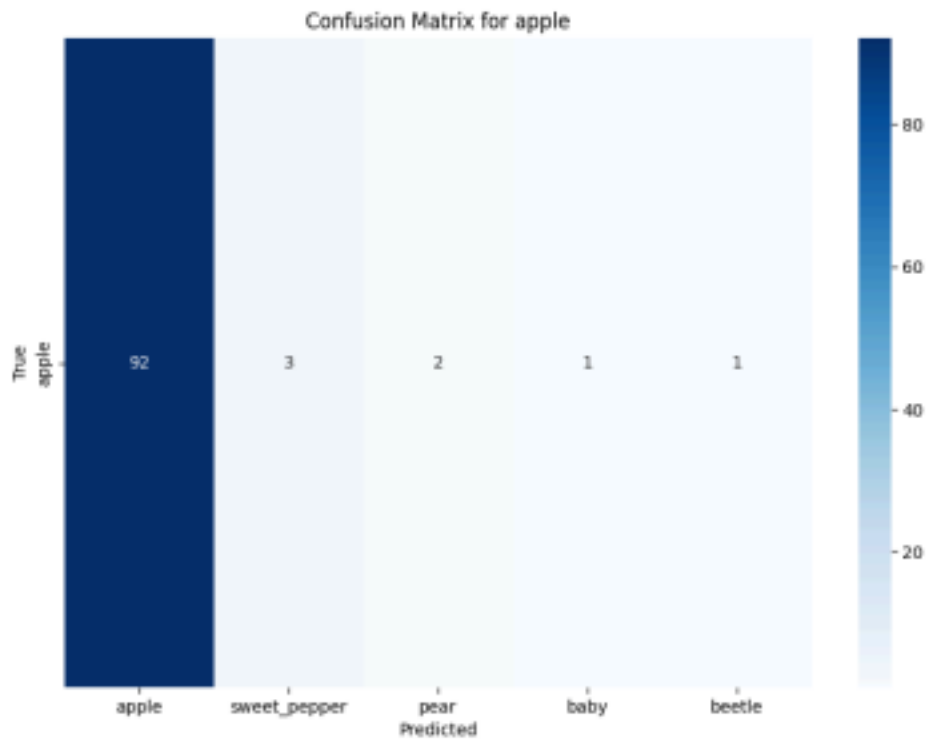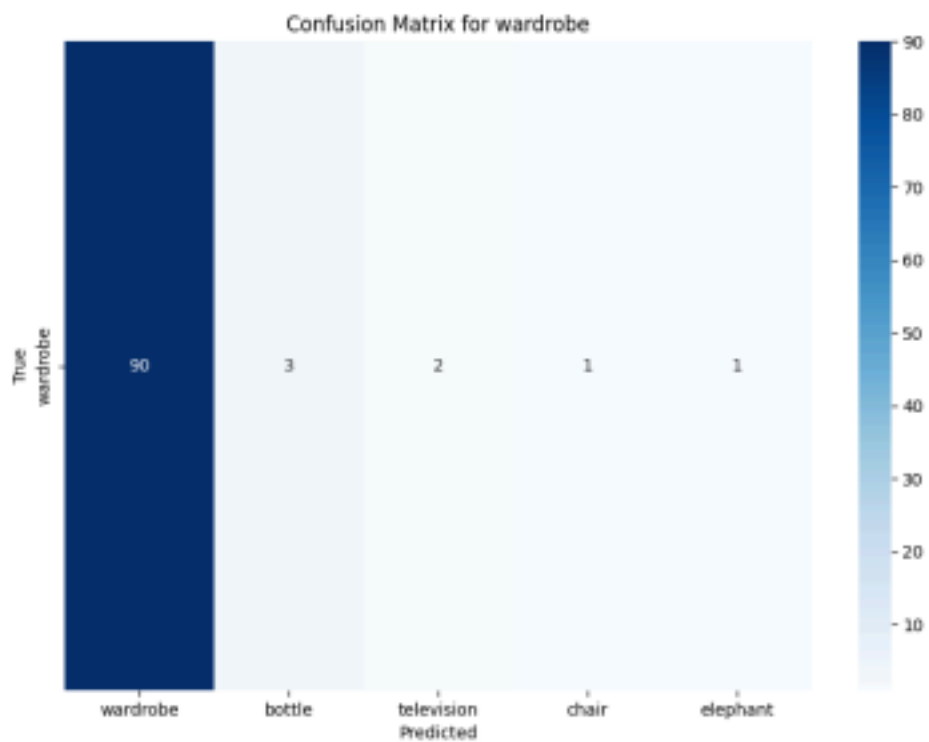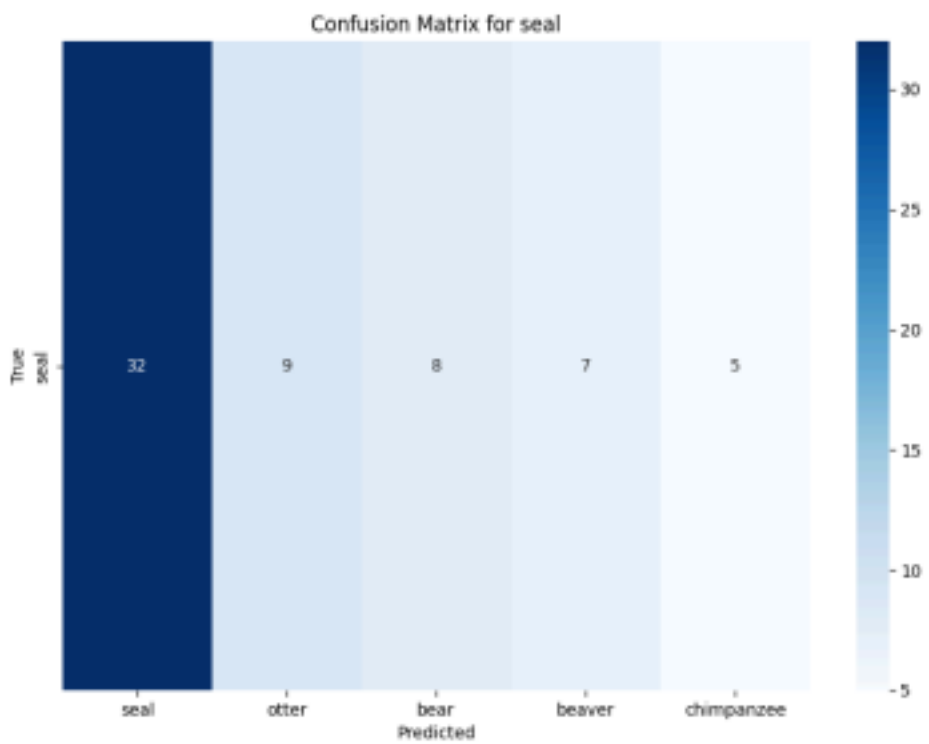# Best model - 66,30% accuracy

Confusion Matrices

The confusion matrices for the three best-predicted and three worst-predicted classes are presented. The matrices show the 5 most frequently predicted classes by the model for a given class, which also represents FN.

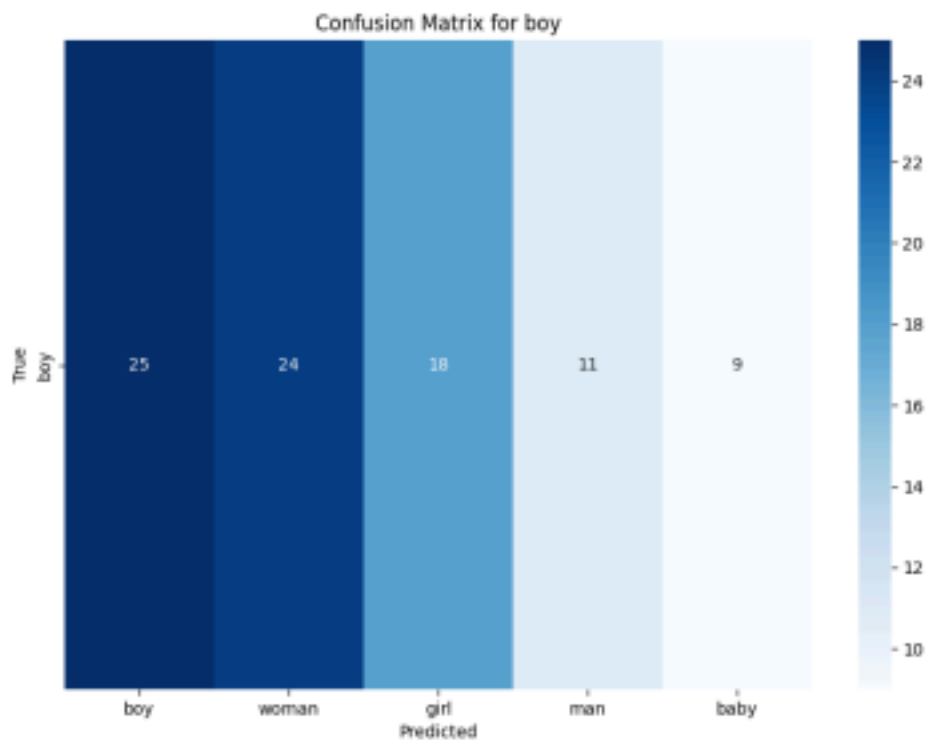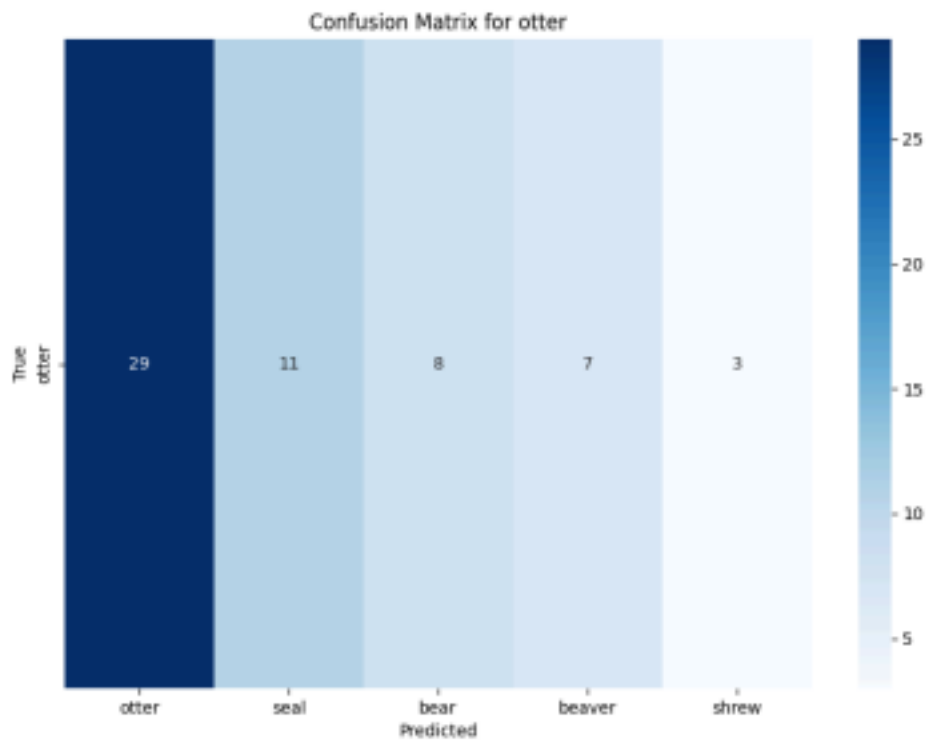The best-predicted classes: **Apple (92%), Sunflower (91%), Wardrobe (90%)**.

Confusion Matrix for apple

Confusion Matrix for sunflower

Confusion Matrix for wardrobe

The worst-predicted class: **Seal(32%), Otter(29%), Boy(25%)**

Confusion Matrix for seal

Confusion Matrix for otter

Confusion Matrix for boy

Metrics

**Precision Score - 0,675**

It defines how many of the cases predicted as positive are actually positive. To calculate precision, we used the weighted variant, meaning precision was calculated for each class and weighted according to the number of examples of that class in the entire dataset.

The class with the best precision: **Trout (0.94)**

The class with the lowest precision: **Woman (0.34)** – in the chart below, it can be

seen that the model frequently makes false positive predictions for the Woman class:


Confusion Matrix for woman
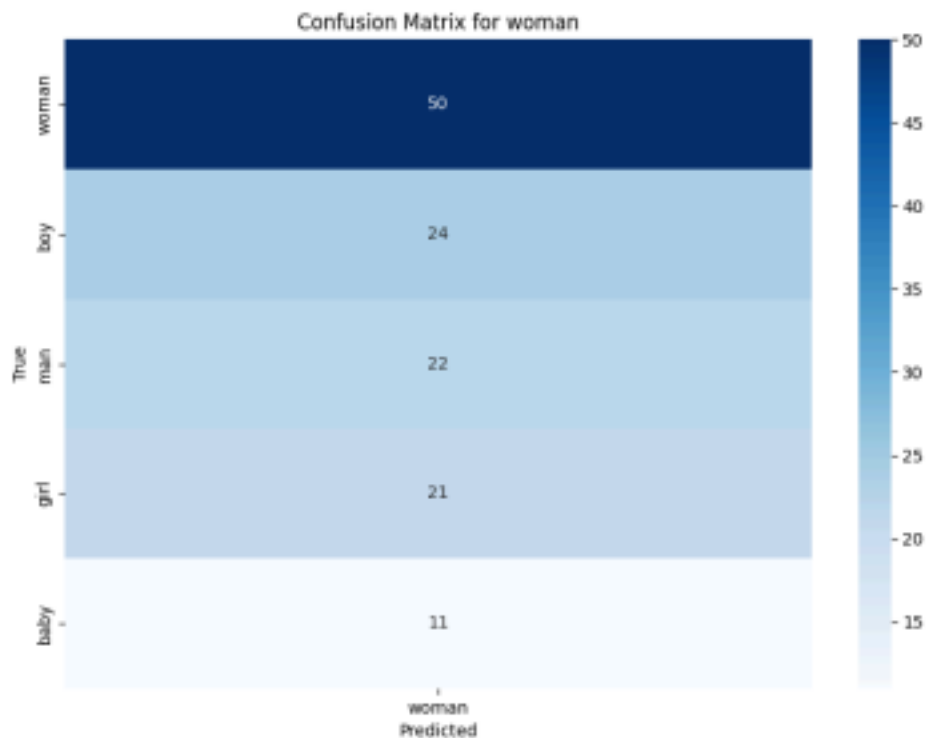
Recall - 0,663
It measures what portion of all truly positive cases the model detected. This is a key metric when missing a positive case could have serious consequences. We also used the weighted variant to calculate sensitivity.

The class with the best sensitivity: **Apple (0.92)**

The classes with the lowest sensitivity: **Boy (0.26)** – as seen in the chart below, the model predicts many false negatives (FN) for the Boy class:

Confusion Matrix for boy
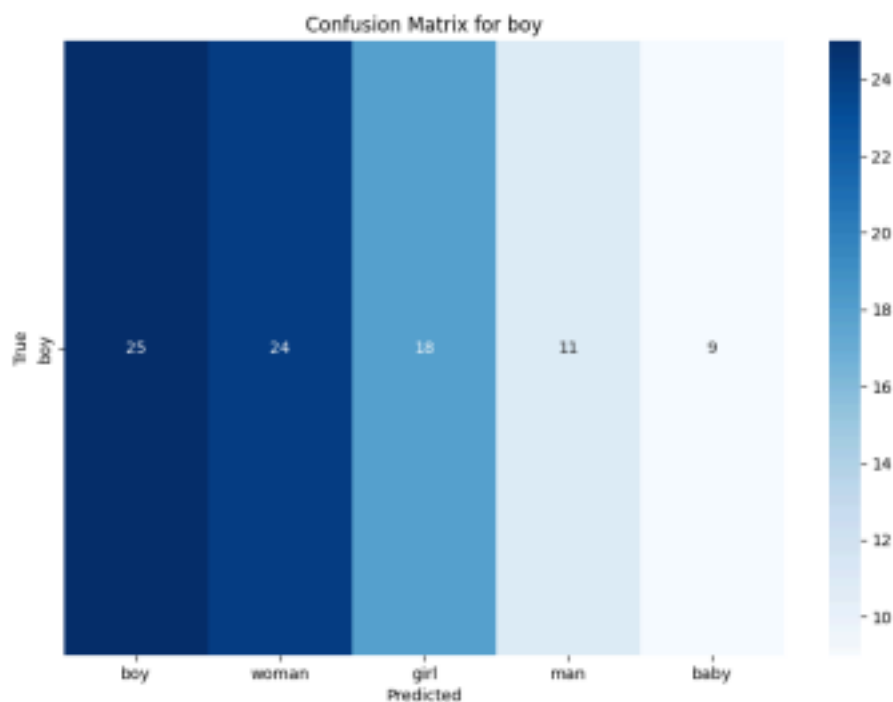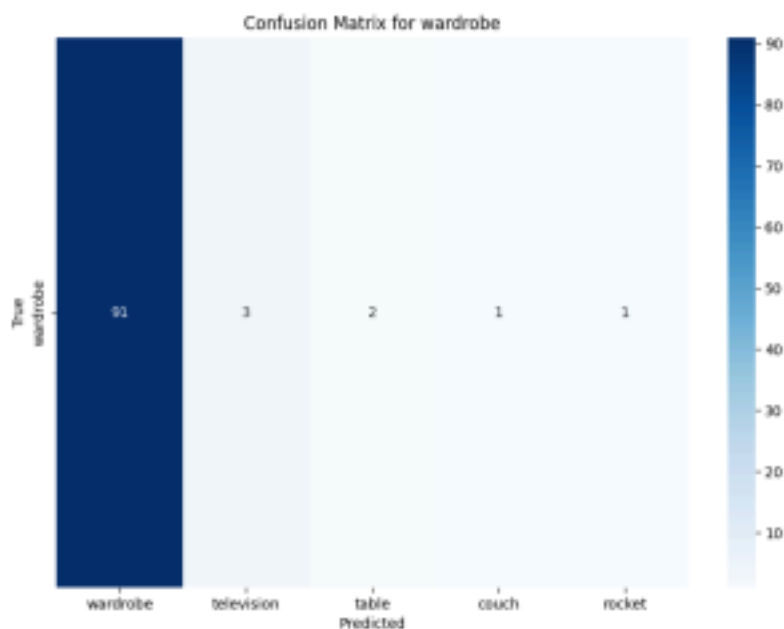
F1-Score - 0,665

It is the harmonic mean of precision and sensitivity, which takes both aspects into account. We also used the weighted variant to calculate the F1-score. The F1-score is useful when we want a balance between precision and recall, for example, in situations where both false positives and false negatives are costly.
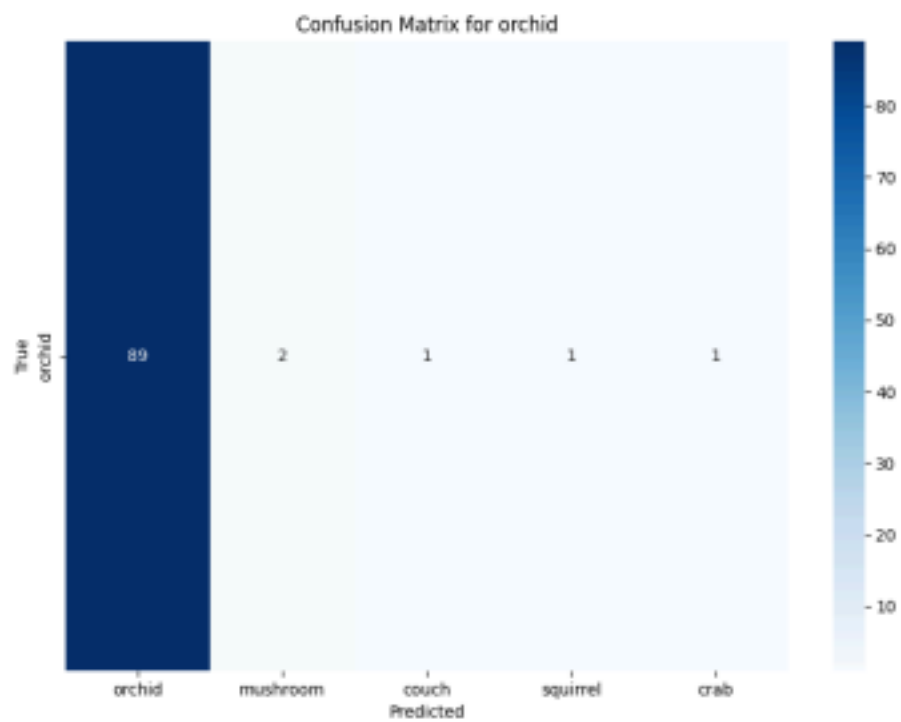
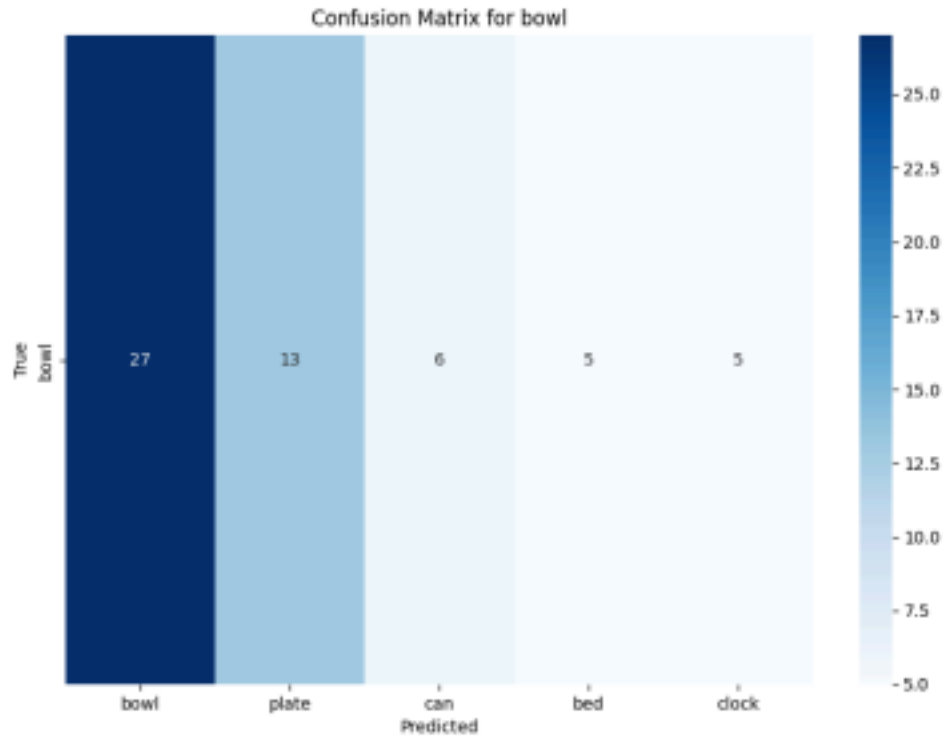## Second best model - 64,32% accuracy

Confusion Matrices

As before, confusion matrices are shown where a certain class was expected, but the model classified the image under a different label. The matrices are presented for the 3 best-predicted classes and the 3 worst-predicted classes.
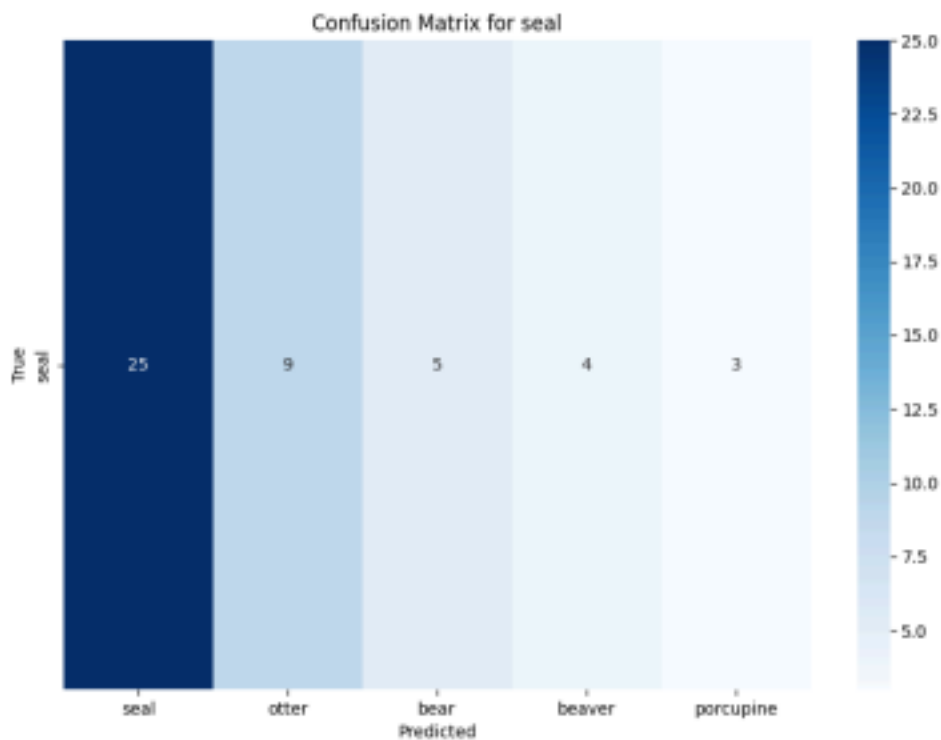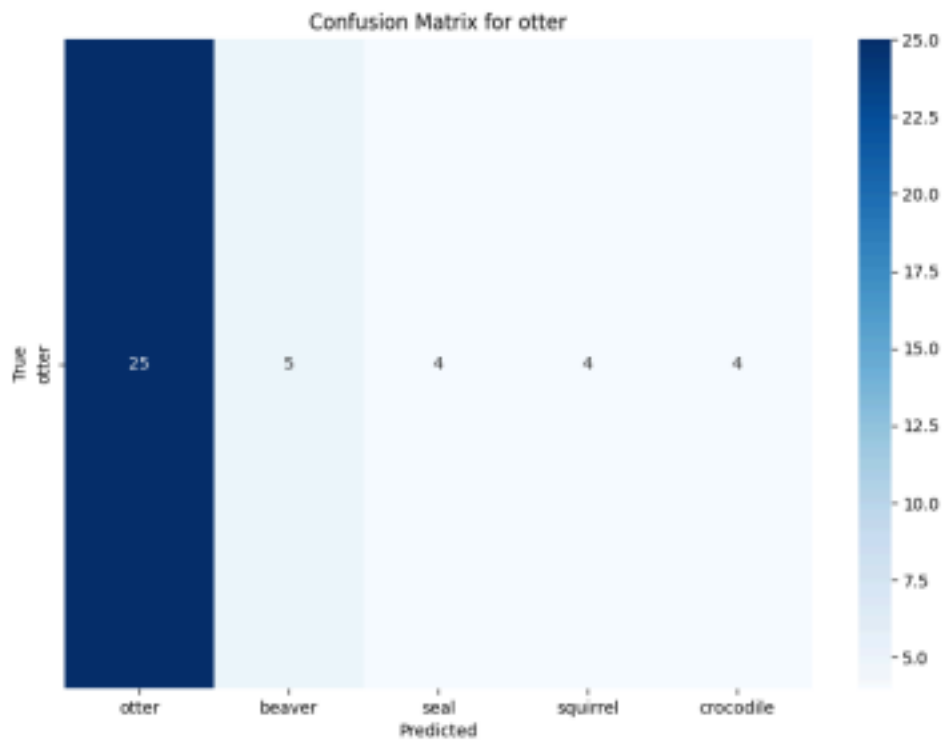The best-predicted classes: **Wardrobe (92%), Orchid (91%), Road (90%)**.

Confusion Matrix for orchid

Worst-predicted classes: **Bowl(27%), Otter(25%), Seal(25%)**

Confusion Matrix for bowl

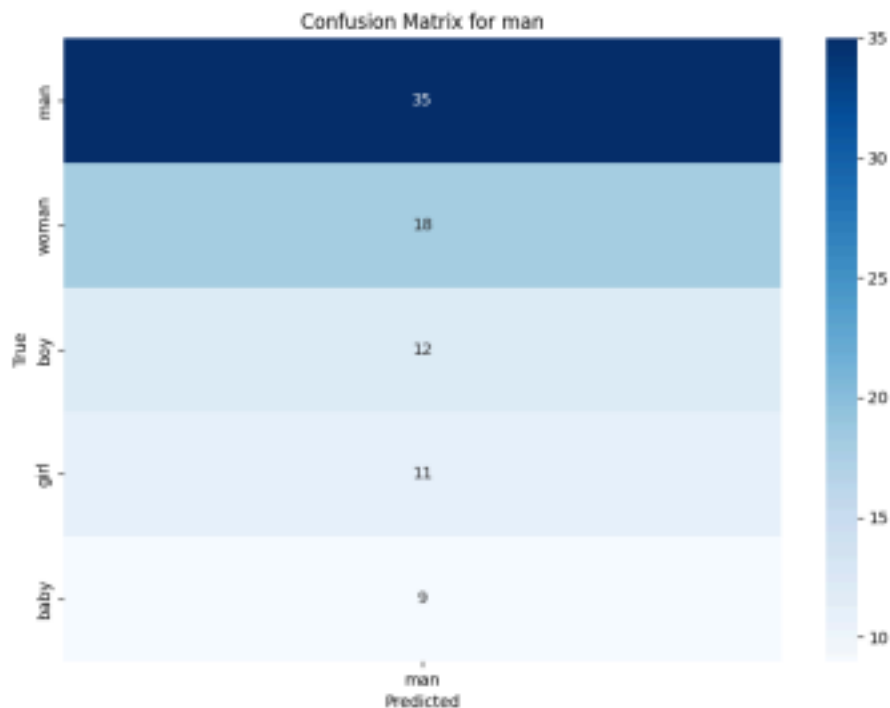Confusion Matrix for otter

Confusion Matrix for seal

Metrics
Precision score: 0.656

It was calculated in the same way as for the best model.

The class with the best precision: **Skunk (0.90)**

The class with the lowest precision: **Man (0.27)** – in the chart below, it can be seen that the model frequently makes false positive predictions for the Man class:
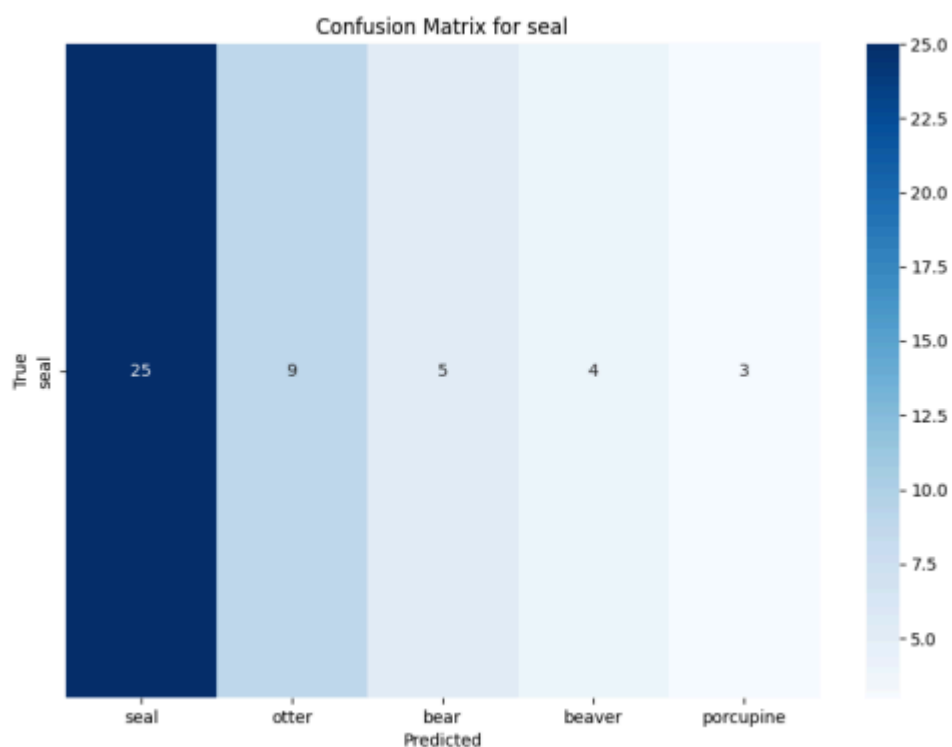
Confusion Matrix for man

Recall: 0.643

Recall was calculated in the same way as for the best model.

The class with the best sensitivity: **Wardrobe (0.91)**

The classes with the lowest sensitivity: **Seal (0.25)** – as seen in the chart below, for the Seal class, the model frequently predicts many false negatives (FN):

Confusion Matrix for seal

F1-Score: 0.643

The harmonic mean of precision and sensitivity for this model is also similar to the values of precision and sensitivity. Therefore, the model achieves balanced

performance and does not favor any of the metrics, which is also due to the well-balanced data, as each class in the CIFAR100 dataset has the same number of images.