Categorization of emotions in dog behavior based on the deep neural network

Zdzisław Kowalczuk | Michał Czubenko | Weronika Żmuda-Trzebiatowska

[1]Faculty of Electronics Telecommunications and Informatics, Gdansk University of Technology, Pomorskie, Poland
Correspondence
*Zdzisław Kowalczuk Email: kova@pg.edu.pl

Summary

The aim of this article is to present a neural system based on stock architecture for recognizing emotional behavior in dogs. Our considerations are inspired by the original work of Franzoni *et al.* on recognizing dog emotions. An appropriate set of photographic data has been compiled taking into account five classes of emotional behavior in dogs of one breed, including joy, anger, licking, yawning and sleeping. Focusing on a particular breed makes it easier to experiment and recognize the emotional behavior of dogs. To broaden our conclusions, in our research study we compare our system with other systems (DNNs) of different architectures. In addition, we also use modern transfer learning with augmentation and data normalization techniques. The results show that VGG16 and VGG19 are the most suitable backbone networks. Therefore, a certain deep neural network, named mVGG16, based on the sub-optimal VGG16 has been created, trained and fine-tuned with transfer (without augmentation and normalization). The developed system is then tested against an internal test dataset. In addition, to show the robustness of the system, a set of external data outside the breed is also taken into account. Being able to detect unsafe dog behavior and rely on a generalization for other breeds is worth popularizing. Equally important are the possible applications of the system to monitor the behavior of pets in the absence of their owners.

KEYWORDS:
neural networks; anthropomorphism; animals; dogs; emotions; evaluation of emotions.

1 | INTRODUCTION

There is a tendency among people to anthropomorphize their pets (cats and dogs), that is, to assign them human characteristics[1]. The consequences of treating animals in this way can be comical, harmful and, in the worst case scenario, even dangerous. Dog owners, by treating them as their own children, may unknowingly contribute to exacerbating existing, albeit hidden, behavioral problems.

Certainly, the most effective action is always to prevent the dog's misbehavior and inadequate education (anthropomorphically speaking). Therefore, a remedy for reducing the number of bites is the use of legally imposed safeguards (e.g. the use of a leash and a muzzle) and proper dog training combined with the education of the owner and the help of dog behaviorists. To some extent, this can be achieved by better understanding the dog's behavior, especially its emotions. It should be emphasized, however, that an unprepared person may misinterpret the signals sent by the animal (sometimes even the warning ones)[2].

The introduction of a technology that allows the recognition of an animal's emotions or emotional behavior may increase the awareness of dog owners, in this regard. In the future, such technology may also reduce the chance of being bitten before a child makes contact with an unfamiliar dog. It is worth noting that, currently, police officers are already equipped with a video camera that can perform the function of recognizing emotions.

At this point, we should mention the recent scientific report by Franzoni *et al.*[3] on the problem under consideration, who present an AlexNet system for recognizing the three postures of dogs: growling, smiling and sleeping, where the first two (that is, anger and joy) correspond to the basic human emotions of Paul Ekman's model[4].

Dog monitoring systems (or dog/pet cameras) available on Amazon (like the Furbo dog camera or the WOpet TitBit interactive dog camera) usually have functions, such as (duplex) transmission of audio and video, automatic (motion sensor) wake-up of the device, tracking animals and (in the most advanced model) tossing treats. Adding the option to automatically monitor the dog's behavior/emotions and respond to specific emotional states can improve its well-being. Moreover, this option will allow the owner to be in contact with the pet, which may also lead to, for example, the detection of its potential depression.

In the case of humans, there are many different systems for recognizing[5,6] and modeling[7,8] emotions, ranging from the (psychological) FACS (Facial Action Coding System)[9,6] to Affectiva applications, the latter being one of the most popular solutions for detecting and recognizing emotions. The recognition of human emotions can be based on the analysis of a static image (in the mainstream – based on facial expression)[10] or dynamic (based on the assessment of movements), sound (intonation), text (vocabulary), or even physiological signals[11], including EEG (*ElectroEncephaloGram*)[12]. However, in the case of dogs, we do not have to deal with such a variety of systems.

1.1 | Motivation

The primary motivation of this work stems from the authors' passions, including human sociology, cynology, canine behaviorism, and various types of artificial intelligence techniques (including neural networks). As a result, we would like to contribute to the, as yet, not-so-popular field of automatic categorization or even recognition of dogs' emotions. This technical paper also aims to indicate which stock architectures provide a solid basis for solving the canine problem.

The main purpose of this work is to show the potential of deep transfer learning (especially tensorflow stock architectures) for the emotional behavior recognition task of dogs. The aspects of behavior monitoring and threat detection presented in the article indicate a high application potential of the presented results. Nevertheless, at this level of development, it is clear that this issue is still a specific research niche and such a target behavior-monitoring system is rather dedicated to pet lovers, especially with a view to the periods of their absence from the dog. Overall, the initial results are promising and could be turned into an automated monitoring system that is a useful option for human monitoring of a dog.

From a technical point of view, our goal is to compare different stock architectures in a classification task when dealing with a small dataset in an unpopular application. In addition, to increase the relevance and generality of the conclusions, we extended the spectrum of our experiments to data augmentation and normalization – in order to verify the learning results with the use of many different architectures.

We are currently seeing a growing interest in the subject of emotional recognition in dogs. For example, two new articles[13,14] appeared last year. The first approach is to convert the FACS method to a dog's case[15] and use visual features to discover a dog's emotions using a neural network and a Vision Transformer. According to the authors, they achieved the accuracy of validation at the level of 0.853. Whereas the authors of the second article[14] used keypoints in the dog's body (some kind of poselet) to predict its emotions. Therefore, this topic seems to be timely and potentially interesting for the readers.

The approach and monitoring system proposed in this article has several goals. Its use will enable dog owners to better understand their pets, learn about the dog's habits and the sources and extent of stress in a dog left alone at home. Our long-term goal is to create a dog monitoring system that should consist of a behavioral analysis module that includes emotions.

1.2 | Leading idea and structure of work

In order to achieve the intended goal, after a short introduction and basic literature review, we present in section 3 a neural architecture based on VGG16 (from the Visual Geometry Group, Oxford)[16], which allows us to classify the five emotional states of Shetland Sheepdogs. The experiments performed (section 4) can also be treated as additional proof of the validity of the idea[3] regarding the possibility of machine learning to recognize emotions in dogs. Note that this study uses AlexNet (CNN – Convolutional Neural Network) to analyze the three aspects that make up the overall picture of dogs' emotions: joy, anger and neutrality. In section 4, we present a direct comparison of both approaches and present conclusions resulting from the technological research carried out.

1.3 | Possible applications

It is worth pointing to the possible application of the presented network. The main potential audience is dog owners who have to leave them home alone. A microcomputer with a camera will allow not only remote monitoring of the pet, but also describing and summarizing recordings from such supervision. Owners will have the opportunity to find out how often their pet is bored, happy or angry. This type of application should also be equipped with an appropriate tracking tool, which should monitor the movements of the pet and identify where it spends time (in what room and possibly in what place).

In addition, such a robotic application can be equipped with a set of actuators for the interaction of the robot with the dog. For example, it can be a treat dispenser supported by the aforementioned interactive WOpet TitBit camera. The system concept in question can also be provided with various interactive toys, such as a light ball cannon for retrieving toys. A properly adapted bullet magazine could also function as a cannon feeder. The robot can also be equipped with the owner's 'voice' for verbal communication with the pet. Of course, the propotype of such a device should be properly tested by the canine behaviorist.

Another use is for alien animal monitoring in public places. You can use a dedicated standalone device approach that will guarantee adequate processing time – then such a device can be used on the path of the child's movement (to or from school) to monitor foreign animals.

After suitable optimization, you can also produce an application for mobile devices. The main disadvantage of this solution is the need to pull out the phone, run the application and take a photo. However, it can serve to satisfy owners' curiosity and encourage them to use the target device (instead of having to warn of a dangerous dog).

2 | DOGS AND EMOTIONS

It is known that dogs have accompanied humans for tens of thousands of years, although the exact time and place of their domestication have yet to be determined. One of the most popular theories about canine genetic variability in wolves that has evolved in the Middle East indicates that dogs are of Asian origin[17]. On the other hand, some theories define East Africa as the area of origin of dogs[18]. Other studies locate parentage and domestication in Europe or Asia in general. Certainly, in Europe, the domestication of these animals took place over 15,000 years ago[19].

In connection with the above, it can be concluded that the current relationship between a human and a dog (as a family friend) has been shaped over the last several decades. Which results in high awareness of the dog's needs and, consequently, also in the creation of the canine behaviorist profession.

Studies have shown that with age, the dog's intelligence increases to the level of a 2-2.5-year-old child and also includes emotional development[20]. Initially, in human development, infants feel only excitement, which changes over time to positive and negative. In the following months, disgust, anger and fear develop, and only after six months do children feel joy[21]. In the early stages, the dog's brain grows faster than the human brain, and the areas of the brain dedicated to emotional recognition fully develop within the first 4-6 months in dogs, and in humans within 2.5 years[22].

Humans anthropomorphize animals excessively and in many ways. The same goes for dogs' emotions. But it is not entirely clear whether they can be directly compared to human states[23]. On the other hand, it is known that human primitive emotional responses such as fear originate in the amygdala and are amplified in the visceral (old mammal) brain, while the neocortex connects emotions and cognition[24,25]. In this way, mammals share certain emotions and emotional responses (note that the reptilian brain theory is now rather obsolete[26]).

There are already studies (e.g.[27]) comparing emotions that take place in dogs and in humans. It seems that the basic emotions in a dog can be partially translated into human language, but rather with the proviso that they may not fully correspond to our imagination[28]. For example, the sense of smell plays an important role in dogs and affects their emotions very strongly[29]. In addition, because emotions relate to the senses and the senses of dogs and humans are completely different, generating emotions can have different effects. Thus, we can only try to guess the dog's emotions, based on the context we know. However, taking into account, for example, the dog's growling alone, slight differences in such a sound can mean significantly different emotions[30]. The more that even our emotions can also have different meanings, depending on the culture in which we grew up[31]. Nevertheless, there are (6) basic emotions that are common to the whole world[4].

While we have many definitions of emotions worth citing, we'll only cover a few of them here[8]:

- human point of view: "Emotions are organized psychophysiological reactions to news about ongoing relationships with the environment"[32].
- human point of view: "Emotion is a complex chain of loosely connected events that begins with a stimulus and includes feelings psychological changes, impulses to action and specific, goal-directed behavior"[33].
- canine case: "Emotion, as described by neurobiologist John Ratey, is a movement outward, a way of communicating our most important internal states and needs"[34].

Note that in all definitions, emotions encompass certain goal-directed behaviors (and, therefore, are followed by certain behaviors). One method of studying dogs' emotions is to study their behavior or photographs. An example of an emotional dog can be found in[2]. According to work[34] supported by[2], several dog behaviors can be distinguished:

- tongue flicking, related to anxiety
- smiling by lifting upper lips, connected to joy
- body freezing, always giving a warning
- 'slightly open mouth', meaning relaxation
- teeth grinding, related to anger.

The expressions presented above do not exhaust all the possibilities.

2.1 | Similar work

Great progress has been made in the areas of computer systems, machine learning and robotics. The situation is different in the field of the automatic monitoring of animal behavior and other aspects related to their emotions, where, apart from one seminal work described in[3], it is difficult to find useful research results in the field of monitoring dogs' emotions.

The studies described in[35] mention the monitoring of an animal in a cage with an automated feeding system. However, such a system is only intended for cages that are not the most suitable environment for dogs. On the other hand, following the message[36], it can be concluded that autonomous monitoring systems that allow some interaction with the animal can also de-stress it. It is also worth citing work [11], in which a mobile robot interacts with a dog, and the applied pet behavior analysis module detects three different behaviors, depending on the distance from the robot. In addition, you can also find preliminary studies for creating a companion for a dog left at home[37]. In this context, it is worth considering that from a machine learning (ML-DNN) perspective, there is only one publicly available dataset, namely the Stanford Dog Breed Data Set[38,39], which focuses on the task of breed identification.

There are also several other works on similar topics. They focus, for example, on classifying an image in terms of assessing the suffering experienced by a specific animal, e.g. laboratory mice, horses or sheep[40,41]. Moreover, there is also a lot of work on population-based animal tracking, counting and segmentation systems[42].

It seems that the most important report directly related to the recognition of emotions in photos of dogs is the conference report[3], the authors of which focused on the classification of canine emotions from photos using the learning transfer and AlexNet (CNN from Alex Krizhevsky)[43]. For fine-tuning, this work uses a dataset of 231 images collected from the Internet. The images were classified by experts (e.g. veterinarians and dog breeders) into one of three different stances ('Growling', 'Smile' and 'Sleep'), with the first two corresponding to basic human emotions (anger/aggressiveness and joy/friendliness), and the third is considered a neutral state (not the first two emotions)[3]. The implementation of emotions is based on the Ekman model[4]. The dataset used for fine-tuning was unbalanced[1], with the *smile* class dominating (92 samples, covering 40% of the set). The input to AlexNet, used in[3], automatically scales the images to a size of 256x256 pixels. The AlexNet CNN structure was expanded with two layers, dense and softmax. In this work, the authors proved that the effectiveness of such a network in recognizing canine emotions can reach the level of about 95%.

[1] Which does not necessarily contribute to bias (our dataset is also unbalanced).

There is also an approach to recognizing the emotions expressed by dogs through sound analysis. For example, the authors of[44] argue that four animal emotions can be identified from sound cochleagrams: anger, loneliness, crying, and happiness. Using a bagged decision tree in classification, they achieved the accuracy in recognizing, anger - 88%, loneliness - 74%, crying - 56% and happiness - 35%[2].

3 | THE METHOD

Modern technology will be used to achieve the utilitarian purpose of this work and to re-ensure that deep learning, combined with transfer learning, allows for the construction of a modern classifier of emotional behavior in animals, e.g. horses[45] or dogs[3].

By reviewing literature, it is easy to see that many neural architectures are used in transfer learning. A comparison of the most interesting of them can be found in[46]. Different types of architectures in the emotion recognition task have been compared in[47]. Based on these results, it is safe to assume that the VGG16 backbone[16], pre-trained on the ImageNet dataset[39], is one of the most robust architectures worth employing for emotional recognition. To find our own effective solution (mVGG16) to the problem under consideration, VGG16 architecture will be redesigned and fine-tuned, based on our dataset (AAFCI).

In order to broaden and confirm the message of this article, a comparison of several stock neural networks relevant to the discussed issue is made in 4.3.

3.1 | Dataset

In general, it is not easy to pinpoint the exact set of basic emotions dogs may be feeling. Canine behaviorism is only an emerging field. Dogs communicate emotionally with their surroundings through body language. However, classifying dog photos from different angles is quite a complicated problem. Therefore, in this work we will only focus on the emotions expressed in facial expressions.

After analyzing the signals sent by the dogs, including the head and ears, five behaviors (including showing emotions) were selected: smiling (the emotion of joy), growling (the emotion of anger), licking (anxiety), yawning (relaxation/trust), and sleeping. The dataset was compiled specifically for the recognition of emotional behavior in dogs. On this basis, the behavior was assessed by dog keepers, as it was more reliable than self-inferring emotions from the photos. Other behaviors identified in this experiment, such as licking or yawning, were included, in order to draw attention to the possible existence of other types of emotions than the classic ones, such as anger and joy. At the same time, it should be noted that licking is often performed by dogs in various emotional situations, e.g. stressful or related to anxiety. The yawning signal can be a symptom of relaxation and confidence (trust)[34]. Meanwhile, sleep, which controls emotional reactivity[48], can be treated as a neutral state (as in[3]). Note that emotions will henceforth be used to describe the dog's signal classes of behavior, each of which has a list of behavioral traits shown in Tab. 1.

TABLE 1 Characteristics of dog signal classes.

behavior	emotion	head	Eyes	ears	muzzle
smiling	joy	raised	open	raised	open, visible tongue
growling	anger	tilted	open, squared	tilted sideways	open, visible teeth
licking	anxiety	raised	open, squinted	raised or placed	closed, tongue on nose
yawning	relaxation/trust	raised	closed, squinted	tilted back	wide open
sleep	neutral	located	closed	relaxed	closed

We applied the AAFCI dataset of photos of Shetland Sheepdogs coming from different kennels[3]. This AAFCI dataset consists of 360 collected photos of dogs in one of these 5 behavioral classes (emotional states signaled by a dog). Note that the analyzed breeding dogs are of different age and color. Sample images from the training dataset are shown in Fig. 1.

[2]Which means that sound does not seem to be an ideal symptom, nor is it a medium for signals of happiness.

[3]Mainly from the kennel "*Ponad Wszystko FCI*", whose name can be translated into English as "Above All, FCI", where FCI stands for the Canine Federation International.

Importantly, the AAFCI dataset was compiled for this study by professional keepers of various kennels and then validated by a dog behavior specialist who is also a dog breeder. The procedure of collecting photos was prepared in cooperation with the caretakers, behaviorists and owners of the dog kennel. During the session, the photographer, owner or guardian, waited for the dog to express certain emotions. Note that such a close person knows his/her dogs and can easily recognize the emotion that has arisen. In many cases of taking a picture, the photographer waited a long time with the camera for the dog to adopt the appropriate posture (behavior). The essence of the entire operation was to ensure that the owner was aware of both the dog's personality and its current environmental context. The photos have been manually cropped to cover the dog's entire head.

(a) Anger. **(b)** Joy. **(c)** Anxiety.

(d) Neutral. **(e)** Trust.

FIGURE 1 A few examples from the test dataset.

As mentioned, the AAFCI dataset is not perfectly balanced. While there are methods for dealing with unbalanced datasets[49,50,51], such methods were not used in this study. So, our unbalanced AAFCI dataset may be causing extra bias in the classes.

The AAFCI (validated training) dataset is used in the learning and validation process, with the latter only being applied after each epoch to verify system hyperparameters (such as structure and learning parameters). On the other hand, the internal/project test dataset is used for the final evaluation of the trained network, in terms of its desired effectiveness (within the adopted assumptions).

In the 360-element, basic photographic dataset, the learning part (for validated learning) covers 93% of the data set, while the remaining 7% (5 classes of 5 images) are left for the internal test part. The training part is further broken down in an 8: 2 ratio into regular training data (270 images) and validation data (65 images). Some details of the parametrization of the datasets used are presented in Tab. 2.

As an external test source used for generalized and practical proof of concept, we considered dogs of different breeds collected from other kennels in the same procedure as for the AAFCI dataset.

TABLE 2 Details of the datasets.

	Shetland Sheepdogs (AAFCI dataset)			External **(different breeds)**
	train	**validation**	**test**	**test**
joy	89	22	5	10
anger	41	10	5	10
anxiety	52	12	5	10
trust	37	9	5	10
neutral	51	12	5	10
sum	270	65	25	50

3.2 | mVGG16 network structure

The structure of the mVGG16 network used is shown in Fig. 2. The input layer of the model accepts RGB (Red Green Blue) photos with a size of $(3 \times 256 \times 256)$ pixels. Then, the *Conv* (convolution) operations are performed with the 3×3 filters and the *LeakyReLu* activation function. The yellow blocks with convolution layers are separated, in Fig. 2, by the *Max-Pooling* layers, which are built on the basis of 2×2 pixel windows, with a step equal to 2. After these five (basic) blocks, two *dense* layers with 256 neurons and the *ReLu* activation function are applied, the last block being the output layer with the *Softmax* activation function. Note that the standard VGG16 network uses three dense layers in its output.

The output layer with five marked classes is responsible for the final classification. Dense layers are woven over *Dropout* segments with a probability of 0.5 to prevent network overload.

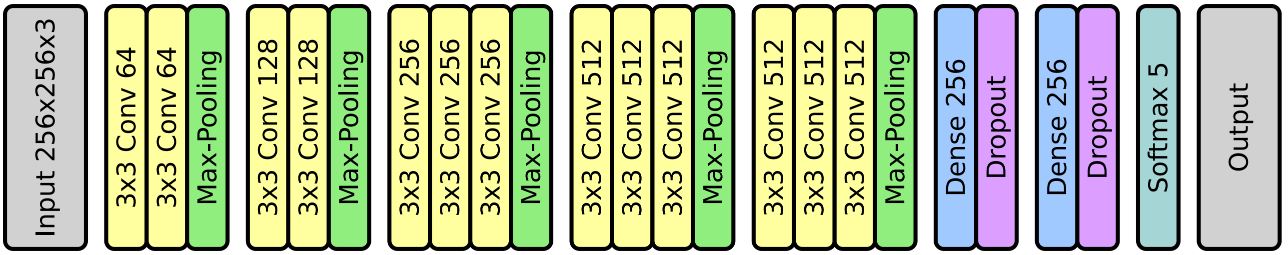


FIGURE 2 Modified VGG16 network structure.

4 | TEST RESULTS

An ADAM (*ADAptive Momentum*) optimizer was used to teach the neural architecture[52] we have adopted. This algorithm used standard parameter values, with a slight change in the learning rate (0.00001). The network was trained for approximately 80 epochs on the validated training dataset, using a batch size[4] of 8. No augmentation or normalization methods were used in this particular training setup. This training used a callback operation (similar to the elitism mechanism used in evolutionary computations or genetic algorithms) to record the best optimized network in real time, during the training runs of each era. As a final observation, it can be stated that the results obtained after the 19th epoch achieved a validation accuracy of 0.94 and turned out to be optimal for the entire learning process. This, in turn, means that further training of the network was excessive.

As mentioned earlier, the trained (learned and validated) neural network was assessed on the prepared test dataset containing 25 images. Moreover, for wider verification, the effectiveness of the network was also assessed on an external dataset collected from a set of dogs of different breeds: dalmatian, samoyed, welsh corgi cardigan, miniature poodle, pug, husky, Irish setter, and

[4]The batch size is the number of training images used in one iteration, and one epoch contains a fixed number of iterations – in our case, on a PC with an NVIDIA 1050 Ti and 3GB of memory, this number was set to 9.

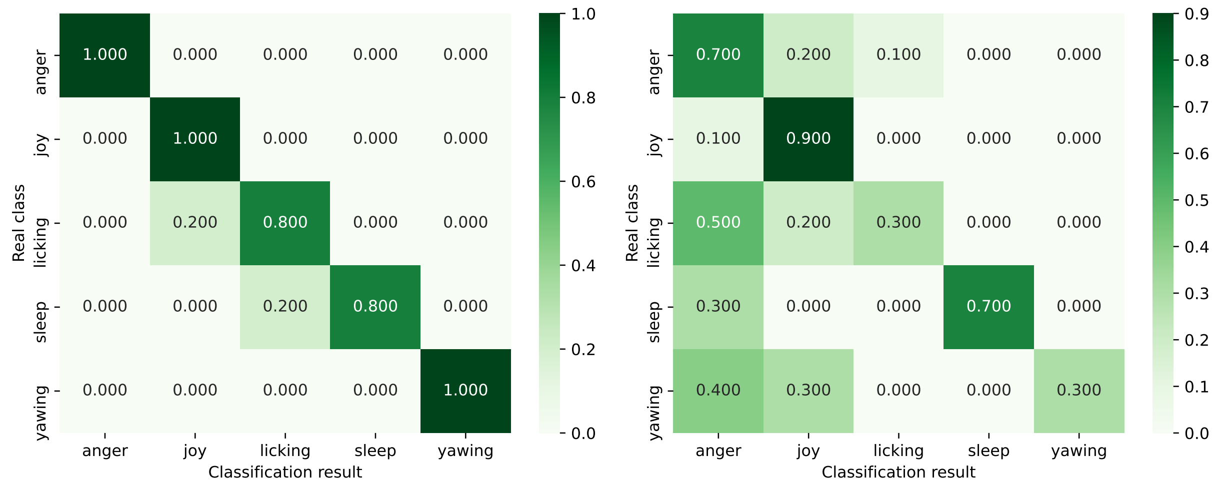
mongrels (50 images, 10 for each class). This external dataset was obtained in the same way as the training set. It is also important that these breeds belong to different groups, characterized by a completely different anatomical structure. This procedure was intentionally used to make the classification results even more reliable.

The effect of such an assessment is presented in Fig. 3, which shows the confusion matrix obtained for the two (internal and external) types of datasets. For example, note that in these demanding user tests, the most important component of anger (1, 1) in the confusion matrix shown is 0.7, which means that there were three mismatches in the anger rating. As previously mentioned, the internal test was performed on 25 samples and the external test on 50.

In turn, Tab. 3 contains a measure of the classification quality assessment for each identified class. Considering, for example, the internal test dataset and the *F1-Score* metric[53] used, which is considered a fairly reliable measure[5], the best net training effect was achieved for the anger/growl and trust/yawn classes. Interestingly, in the case of external testing, the best effects are associated with the canine emotional states of joy and neutral/sleep. Overall[6], however, it should be noted that a high average accuracy of 92% was achieved for the internal test dataset and a pretty good result (58%) for the external set. In addition, it should be noted that the developed algorithm most often classifies images with visible teeth as anger (meanwhile, some images of licking or yawning also show teeth). Since the causes of misclassification are similar to the case of[3], it can be assumed that one of the possible abstract traits trained by ImageNet should be teeth.

TABLE 3 Testing mVGG16: F1-score, accuracy and precision for each class.

	Internal (AAFCI) dataset				*External (other breeds) dataset*			
	accuracy	precision	recall	f1-score	accuracy	precision	recall	f1-score
anger	1.00	1.00	1.00	1.00	0.70	0.35	0.70	0.47
joy	1.00	0.83	1.00	0.91	0.90	0.56	0.90	0.69
anxiety	0.80	0.80	0.80	0.80	0.30	0.75	0.30	0.43
neutral	0.80	1.00	0.80	0.89	0.70	1.00	0.70	0.82
trust	1.00	1.00	1.00	1.00	0.30	1.00	0.30	0.46
average	0.92	0.93	0.92	0.92	0.58	0.73	0.58	0.57



(a) Internal/designer test dataset.

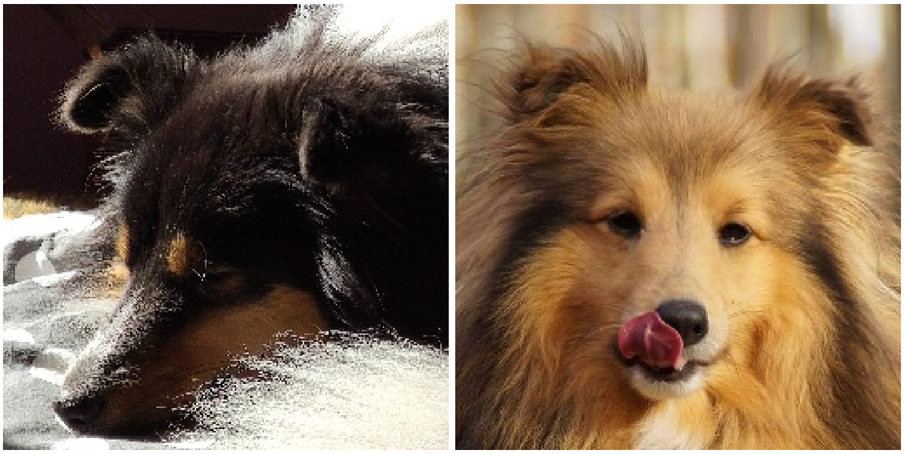
(b) External/user test dataset.

FIGURE 3 Confusion matrix for two types of test dataset.

[5]Note that there are studies that suggest better multi-class classification indicators, such as the Matthews Correlation Coefficient (MCC)[54].
[6]In another case, such as the MCC, it was 90.19% on the internal test dataset, while on the external set, the network reached 50.49%.

4.1 | Test misclassifications

Recall that we take into account two different tests: the first/internal carried out on the test part (25 images) of the Shetland Sheepdog dataset (AAFCI), and the second/external conducted on the test dataset (50 images) of completely different breeds. Importantly, of the entire set of Shetland tests (25 photos), only two images were misclassified, shown in Fig. 4. While there is no simple explanation for the misclassification of image 4a, it is easy to suppose that the dog's smile and its relatively serene gaze in Fig. 4b should be classified as joy.



(a) Label: neutral; Prediction: anxiety. **(b)** Label: anxiety; Prediction: joy.

FIGURE 4 Misclassified images from the internal test dataset.

4.2 | External misclassifications

The situation is much worse in the external (data set) test. Fig. 5 shows four examples of misclassification by mVGG16 on the external dataset (for different breeds), which led to 21 errors when analyzing 50 images. Only one image has been misclassified as an emotional state of joy, while there have only been three misclassified with anger. Note that the worst-case scenario is classifying anger as joy (it only happened once). Such misclassification is the least desirable, as it may be a hazard. In 5a and 5b, we show two of the three cases with the most dangerous confusion of canine emotions, that is, not recognizing anger.

Most of the misclassifications belong to the classes of licking and yawning, for which – due to visible teeth – are easily mistaken for the anger emotion (e.g. Fig. 5d). Nevertheless, when dealing with completely different races, the classifier does quite well (at least in the case of the two distinct classes).

As mentioned above, only one breed (Shetland Sheepdogs) was used in the internal data, while external test photos were obtained on other breeds. Therefore, the difference in system performance in both cases is understandable. However, the positive results obtained in external tests – as with joy and neutral emotions – also lead to the conclusion that training on one breed can be a good basis for tuning the model to the dataset of other breeds (consistent with learning transfer).

4.3 | Comparison of neural structures

In general, it can be concluded that the generalized feature maps (weights of the convolutional layers) are well trained, and the obtained effect of recognizing the adopted classes of canine emotional states (demonstrated on the basis of the internal test) can be considered a promising utilitarian result.

On the other hand, to prove the correctness of the chosen mVGG16 architecture, a comparison of several arbitrarily selected base architectures was made. In particular, the following well-known stock DNN architectures have been adopted (with ImageNet-trained weights): VGG16, VGG19[16], DenseNet121[55], InceptionResNetV2[56], ResNet50V2, ResNet101V2[57] and Xception[58].

(a) Label: anger; Prediction: anxiety.
(b) Label: anger; Prediction: joy.
(c) Label: joy; Prediction: anger.
(d) Label: trust; Prediction: anger.

FIGURE 5 Misclassified images from the external dataset.

Taking into account our previous experience with the modification of VGG16, each of the stock nets was supplemented with two dense layers and one softmax (therefore, the original name VGG16 will now be used). Each of the DNNs was trained using the same (ADAM) algorithm with the same hyperparameters. However, a different dataset seed type than before was used to split the data into training and validation. Note that in the initial test, a random seed was used, while in this comparative study, the baseline conditions should be the same, so this parameter had to be pre-determined. This explains why some of the training results obtained here may be worse or better than in the case of the previous (random) setting.

The performance results obtained by learning the above-mentioned neural architectures are presented in Tab. 4. Note that both VGG architectures achieve the highest accuracy on the validation set. At the same time, it should be noted that both networks are also characterized by relatively high efficiency on external datasets (in relation to the others). Under the initial conditions used, VGG19 performs better than VGG16. However, the VGG19 (with 20,024,384 trainable parameters) has 16 convolution layers, while the VGG16 (with 14,714,684 trainable parameters) has only 13[59]. It should be emphasized that the smaller VGG16 network was previously selected as the template for the detailed study of the mVGG16 classifier for recognizing canine emotions.

The contents of Tab. 4 present a comparative study based on 5 criteria and using the techniques of input data augmentation and normalization (to the range $< -1; 1 >$). Random horizontal flip, random rotation (10%), random translation ($< -20\%, 20\% >$) and random zoom (10%) were selected for augmentation.

As with learning without data augmentation, both VGG networks show fairly high results. Augmentation positively influenced other architectures (like ResNet101V2), significantly increasing their performance on the validation set. However, this effect did not translate into our experimentation with external tests.

Particularly noteworthy is the work[3], which is an important reference for this research and which is based on a different backbone architecture (AlexNet), also with two dense layers added. In the two approaches AlexNet and VGG, unbalanced datasets are used for a different number of recognized emotions (3 and 5) and breeds (multiple and single). Due to these differences, it is not easy to make a direct comparison of the two approaches (although it is worth considering).

TABLE 4 Training outcomes of stock architectures, with and without data augmentation and normalization (best results are in bold). The shortcuts in the first column are: tr. – training, val. – validation, ext. – external, acc. – accuracy (the highest rate obtained during training), and f1 means the average f1 score.

Results without data augmentation and normalization								
	VGG16	VGG19	DenseNet121	ResNetV2	ResNet50V2	ResNet101V2	MobileNetV2	Xception
val. acc.	0.6866	**0.7164**	0.5075	0.4776	0.4776	0.5373	0.5970	0.5373
test acc.	0.44	**0.52**	0.44	0.36	0.36	0.24	0.40	**0.52**
test f1	0.38	**0.52**	0.40	0.36	0.30	0.20	0.37	0.48
ext. acc.	0.46	**0.50**	0.40	0.22	0.28	0.34	0.28	0.20
ext. f1	0.44	**0.48**	0.37	0.19	0.24	0.30	0.24	0.17
Results with data normalization								
val. acc.	0.7015	0.6716	0.7462	0.6716	**0.7761**	**0.7761**	0.7612	**0.7761**
test acc.	0.56	0.68	0.64	0.52	0.60	0.60	0.64	**0.72**
test f1	0.48	0.65	0.61	0.47	0.56	0.57	0.60	**0.72**
ext. acc.	**0.58**	0.56	0.44	0.40	0.46	0.52	0.44	0.36
ext. f1	**0.55**	0.52	0.35	0.33	0.43	0.47	0.36	0.30
Results with data augmentation and normalization								
val. acc.	0.7612	0.7015	0.7761	0.6866	0.8209	**0.8508**	0.7910	0.7313
test acc.	0.64	0.64	**0.68**	0.48	0.56	0.60	0.60	0.64
test f1	0.61	0.63	**0.68**	0.45	0.54	0.57	0.58	0.66
ext. acc.	0.56	0.56	0.48	0.40	**0.60**	0.42	0.48	0.44
ext. f1	0.51	0.52	0.44	0.35	**0.59**	0.41	0.41	0.36

We suppose that the better performance of the VGG architectures is due to the limited number of samples in the dataset used. Taking this fact into account, network learning supported by the technique of data augmentation should result in greater accuracy. Such a phenomenon can be observed for all architectures in Tab. 4. It should be noted that VGG backbones tested against the external dataset (for breeds other than Shetland Sheepdogs) yield surprisingly high performances (e.g. with data normalization). This may lead to the conclusion that VGG has greater generalizing abilities. It should be noted that the so-called explainable AI with color maps can provide a more detailed analysis of the performance of such a classification. This particular task is also on our future research path.

5 | FINAL REMARKS

In addition to the presented VGG technique of neural network design and its positive verification, it is worth adding that other learning methods and structures were also tested, including the reduction of dense layers and the reduction of the number of their neurons. However, such actions did not bring the intended results. There have also been several different attempts to use data augmentation to reduce the impact of a small data set. However, this action also did not lead to the expected improvement in the effectiveness of identifying canine emotions.

The designed network was expected to score very low for the external test dataset, that is, it was assumed that it might react in a way that is closer to random action than to give a truly correct classification. However, even with the training set limited to one breed, the results turned out to be very promising – the average accuracy for all classes was 0.58, which can be attributed to a well-trained network (pre-trained, based on transfer learning, in terms of weights of convolution layers).

Our research using the VGG16 architecture and the work with AlexNet [3] show that DNNs are effective in identifying specific emotional states and behaviors in dogs. The breed of dog also plays a role here. The most useful finding is the ability to detect these most salient emotional states in a dog (especially the dangerous ones), as well as the ability to generalize a given network structure to detect such conditions in other breeds.

5.1 | The discussion

From the technical side, our work presents a comparison of different backbone networks for a selected application. We used the classic transfer learning approach with optional augmentation and normalization. Taking into account the strong limitation of the training dataset, the obtained results seem quite surprising. It turns out that pure augmentation for a small dataset with very similar abstract concept images does not help. On the other hand, normalization techniques improve the quality of most of the models considered. However, both techniques (augmentation and normalization) are ineffective when it comes to external data (other races in our case). Such a situation can also be interpreted in terms of overfitting. However, the main goal was to recognize the emotions of dogs (specific breed), therefore the dataset used was only obtained from Shetland Sheepdogs. Here, networks without augmentation and standardization gave better results in terms of generalization.

5.2 | Future work

As part of future work, we plan to make our dataset public. At the same time, seeing a large deficiency in the obvious poverty of the dataset, it is necessary to work on the development of a larger set of data on the different emotional behaviors of dogs, from the largest possible set of breeds (especially, the dangerous ones). Additionally, a dataset consisting of short video sequences is needed, as emotional expressions are quite short[7] [60]. Such a set is necessary to create a very resistant application for monitoring animal behavior (including emotions). The main engines for such applications can be neural networks, capable of tracking a pet and determining its emotional state (which may also involve changing an emotion into a sentiment, which is simply a positive or negative emotion).

On the other hand, given the small size of the dataset used, it would be interesting to test a variety of approaches, including the few-shots methodology and other ways to deal with the critical case of small datasets. An interesting and modern direction may be the use of text-to-image art generators such as DALE-E2 [61] to create more images according to the so-called synthetic approach (see also [62]). Moreover, 'extreme' emotion transfer can also be a good idea where you can train the model to recognize human emotions and then transfer this knowledge to the application being considered here.

By consistently continuing this research direction and looking for further evidence of the effectiveness of deep learning in the context of recognizing canine emotions, it is also possible to consider, in the future, deepening the general research on transfer learning itself. Assuming that the weights trained according to ImageNet are not yet optimal for this task, another very time-consuming experiment can be proposed. Namely, taking the previously selected stock network (VGG16) as a starting point, it is worth training it to categorize all (120) dog breeds on a set (20,580) of Stanford Dogs photos [38]. And, only then (due to transfer learning) can the network be trained to recognize the emotions of dogs of all 120 breeds.

In addition, bearing in mind the efficiency and complexity of the designed system, you can also give up the advantages of the learning transfer paradigm and try to design a neural network yourself, using a more adequate database for training (e.g. the Stanford dog breed dataset).

Note that current approaches relies only on selected key points and specific dog behavior. However, as with humans, most of the current models are based on deep learning from a single image, rather than on so-called poselets or FACS. In the analyzed case of recognizing emotions in living creatures (humans and dogs), achieving good results on the basis of one frame may not be easy. Thus, a new interesting alternative research direction appears here, which we would also like to continue in our future work. Note that a dynamic image (movie) may well fit the emotional recognition task as the most specific dog behaviors are not static, but dynamic.

[7] We assume that the duration of emotions in dogs is similar to that in humans, which lasts less than 4 seconds.

At the very end, it is also worth recommending the use of the so-called embodiment of intelligence[63,64], in which, for example, we create several binary classifiers for each of the emotional states, and then combine them into one mechanism.

References

1. Brown Christina M, McLean Julia L. Anthropomorphizing dogs: Projecting one's own personality and consequences for supporting animal rights. *Anthrozoös.* 2015;28(1):73–86.

2. Bloom Tina, Friedman Harris. Classifying dogs'(Canis familiaris) facial expressions from photographs. *Behavioural processes.* 2013;96:1–10.

3. Franzoni Valentina, Milani Alfredo, Biondi Giulio, Micheli Francesco. A Preliminary work on dog emotion recognition. In: :91–96; 2019.

4. Ekman Paul. Basic emotions. In: Dalgleish Tim, Power Mick, eds. *Handbook of Cognition and Emotion*, Chichester England: John Willey & Sons 1999 (pp. 45-60).

5. Kołakowska Agata, Szwoch Wioleta, Szwoch Mariusz. A review of emotion recognition methods based on data acquired via smartphone sensors. *Sensors.* 2020;20(21):6367.

6. Kowalczuk Zdzisław, Chudziak Piotr. Identification of emotions based on human facial expressions using a color-space approach. In: Koscielny Jan Maciej, Syfert Michal, Sztyber Anna, eds. *Advanced Solutions in Diagnostics and Fault Tolerant Control*, Advances in Intelligent Systems and Computing, vol. 635: Cham: Springer IP AG 2018 (pp. 291–303).

7. Kowalczuk Zdzisław, Czubenko Michał. Computational approaches to modeling artificial emotion–an overview of the proposed solutions. *Frontiers in Robotics and AI.* 2016;3:21.

8. Kowalczuk Zdzisław, Czubenko Michał, Merta Tomasz. Interpretation and modeling of emotions in the management of autonomous robots using a control paradigm based on a scheduling variable. *Engineering Applications of Artificial Intelligence.* 2020;91:103562.

9. Ekman Rosenberg. *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS).* Oxford University Press, USA; 1997.

10. Lu Lihua. Multi-angle face expression recognition based on generative adversarial networks. *Computational Intelligence.* 2022;38(1):20–37.

11. Shu Lin, Xie Jinyan, Yang Mingyue, et al. A review of emotion recognition using physiological signals. *Sensors.* 2018;18(7):2074.

12. Kim Youmin, Choi Ahyoung. EEG-Based Emotion Classification Using Long Short-Term Memory Network with Attention Mechanism. *Sensors.* 2020;20(23):6727.

13. Boneh-Shitrit Tali, Amir Shir, Bremhorst Annika, et al. Deep Learning Models for Automated Classification of Dog Emotional States from Facial Expressions. *arXiv preprint arXiv:2206.05619.* 2022;.

14. Ferres Kim, Schloesser Timo, Gloor Peter A. Predicting Dog Emotions Based on Posture Analysis Using DeepLabCut. *Future Internet.* 2022;14(4):97.

15. Bremhorst Annika, Sutter Nicole A, Würbel Hanno, Mills Daniel S, Riemer Stefanie. Differences in facial expressions during positive anticipation and frustration in dogs awaiting a reward. *Scientific reports.* 2019;9(1):1–13.

16. Simonyan Karen, Zisserman Andrew. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556.* 2014;.

17. Pollinger John P, Lohmueller Kirk E, Han Eunjung, et al. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature.* 2010;464(7290):898–902.

18. Boyko Adam R, Boyko Ryan H, Boyko Corin M, et al. Complex population structure in African village dogs and its implications for inferring dog domestication history. *Proceedings of the National Academy of Sciences.* 2009;106(33):13903–13908.

19. Frantz Laurent AF, Mullin Victoria E, Pionnier-Capitan Maud, et al. Genomic and archaeological evidence suggest a dual origin of domestic dogs. *Science.* 2016;352(6290):1228–1231.

20. Association American Psychological. Dogs' Intelligence On Par With Two-year-old Human, Canine Researcher Says. *ScienceDaily.* 2009;.

21. Tronick Edward Z. Emotions and emotional communication in infants.. *American psychologist.* 1989;44(2):112.

22. Coren Stanley. Which Emotions Do Dogs Actually Experience?. *Psychology today.* 2013;.

23. Gácsi Márta. Beyond the provable?. *Animal Sentience.* 2017;2(14):12.

24. Dalgleish Tim. The emotional brain. *Nature Reviews Neuroscience.* 2004;5(7):583–589.

25. LeDoux Joseph E. Coming to terms with fear. *Proceedings of the National Academy of Sciences.* 2014;111(8):2871–2878.

26. Cesario Joseph, Johnson David J, Eisthen Heather L. Your brain is not an onion with a tiny reptile inside. *Current Directions in Psychological Science.* 2020;29(3):255–260.

27. Kujala Miiamaaria V. Canine emotions: Guidelines for research. *Animal Sentience.* 2018;2(14):18.

28. Konok Veronika, Nagy Krisztina, Miklósi Ádám. How do humans represent the emotions of dogs? The resemblance between the human representation of the canine and the human affective space. *Applied Animal Behaviour Science.* 2015;162:37–46.

29. Bräuer Juliane, Silva Karine, Schweinberger Stefan R. Communicating canine and human emotions: Commentary on Kujala on Canine Emotions. *Animal Sentience: An Interdisciplinary Journal on Animal Feeling.* 2017;14(7):1–5.

30. Zentall Thomas R. The study of emotion in animals. *Animal Sentience.* 2017;2(14):3.

31. Van Hemert Dianne A., Poortinga Ype H., Vijver Fons J.R.. Emotion and culture: A meta-analysis. *Cognition and Emotion.* 2007;21(5):913–943.

32. Lazarus Richard S, Lazarus Bernice N. *Passion and reason: Making sense of our emotions*. Oxford University Press, USA; 1994.

33. Plutchik Robert. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist.* 2001;89(4):344–350.

34. McConnell Patricia. *For the love of a dog: Understanding emotion in you and your best friend*. New York, USA: Ballantine Books; 2009.

35. Suksaengjun Permsub, Thanapatay Dusit, Nobuhiko Sugino, Chinrungrueng Jatuporn. The design of smart dog cage system. In: :MIT–161IEEE; 2016.

36. Paldanius Mikko, Kärkkäinen Tuula, Väänänen-Vainio-Mattila Kaisa, Juhlin Oskar, Häkkilä Jonna. Communication technology for human-dog interaction: exploration of dog owners' experiences and expectations. In: :2641–2650; 2011.

37. Sun Zhe, Feng Wei, Jin Jintao, et al. Pet Companion System Based on Image Recognition. In: :303–309IEEE; 2021.

38. Khosla Aditya, Jayadevaprakash Nityananda, Yao Bangpeng, Li Fei-Fei. Novel dataset for fine-grained image categorization: Stanford dogs. In: :1–2Citeseer; 2011.

39. Deng Jia, Dong Wei, Socher Richard, Li Li-Jia, Li Kai, Fei-Fei Li. ImageNet: A large-scale hierarchical image database. In: :248–255IEEE; 2009.

40. Andresen Niek, Wöllhaf Manuel, Hohlbaum Katharina, et al. Towards a fully automated surveillance of well-being status in laboratory mice using deep learning: Starting with facial expression analysis. *Plos one.* 2020;15(4):e0228059.

41. Häger C, Biernot S, Buettner M, et al. The Sheep Grimace Scale as an indicator of post-operative distress and pain in laboratory sheep. *PloS one.* 2017;12(4):e0175839.

42. Norouzzadeh Mohammad Sadegh, Nguyen Anh, Kosmala Margaret, et al. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences.* 2018;115(25):E5716–E5725.

43. Krizhevsky Alex, Sutskever Ilya, Hinton Geoffrey E.. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems.* 2012;25:1–9.

44. Maskeliunas Rytis, Raudonis Vidas, Damasevicius Robertas. Recognition of emotional vocalizations of canine. *Acta Acustica united with Acustica.* 2018;104(2):304–314.

45. Corujo Luis A, Kieson Emily, Schloesser Timo, Gloor Peter A. Emotion Recognition in Horses with Convolutional Neural Networks. *Future Internet.* 2021;13(10):250.

46. Kornblith Simon, Shlens Jonathon, Le Quoc V. Do better ImageNet models transfer better?. In: :2661–2671; 2019.

47. Vorontsov Anton O., Averkin Alexey N.. Comparison of different convolution neural network architectures for the solution of the problem of emotion recognition by facial expression. In: :35–40; 2018.

48. Walker Matthew P. The role of sleep in cognition and emotion. *Annals of the New York Academy of Sciences.* 2009;1156(1):168–197.

49. Al-Stouhi Samir, Reddy Chandan K. Transfer learning for class imbalance problems with inadequate data. *Knowledge and information systems.* 2016;48(1):201–228.

50. Ngo Quan T, Yoon Seokhoon. Facial Expression Recognition Based on Weighted-Cluster Loss and Deep Transfer Learning Using a Highly Imbalanced Dataset. *Sensors.* 2020;20(9):2639.

51. Pan Sinno Jialin, Yang Qiang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering.* 2009;22(10):1345–1359.

52. Kingma Diederik P, Ba Jimmy. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980.* 2014;.

53. Grandini Margherita, Bagli Enrico, Visani Giorgio. Metrics for multi-class classification: an overview. *arXiv preprint arXiv:2008.05756.* 2020;.

54. Chicco Davide, Jurman Giuseppe. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics.* 2020;21(1):1–13.

55. Huang Gao, Liu Zhuang, Van Der Maaten Laurens, Weinberger Kilian Q. Densely connected convolutional networks. In: :4700–4708; 2017.

56. Szegedy Christian, Ioffe Sergey, Vanhoucke Vincent, Alemi Alexander A. Inception-v4, inception-resnet and the impact of residual connections on learning. In: ; 2017.

57. He Kaiming, Zhang Xiangyu, Ren Shaoqing, Sun Jian. Identity mappings in deep residual networks. In: :630–645Springer; 2016.

58. Chollet François. Xception: Deep learning with depthwise separable convolutions. In: :1251–1258; 2017.

59. Setiawan Wahyudi, Damayanti Fitri. Layers Modification of Convolutional Neural Network for Pneumonia Detection. *Journal of Physics: Conference Series.* 2020;1477:052055.

60. Ekman Paul. Approaches to emotion. In: Scherer Klaus, Ekman Paul, eds. *Expression and the nature of emotion*, New York, USA: Psychology Press 1984 (pp. 319–344).

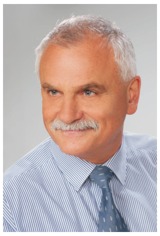
61. Wu Jeff, Ouyang Long, Ziegler Daniel M, et al. Recursively summarizing books with human feedback. *arXiv preprint arXiv:2109.10862.* 2021;.

62. Kowalczuk Zdzisław, Glinko Jan. Training of deep learning models using synthetic datasets. In: Kowalczuk Zdzislaw, ed. *Intelligent and Safe Computer Systems in Control and Diagnostics*, Lecture Notes in Networks and Systems, vol. 545: Cham: Springer IP AG 2023 (pp. 141–152).

63. Anderson Michael L. Embodied cognition: A field guide. *Artificial intelligence.* 2003;149(1):91–130.

64. Kowalczuk Zdzisław, Czubenko Michał. Embodying intelligence in autonomous and robotic systems with the use of cognitive psychology and motivation theories. In: Gawęda A.E., Kacprzyk J., Rutkowski L., Yen G., eds. *Advances in Data Analysis with Computational Intelligence Methods*, Studies in Computational Intelligence. Berlin–Heidelberg: Springer 2018 (pp. 335–352).

AUTHOR BIOGRAPHY



Zdzisław Kowalczuk (ORCID: 0000-0001-9174-546X) – Prof. DSc PhD MScEE (2003, 1993, 1986, 1978). Since 1978 he has been with the Faculty of Electronics, Telecommunications and Informatics at the Gdańsk University of Technology, where he is a Full Professor in automatic control and robotics, and the Chair of the Department of Robotics and Decision Systems. He held visiting appointments at University of Oulu (1985), Australian National University (1987), Technische Hochschule Darmstadt (1989), and at George Mason University (1990-1991). Main scientific interests include robotics, control theory, adaptation, system modeling, identification and estimation, diagnostics, failure detection, signal processing, artificial intelligence, control engineering and computer science. He has authored and co-authored about 30 books (incl. WNT 2002, PWNT 2007-2021, Springer 2004-2023), about 120 journal papers (about 50 on JCR) and over 330 conference publications and book chapters. His citation index on Google Scholar exceeded 30 hundreds. He has also about 800 JCR quotes and his Hirsch indexes range from 12 to 18). He is the President of the Polish Consultants Society (TKP) and of the Polish Society for Measurements, Automatic Control and Robotics (POLSPAR, the NMO of IFAC), and a member of the Automation and Robotics Committee of the Polish Academy of Sciences. Since 2003 professor Kowalczuk is the founder and chief editor of the publishing house PWNT – the Pomeranian Science and Technology Publishers. He is also a recipient of 1990 and 2003 Research Excellence Awards of Polish National Education Ministry, and the 1999 Polish National Science Foundation Award in automatic control, as well as the 2014 Medal of the Association of Polish Electrical Engineers SEP named after Professor Paweł Jan Nowacki.

Czubenko Michał (ORCID: 0000-0002-9483-4983) – PhD, MScEE is a 2009 graduate of the Faculty of Electronics, Telecommunications and Informatics at the Gdańsk University of Technology, in the discipline of automatic control and robotics. He is currently an adjunct in the Department of Robotics and Decision Systems of the same Faculty. In 2012, he completed a three-month internship at Kingston University London. He was a Vice-Chair of the Organizing Committee for the 12th International Conference on Diagnostics of Processes and Systems in 2015. In 2017 he defended his dissertation titled „Anthropoid Model of Intelligent Decision System for Autonomous Units" with distinction. In 2020 he was a recipient of Space and Autonomy award of the Polish Ministry of National Defence. He is currently analyzing and studying the concepts of qualia and emotions and their applications in robotics and autonomous systems. The area of his research lies mainly within the fields of artificial intelligence, robotics, and psychology. His main publications include three articles on the JCR index and eight articles indexed by the Web of Science. His quotation index at Google Scholar has exceeded 260. His non-scientific interests concern chess, playing the guitar, hiking and sailing.

Żmuda-Trzebiatowska Weronika (ORCID: 0000-0000-0000-0000) – She completed her engineering studies in January 2021 in the discipline of Automation and Robotics at the Department of Decision Systems and Robotics of the ETI Faculty of the Faculty of Electronics, Telecommunications and Informatics of the Gdańsk University of Technology. Currently, he is a graduate student at the Faculty of Electronics, Telecommunications and Informatics of the Gdańsk University of Technology.

How to cite this article: Kowalczuk Z., M. Czubenko, and W. Żmuda-Trzebiatowska (2022), Categorization of emotions in dog behavior based on the deep neural network, *Computational Intelligence*, *2022;00:1–13*.