DogFLW: Dog Facial Landmarks in the Wild Dataset

Abstract

Affective computing for animals is a rapidly expanding research area that is going deeper than automated movement tracking to address animal internal states, like pain and emotions. Facial expressions can serve to communicate information about these states in mammals. However, unlike human-related studies, there is a significant shortage of datasets that would enable the automated analysis of animal facial expressions. Inspired by the recently introduced Cat Facial Landmarks in the Wild dataset, presenting cat faces annotated with 48 facial anatomy-based landmarks, in this paper, we develop an analogous dataset containing 3,274 annotated images of dogs. Our dataset is based on a scheme of 46 facial anatomy-based landmarks.

The DogFLW dataset is available from the corresponding author upon a reasonable request.

1. Introduction

Due to the internal nature of affective states, including emotions and pain, their recognition in animals is very challenging in light of the lack of a verbal basis for communication. Nevertheless, observing subtle changes in their facial expressions and body language shows promise as a nonintrusive method for studying these states. Mammals produce facial expressions, which have been linked emotional states in a variety of species [14]. Therefore, it is unsurprising that the number of works addressing automation of animal affect recognition tasks is rapidly growing; Broome et al. [4] provides a comprehensive review of state-of-theart works using computer vision to address such recognition in animals. Despite dogs being one of the most well-studied animals in behavior studies, only a few of these works address these species, with solely one work [1] focusing on facial expressions and using a dataset collected in a controlled experimental setting.

The interest in cognitive and behavioral aspects of dogs

has been increasing dramatically [35]. Dogs can serve as valuable clinical models for numerous human disorders, owing to their large size: many canine conditions mirror human diseases, including diabetes, cancers, epilepsy, eye diseases, autoimmune diseases, and other rare diseases [27]. Other factors contributing to the widespread interest in dogs include fascination with their origins in the context of domestication, as well as their behavior and cognitive abilities. Moreover, we must improve our understanding of and regulate dog-human interactions and welfare impacts, including working dogs and shelter dogs [35].

The objective measurement of dog facial expressions is crucial for their investigation as indicators of emotional states in different contexts [3,41]. Facial expressions have also been studied in the context of understanding doghuman communication (e.g., the impact of dog facial phenotypes on their communication abilities with humans [44], the impact of facial features on the ability of humans to understand dogs [16], etc.)

The gold standard for objectively assessing changes in facial expressions in human emotion research is the Facial Action Coding System — FACS [15]. FACS has recently been adapted for different non-human species, including dogs [48]. DogFACS has been applied in several studies [2, 3, 8, 41, 44] to objectively measure facial changes. However, using this method for facial expression analysis depends on laborious manual annotation, which also requires extensive human training and certification and may be prone to human error or bias [24]. Some first steps to automated DogFACS were taken by Boneh et al. [1].

Geometric morphometrics offers an appealing alternative approach, successfully applied in analyzing cat facial features [21, 22]. This method uses points (landmarks) on objects as proxies for shape, allowing for the quantification of facial shape changes. This concept is closely related to landmarks extensively studied in the human domain.

Indeed, numerous fundamental methods for detecting, labeling, and aligning facial and body landmarks in humans have emerged [6,11,23,25,28,32,43,49]. Typically, datasets with human facial landmarks consist of thousands of images with dozens of landmarks. This abundance of data leads to better model performance, even in challenging scenarios

¹ University of Haifa, Israel

² Dogs and Science, Switzerland

³ University of Parma, Italy

Dataset	Animal	Size	Landmarks
Khan et al. [30]	Various	21,900	9
Zhang et al. [51]	Cat	10,000	9
Liu et al. [33]	Dog	8,351	8
Cao et al. [10]	Various	5,517	5
Mougeot et al. [40]	Dog	3,148	3
Martvel et al. [36]	Cat	2,091	48
Sun et al. [45]	Cat	1,706	15
Pessanha et al. [42]	Horse (tilted)	952	44
Hewitt et al. [26]	Sheep	850	25
Yang et al. [50]	Sheep	600	8
Pessanha et al. [42]	Horse (frontal)	370	54
Pessanha et al. [42]	Horse (side)	348	45
DogFLW	Dog	3,274	46

Table 1. Comparison of animal facial landmarks datasets

such as occlusions or low-quality images.

The animal domain, on the other hand, severely needs landmark-related datasets. Table 1 shows the available datasets, including their size and number of landmarks, which generally are extremely small compared to the human domain.

In the case of domesticated species, which are often artificially selected for specific features, there is a great variation in facial morphology and appearance. It is particularly noticeable in the case of cats and even more so in dogs, making it challenging to create versatile landmark detection models [9, 12, 48]. With the increase in quantity and quality of animal facial landmark datasets, it has become possible to create cross-view models [5, 10, 34], but the number of landmarks in such works is usually relatively small. To solve the problems of determining the emotions or states of animals, several dozens of landmarks are needed [17,19,20], which are unique for a particular animal.

To tackle the shortage of datasets, the Cat Facial Landmarks in the Wild (CatFLW) dataset was recently introduced by Martvel et al. [36]. The dataset comprises over 2,000 images of cat faces in various environments, each with face bounding boxes and 48 facial landmarks as detailed in [21]. Notably, the utilized landmark scheme is based on the cat's facial anatomy and the CatFACS annotation system [9]. This approach has facilitated the development of models for automated identification of pain in cats [18, 37, 38].

Dogs exhibit a wide range of facial features, making them more challenging than cats for automated facial analysis. To fill this gap, this paper introduces the Dog Facial Landmarks in the Wild (DogFLW) dataset, comprising a set of 46 facial landmarks that were defined, based on and guided by the facial anatomy of dogs and the Dog-

FACS method. We utilized the Ensemble Landmark Detector (ELD) [39] to provide a benchmark for this dataset.

2. Landmark Scheme

The development of the dog facial landmark scheme was undertaken with the aim of creating a comprehensive framework for analyzing canine facial expressions with the help of three experienced dog behavior researchers certified in DogFACS coding [48]. To establish the number of landmarks and each landmark location, the experts worked independently and then converged using expert consensus. The obtained final landmark scheme is shown in Figure 1. All details and landmark descriptions are available in the spreadsheet.

3. Dataset Properties

As a source for our dataset, we used the Stanford Dog dataset [31], which contains 20,580 images, 120 breeds and bounding boxes for dogs. Since we are interested in bounding boxes for the dog's face rather than the entire body, we did not utilize the latter in the current study.

First, we selected a random subset with an equal amount of images per breed. Then, we filtered the images according to the following criteria: the image contains a single visible dog face, where the dog is in non-laboratory conditions ('in the wild'). Other dogs could be present, but their faces shouldn't be visible for unambiguity of detection. The resulting subset contains 7-40 images per breed (25 on average) of all 120 breeds (3,274 images total), ranging in size from 100×103 to 1944×2592 pixels. Dogs in images have different sizes, colors, body and head poses, as well as different environments.

Each image is annotated with 46 facial landmarks using the CVAT platform [13] according to the scheme described



Figure 1. **Annotated Dog's Face.** Image of a dog with a face bounding box and 46 facial landmarks.

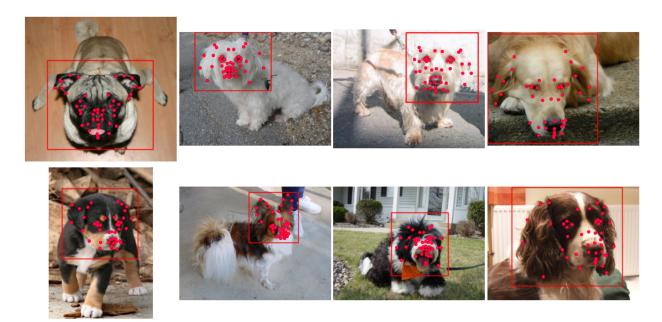


Figure 2. Examples of annotated images from the DogFLW.

above. Unlike the CatFLW [36], our dataset has occluded landmarks, all annotated landmarks have a visibility indicator ("0" — occluded, "1" — visible). Each image is also annotated with a face bounding box, which encompasses the entire face of the animal, along with approximately 10% of the surrounding space. This margin has proven crucial for training face detection models, as it prevents the cropping of important parts of the dog's face, such as the tips of the ears or the mouth. Figure 2 shows some examples of annotated images from the DogFLW dataset.

4. Benchmarks

For the evaluation, the dataset was randomly divided into the train (2,794 images) and test (480 images) sets. To provide correct metrics, we cropped all the faces by their detected bounding boxes in the preprocessing stage. All the models in this section were trained on a train set, and the final results are provided for the test set.

Metrics. We use Normalized Mean Error (NME_{iod}) that preserves the relativity of the error regardless of the size of the image or the scale of the face on it. It is commonly utilized in landmark detection [7, 46, 52], and uses MAE as the basis and inter-ocular distance (distance between the outer corners of the two eyes, IOD) for normalization:

$$NME_{iod} = \frac{1}{M \cdot N} \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{\left\| x_{i}^{j} - x_{i}^{\prime j} \right\|_{1}}{iod_{i}},$$

where M is the number of landmarks in the image, N is the number of images in the dataset, $x_i{}^j$ and $x_i'{}^j$ — the coor-

dinates of the predicted and ground truth landmark respectively.

Experimental Setup. We used the Ensemble Landmark Detector (ELD) [39] model as a landmark detector, which is justified by its high performance on the CatFLW. We used the YOLOv8 [29] model as a face detector and the EfficientNetV2S [47] model as a backbone for the ensemble. All other parameters are similar to the original article.

The face detection model was trained using standard YOLOv8n training parameters for 100 epochs. The ensemble models were trained for 300 epochs using the mean squared error loss and the ADAM optimizer with a learning rate of 10^{-4} and a batch size of 16.

To train the models, we used the Google Colab cloud service with the NVIDIA TESLA V100 GPU.

Augmentation. We randomly applied different augmentations (rotation, color balance adjustment, brightness and contrast modification, sharpness alteration, application of random blur masks, and addition of random noise) to the training data, doubling the size of the training set. Additionally, when training ensemble models, we mirrored symmetrical regions (ears and eyes) in order to further artificially increase the amount of training data.

4.1. Dataset Size

In order to examine how the size of the training dataset affects the accuracy of landmark detection, we carried out a series of experiments, training the ELD on random subsets of the training data (averaging the results for 5 iterations) and gradually increasing their size. Figure 3 indicates that the detection accuracy is approaching a plateau when the

model is trained on 65% of the data (1800 images). However, the accuracy ceiling has yet to be reached, indicating potential for further work into diversification and expanding the dataset size. The ELD model trained on the whole training set has a NME of 6.52. The detection examples are shown on Figure 4.

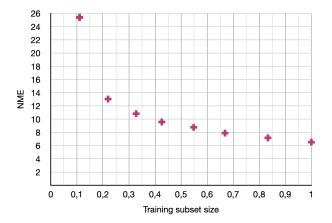


Figure 3. Impact of the training subset size on the normalised mean error. The size of the subset is measured as a fraction of the total size of the training data.

4.2. Ear Types

While conducting experiments, it was observed that the accuracy of localizing landmarks in the ears is noticeably lower compared to other regions (Table 2 shows the NME for different regions). This is presumably due to the diverse shapes and lengths of ears across different breeds, which may include previously unseen ear types and positions in the test set, particularly for floppy or half-floppy ear types.

To test our hypothesis, we divided the data into three subsets: one with erect ears (pointy), another with hanging ears (floppy), and the third with other types (half-floppy). When

Region	Eyes	Nose	Ears
NME	2.24	4.22	13.19

Table 2. Normalised mean error of landmark detection on the test set for different face regions

dividing, we were guided by the average breed ear type in a relaxed state, obtaining a ratio of 31:50:19 for the training set and 29:52:19 for the test set.

Despite the dataset's predominance of breeds with floppy ear types, Table 3 shows that the accuracy of landmark detection on a test set for such dogs is less than for dogs with erect and half-floppy ears. This can be explained by the variety of lengths and positions of such ears.

4.3. Breeds

Due to the uneven presence of breeds in the training set, the accuracy of landmark detection for rare breeds may be lower than for frequent breeds. However, it would be incorrect to say that only the number of samples affects the detection error value. The influence of the breed as an additional feature on the accuracy of detection is apparent — many breeds have distinct fur length and texture, as well as facial anatomy.

Figure 5 shows the image distribution of dogs of each breed in the training set and the detection error for the same

Ear Type	% in train set	NME total	NME ears only
Pointy	31	5.31	9.79
Floppy	50	7.18	15.28
Half-floppy	19	6.56	12.77

Table 3. Normalised mean error of landmark detection on the test set for dogs with different ear types







Figure 4. Landmark detection examples on pre-cropped images from the test set. Red — ground truth, blue — predictions.

breeds in the test set. Based on the distribution, it can be observed that breeds with long fur covering facial features or extending on ears are the most difficult to detect accurately (*Irish Water Spaniel, Briard, Standard Poodle, Bedlington Terrier, Scottish Deerhound, Komondor, Kerry Blue Terrier*). Moreover, in the Ear Types subsection, we demonstrated that detection has a high error on dogs with long, floppy ears, resulting in low accuracy in detecting facial landmarks for breeds with such ears (*Basset, Redbone*). It is also worth noting that some breeds have non-obvious detection results. In some cases, this could be explained by a significant variation in the position of the ears (*Ibizan Hound, Collie, Great Dane*) or they being obscure (*Chow*), which can cause a significant detection error.

Breeds with short smooth facial fur (*Cardigan, Kelpie, Miniature Pinscher, Dingo, Chihuahua, Malinois*, etc.), for which facial features are clearly distinguishable, have the lowest detection error on average.

5. Conclusion

The domestic dog is a highly social animal with a complex, elaborate communication system via facial signaling, which also underwent the most extreme morphological changes due to domestication and selective breeding. In this paper, we introduce a landmark scheme that is grounded in dog facial musculature. We further present an annotated dataset and benchmark detection results aimed at advancing the field of dog facial analysis. As expected, the model

has the most difficulties with ear landmarks, which can be improved by enriching the dataset with more samples with diverse ear types. A similar strategy can be taken to improve the performance on certain breeds. Our future work includes utilizing the presented scheme and dataset to classify the internal states of dogs.

The presented dataset holds great promise for canine cognition, emotion, health, and welfare research. We hope that it will aid the scientific community in utilizing AI-driven methods to deepen our understanding of our best friends' behavior and emotional world.

Acknowledgements

The research was supported by the Data Science Research Center at the University of Haifa. We thank Ephantus Kanyugi for his contribution with data annotation and management. We thank Yaron Yossef for his technical support.

References

- [1] Tali Boneh-Shitrit, Marcelo Feighelstein, Annika Bremhorst, Shir Amir, Tomer Distelfeld, Yaniv Dassa, Sharon Yaroshetsky, Stefanie Riemer, Ilan Shimshoni, Daniel S Mills, et al. Explainable automated recognition of emotional states from canine facial expressions: the case of positive anticipation and frustration. *Scientific reports*, 12(1):22611, 2022. 1
- [2] A Bremhorst, DS Mills, H Würbel, and S Riemer. Evaluating the accuracy of facial expressions as emotion indicators across contexts in dogs. *Animal cognition*, pages 1–16, 2021.

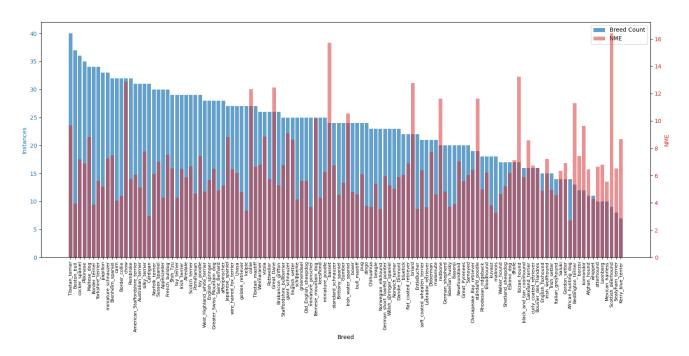


Figure 5. The distribution of images in the training set for each breed of dogs and the detection errors for the same breeds in the test set.

- [3] Annika Bremhorst, Nicole A Sutter, Hanno Würbel, Daniel S Mills, and Stefanie Riemer. Differences in facial expressions during positive anticipation and frustration in dogs awaiting a reward. Scientific reports, 9(1):1–13, 2019.
- [4] Sofia Broomé, Marcelo Feighelstein, Anna Zamansky, Gabriel Carreira Lencioni, Pia Haubro Andersen, Francisca Pessanha, Marwa Mahmoud, Hedvig Kjellström, and Albert Ali Salah. Going deeper than tracking: A survey of computer-vision based recognition of animal pain and emotions. *International Journal of Computer Vision*, 131(2):572–590, 2023. 1
- [5] Erickson R. Nascimento Bruna Vieira Frade. A two-step learning method for detecting landmarks on faces from different domains. In Proceedings of the 2018 IEEE International Conference on Image Processing (ICIP), 2018. 2
- [6] A. Bulat and G. Tzimiropoulos. Binarized convolutional landmark localizers for human pose estimation and face alignment with limited resources. *ICCV*, 2017. 1
- [7] Xavier P. Burgos-Artizzu, Pietro Perona, and Piotr Dollár. Robust face landmark estimation under occlusion. In 2013 IEEE International Conference on Computer Vision, pages 1513–1520, 2013. 3
- [8] Cátia Caeiro, Kun Guo, and Daniel Mills. Dogs and humans respond to emotionally competent stimuli by producing different facial actions. *Scientific reports*, 7(1):1–11, 2017. 1
- [9] Cátia C Caeiro, Anne M Burrows, and Bridget M Waller. Development and application of catfacs: Are human cat adopters influenced by cat facial expressions? *Applied Ani*mal Behaviour Science, 2017. 2
- [10] Jinkun Cao, Hongyang Tang, Hao-Shu Fang, Xiaoyong Shen, Cewu Lu, and Yu-Wing Tai. Cross-domain adaptation for animal pose estimation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9498– 9507, 2019. 2
- [11] X. Chu, W. Ouyang, H. Li, and X. Wang. Structured feature learning for pose estimation. *CVPR*, 2016. 1
- [12] Catia Correia-Caeiro, Kathryn Holmes, and Takako Miyabe-Nishiwaki. Extending the MaqFACS to measure facial movement in japanese macaques (macaca fuscata) reveals a wide repertoire potential. *PLOS ONE*, 16(1):e0245117, 2021. 2
- [13] CVAT.ai Corporation. Computer Vision Annotation Tool (CVAT), Nov. 2023. 2
- [14] Kris A Descovich, Jennifer Wathan, Matthew C Leach, Hannah M Buchanan-Smith, Paul Flecknell, David Farningham, and Sarah-Jane Vick. Facial expression: An underutilized tool for the assessment of welfare in mammals. ALTEX-Alternatives to animal experimentation, 34(3):409–429, 2017.
- [15] Paul Ekman and Wallace V. Friesen. *Facial Action Coding System: Manual.* Palo Alto, Calif: Consulting Psychologists Press, 1978. 1
- [16] Petra Eretová, Quanxiao Liu, Lucie Přibylová, Helena Chaloupková, Viktória Bakos, Rita Lenkei, and Péter Pongrácz. Can my human read my flat face? the curious case of understanding the contextual cues of extremely brachycephalic dogs. Applied Animal Behaviour Science, 270:106134, 2024.

- [17] Marina C Evangelista, Ryota Watanabe, Vivian SY Leung, Beatriz P Monteiro, Elizabeth O'Toole, Daniel SJ Pang, and Paulo V Steagall. Facial expressions of pain in cats: the development and validation of a feline grimace scale. *Scientific* reports, 9(1):1–11, 2019. 2
- [18] Marcelo Feighelstein, Lea Henze, Sebastian Meller, Ilan Shimshoni, Ben Hermoni, Michael Berko, Friederike Twele, Alexandra Schütter, Nora Dorn, Sabine Kästner, et al. Explainable automated pain recognition in cats. *Scientific re*ports, 13(1):8973, 2023. 2
- [19] Marcelo Feighelstein, Ilan Shimshoni, Lauren Finka, Stelio P. Luna, Daniel Mills, and Anna Zamansky. Automated recognition of pain in cats. *Scientific Reports*, 12, 2022. 2
- [20] Kim Ferres, Timo Schloesser, and Peter A Gloor. Predicting dog emotions based on posture analysis using deeplabcut. Future Internet, 14(4):97, 2022. 2
- [21] Lauren R Finka, Stelio P Luna, Juliana T Brondani, Yorgos Tzimiropoulos, John McDonagh, Mark J Farnworth, Marcello Ruta, and Daniel S Mills. Geometric morphometrics for the study of facial expressions in non-human animals, using the domestic cat as an exemplar. *Scientific reports*, 9(1):1–12, 2019. 1, 2
- [22] Lauren R Finka, Stelio PL Luna, Daniel S Mills, and Mark J Farnworth. The application of geometric morphometrics to explore potential impacts of anthropocentric selection on animals' ability to communicate via the face: The domestic cat as a case study. Frontiers in Veterinary Science, page 1070, 2020.
- [23] X. Guo, S. Li, J. Zhang, J. Ma, L. Ma, W. Liu, and H. Ling. Pfld: A practical facial landmark detector. *CoRR*, 2019.
- [24] Jihun Hamm, Christian G Kohler, Ruben C Gur, and Ragini Verma. Automated facial action coding system for dynamic analysis of facial expressions in neuropsychiatric disorders. *Journal of neuroscience methods*, 200(2):237–256, 2011.
- [25] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. CVPR, 2016.
- [26] Charlie Hewitt and Marwa Mahmoud. Pose-informed face alignment for extreme head pose variations in animals. In 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), pages 1–6. IEEE, 2019. 2
- [27] Marjo K Hytönen and Hannes Lohi. Canine models of human rare disorders. *Rare Diseases*, 4(1):e1006037, 2016. 1
- [28] H. Jin, S. Liao, , and L. Shao. Pixel-in-pixel net: Towards efficient facial landmark detection in the wild. *International Journal of Computer Vision*, 2021. 1
- [29] G. Jocher, A. Chaurasia, and J. Qiu. YOLO by Ultralytics, 2023. 3
- [30] Muhammad Haris Khan, John McDonagh, Salman H Khan, Muhammad Shahabuddin, Aditya Arora, Fahad Shahbaz Khan, Ling Shao, and Georgios Tzimiropoulos. Animalweb: A large-scale hierarchical dataset of annotated animal faces. CVRR, abs/1909.04951, 2019. 2
- [31] Aditya Khosla, Nityananda Jayadevaprakash, Bangpeng Yao, and Li Fei-Fei. Novel dataset for fine-grained image categorization. In First Workshop on Fine-Grained Visual Categorization, IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, June 2011. 2

- [32] W. Li, Y. Lu, K. Zheng, H. Liao, C. Lin, J. Luo, C.-T. Cheng, J. Xiao, L. Lu, C.-F. Kuo, and S. Miao. Structured landmark detection via topology-adapting deep graph learning. *ECCV*, 2020. 1
- [33] Jiongxin Liu, Angjoo Kanazawa, David Jacobs, and Peter Belhumeur. Dog breed classification using part localization. In European conference on computer vision, pages 172–185. Springer, 2012. 2
- [34] Yong Jae Lee M. Rashid, Xiuye Gu. Interspecies knowledge transfer for facial keypoint detection. CVPR, 2017. 2
- [35] Evan L MacLean, Aubrey Fine, Harold Herzog, Eric Strauss, and Mia L Cobb. The new era of canine science: reshaping our relationships with dogs. *Frontiers in Veterinary Science*, page 762, 2021.
- [36] George Martvel, Nareed Farhat, Ilan Shimshoni, and Anna Zamansky. Catflw: Cat facial landmarks in the wild dataset. arXiv preprint arXiv:2305.04232, 2023. 2, 3
- [37] George Martvel, Teddy Lazebnik, Marcelo Feighelstein, Lea Henze, Sebastian Meller, Ilan Shimshoni, Friederike Twele, Alexandra Schutter, Nora Dorn, Sabine Kastner, Lauren Finka, Stelio Luna, Daniel S Mills, Holger A Volk, and Anna Zamansky. Automated pain recognition in cats using facial landmarks: Dynamics matter. 2023. in review. 2
- [38] George Martvel, Teddy Lazebnik, Marcelo Feighelstein, Sebastian Meller, Ilan Shimshoni, Lauren Finka, Stelio Luna, Daniel S Mills, Holger A Volk, and Anna Zamansky. Automated landmark-based cat facial analysis and its applications. 2023. in review. 2
- [39] George Martvel, Ilan Shimshoni, and Anna Zamansky. Automated detection of cat facial landmarks. *International Jour*nal of Computer Vision, pages 1–16, 2024. 2, 3
- [40] Guillaume Mougeot, Dewei Li, and Shuai Jia. A deep learning approach for dog face verification and recognition. *Lecture Notes in Computer Science*, 2019. 2
- [41] Giulia Pedretti, Chiara Canori, Sarah Marshall-Pescini, Rupert Palme, Annalisa Pelosi, and Paola Valsecchi. Audience effect on domestic dogs' behavioural displays and facial expressions. *Scientific Reports*, 12(1):9747, 2022. 1
- [42] Francisca Pessanha, Albert Ali Salah, Thijs van Loon, and Remco Veltkamp. Facial image-based automatic assessment of equine pain. *IEEE Transactions on Affective Computing*, 2022. 2
- [43] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. CVPR, 2018.
- [44] Courtney L Sexton, Colleen Buckley, Jake Lieberfarb, Francys Subiaul, Erin E Hecht, and Brenda J Bradley. What is written on a dog's face? evaluating the impact of facial phenotypes on communication between humans and canines. *Animals*, 13(14):2385, 2023. 1
- [45] Yifan Sun and Noboru Murata. Cafm: A 3d morphable model for animals. *IEEE Winter Applications of Computer Vision Workshops (WACVW)*, 2020.
- [46] Yi Sun, Xiaogang Wang, and Xiaoou Tang. Deep convolutional network cascade for facial point detection. In 2013 IEEE Conference on Computer Vision and Pattern Recognition, pages 3476–3483, 2013. 3

- [47] Mingxing Tan and Quoc V Le. Efficientnet: Rethinking model scaling for convolutional neural networks. *CoRR*, abs/1905.11946, 2019. 3
- [48] Bridget M Waller, Kate Peirce, Cátia C Caeiro, Linda Scheider, Anne M Burrows, Sandra McCune, and Juliane Kaminski. Paedomorphic facial expressions give dogs a selective advantage. *PLoS one*, 8(12):e82686, 2013. 1, 2
- [49] Z. Xu, B. Li, Y. Yuan, and M. Geng. Anchorface: An anchorbased facial landmark detector across large poses. Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Ed- ucational Advances in Artificial Intelligence, EAAI 2021, 2021.
- [50] Heng Yang, Renqiao Zhang, and Peter Robinson. Human and sheep facial landmarks localisation by triplet interpolated features. CVRR, abs/1509.04954, 2015. 2
- [51] Weiwei Zhang, Jian Sun, and Xiaoou Tang. Cat head detection how to effectively exploit shape and texture features. ECCV, 2008. 2
- [52] Zhanpeng Zhang, Ping Luo, Chen Change Loy, and Xiaoou Tang. Learning and transferring multi-task deep representation for face alignment. *CoRR*, abs/1408.3967, 2014. 3