Semantic Segmentation for Pet Detection Using Convolutional Neural Network

CVVR2023

Sanya, China

Chibuzo Valentine Nwadike

Artificial Intelligence for Computer vision and control,
Illinois Institute of Technology, Chicago, Illinois, 60616, United States

Scope

Categories and Subject Descriptors
Computer vision, Image processing, deep learning

Keywords
Pet detection, semantic segmentation, data augmentation

Abstract

This research project encompasses an in-depth exploration of semantic segmentation methods for pet detection using the Oxford Pets Dataset. The primary objective involves the development of a convolutional neural network model rooted in deep-learning principles, designed to achieve precise segmentation and detection of pets within images. The approach integrates advanced image processing techniques, leveraging deep learning methodologies, and dataset augmentation strategies to enhance pet detection accuracy substantially. The outcomes underscore the considerable potential of semantic segmentation in elevating the effectiveness of pet detection applications. This study offers promising avenues for practical integration in real-world contexts such as pet care and surveillance systems. The achieved advancements underscore the proposed technique's viability and contribute to the broader discourse on enhancing object detection through sophisticated segmentation strategies.

INTRODUCTION

Pet detection plays a vital role in various domains, including pet care and security systems. This project introduces an approach utilizing semantic segmentation to accurately identify and delineate pets within images. The Oxford Pets Dataset, a widely recognized benchmark for animal image analysis, is utilized for the development and evaluation of the model.

Proposed Method

The approach presented in this study encompasses the following key steps. :

1. Data Preprocessing and Augmentation:
The input images and segmentation masks from the Oxford Pets Dataset are pre-processed and normalized to ensure consistent quality and compatibility with the model. Augmentation techniques such as resizing and flipping are applied to enhance the dataset's diversity and robustness.

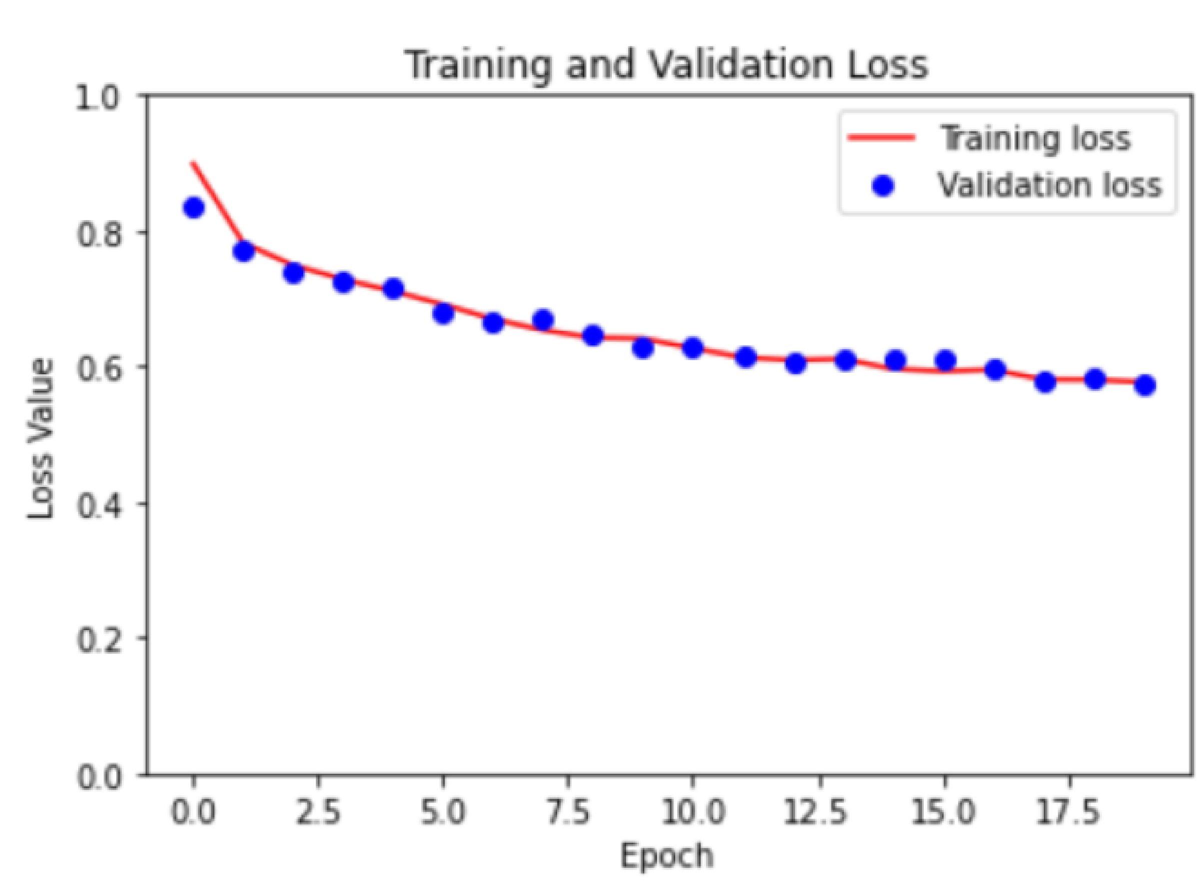
2. Model Architecture:
A Convolutional Neural Network (CNN) architecture was designed for semantic segmentation. The model consists of convolutional layers, transposed convolutional layers, and pooling layers, each tailored to capture intricate features and patterns crucial for accurate pet detection. The architecture is structured to handle varying scales and complexities within the pet images.

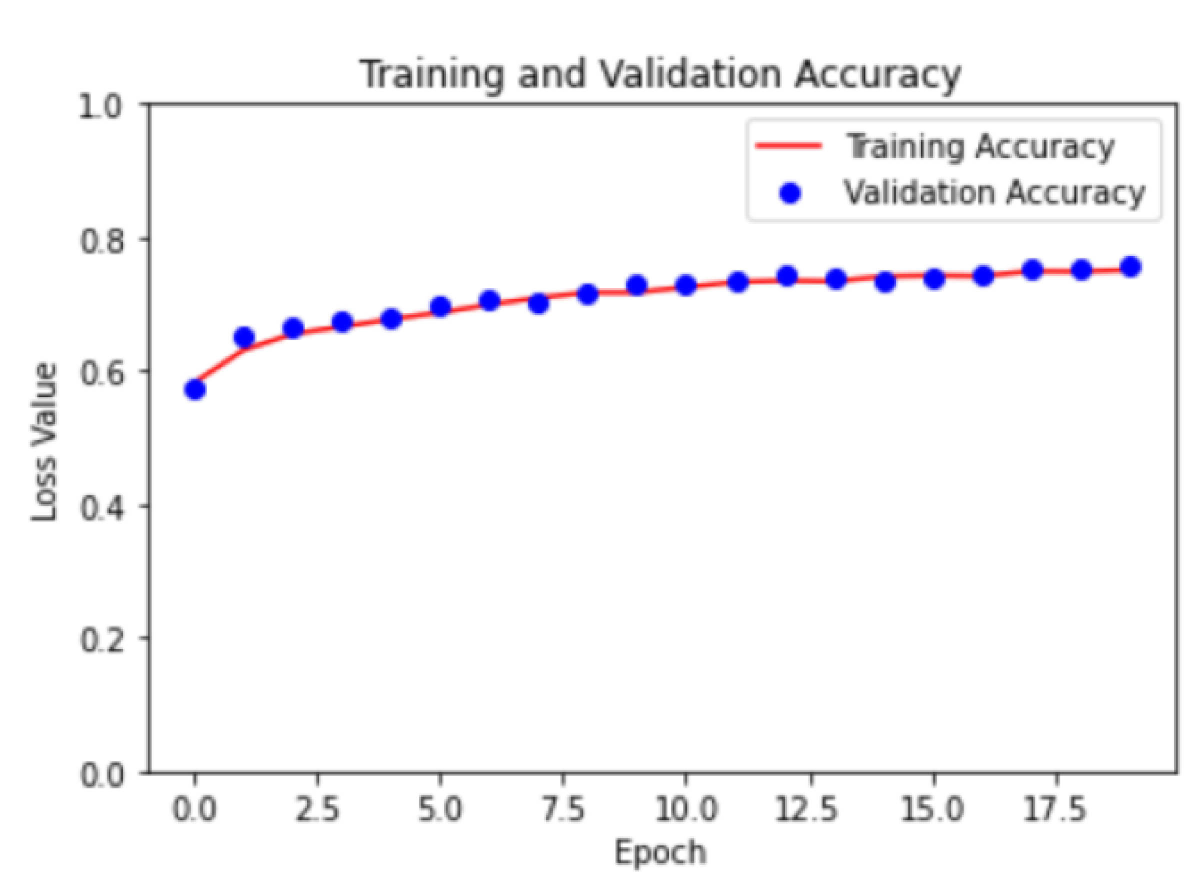
```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 64, 64, 16)        448

 conv2d_1 (Conv2D)           (None, 64, 64, 32)        4640

 max_pooling2d (MaxPooling2D  (None, 32, 32, 32)       0
 )

 conv2d_2 (Conv2D)           (None, 32, 32, 64)        18496

 conv2d_transpose (Conv2DTra  (None, 64, 64, 64)       36928
 nspose)

 conv2d_transpose_1 (Conv2DT  (None, 128, 128, 3)      1731
 ranspose)

=================================================================
Total params: 62,243
Trainable params: 62,243
Non-trainable params: 0
_________________________________________________________________
```

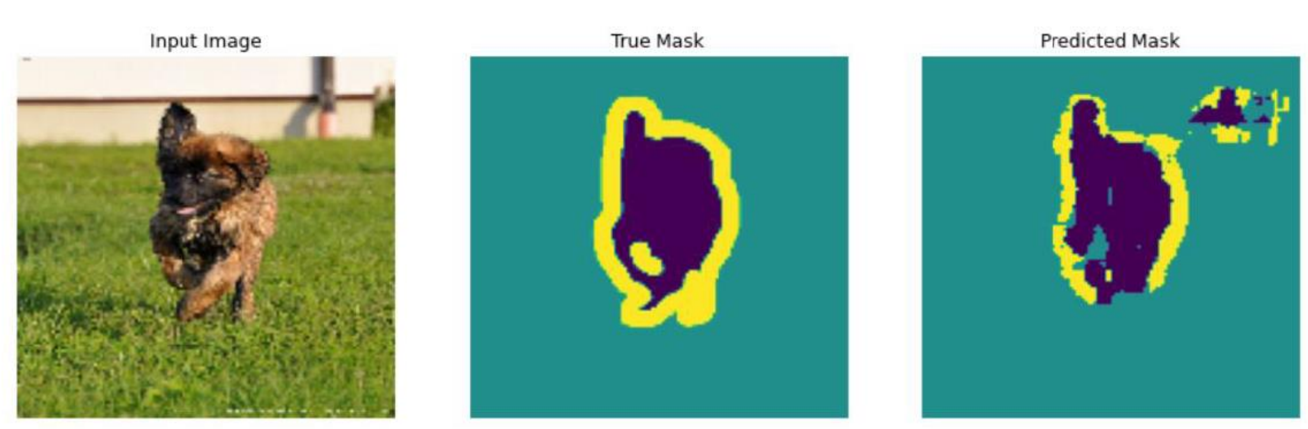
3. Training and Optimization:
The model is trained using the pre-processed dataset, with the loss function set as sparse categorical cross-entropy. The optimization process utilizes the Adam optimizer. The training process involves multiple epochs, with monitoring of training and validation accuracy and loss.

Experimental Results

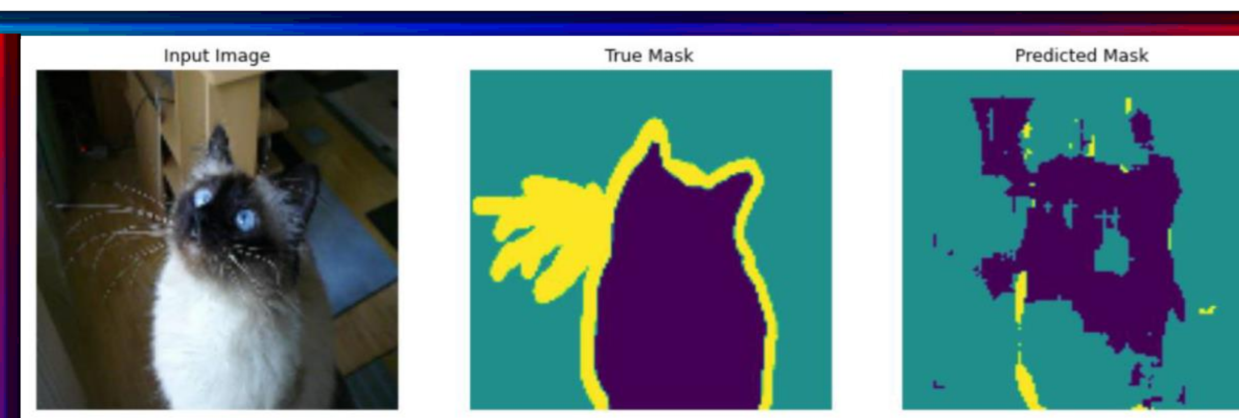
The performance of semantic segmentation-based pet detection model is evaluated comprehensively. Quantitative metrics, including accuracy and loss, are analyzed and compared across both training and validation datasets. The model's ability to accurately segment pets within images is demonstrated through visualizations of predicted masks.



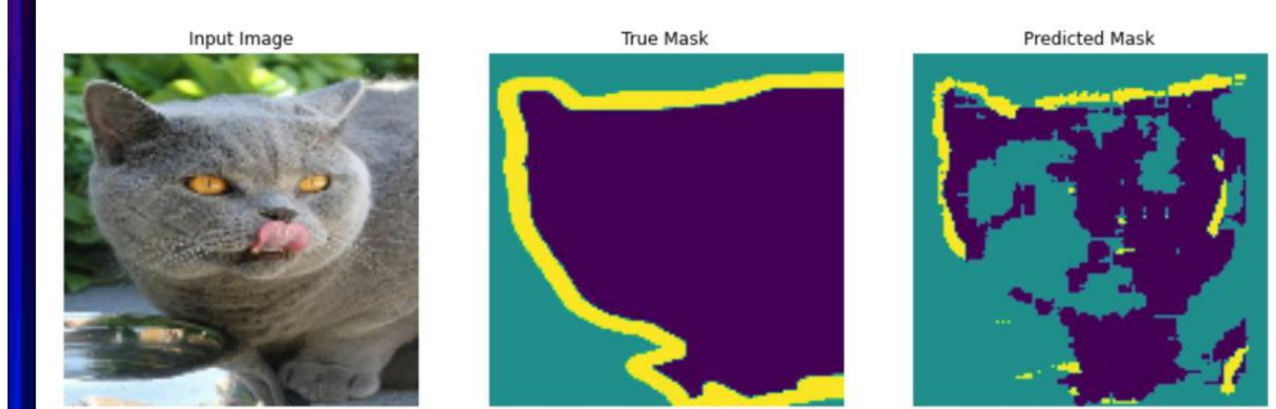




Output-1 with semantic segmentation label



Output-2 with semantic segmentation label



Output-3 with semantic segmentation label

DISCUSSION

The findings underscore the effectiveness of semantic segmentation in enhancing pet detection accuracy. The CNN architecture's ability to capture nuanced features contributes to the model's robust performance. The potential applications of the model extend to pet care, behaviour analysis, and surveillance systems, contributing to advancements in these domains.

CONCLUSSION

This research project presents an approach to pet detection using semantic segmentation, showcasing its potential to accurately identify and segment pets within images. Leveraging the Oxford Pets Dataset and advanced deep learning techniques, the model's accuracy and robustness stand as a testament to its capabilities. The success of this project paves the way for real-world implementations in pet care and surveillance, emphasizing the role of advanced image analysis in modern applications.

Ways to Improve Model Accuracy

Hyperparameter Tuning: Systematically search for optimal hyperparameters, such as learning rate, batch size, and optimizer settings, to fine-tune the model's training process and improve convergence.

Ensemble Learning: Train multiple instances of the model with different initializations or architectural variations and combine their predictions through ensemble techniques to boost accuracy and robustness.

Evaluation Metrics: Utilize a variety of evaluation metrics beyond accuracy, such as Intersection over Union (IoU), Mean Average Precision (mAP), and Dice coefficient, to gain a comprehensive understanding of the model's performance.

ACKNOWLEDGEMENT

I extend my sincere appreciation to the creators of the Oxford Pets Dataset, whose valuable contribution has been pivotal to my solo research project. Furthermore, the unwavering support and collaborative efforts of my mentors and colleagues throughout the course of this endeavour are deserving of recognition.