Computer vision fundamentals

Wild cats detection and classification

Leo Škvorc

Zagreb, June 9. 2024.

Content

[1 Introduction 4](#_Toc168862419)

[2 Data 4](#_Toc168862420)

[3 Training models 4](#_Toc168862421)

[3.1 My model 4](#_Toc168862422)

[3.2 EfficentNetB0 and ResNet50 4](#_Toc168862423)

[3.3 Training 4](#_Toc168862424)

[4 Results 4](#_Toc168862425)

[4.1 Accuracy loss f1 socres 5](#_Toc168862426)

[4.2 Confusion matrix 5](#_Toc168862427)

[5 Conclusion 5](#_Toc168862428)

Tables

**No table of figures entries found.**

Images

[Figure 1 Model scores 5](#_Toc168946817)

[Figure 2 Metrics graph 6](#_Toc168946818)

[Figure 3 Confusion matrix 7](#_Toc168946819)

# Introduction

This project focuses on the domain of computer vision, specifically targeting the challenge of detecting and classifying cats in wild settings. The purpose of this project is to develop a robust model that can accurately identify and classify cats in various outdoor environments.

# 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

Four models have been tested, the first one is custom made, the rest are state of the art models: EfficentNetB0, ResNet50 and MobNet.

## Custom model

Custom model has been created using multiple convolutional and pooling layers. The first layer additionally has input shape of the image what is 256x246x3, meaning that the height and width of the image are 256, and it has RGB colours. All convolution layers have relu activation functions. Before the end data is flattened and densed. The final densing function has the output of the number of classes and activation function softmax.

## EfficentNetB0, ResNet50 and MobNet

State of the art models are adjusted to fit desired classes using pooling and dense function.

## Training

All models are compiled using “Adam” optimiser, “categorical\_crossentropy” loss and accuracy metrics. Additionally, each model had custom metrics callback to record f1 scores over each epoch, with Custom model having checkpoint callback to save it to the file system.

Models are trained over 40 epochs and created model is saved to the file system.

# 

# Results

## Accuracy loss f1 socres

## 

Figure 1 Model scores

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

### Accuracy

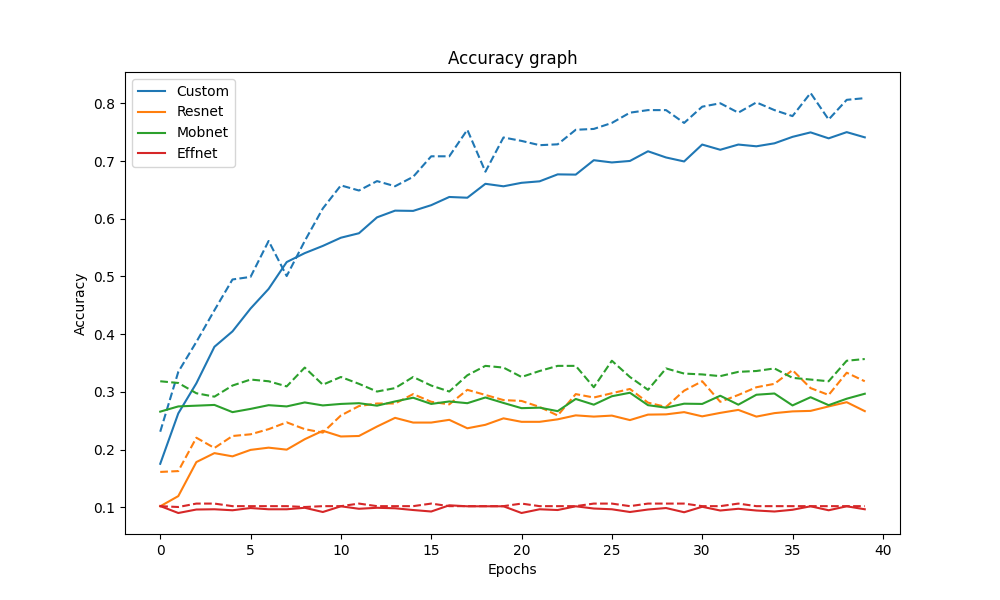


Figure 2 Accuracy graph

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

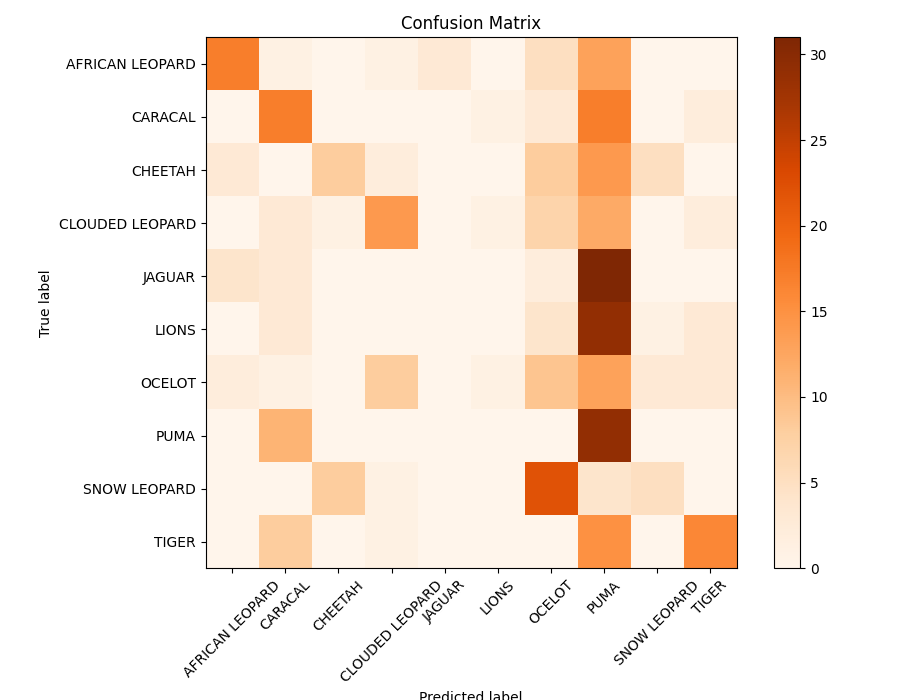


Figure 5 Confusion matrix

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

## Classification reports

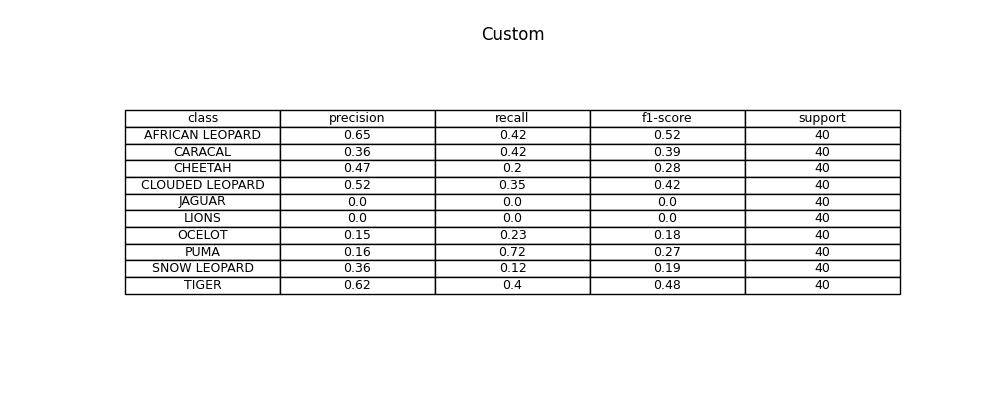


Figure 6 Custom model classification report

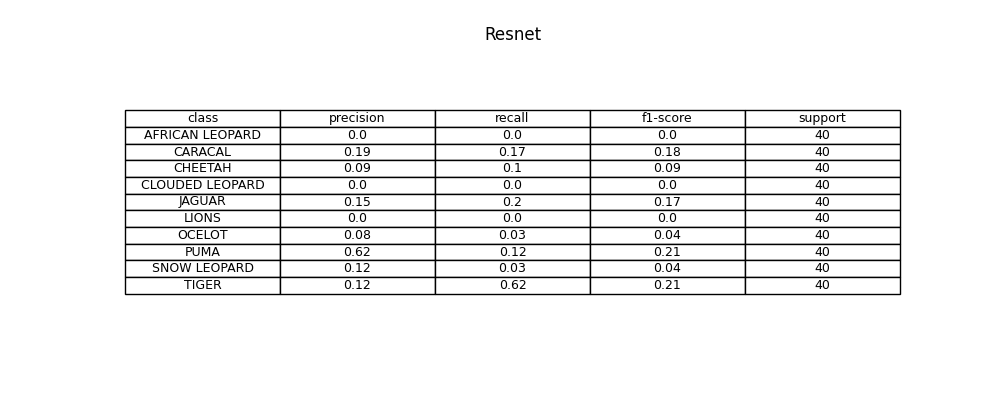


Figure 7 ResNet model classification report

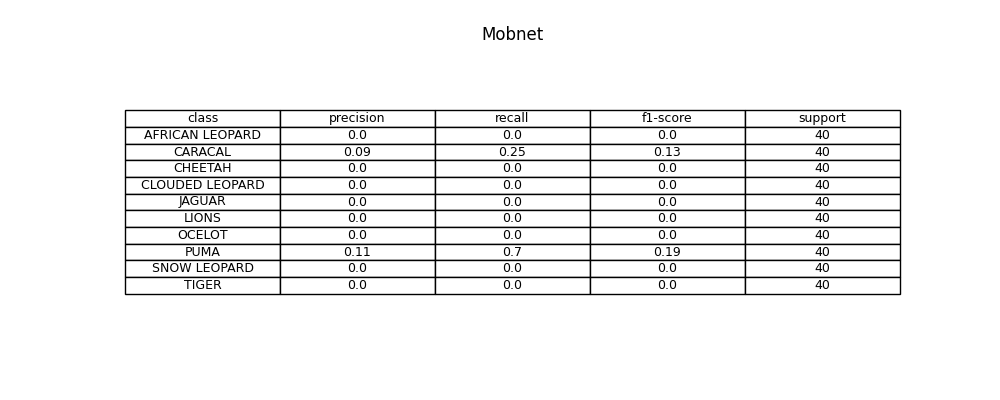


Figure 8 MobNet model classification report

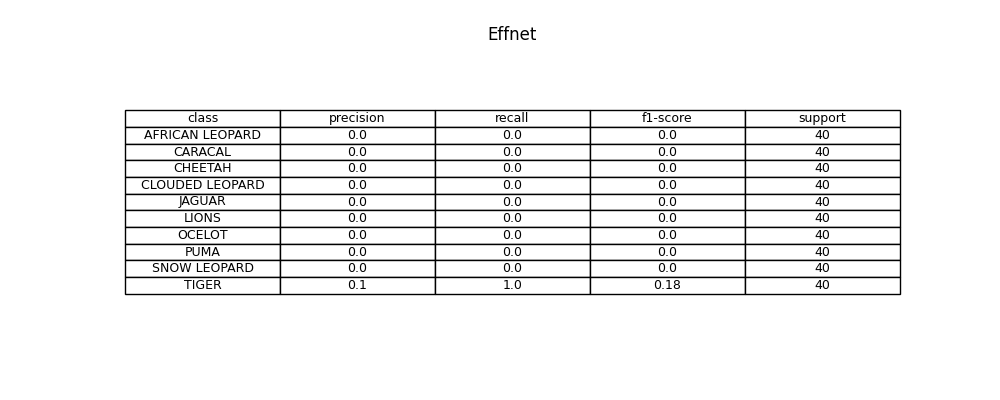


Figure 9 EffNet model classification report

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

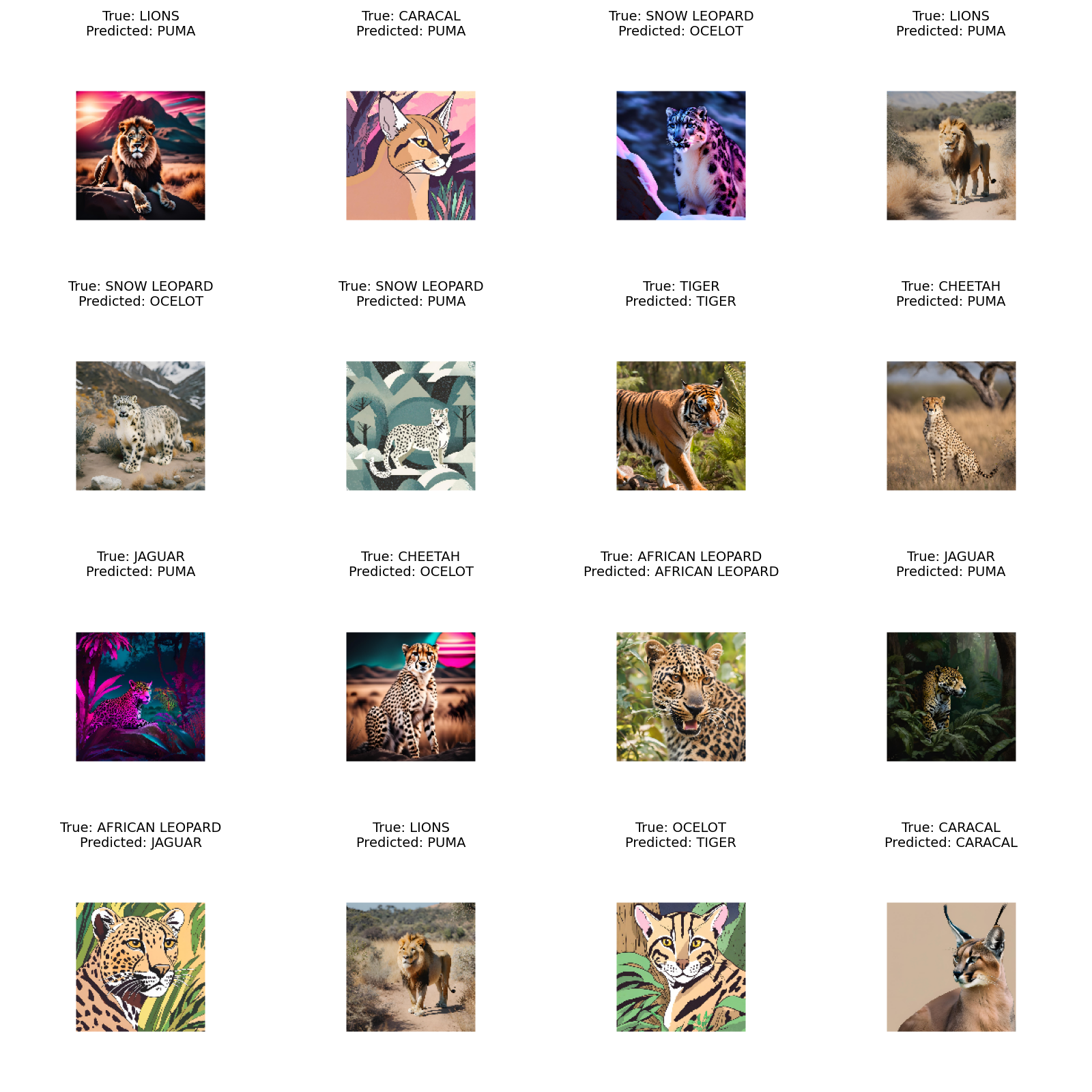


Figure 10 Test prediction visualization

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.