🔍Bridging the Gap: Pet Facial Expression Recognition for Enhanced XR Human-Pet Interactions

🤔Motivation: what problems are we tackling

In extended reality (XR), the integration of pets into virtual spaces is has created for a wide range of new possiblities. However, there's still a significant void in understanding the emotianal nuances of pets in XR. We will try to enhance the technology by developing a model that (a) identifies between two different pet types (cats and dogs) and (b) classify between different facial expressions (happy, sad, and angry)[1]. Additoinally, the article [2] has presented the importance of accurate recognition and assessment of pain in animals that is crucial in clinical contexts for pain management and welfare assessment, especially in the veterinary clinical studies.

Motivated by the growth of XR based technologies and the value of pets in human mental health, we will aim to integrate pets into XR. This will not only accommodate the preferences of the user in question but also enhances the overall immersive experience in human-pet interaction.

Potential use case

Our model could find it's application in remote pet monitoring for healthcare assessment. For instance: by using trail cameras biologists, researchers, or even hobbyist could remotely look at the emotial state of an animal; conclude distress; behavioural change and so forth. We think this model has huge potential, especially when this model could get extended to multiple animals and multiple emotions.

💡 Solution

Pre-trained model: Using a pre-trained CNN regression as base model (e.g. ResNet algorithm applied in the master thesis of Joerie). The CNN models have shown their success in image classifcation to capture feature vectors as explained in the Machine Learning (ML) Course provided by Faculty of Engineering Technology at KU Leuven.

Custom model: Using ML techinques/principles from the course we can make a customized CNN architecture with added convulutioonal layers and pool layers for feature extraction and spactial recudtion respectivily

🧮 Methods

Baseline model: One requirement for this project is to develop based on the taught/used code snippets. In our case, as the emotion classification on top of the cat and dog face recongition is a multicalss classification problem and that the nature of emotion having a zero state where there is no emotion and an continuous range of values, an activation function of **Rectified Linear Unit (ReLU)** in the **Neural Networks** would suit the purpose of the classifier better than of sigmoid. **sigmoid** is best for on/off binray situations. **The ReLU provides a continuous linear relationship**. Additionally it has an **'off' range** where the output is zero. The **"off"** or disable feature of the ReLU activation enables models to **stitch together linear segments to model complex non-linear functions**.

ReLU Activation Definition $$a = \max(0,z)\tag{1}$$

```
def relu(z):
    return max(0, z)
```

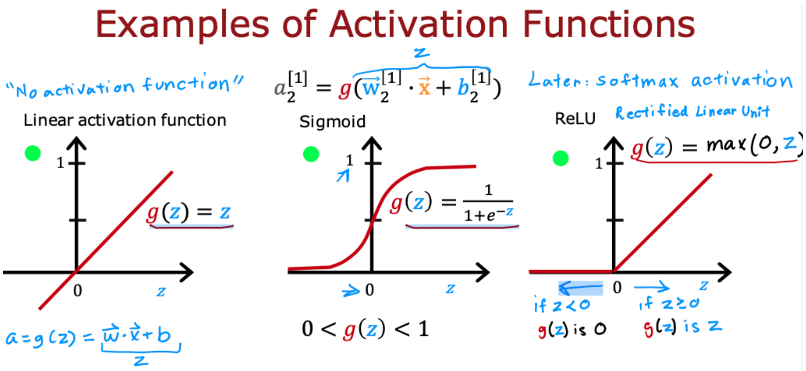

Fig. 1: Sigmoid and ReLU activation function[3]

`![Alt text](image.png)`

Softmax Regression: Given that one output is selected as the predicted answer out of N outputs yielded by the multiclass neural network, softmax function is exploited to process the N outputs as a **linear vector z** and converts **z** into a **probability distribution**. Each output after softmax application will range $\in[0,1]$ and will sum to 1. They can be interpreted as probabilities.

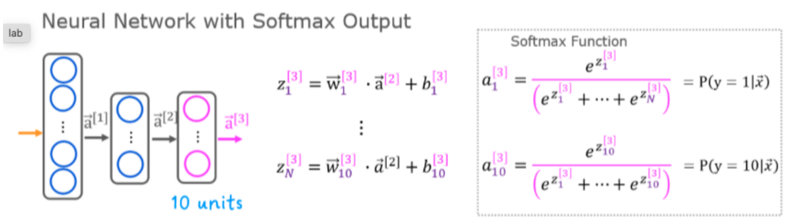

Fig. 2: Coursera: Advanced Learning Algorithms[3]

The softmax function can be written: $$a_j = \frac{e^{z_j}}{ \sum_{k=0}^{N-1}{e^{z_k} }} \tag{2}$$

Where $z = \mathbf{w} \cdot \mathbf{x} + b$ and N is the number of feature/categories in the output layer.

```
def softmax(z):
    ez = np.exp(z)       #element-wise exponenial
    sm = ez/np.sum(ez)
    return(sm)
```

Optimization: Adaptive Moment estimation (Adam): ADAM is an optimization algorithm that by holding the core principle of **maintianing a moving avergae** of both the 1) **gradients** and 2) the **squares** of the gradients of the loss function, computing each firstly the gradient of the loss function for a batch of data, update the bias, then again compute the squares of the gradients and update second time the bias to compensate and maintain a certain average, or momenteum.
In our case, gradient optimization using ADAM to iteratively and **adaptively** update the learning rate of the loss function can ensure us a steady progress towards the converging point (minimum) when working with **sparse and/or noisy gradeint problems**.

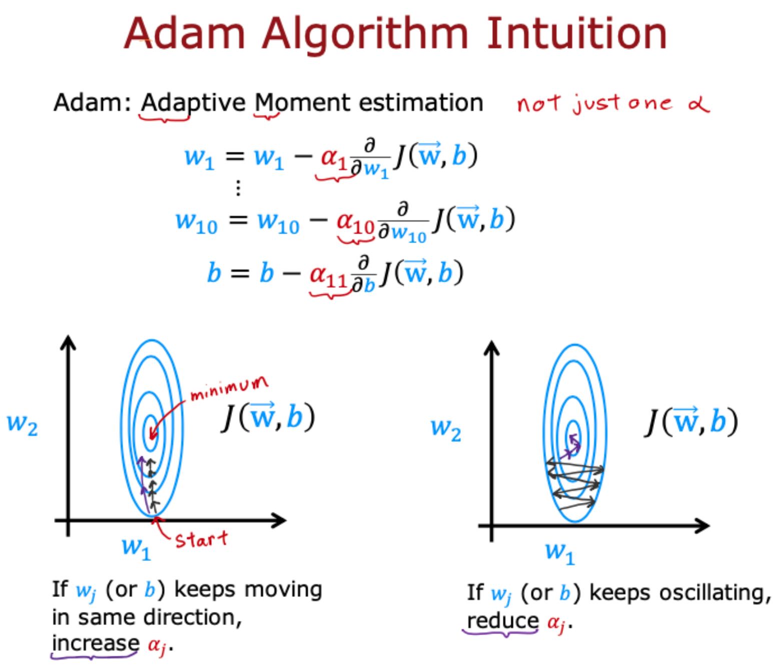

Fig. 3: ADAM Algorithm Intuition[3]

Proposed Model will apply cronologically the following steps

1. Load data by storing each image path (e.g. "☑️ = "/images/happy/dog15.png") in a list and it's corresponding label in another (☑️ = "happy")
2. Transform the lists into a dataframe
3. Exokiratiry Data Analysis (EDA) and analyze data to get more insights (just like we did in all the labs, explore data get familiar with it)
4. Create labels for each emotion category
5. Normalize pixel values to range [0, 1] for ReLU
6. encode the labels
7. Make train, test, and validate sets
8. Make Data Generator (DG) for Train, Test and Validate set. We can use Tensorflow Generator for it.
9. Build the CNN regression model: (inject lab code as baseline)
    - add layers of ReLU and Softmax.
    - compile with Adam optimizer.
10. Calculated class weights and pre-train the model with computed class weights.
11. Check accuracy by evaluating the pre-trained model on test data.
12. Load test images, reshapte the iage to match the model input shape, map.
13. Evaluate the model prediction result on validate data by plotting the predicted label, images, and original images.

📊 Dataset and Features

Kaggle 🐶 Pet's Facial Expression Image Dataset😺
This dataset contains 1000 face images of various pets, such as dogs, cats, rabbits, hamsters, sheep, horses, and birds, which will be used in our projects for emotion classification training and testing

Features:
The images capture the diversity of expressions these animals can display, such as happiness, sadness, anger etc, and provides a **classified and labeled** emotions features.
The dataset can be processed using machine learning techniques to learn an algoirhtm for pet species identificatoin and gain insights into pet emotions and personalities, enabling the creation of projects with pet face images whlie contributing to pet face recognition research and animal welfare.

Kaggle Animal faces
If one were to evalute subjectively the performance of our proposed model, this dataset, also known as Animal Faces-HQ (AFHQ), consists of 16,130 high-quality images at 512×512 resolution.

Features:
There are three domains of classes, each providing about 5000 images. By having multiple (three) domains and diverse images of various breeds per each domain, AFHQ sets a challenging image-to-image translation problem. The **provided classes labels** are: Cat, Dog, and Wildlife.

📑 Results/Discussion

Preliminary Result Pending (estimated time of arrival (ETA): 2023.11.29)

📚 References

📑 Literature

[1] *Mao, Y., Liu, Y.* ***Pet dog facial expression recognition based on convolutional neural network and improved whale optimization algorithm.*** *Sci* Rep 13, 3314 (2023). DOI 10.1038/s41598-023-30442-0

[2] *Feighelstein, M., Shimshoni, I., Finka, L.R. et al.* ***Automated recognition of pain in cats***. *Sci* Rep 12, 9575 (2022). DOI 10.1038/s41598-022-13348-1

[3] *A. Wu, "**Online courses & credentials from top educators. join for free**," Coursera,* Website Visit (accessed Nov. 12, 2023).

💻 Related Work

- CNN | Beginners | 🐶 Pet's Expression Recognition
- Facial Landmark and Image Morhphine: Species