CMPE 258 Final Project, Milestone 1

Name of Leader: Khai Nguyen
Other team members: Suh Young Joon, Tin Pho

One sentence summary

We propose to build an app that can better recognize emotion of animals for people who own their pets and want to understand their pets more advanced and easier way.

Proposed application and its novelty

Application

The proposed project is an Emotion Detection Tool for Pets, which helps pet owners better understand the emotional state of their pets through analyzing images and live videos of their pets. Unlike traditional pet-monitoring tools that major in activity tracking or location, this application attempts to go further and deeper into the pet's emotional well-being.

Pet owners can take a picture or live record a video of their pets using a smartphone camera. The tool will identify and categorize emotions like happiness, anxiety, boredom, or discomfort, which will significantly enhance their knowledge about their pets' emotional state and needs.

This app can be developed further to serve pet owners that want to monitor their pets even when they are not at home. It can also act as an assistant for pet behaviorists and veterinarians by enabling an extra layer of information to assess a pet's emotional health. Users can take the gained knowledge to make better decisions for their pets: Either changing the environment or seeking professional help if needed.

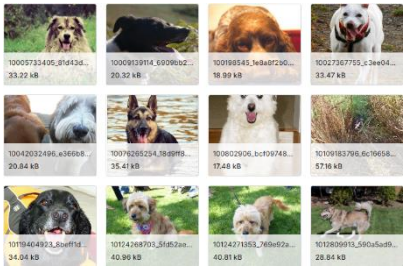
Novelty

While most of the existing solutions for pet monitoring focus either on activity tracking, location monitoring, or even basic health indicators, they mostly consist of wearable devices or demand manual input from owners. This manual method is often a hassle and requires a lot of materials to acquire the desired results. Also, while there are devices to measure movement or physical activity, they cannot usually capture the emotional condition of the pet, which is another important part of their daily lives. Currently, pet owners can only use their judgment about pet behavior based on guesses or empirical understanding about their pets' emotions. However, this

activity may be subjective and can possibly lead to misinterpretation by people who are first-time pet owners or the ones who do not do good research on pet behaviors. More professionally, there are behaviorists who analyze pets; however, these services are so limited that they may not be available or afforded by all. The novelty of the solution proposed lies in the fact that it focuses on detecting and classifying the emotional state of pets in an affordable way, which can totally be compressed into mobile applications. Using computer vision and deep learning techniques, this model will be able to give more objective and accurate assessments of pet emotions. This can build on already existing solutions, filling the gap in emotional monitoring and hence furnishing pet owners with real-time insights into the feelings of their pets.

In conclusion, this tool can add an ability to detect minute changes in facial expressions, body language, and behavior that may bring a deeper level of understanding and is also affordable to pet owners in comparison with the existing modes of approach.

Possible datasets, 25% of grade

Dataset name	**Contents with examples**	**Pros**	**Cons**
Dog Breed Image Dataset		Variety of Breeds, High-Quality Images, Benchmarking, Research and Development	Image Quality Variability, Limited Context, Potential Bias
Dog Emotions Prediction		Emotion Labels, Real-World Data, Interdisciplinary Research, Open Access	Imbalance in Data, Limited Context, Data Privacy, Generalization

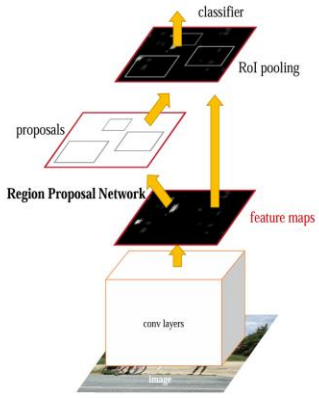
Possible models, 25% of grade

- Three models were investigated: 2 deep learning and 1 non-deep learning
- Two evaluation metrics were used: accuracy and amount of compute.
- Performance values used are mean and max.

Model name	Description with model diagrams/results	Classification accuracy (top 1)	Pros	Cons	Notes
Faster R-CNN	Model architecture is shown below. Two stages object detector	73.2 % (max)	High accuracy, end to end training, and flexible	Computationally intensive, may struggle with overlapping objects	Introduced in 2015
YOLOv11 (n-cls)	Model architecture is shown below. One stage object detector	76.86% (mean)	High performance and speed. It is also efficient and can scale	Struggle with small objects, might lead to lower accuracy than two stage detectors in complex scenarios	It is a newer technique, introduced in Sep 2024
SVM	Support Vector Machine	82% (max)	Effective in high dimensions, it is memory efficient and robust to overfitting	It is complex in training, need to choose the right kernel and sensitive to noisy data	First introduced in 1963, and widely used from 1992

Models description

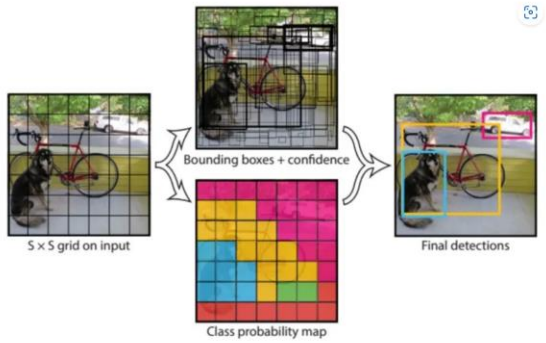
Faster R-CNN



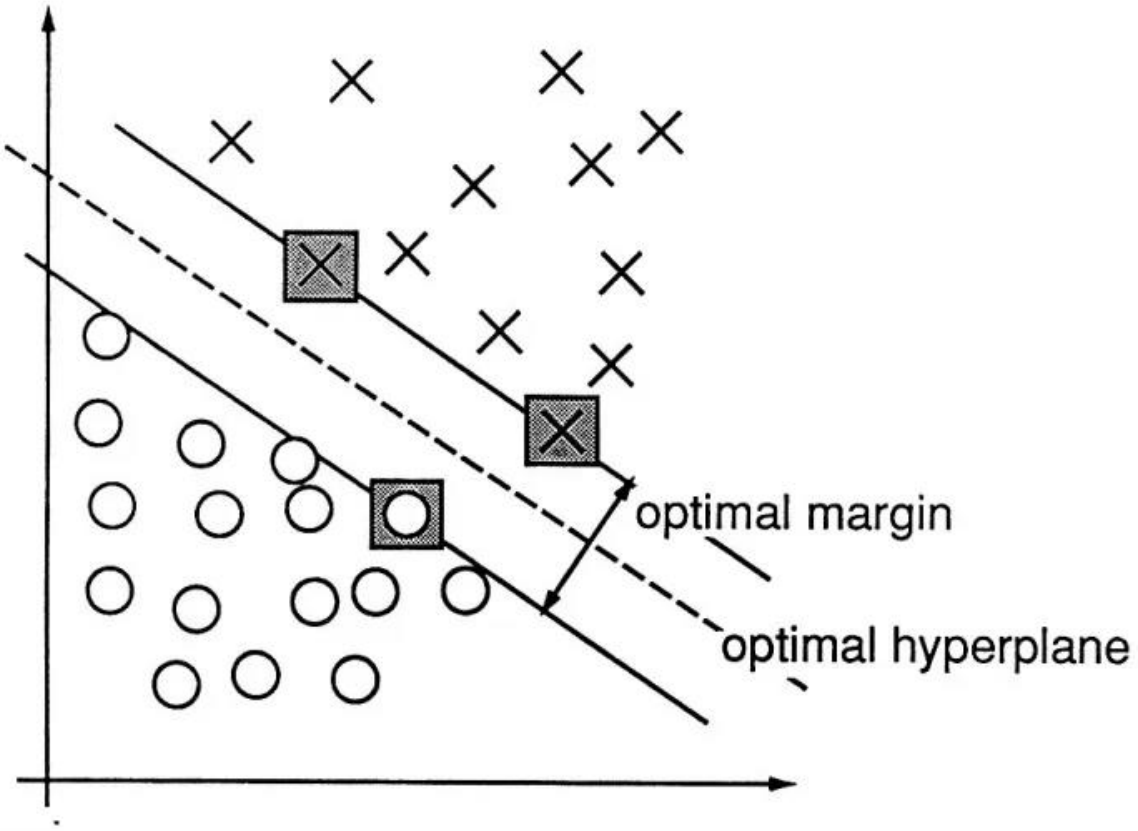
Picture 1. Faster R-CNN with region proposal network [1]

Faster R-CNN is a two-stage object detector. Region proposal network generates proposals, part of the image, that has potential objects. These proposals are then given to Fast-RCNN to classify.

Yolo
[YOLO for Object Detection, Architecture Explained! | by Sairaj Neelam | Analytics Vidhya | Medium]



YOLO is a one stage object detector, which has faster performance than Faster R-CNN. It classifies objects by dividing the object into S x S grid, then predicts bounding boxes and confidence scores for each grid.



SVM

SVM is not a deep learning model, it classifies object by finding the optimal hyperplane that separates data points in a high dimensional space.

[Support Vector Machine Explained. Theory, Implementation, and… | by Zixuan Zhang | Towards Data Science]

Summary of work plan, 20% of grade

In terms of dataset, we will use Chest X-Ray images from Kaggle to train our models.

For models, we will modify SVM, YOLO and Faster R-CNN and will compare their performance and training and testing.

We will look at different ways to improve the models such as hyperparameter tuning and other ways to improve the models.

References

[1] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Trans. Pattern Anal. Mach. Intell. 39, 6 (June 2017), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031