5COSC009C.2 Software Development

Group Project

Literature Review

Module Leader

Mr. Banuka Athuraliya

Team Name

DEBUG PIRATES

Team Member	Registration ID
H. Thiwanki Dias Hettiarachchi	2019741 / w1790191
Srimali Udayangani	20191114/ w1790361
Ahmed Isthaffa	2018497 / w1761936
K.N. Kumar	2018432 / w1761759
Akthab Bifaz	2018060 / w1761866
H.M.A.D. Herath	20191180 / w1790023

Table of Content

Contents

List of figures

List of Abbreviations

Abbreviation	Explanation
AI	Artificial Intelligent
SVM	Support Vector Machine
LSTM	Long Short-Term Memory
DBN	Deep Belief Network
RNN	Recurrent Neural Network
HRV	Heart Rate Variability
CNN	Convolutional Neural Network
ML	Machine Learning
DCNN	Deep Convolutional Neural Network
SCR	Silicon Controlled Rectifie
BVP	Boundary Value Problems
ICM	Independent Chip Model
NLP	Neural Language Processing
MAN	Metropolitan Area Network
Map	Mean Arterial Pressure
QRS	Quick Reaction Section
GSR	General Statutory Rules
DEAP	Diagnostic Evaluation of Articulation and Phonology

1. Chapter Overview

This chapter is the chapter of the Literature Review and the main aim of this chapter is to analyze the existing knowledge of our product. Our product is mainly focused on to doggy lovers because using this Ooggy doggy product they can identify their puppy or a dog's emotions easily and fast. For these purposes, we will look at the latest ideas, the rules on the topic we have Chosen, get an idea of the current technologies we are using in order to implement our emotion recognition system for dogs. Then it will discuss the improvements of previous works and current structures, advantages and disadvantages of our technologies and Also, this chapter explains our product's new features.

2. Research

2.1 Review of Related Researches on domain

This part discusses the literature review of the research topic that is being addressed. This part is aligned in a way that it focuses on the important aspects of the research to be understood with regard to previous studies. There are limited number of research on the based topic, therefore referring to similar studies.

2.1.1 What is emotion?

Emotion is the key for judgments in the human world. As Solomon, (2019) stated in his article emotion is the gateway of a complex conscious experience which is the reflection of a personal significance of an event or a state. Further clarifying that emotions are those which create the sense of judgment in men with varied expressions such as “Anger, Pity, Happy, Fear and so on”. But emotions are not always expressed and physically visible through facial expression, they can also be invisible.

2.1.2 Humans emotion vs dog’s emotion

Humans have been evolving with lot of emotions expressed or unexpressed but duly configured by other humans. It is not easy but possible for most of the humans to read another humans emotion. However, an emotion of a pet (dog) companion domestically grown with human interaction can be misread by its owner itself. A dog’s emotion is mostly misinterpreted by their owner in certain conditions where they confuse with dog’s inability to express itself (Correia-Caeiro, Guo and Mills, 2020).

Similar to human’s facial expression, dog’s expression is also fixated mostly on internal facial expressions that is mostly dealt with facial elements. In various studies by Somppi et al., (2013) dogs’ expressions are limited as they deliver negative expression through eyes and mouth while positive expressions are exposed by forehead, while in some other studies he has stated that negative expressions are mostly cheeks and sidelines while positive expressions come from eyes and mouth muscles.

2.1.3 What is data science?

This is the science behind the use of complex tools in the information technology. Using various tools, machine learning principles and algorithms to interpret hidden patterns from huge pile of

raw data. This can also be used as tool to discover and create new patterns to emerge new discoveries (Hemant Sharma, 2017).

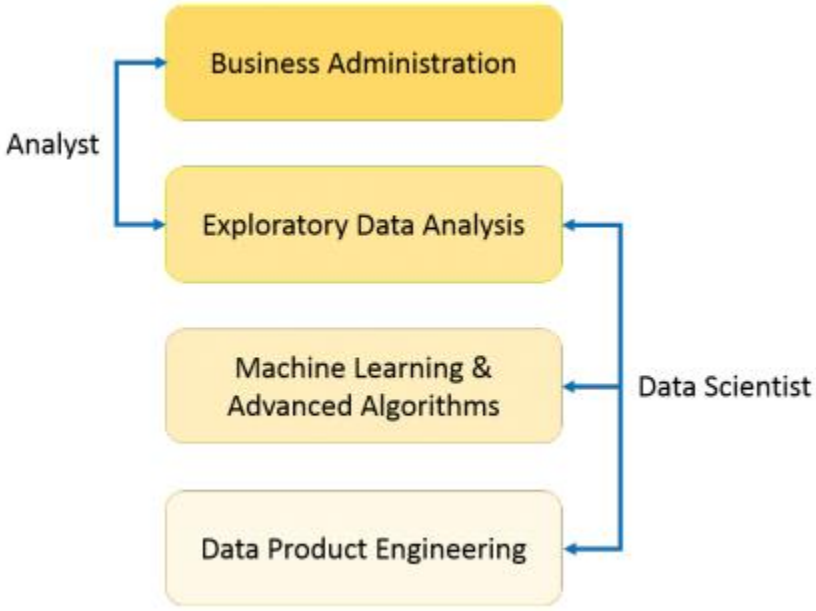

Figure 1: What is Data Science Source: Edureka, (2017)

2.1.4 What is Machine learning

Machine learning is the way of predicting patterns and links between a group of data. That process helps us understand, extract and learn useful patterns. Thus, helping us to create new discoveries from the emerged patterns (Srihari Sasikumar, 2017).

2.1.5 Facial image Emotion detection using machine learning

In a previous study done by Maglogiannis, Vouyioukas and Aggelopoulos, (2007) using humans facial emotions, they have stated that human emotions are widely expressed with eyes and cheeks muscles. Their detection algorithm uses skin muscles as a base to detect facial changes that eventually categorizes the patterns and match them with previously input data set which contains basic emotions. This process multiplies the accuracy and tendency for various human emotions with machine learning approach.

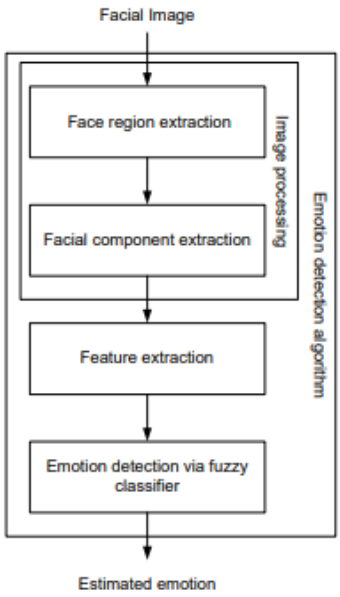

Figure 2: Overall Procedure of Emotion Detection. Source: (Kim, Joo and Park, 2005)

The algorithm is composed of three main stages as shown in figure 2. Image processing stage and facial feature extraction stage, and emotion detection stage. In image processing stage, the face region and facial component is extracted by using fuzzy color filter, virtual face model, and histogram analysis method. The features for emotion detection are extracted from facial component in facial feature extraction stage. In emotion detection stage, the fuzzy classifier is adopted to recognize emotion from extracted features. It is shown by experiment results that the proposed algorithm can detect emotion well (Kim, Joo and Park, 2005).

2.2 Research Gap

Humans and dogs have been in domestic relation for centuries and even more. Understanding them and their patterns are an important thing in today's world. It is very important for human beings to understand the activity pattern of the dog, and its emotional behavior. A wearable, sensor-based system is suitable for such ends, as it will be able to monitor the dogs in real-time. However, the question remains unanswered as to what kind of data should be used to detect the activity patterns and emotional patterns (Aich et al., 2019).

Past studies have been done on detecting emotions from humans using machine learning approach. Need for such method to understand dog's emotion is also very crucial as learning human emotions. Generally, a pet is a companion for a human where they interact with their pet with all those emotions. But, as stated above with the decreasing interaction with the pets it has been difficult for owners to understand their pet in times of need, thus creating a vast distance between each other which lead to loss of relationship. So, here we are focusing on

developing a relationship with the help of technology to understand the emotions of a pet for the owner to understand their emotions.

Therefore, we propose a system which is based on the data collected from dogs, including varied breeds of various sizes and ages, and both genders an using machine learning classification techniques to automate the detection of various emotional expressions of dogs.

3. Research on approaches and Techniques

3.1 Algorithms and Technologies

There are two main strategies for emotion detection: facial recognition and semantic analysis. Facial detection analyzes facial expressions in video and photos, detecting micro expressions which determine common emotions such as surprise, joy, anger, sadness, disgust and more. These powerful algorithms can detect expressions by plotting points on a face and reading their relationships to one to another, with the help of facial databases. The second type of machine learning-based sentiment analysis-and one likely encountered online is semantic analysis includes algorithms that detect emotion in language, whether keywords in a text are positive or negative in connotation, through which an overall tone emerges. Algorithms may detect multiple examples of tone in a single statement, offering a comprehensive look into what the speaker or writer is thinking. The algorithm is composed of three main stages: image processing stage and facial feature extraction stage, and emotion detection stage. In image processing stage, the face region and facial component is extracted by using fuzzy color filter, virtual face model, and histogram analysis method. The features for emotion detection are extracted from facial component in facial feature extraction stage. In emotion detection stage, the fuzzy classifier is adopted to recognize emotion from extracted features. It is shown by experiment results that the proposed algorithm can detect emotion well.

Some studies for the affect recognition of have implemented supervised classification approaches such as k-Nearest Neighbor, and Support Vector Machine (SVM). The researchers defined keyword to valid the user's emotional response through the valence and excitation model. On the other hand, Deep learning approach applies non-linear transformation to physiological signs for the detections of features of dog's emotional behavior. In this context, CNN techniques have been used for the automatic extraction of SCR and BVP features and 70 to 75% accuracy results have been obtained in the prediction of emotion (relaxation, anxiety,

excitement, and fun). Other investigations validated the performance of affection models with deep learning using the multimodal DEAP database and adopted a multiple-fusion-layer-based ensemble classifier of stacked autoencoder (MESAE) framework, to extract the physiological features that were merged into an SAE network. The accuracy results in arousal and valence were 0.83 and 0.84 respectively. Regarding to semi-supervised learning methodologies SAE was integrated with Deep Belief Network (DBN) using a Bayesian inference classification based decision fusion method, results of arousal were obtained in 73.1% and valence in 78.8%. In they defined a hybrid model composed of a CNN and a Recurrent Neural Network (RNN). As a requirement for the sequential processing in the CNN, the features were extracted and the prediction was made in the Long Short-Term Memory (LSTM) unit of the RNN. This model obtained an accuracy of 74.1% for arousal and 72.1% for valence. The models based on CCN and DNN showed better results in the affective classification when using the image domain of the EEG signals .

. In this case, the set of convolution formula is given by,

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau$$

Figure 3: convolutional formular

(f *g)(t) = functions that are being convoluted

t = real number variable of functions f and g

g (T) = convolution of the function f(t)

T' =first derivative of g (tau) function

convolution is a mathematical operation on two functions (*f* and *g*) that produces a third function (f*g) that expresses how the shape of one is modified by the other. The term convolution refers to both the result function and to the process of computing it. It is defined as the integral of the product of the two functions after one is reversed and shifted. And the integral is evaluated for all values of shift, producing the convolution function.

Some features of convolution are similar to cross-correlation for real-valued functions, of a continuous or discrete variable, it differs from cross-correlation (f*g) only in that either $f(x)$ or $g(x)$ is reflected about the y-axis; thus it is a cross-correlation of $f(x)$ and $g(-x)$, or $f(-x)$ and $g(x)$. For complex-valued functions, the cross-correlation operator is the adjoint of the convolution operator.

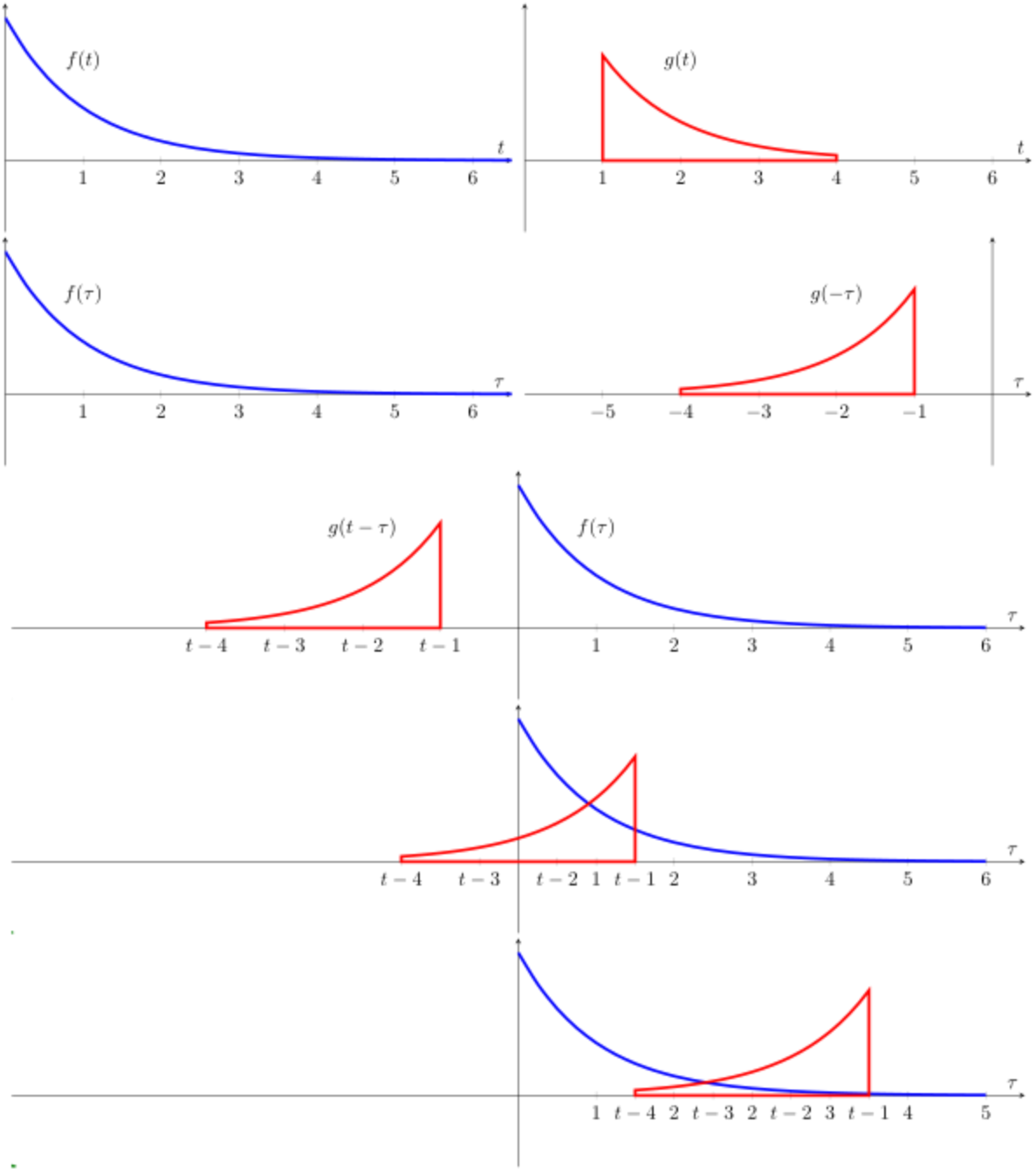

Figure 4: visual explanation of convolution

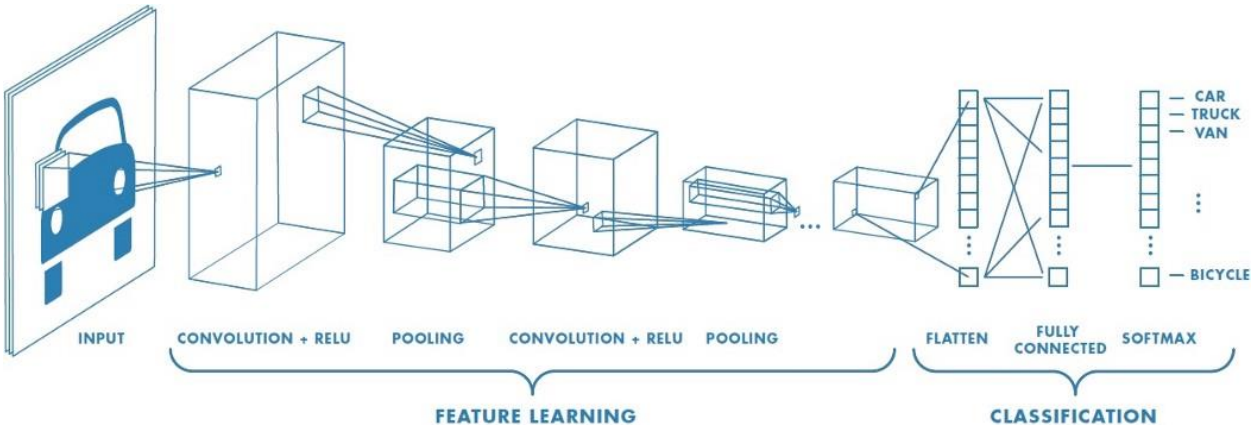

Figure 5: Step by step of convolutional Neural Network (Deep learning)

A. MACHINE LEARNING

As a previous step to the features extraction of the physiological signals, the detection of peaks of the ECG and GSR signals is performed, because the emotions generate significant changes in these segments. The Heart Rate Variability (HRV) analysis is an affective diagnostic tool to determine the beat to beat interval (RR interval) . The values between a RR interval correspond to the time between two peaks R, which is calculated through a standard wave of the QRS complex. The ECG signal is transformed with the PanTomkins QRS detection algorithm proposed in . The signal is filtered to reduce the noise with cutoff frequencies of 0.5 and 15 Hz and uses an adaptive threshold for the detection of the QRS complex . Similarly, the GSR signal is preprocessed using bandpass filters to reduce noise with cutoff frequencies of 0.05 and 19 Hz . Then it is resampled with a digital phase filter of 10 Hz. During SCR peak detection a standard method is used that identifies the max, min and offset indexes of the signal GSR . So, the threshold of the amplitude is determined and the features between SCR peaks are calculated.

DEEP CONVOLUTIONAL NEURAL NETWORK

Deep learning is an area of machine learning based on algorithms and techniques for modeling high-level abstractions in datasets , such as the patterns recognition in images, text or emotions. The learning levels take as input the results of the previous levels, which are transformed into insights, to train and validate the classification model. The DCNN architecture proposed for the emotion detection system was adapted from the work of Pyakillya et al, with the Keras framework . The DCNN involves a sequence of CNN layers and pooling layers to

automatically extract features from the physiological signals. Fully connected layers are located in front of CNN, operate on all nodes and are used to predict the affective state.

PROPOSED FACE DETECTION ALGORITHM

The proposed face detection algorithm contains two major modules: (1) face localization for finding face candidates; and (2) facial feature detection for verifying detected face candidates. The algorithm first detects skin regions that possibly contain a human face. The skin detection algorithm is a segmentation technique that first transforms the image from the RGB to YCbCr color space.

The skin detection algorithm is based on statistical imageprocessing model using Bayesian estimation [2]. A segmentation technique has to be adopted to track the skin regions. We used a Markov Random Field and MAP image segmentation [5], for assuming an image which contains a skin region that can be divided into two classes (skin vs. non-skin). The discrete steps are depicted in Fig. 1. An MRF is used for skin detection because it is a non casual model, it is strongly attached to an isotropic behavior and MRFs use only local dependencies. This feature makes them very flexible, while in conjunction with MAP, this approach may be more accurate than many other estimation techniques and statistic models.

Figure 6: facial recognition of dogs

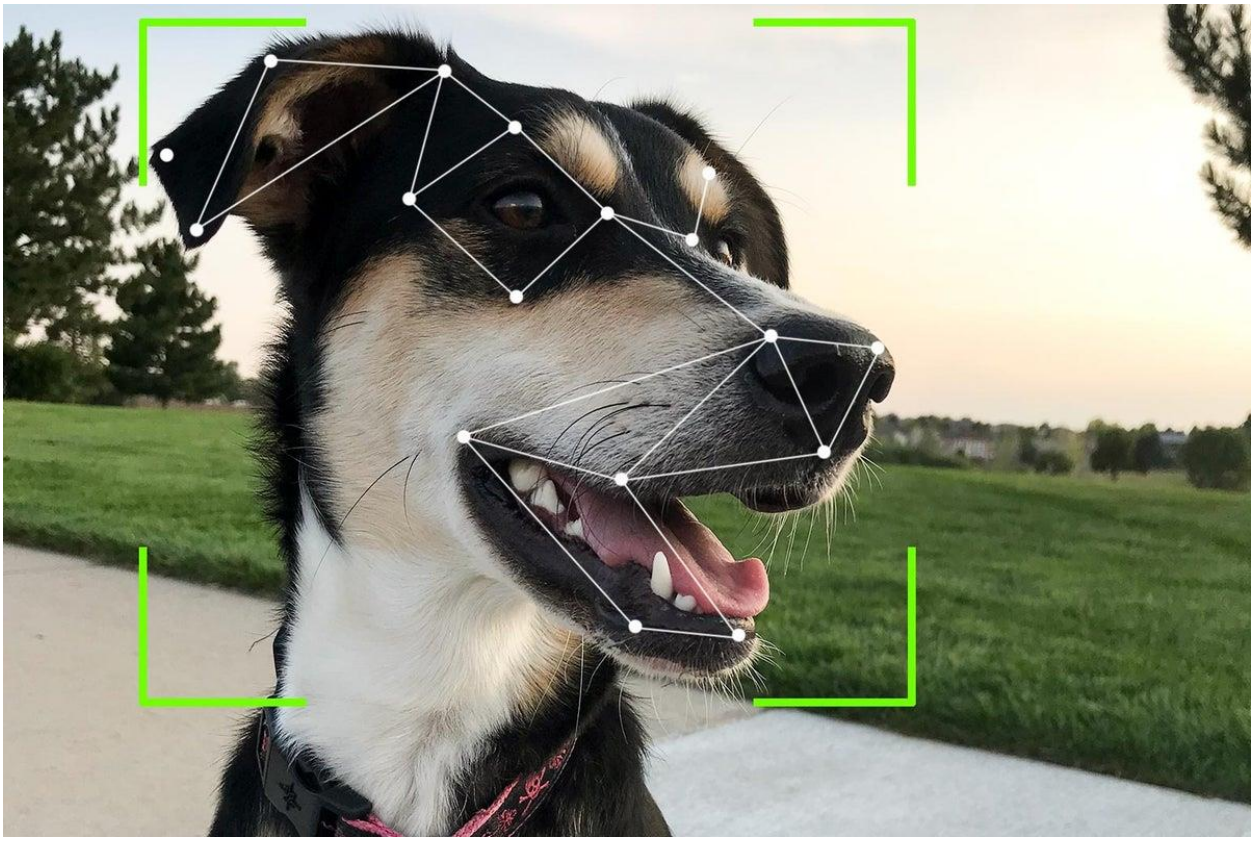

Figure 7: Facial recognition of dogs

Past studies have been done on detecting emotions from humans using machine learning approach. Need for such method to understand dog's emotion is also very crucial as learning human emotions. Generally, a pet is a companion for a human where they interact with their pet with all those emotions. But, as stated above with the decreasing interaction with the pets it has been difficult for owners to understand their pet in times of need, thus creating a vast distance between each other which lead to loss of relationship. So, here we are focusing on developing a relationship with the help of technology to understand the emotions of a pet for the owner to understand their emotions.

Therefore, we propose a system which is based on the data collected from dogs, including varied breeds of various sizes and ages, and both genders an using machine learning classification techniques to automate the detection of various emotional expressions of dogs.

3.2 Advantages and disadvantages of algorithms and technologies

Advantages of the technologies used

- CNNs provide better performance in image resolution as compared to traditional sparse representation because it possesses higher representation capability [21]. So there is no need of feature extraction it automatically detects the important features without any human supervision
- Moreover, if CNN is trained once, it's weights can be learned by another network by transfer learning [25].
- K- nearest neighbor algorithm requires no training before making prediction new data can be added seamlessly
- An advantage of the decision tree over other methodologies, such as neural networks, is that it can provide understandable English-like rules or logic statements. For instance, if the gray level of a given pixel ranges between 180 and 240 and its entropy is greater than 0.5, then it is a pixel of interest, POI.[26]

Disadvantages of the technologies used

- K nearest neighbor need feature scaling (standardization) and normalization before applying KNN algorithm to any dataset . If we don’t do so, KNN may generate wrong predictions

- Does not work well with large dataset. In large dataset the cost of calculating the distance between the new point and each existing point is huge which degrades the performance of the algorithm

- One of the main disadvantage of using CNN is that it may take longer to train data[28]

- Another important disadvantage is the need for large data sets (i.e. hundreds or thousands of images), and their proper annotation, which is sometimes a delicate procedure that must be performed by domain experts.[28]

- A small change in the data can cause a large change in the structure of the decision tree causing instability

- Decision tree often involves higher time to train the model

3.3 Methods, Steps and techniques

There is no lot of direct researches for detecting dogs' emotions using machine learning. But in the approaches and technical side world has more researches for machine learning, deep learning, artificial intelligent and convolutional neural networks etc. As mentioned before this product is basically focused on to dog lovers to identify their dogs' emotions. So this chapter explains this Ooggy doggy product's methods and techniques with other researches.

- Data preparation and preprocessing
  - **Data collection**

    Data is the foundation of any machine learning project. Data collection is a major bottleneck in machine learning and an active research topic in multiple communities. (Roh, Heo and Whang, 2019). This "*Ooggy doggy*" project is used kaggle's data set to Machine Learning part.

  - **Data visualization**

    The large amount of data represent in graphical ways is easier to understand and analyze. "Data visualization plays a crucial role in identifying interesting patterns in exploratory data analysis." (Leban et al., 2006).

  - **Labeling**

    What is data labeling in machine learning? Let's say as example this "*Ooggy doggy*" project is image recognition system and has already collected thousands of photos of dogs. Label would be telling the AI, that dog is in happy mood, dog is in angry mood and that dog is in neutral and so on. The main challenge of data labeling is labeling approaches, techniques and tools.

    - Acquiring domain expertise- veterinary doctors/dog's trainers
    - CAPTCHA challenge
    - Transfer learning

    Transfer learning is a strategy where we use a model for one task as the starting model point for a model on a second task. Transfer learning is mostly applied for training neural network- models used for image recognition, image segmentation and human motion modeling etc.

  - **Data selection**-

After getting data, subgrouping them to solve the problem of detecting dog's emotions

- **Data preprocessing**

The value of preprocessing is to convert raw data into a form that fits machine learning. In other words "Data Preprocessing is a technique that is used to convert the raw data into a clean data set." (Data Preprocessing for Machine learning in Python, 2017)

  - Data formatting – Data formatting is like changing images name, their services and formats etc.
  - Data cleaning - This technique remove useless and incomplete data
  - Data Anonymization
  - Data sampling-big data get more time and more computational power. This technique helps to build and run models much faster.

- **Data transformation**

Data transformation can identify outcome of the machine learning project.

  - Scaling
  - Data decomposition
  - Aggregation

- Dataset splitting

Train and test

Dataset used in machine learning partitioned in to three subsets.

- **Training set** - "The sample of data used to fit the model" (Brownlee, 2017).Model will train over and over from the training set.
- **Test set**- Test data is the set of data that is used to test the model actually model is already trained. This test set already separate from both the test set and validation set. And also test set is unlabeled.
- **Validation set**-validation set is the set of data separate from the training set. That is used to validate our model during training.

- Modeling
  - **Model training**

    The purpose of model training is to develop a model. There are common two model training styles called supervised learning and unsupervised learning.

    - Supervised learning

      "Supervised learning entails learning a mapping between a set of input variables X and an output variable Y and applying this mapping to predict the outputs for unseen data" (Cunningham, Cord and Delany, 2008).With supervised learning , can solve classification and regression problems.

    - Unsupervised learning

      In Unsupervised learning training style, an algorithm can analyze unlabeled data. The goal of model training is to find hidden interconnection between data objects and structure objects by similarities and or differences. Unsupervised learning aims at solving such problems as clustering, association rule learning, and dimensionality reduction.

  - Model Evaluation and Testing

    The goal of this step is to develop the simplest model. This task can do by model tuning. So one of the method to model evaluation and tuning is cross validation.

    - Cross validation

      Cross validation entail splitting a training dataset. It can shuffle the data and split into k partitions called folds. As an example if k is 5, then each time cross validation takes 4 folds as the training set and the remaining one as the validation set. This cross validation can do by keras deep learning frame work and this can run by Tensorflow and numpy.

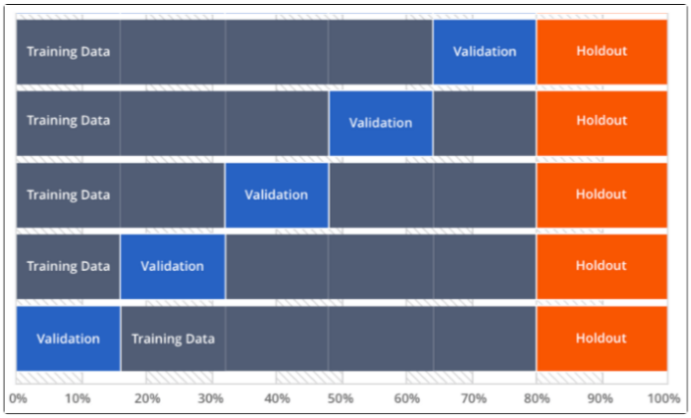

Figure 8: cross validation explanation

- Improving prediction with ensemble method

  The ensemble methods produce more accurate solutions than single model would. "Ensemble methods aim at improving the predictive performance of a given statistical learning or model fitting technique" (Bühlmann, 2012). The common ensemble methods are,

  - Stacking –This stacked generalization approach suggests developing a meta-model or higher level learner by combining multiple base models. The purpose of this technique is to reduce generalization error.
  - Bagging – This sequential model ensemble method first slit the training data into subsets then model are trained on each of these subsets. This bagging helps to avoid model overfitting and reduce the variance error.
  - Boosting – In this technique first use subset of original data set and to develop several averagely performing models and then combines them to increase their performance using majority votes.

- Model Deployment

  Deployment is the method by which you merge a machine learning model into an existing production environment to make practical business decisions based on data. This Ooggy doggy product is deploy from mobile application.

4. Chapter Summary

This chapter gives a clear idea about the current approaches, technologies, methods which are used to implement our product and will get a good idea about our products advantages and disadvantages as well as recognizing the main features that we have to emphasis on in order to have a better product. The existing products have also been thoroughly explored in depth. The next section will cover the project management process of the project and explain the methodologies used in this project.

5. References

- Aich, S., Chakraborty, S., Sim, J.-S., Jang, D.-J. and Kim, H.-C. (2019). The Design of an Automated System for the Analysis of the Activity and Emotional Patterns of Dogs with Wearable Sensors Using Machine Learning. *Applied Sciences*, 9(22), p.4938.

- Correia-Caeiro, C., Guo, K. and Mills, D.S. (2020). Perception of dynamic facial expressions of emotion between dogs and humans. *Animal Cognition*, 23(3), pp.465–476.

- Hemant Sharma (2017). *What Is Data Science? A Beginner's Guide To Data Science*. [online] Edureka. Available at: https://www.edureka.co/blog/what-is-data-science/.

- Kim, M., Joo, Y. and Park, J. (2005). *Emotion Detection Algorithm Using Frontal Face Image*.

- Maglogiannis, I., Vouyioukas, D. and Aggelopoulos, C. (2007). Face detection and recognition of natural human emotion using Markov random fields. *Personal and Ubiquitous Computing*, 13(1), pp.95–101.

- Solomon, R.C. (2019). emotion | Definition, Scope, Variety, & Structures. In: *Encyclopædia Britannica*. [online] Available at: https://www.britannica.com/science/emotion.

- Somppi, S., Törnqvist, H., Hänninen, L., Krause, C.M. and Vainio, O. (2013). How dogs scan familiar and inverted faces: an eye movement study. *Animal Cognition*, 17(3), pp.793–803.

- Srihari Sasikumar (2017). *Expert Talk: Data Science vs. Data Analytics vs. Machine Learning*. [online] Simplilearn.com. Available at: https://www.simplilearn.com/data-science-vs-data-analytics-vs-machine-learning-article.

- Leban, G. et al. (2006). VizRank: Data Visualization Guided by Machine Learning. *Data Mining and Knowledge Discovery*, 13 (2), 119–136. Available from https://doi.org/10.1007/s10618-005-0031-5.
- How to Structure a Data Science Team: Key Models and Roles to Consider. (no date). *AltexSoft*. Available from https://www.altexsoft.com/blog/datascience/how-to-structure-data-science-team-key-models-and-roles/

- Milicevic, M. et al. (2019). Application of Transfer Learning for Fine-Grained Vessel Classification Using a Limited Dataset. In: Ntalianis, K. Vachtsevanos, G. Borne, P. et al. (eds.). *Applied Physics, System Science and Computers III*. Cham: Springer International Publishing, 125–131. Available from https://doi.org/10.1007/978-3-030-21507-1_19.

- [25] Yosinski, J., Clune, J., Bengio, Y., et al.: 'How transferable are features indeep neural networks?'. Proc. Advances in Neural Information ProcessingSystems, Montréal, Canada, 2014, pp. 3320–3328

- [21] Dong, C., Loy, C., He, K., et al.: 'Image super-resolution using deepconvolutional networks', IEEE Trans. Pattern Anal. Mach. Intell., 2016, 38,(2), pp. 295–307
- [26] Sarma, Mani & Abate, Yoseph. (2014). A New Decision Tree Approach to Image Data Mining and Segmentation.

- [28] A. Kamilaris [(a1)] F. X. Prenafeta-Boldú [(a2)] The Journal of Agricultural Science, Volume 156, Issue 3 April 2018 , pp. 312-322
- Shin, et al. "Detection of Emotion Using Multi-Block Deep Learning in a Self-Management Interview App." *Applied Sciences*, no. 22, MDPI AG, Nov. 2019, p. 4830. *Crossref*, doi:10.3390/app9224830.

- Anbarjafari, Gholamreza, et al. "Machine Learning for Face, Emotion, and Pain Recognition." *Machine Learning for Face, Emotion, and Pain Recognition*, SPIE, http://dx.doi.org/10.1117/3.2322572.ch1.

- Halder, Rituparna, et al. "Real Time Facial Emotion Recognition Based on Image Processing and Machine Learning." *International Journal of Computer Applications*, no. 11, Foundation of Computer Science, Apr. 2016, pp. 16–19. *Crossref*, doi:10.5120/ijca2016908707.

- Singh, Himanshu. "Image Processing Using Machine Learning." *Practical Machine Learning and Image Processing*, Apress, 2019, pp. 89–132, http://dx.doi.org/10.1007/978-1-4842-4149-3_5.

- Amjad, A., et al. "Multiple Face Detection Algorithm Using Colour Skin Modelling." *IET Image Processing*, no. 8, Institution of Engineering and Technology (IET), Nov. 2012, pp. 1093–101. *Crossref*, doi:10.1049/iet-ipr.2012.0167.

- Suzuki, Yasufumi, and Tadashi Shibat. “A Directional-Edge-Based Face Detection Algorithm Adapted to the VLSI Image Recognition System.” *Image Fusion and Its Applications*, InTech, 2011, http://dx.doi.org/10.5772/18983.

6. Participation List

- Chapter Overview and chapter summary- Srimali Udayangani : w1790361
- Research part - Ahmed Isthaffa : w1761936
- Research part - K.N. Kumar : w1761759
- Algorithms and Technologies, list of abbreviations - H.M.A.D. Herath : w1790023
- Advantages and disadvantages of algorithms and technologies - Akthab Bifaz : w1761866
- Methods, Steps and Techniques part , Participation list and merge all the parts for report-Thiwanki Dias Hettiarachchi : w1790191
- When question arises help to other member- all the group members