Detection and Classification using CNN and Fastr-RCNN Model

Zhaoyuan Wang

I. Introduction

USing some Deep Learning method to solve the problem brings us more accuracy and efficiency.

I-A. Task1

For task1, we use FasterRCNN, which is a popular deep learning method with high accuracy and high testing-speed compared to other methods.

I-A1. Task2: For task 2, The CNN model was trained and evaluated on the turtles and penguins dataset. The final result of CNN model demonstrate the effectiveness of CNNs' skill of solving image classification problems.

II. Data Preprocessing and Model Construction

The raw images were preprocessed to standardlize the input for the CNN model. The preprocessing steps involved resizing all images to a common resolution, also normalizing pixel values, and data augmentation to increase the diversity of the training data.

II-A. Sample images of turtles and penguins

II-B. Fast-RCNN Model Construction

The MobileNet-Large Backbone we used in the model is designed for efficient processing on to achieve a good trade-off between accuracy and the model size. Also it combines FPN to enhance the performance of object detection network. This image is referred from

Figura 1. Turtles and Penguins

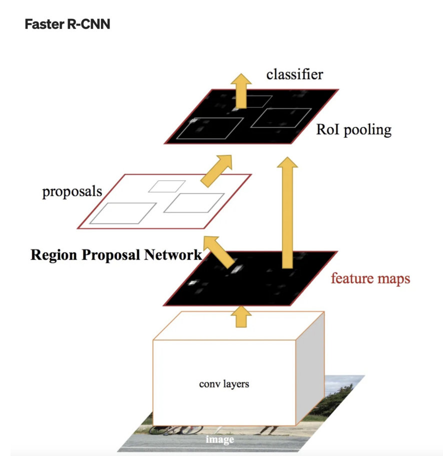


Figura 2. Model structure of Faster-RCNN

II-C. CNN Model Construction

My CNN model used several convolutional laters with ReLu activation for non-linearity. Also a max-pooling layer with size (2, 2) to reduce the spatial dimensions the feature maps by selecting the maximum value over a(2,2)region. Then I used a dropout layer to prevent over-fitting and a Flatten layer to flatten the data to one-dimensional vector. By using the last three fully connected layers, I mapped the flattened features to out final classification. By tuning the size of the epoch, learning rate and the batch sizes, the model can be improved further.

Model: "sequential"

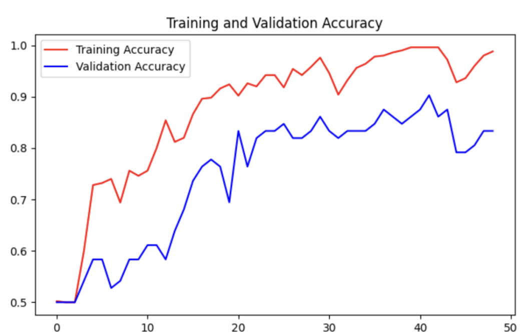
Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 198, 198, 32)	896
max_pooling2d (MaxPooling2D)	(None, 99, 99, 32)	0
conv2d_1 (Conv2D)	(None, 97, 97, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 48, 48, 64)	0
conv2d_2 (Conv2D)	(None, 46, 46, 128)	73856
max_pooling2d_2 (MaxPooling2D)	(None, 23, 23, 128)	0
conv2d_3 (Conv2D)	(None, 21, 21, 256)	295168
max_pooling2d_3 (MaxPooling2D)	(None, 10, 10, 256)	0
conv2d_4 (Conv2D)	(None, 8, 8, 512)	1180160
max_pooling2d_4 (MaxPooling2D)	(None, 4, 4, 512)	0
dropout (Dropout)	(None, 4, 4, 512)	0
flatten (Flatten)	(None, 8192)	0
dense (Dense)	(None, 256)	2097408
dense_1 (Dense)	(None, 128)	32896
dense_2 (Dense)	(None, 1)	129

Total params: 3699009 (14.11 MB)
Trainable params: 3699009 (14.11 MB)
Non-trainable params: 0 (0.00 Byte)

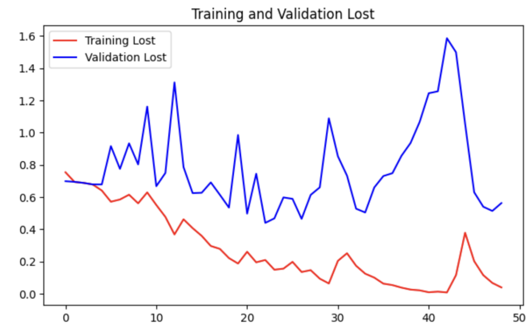
Figura 3. Model structure of CNN

II-D. *Model Training*

Model training After adjusting the parameters and the layers of the model, I obtained the accuracy of nearly 90.28 %, compared to traditional methods, the CNN model has many advantages. Below is the image of Testing and Validation Accuracy and Lost during training.



I monitored the performance of model during training, the accuracy continuously growing higher; at the same time, the Validation and training lost is decreasing with time, but after the 27th epoch, Lost curve started to grow rapidly, that's the time when we encounter the over-fitting problem. To solve this, I added an early stopped parameter, so as to achieve a better performance.



III. Model outputs and Conclusions

III-A. *Detection outputs using Faster-RCNN Model*

In conclusion, the general output is very precise, because it uses Region Proposal Networks (RPN) to propose candidate regions of interest, and then a Region-based CNN (RCNN) for classifying.Even when animals that are partially occluded or in cluttered backgrounds.

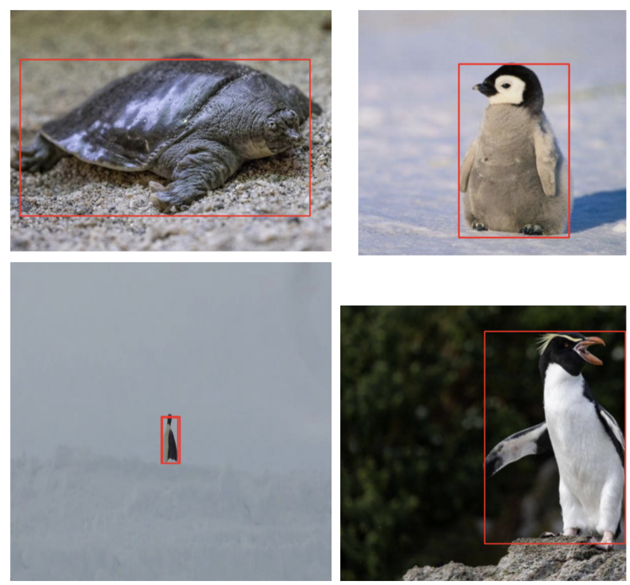


Figura 4. Detection Outputs of Faster-RCNN

III-B. *Classification outputs using CNN Model*

In conclusion, the CNN model employed for the classification of turtles and penguins has demonstrated remarkable performance with an accuracy of 90 %. This achievement highlights the effectiveness of deep learning techniques(CNN) in accurately discerning between penguins and turtles. Except for some extreme cases. The model's success holds significant potential for wildlife conservation and research, aiding experts in studying and preserving these animals' habitats and populations.

Figura 5. Classification Outputs of CNN

III-C. *Confusion Matrix*

According to the confusion matrix, the miss-classification rate of turtles is slightly higher then penguin, this is partly because penguins have a more recognizable colour pattern, makes it easier to be recognised.

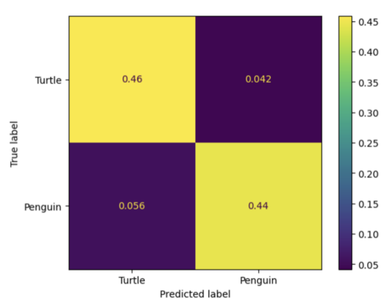


Figura 6. Confusion Matrix

REFERENCIAS

[1] https://towardsdatascience.com