Computer vision fundamentals

Wild cats detection and classification

Leo Škvorc

Zagreb, June 9. 2024.

Content

[1 Introduction 4](#_Toc169260536)

[2 Data 5](#_Toc169260537)

[4 Training models 6](#_Toc169260538)

[4.1 Custom model 6](#_Toc169260539)

[4.2 EfficentNetB0, ResNet50 and MobNet 6](#_Toc169260540)

[4.3 Training 6](#_Toc169260541)

[6 Results 7](#_Toc169260542)

[6.1 Accuracy loss f1 socres 7](#_Toc169260543)

[6.1.1 Accuracy 8](#_Toc169260544)

[6.1.2 Loss 9](#_Toc169260545)

[6.1.3 F1 score 10](#_Toc169260546)

[6.2 Confusion matrix 11](#_Toc169260547)

[6.4 Classification reports 12](#_Toc169260548)

[6.6 Image example 13](#_Toc169260549)

[7 Conclusion 14](#_Toc169260550)

Tables

**No table of figures entries found.**

Images

[Figure 1 Model scores 8](#_Toc169260551)

[Figure 2 Accuracy graph 9](#_Toc169260552)

[Figure 3 Loss graph 10](#_Toc169260553)

[Figure 4 F1 score graph 11](#_Toc169260554)

[Figure 5 Confusion matrix 12](#_Toc169260555)

[Figure 6 Custom model classification report 13](#_Toc169260556)

[Figure 7 ResNet model classification report 13](#_Toc169260557)

[Figure 8 MobNet model classification report 13](#_Toc169260558)

[Figure 9 EffNet model classification report 13](#_Toc169260559)

[Figure 10 Test prediction visualization 14](#_Toc169260560)

# Introduction

The domain of computer vision has advanced significantly, enabling sophisticated models that can identify and classify objects within images and videos. This project targets the challenge of detecting and classifying cats in wild settings, a task complicated by the variability and complexity of outdoor environments. The goal is to develop a robust and accurate model that can effectively identify and classify cats in diverse conditions. The developed model will then be compared to other 3 state of the art models.

# Data

Training and validation data is sourced from [10 Big Cats of the Wild – Image Classification](https://www.kaggle.com/datasets/gpiosenka/cats-in-the-wild-image-classification/data). The data is brought from train, validation and test folder to one master folder and then split 80-20 to train and valid data. Additionally test data is created by [Canva](https://www.canva.com) using prompts “(Cat name) in the wilderness”. Each test class has 40 images of different styles.

To prepare the data for training image generators have been used to rescale all the images to 0 and 1. Additionally train images have randomly been rotated by 20 degrees, shifted by 20%, horizontally flipped.

Image generators have been created with respected datasets, targeted image size of 256x256, batch size of 16 and categorical class model.

Training data contains 1946 images, validation data contains 493 images and test data contains 400 images. The data is split into 10 classes: AFRICAN LEOPARD, CARACAL, CHEETAH, CLOUDED LEOPARD, JAGUAR, LIONS, OCELOT, PUMA, SNOW LEOPARD and TIGER.

# Training models

In this project, four models were tested for detecting and classifying cats in wild settings. The first model is custom-made, while the remaining three are state-of-the-art models: ResNet50, MobileNet, and EfficientNetB0.

## Custom model

Custom model has been created using multiple convolutional and pooling layers. The first layer has an input shape of 256x246x3, representing the height, width and RGB colour channels of the images. All convolution layers utilize ReLU activation functions. Before the end final layer, the data is flattened and then passed through dense layers. The final dense layer outputs the number of classes with softmax activation function, allowing the model to predict the class of the given image.

## EfficentNetB0, ResNet50 and MobNet

The state-of-the-art models EfficientNetB0, ResNet50, and MobileNet were adjusted to fit the desired classes by adding pooling and dense layers. These modifications ensure that the models can handle the specific classification tasks required for this project while leveraging the powerful feature extraction capabilities of these pre-trained networks. To check the plain models training was disabled.

## Training

All models are compiled using the Adam optimiser, categorical cross entropy loss and accuracy metrics. Additionally, each model had a custom metrics callback to record F1 scores over each epoch, providing a more comprehensive evaluation of model performance. The custom model also included a checkpoint callback to save the model to the file system during training.

All models were trained over 40 epochs, and the final created model is saved to the file system.

# Results

## Accuracy loss f1 socres

A graph of a performance

Description automatically generated

Figure 1 Model scores

The model performance graph shows that all of the models perform poorly for the given data.

### Accuracy

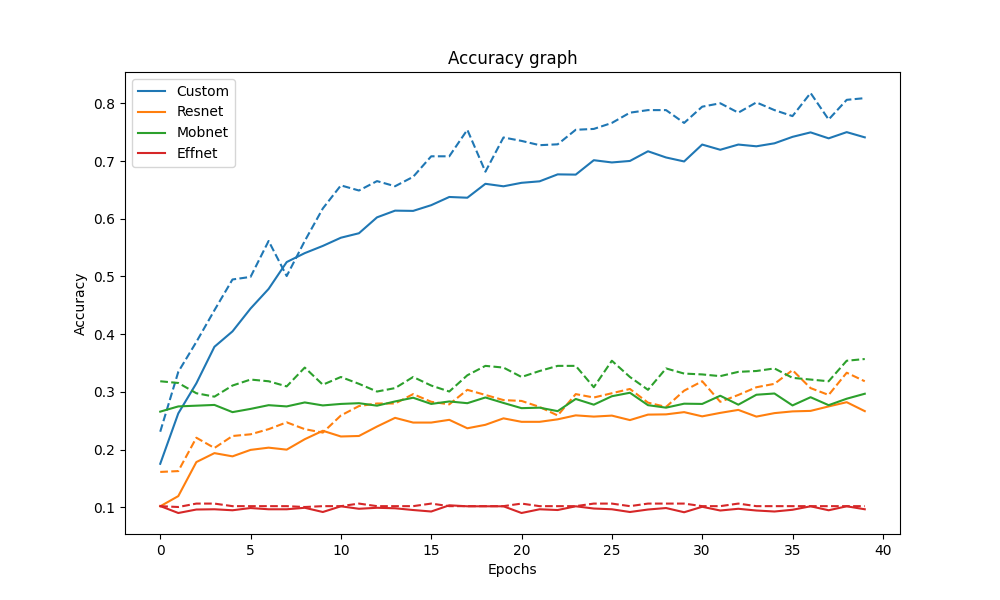


Figure 2 Accuracy graph

Accuracy graph shows that ResNet and MobNet models perform similarly while EffNet model is not useful for given dataset. Custom model has the best score. The accuracy of custom model is increasing thill the last epoch while other models are improving thill fifth epoch. All the graphs don’t have false positives proven by dashed line.

### Loss

A graph of loss of a graph

Description automatically generated with medium confidence

Figure 3 Loss graph

The loss graph values have opposite tendency to the accuracy graph what is desired outcome.

### F1 score

A graph with different colored lines

Description automatically generated

Figure 4 F1 score graph

F1 score graph confirms previous two graphs data giving good results for custom model, bad results for EffNet and below average results for MobNet and ResNet.

## Confusion matrix

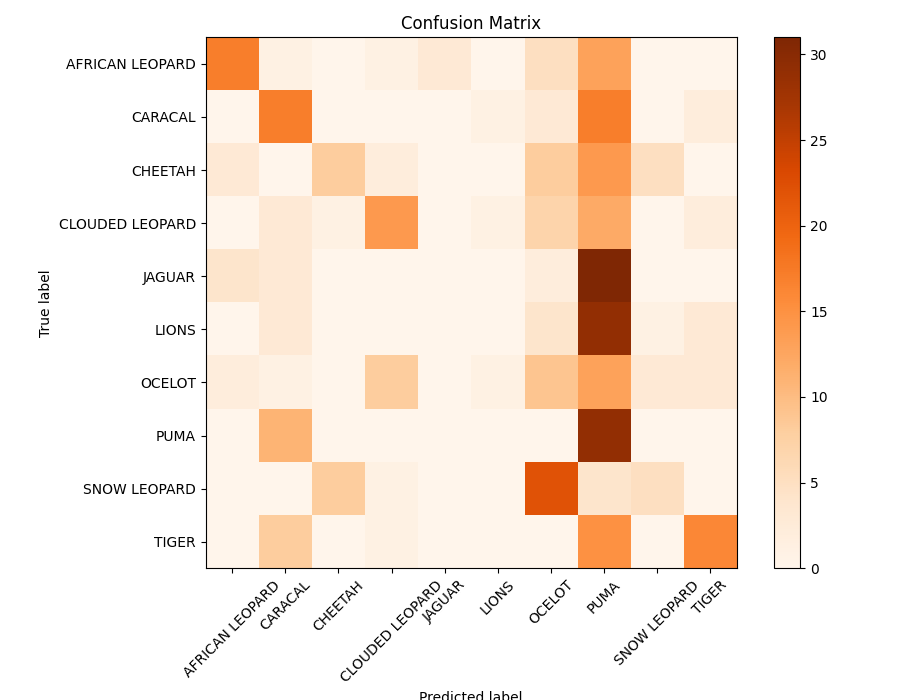


Figure 5 Confusion matrix

Confusion matrix displays the overfitting of the model for pumas and complete lack of detection for jaguars and lions.

## Classification reports

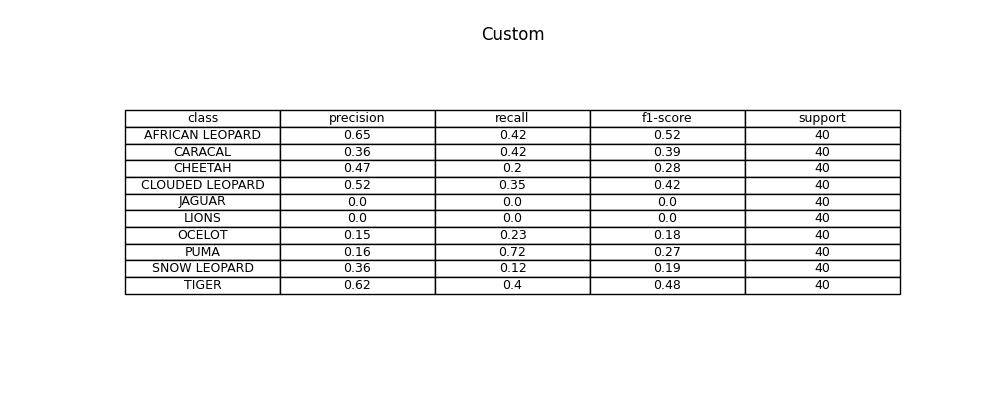


Figure 6 Custom model classification report

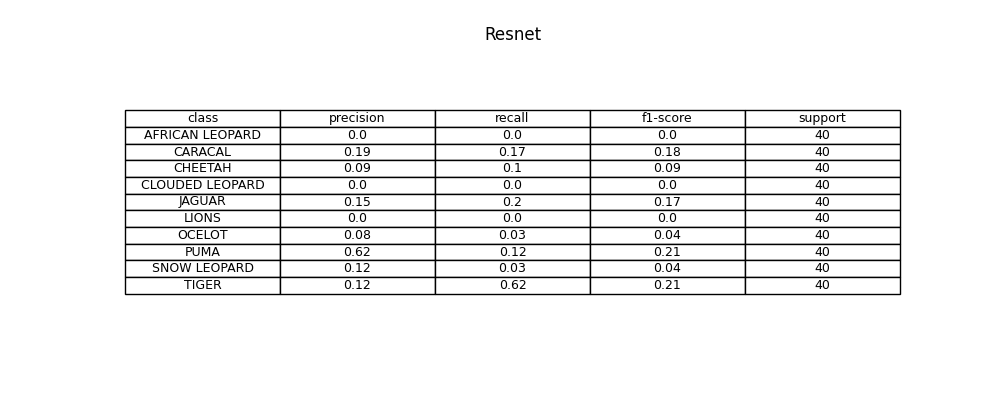


Figure 7 ResNet model classification report

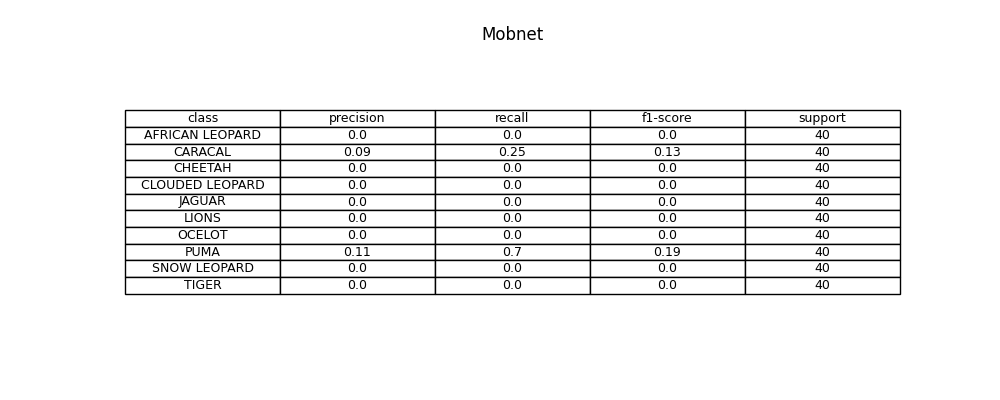


Figure 8 MobNet model classification report

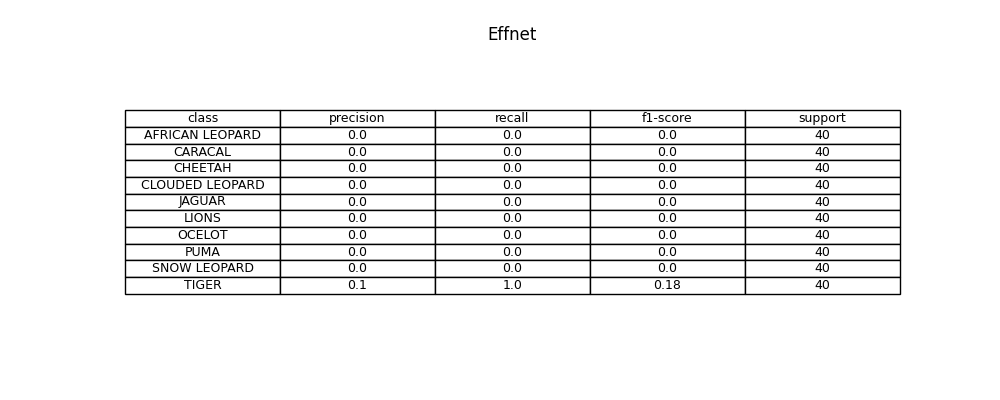


Figure 9 EffNet model classification report

Clasification reports confirm graph data. Jaguars and lions are not being recognized by any model. Custom model has the best scores and EffNet model has the worst scores.

## Image example

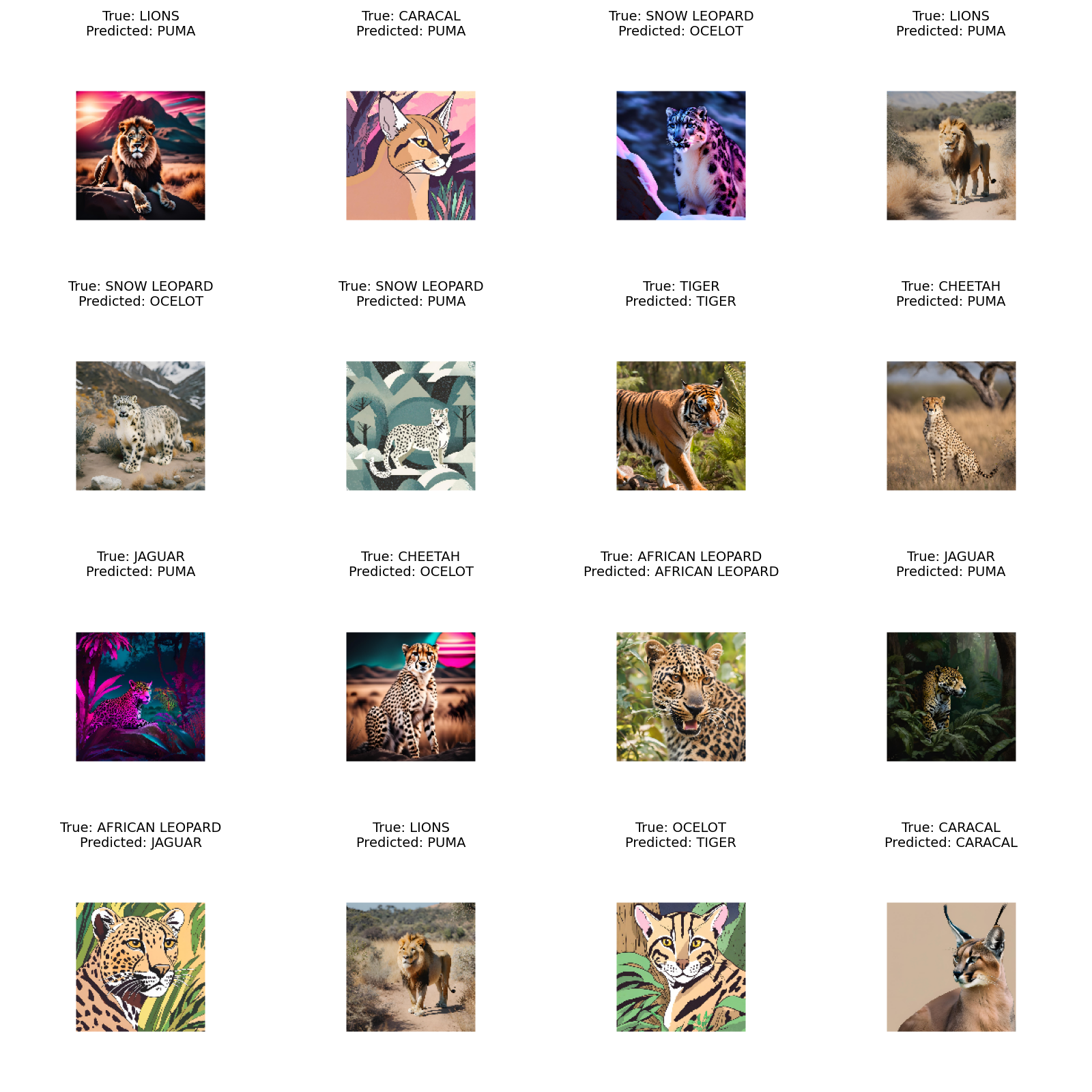


Figure 10 Test prediction visualization

Visualizing the pictures with predicted and true labels it is clear that training scores are better than evaluation scores because AI generated images are not always good representations of real images and models are not trained good enough for classification of such images.

# Conclusion

Custom model is the best way for solving specific requests such as wild cats detection and classification. If high precision is the main requirement we need to provide the biggest dataset possible for training validating and testing the model.