Check for updates

Received: 12 April 2023 | Accepted: 1 May 2023

DOI: 10.1002/cav.2164

SPECIAL ISSUE PAPER

WILEY

Personality expression in cartoon animal characters using Sasang typology

Hamila Mailee[1] | **Sinan Sonlu**[2] | **Uğur Güdükbay**[2]

[1]Department of Computer Engineering, Sharif University of Technology, Tehran, Iran

[2]Department of Computer Engineering, Bilkent University, Ankara, Turkey

Correspondence
Uğur Güdükbay, Department of Computer Engineering, Bilkent University, Ankara, Turkey.
Email: gudukbay@cs.bilkent.edu.tr

Funding information
The Scientific and Technological Research Council of Turkey (TÜBİTAK), Grant/Award Number: 122E123

Abstract

The movement style is an adequate descriptor of different personalities. While many studies investigate the relationship between apparent personality and high-level motion qualities in humans, similar research for animal characters still needs to be done. The variety in animals' skeletal configurations and texture complicates their pose estimation process. Our affect analysis framework includes a workflow for pose extraction in animal characters and a parameterization of the high-level animal motion descriptors inspired by Laban movement analysis. Using a data set of quadruped walk cycles, we prove the display of typologies in cartoon animal characters, reporting the point-biserial correlation between our motion parameters and the Sasang categories that reflect different personalities.

KEYWORDS

cartoon animation, Laban movement analysis, personality expression, Sasang typology

1 | INTRODUCTION

Many studies investigate personality implications in human motion, but a similar focus is not prominent for animals. A few related works on animal personality focus on limited species and use area-specific measurements;[1,2] they focus on the physical aspect of animal motion[3] and the movement of animal groups[4] rather than the psychological nature of the individual animal. We present a framework for personality analysis in animated animal characters. Although future work can extend our approach to real animals, we currently focus on animated animals that exhibit more exaggerated motion.

Applying the human-based solutions to animal characters is non-trivial since the skeletal configurations of the animal characters differ from their human counterparts and vary dramatically. Although a personality recognition system can input images directly, using features derived from joint positions helps reduce input dimensions, prevent bias, and eliminate external features such as the background.[5] While there are numerous human pose estimation systems,[6] the same is not valid for animals. Animal pose estimation requires manual labeling of the feature points, as animal joints vary significantly in shape and texture. We present a workflow for extracting the joints of the animal characters using a package for animal pose estimation, *DeepLabCut* (DLC).[7]

Studies suggest high-level motion qualities, such as the effort components of Laban movement analysis (LMA), influence personality perception.[8] Although LMA focuses on human motion, it is possible to adapt its parameters to generic skeletons following their verbal descriptions.[9] We devise LMA-inspired motion parameters for generic animal skeletons for analyzing animal personality. We establish high-level motion features using the linear combination of our motion

Hamila Mailee and Sinan Sonlu are joint first authors.

 © 2023 John Wiley & Sons Ltd.

parameters yielding the highest personality correlation. Consequently, we report the correlation between our high-level motion features and the personality types of the animal characters using the MGIF dataset.[10]

Various theories investigate the human personality;[11] evidence suggests similar traits are also visible in animals;[12] for example, animal gestures can signal distinct personality traits.[13] We refer to the Sasang typology[14-16] as a theoretical mapping between the biological and psychological aspects of the animal characters. To this end, we classify the animation samples based on their biopsychological features and analyze the motion parameters of each category. We report a high point-biserial correlation between various motion qualities and Sasang types, indicating that our high-level animal motion descriptors can successfully measure the psychological aspect of the character. For example, fast and energetic movements translate into high values of displacement-based motion features associated with the *SoYang* Sasang type.

This article's main contribution is a framework for personality analysis in animal characters based on Sasang typology. We also introduce LMA-inspired motion qualities highly correlated to Sasang types based on generic animal skeletons. The resulting system can help categorize the personality of animal characters. We provide our implementation in a GitHub repository for further studies in animal personality: https://github.com/hamilamailee/AnimatedCharacterPersonality

2 | RELATED WORK

Pose estimation is the initial step of personality analysis, as we perceive most traits through gestures and body movement. There are many pose estimation systems;[6] however, most human pose estimation models use a predefined fixed skeleton for detection, leaving them inadequate for animals. An appropriate model for animal pose estimation should consider the anatomy differences among various species and support user-defined skeletons. Considering the analysis of motion capture models by Mathis et al.,[17] we examine multiple models that work with animals in Appendix A. Among these models, *DeepLabCut*[18] stands out.

While all the models utilize deep learning, their methods result in performance differences. Based on transfer learning and having many pre-trained networks for species other than humans, *DeepLabCut*[7] combines *ResNets*[19] and deconvolutional layers to extract multiple animal poses from video inputs. This model also provides many options such as different frame extraction algorithms (using k-means or uniform selection), data loaders (*imgaug*[20] loader, or *tensorpack*[21]), and network types (*ResNet-50*, *ResNet-101*). These options considerably enhance the flexibility of *DeepLabCut*, making it usable for different study subjects.

Almost all pose estimation models focus on actual human or animal behavior analysis;[22] however, some studies use these models to estimate the pose of comic book characters.[23] This situation encourages us to examine the performance of these pose estimation systems on animated animal characters. Since human-based models have predefined skeletons[24] that cannot fit most animal characters, we use *DeepLabCut* in this work. A similar study by Ferres et al.[25] also uses *DeepLabCut* to predict dog emotions from the posture. While emotions explain short-term mood changes, we focus on personality, which describes long-term characteristics.

Recent studies in human personality analysis follow multi-modal approaches where models utilize sound and image-based features together.[26] The lack of full-body captures in most human-personality datasets causes difficulties for an in-depth analysis of the motion characteristics.[5] Consequently, motion capture animation datasets are more suitable for movement analysis.[27] To this end, a common approach is to label atomic animations with apparent personality traits via crowd sourcing.[8]

While we can observe human-like traits in animals and apply human personality theories to other species using modifications,[12] a standard questionnaire for measuring animal personality is lacking. Existing studies on animal personality usually utilize expert knowledge.[1,2] We refer to Sasang typology as a theoretical basis for categorizing animal characters.

Sasang typology is a categorization system based on biopsychosocial traits used in the traditional Korean medical system. Evidence shows Sasang types are related to psychological traits.[15] Previous studies successfully applied Sasang typology on expressive animated animal characters[16] and non-human robot characters.[28] Furthermore, the Sasang types are related to objective physiological features, which ease the labeling process.

Although applying a personality model that heavily depends on appearance would be unreasonable for human characters, it is relevant for animal characters where exaggeration is dominant. Animals' motion style mainly depends on physiological features, which can explain why some animal characters are related to certain stereotypes. For example,

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

due to a heavy body, the movements of an elephant are often slow, which can be associated with a calm personality. This sustained motion is reflected in the joint-level features as limited limb rotations and displacement. On the other hand, the small body of a mouse enables quick movements, which express energetic behavior. This is reflected in sudden joint position and rotation changes. To this end, our motion experiments validate the motion correspondence of Sasang typology. Sasang types determine different motion styles that we see as distinct personalities. These styles are reflected in distance, rotation, and area parameters based on joints.

3 | FRAMEWORK

Our approach has two main stages: extracting the character's joints from the input animation video and calculating high-level motion descriptors using the extracted joints.

3.1 | Pose estimation using DLC

Figure 1 demonstrates our workflow for pose estimation with DLC. In the *Project Creation* step, we define our custom skeleton (shown in Appendix B) and add the target videos. The *Frame Extraction and Labeling* step includes manually labeling the joints in chosen animation frames. DLC offers two algorithms for frame extraction: uniform and k-means. The uniform algorithm is preferred when the character has sudden movements; otherwise, k-means would be preferable. The default number of extracted frames is 20, but we changed it to 10% of the total frame count for each video. The DLC manual suggests not marking the invisible parts during labeling. It is also essential to be consistent with each joint's location and choose it ensuring visibility at most times.

Network training and evaluation step is crucial for attaining the best model for our case. We test different settings of DLC's network and augmenter types. DLC uses *ResNet-50* and the *imgaug* augmenter by default; we examine different combinations to enhance performance in animal characters. Saving and evaluating different checkpoints help us adjust the iteration count to avoid over-fitting and under-fitting. DLC generates a multi-index Pandas Array containing the joint coordinates in each frame; visualizing the joints on samples gives more insight into the network's performance.

For *outlier frame extraction*, DLC can detect frames with a sign of error in training by going through the generated coordinates, offering three methods for frame picking. *Uncertain* extracts the frame if the likelihood of one or more joints lies below the user-defined *pbound*. *Jump* detects outliers when the disposition of joints exceeds the user-defined *uf* variable. *Fitting* detects missing data and misplaced joints after fitting the auto-regressive integrated moving average model to all or some body parts.

In the *refinement* step, we adjust the labels of selected outlier frames and add the adjusted frames to the initial data set so that training can continue. We can extract outliers and refine them until the results are satisfactory, which expands the training dataset and raises the confidence level of our network's predictions. The initial results are sufficiently accurate for our purpose.

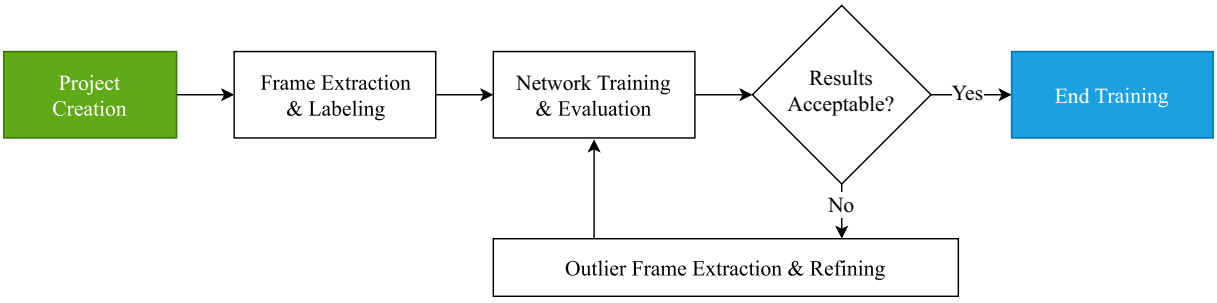

FIGURE 1 DLC's workflow for pose estimation.

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3.2 | Motion personality analysis

We are interested in finding interpretable high-level motion descriptors for animal motion. Studies show that using LMA-based motion descriptors is highly beneficial in computer-assisted movement analysis.[29] Such parameters are usually based on position, velocity, acceleration, distance, angle, and area calculations utilizing a set of joints. This study focuses on a dataset with 2D walk cycles of animated animals.[10] Consequently, we utilize features based on 2D joint distances, angles, and convex hull areas.

While LMA offers a framework for analyzing human motion, we cannot directly use it on animal characters. For example, weight effort in human motion measures the impact of gravity on movement and is highly related to agreeableness.[8] Most studies utilize the spine angle as an identifier of the weight effort; however, this does not translate directly to the animal characters, especially the quadrupeds. For example, in quadrupeds, we can identify the weight effort by the angle of the front and back legs. To this end, we analyze which combination of joints yields the highest correlation to different Sasang categories with a progressive approach.

We calculate high-level motion features using the linear combination of the selected joint pairs for each measurement type (distance, angle, and area). Progressively, we include or exclude joint pairs that increase correlation to each Sasang type. At the end of our analysis, we obtain the combination of joints that yields the best correlation with different Sasang types for each motion feature. The following section explains our motion feature selection process in detail.

4 | EXPERIMENTS AND RESULTS

We group our experiments into two main parts. First, we focus on finding the best DLC settings for extracting joints. The second part includes motion feature analysis using the MGIF dataset.[10]

4.1 | Optimizing DLC settings

Animation's differences from real-world videos (e.g., dramatic coloring changes and exaggerated movements) create difficulties in pose estimation.[22] Most studies using DLC with default settings have test videos with over 3000 frames, while ours do not exceed 500. If trained with enough data, DLC can work on other videos of the same species without re-labeling. However, since our input videos are limited in terms of the number of frames and only cover some parts of the movement, we cannot reuse a trained network for different characters. Creating a network per sample is time-consuming, so we should use settings that fasten the training process to compensate for the time spent creating new models for each character.

DLC has chosen the *ResNet-50* network and *imgaug* augmenter by default, and most projects utilize this configuration. Due to our need for fast training, we decided to test *EfficientNet*, which provides speed and performance simultaneously. On the other hand, considering the properties of our test videos, we decided to add the *tensorpack* augmenter to our tests. It has multi-CPU support, and its requirement for batchsize = 1 will not be a constraint since our training data is minimal. We also include the default augmenter and network and compare these four networks on the same videos to analyze their performance:

- *Shuffle0* (*ResNet-50* and *tensorpack* augmenter),
- *Shuffle1* (*ResNet-50* and *default* augmenter),
- *Shuffle2* (*EfficientNet-b0* and *tensorpack* augmenter), and
- *Shuffle3* (*EfficientNet-b0* and *default* augmenter).

We chose two videos of Simba from the famous animation "The Lion King" (1994). The videos *p1* and *p2* have 85 and 290 frames, respectively. We increase the number of extracted frames from the default value (20) to 29 for *p1* and 37 for *p2*. There were no sudden changes in Simba's movements in these videos, so we used the k-means algorithm for frame extraction.

Figure 2 shows the training and test results. The x-axis shows the number of iterations divided by 1000, and the y-axis represents the mean average Euclidean error (in pixels). We divide the errors into two groups, with and without

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 2 Performance of the compared networks using the same video input.

FIGURE 3 The impact of re-labeling joints. (a) Labeling differences between the first (right) and second (left) tasks. The relocation focused on the knee, elbow, wrist, and ankle joints. (b) The right image is from the first test, and the left shows relocated labels.

p-cutoff. Predicted labels with *likelihood* < *p-cutoff* are ignored in calculating the error with p-cutoff. Based on these, we infer that *Shuffle3* is the most compatible with our task; it converges rapidly and remains mostly stable after around 30k iterations. At the 30k checkpoint, *Shuffle3* exceeds *Shuffle0* in three charts; in the last one, it is outperformed by a close margin.

Using the initial number set by DLC, all videos will have 20 frames extracted regardless of length; however, this number should be chosen based on the frame count for each video separately. Following Nath et al.,[7] adjusting the training set's fraction to 10% results in acceptable performance, on par with the human eye's accuracy. We used this fraction for our second experiment to get 10 and 30 frames out of *p1* and *p2*, respectively. For this experiment, we also change the labeling location of joints, choosing the points that avoid occlusion the most. Figure 3 illustrates the labeling changes and the results of the changes. The new joints were blocked less; thus, the accuracy increased while the number of manually labeled frames decreased.

We depict the results after re-training *Shuffle3* with these modifications in Figure 4. The difference in performance is significant in the test errors, and the train errors follow a clear pattern, which makes it preferable. The results of this experiment emphasize the importance of selecting descriptive positions for the feature points, which is more significant than the number of extracted frames. The figures with "p-cutoff error" are of greater importance since many joints get occluded in an animation segment, so the probability of their absence should be considered. Finally, we determine the ideal training time as 30k iterations (details are available in Appendix C). We further evaluate the best configuration's

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

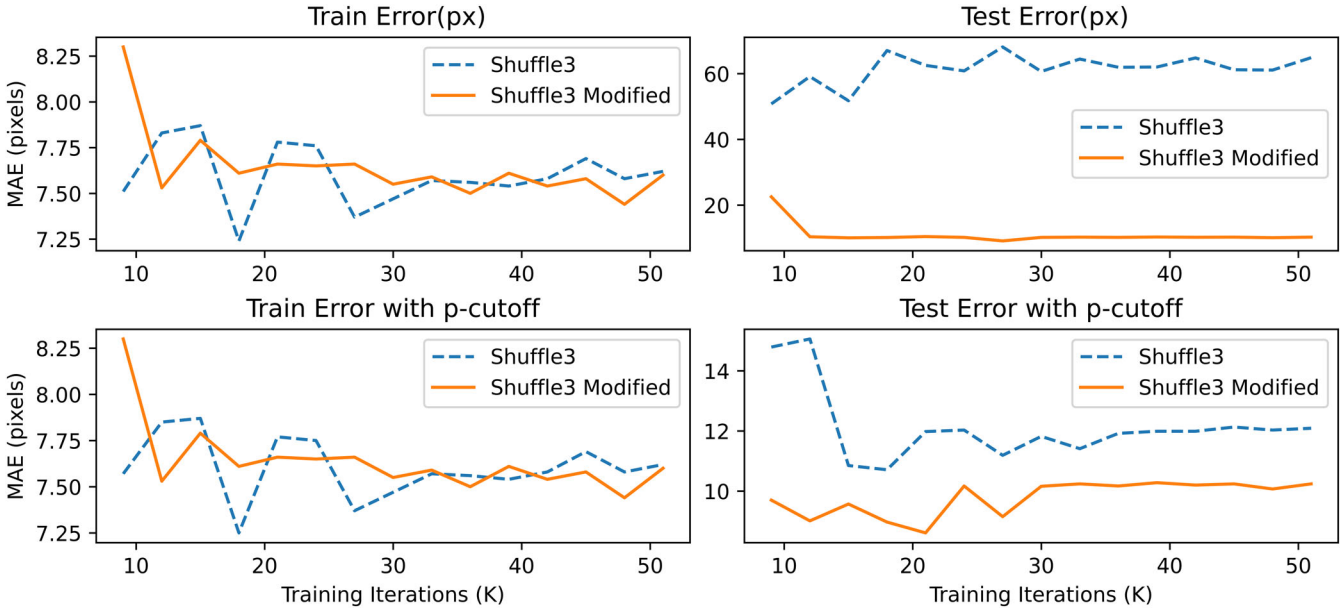

FIGURE 4 Performance of Shuffle3 before and after amendment.

FIGURE 5 Labeled samples for Sasang types. (a) SoEum. (b) SoYang. (c) TaeEum. (d) TaeYang.

performance on four additional video segments (available in Appendix D), where we obtain a test error < 12 in each case. This is the mean average Euclidean error with p-cutoff, meaning non-occluded joints are at most 12 pixels away from their actual positions, which is considerably low.

4.2 | Sasang type correlation

We use a subset of the MGIF dataset in our personality correlation experiment. As a theoretical basis for categorizing the animal character's personality, we refer to the Sasang typology, which consists of *SoEum*, *SoYang*, *TaeEum*, and *TaeYang* types.

Chae et al.[14] summarize the psychological traits of each Sasang type as follows. *SoEum* has developed hips and a weak chest; exhibits selfish, organized, and intelligent traits. *SoYang* has small hips and a developed chest; exhibits unstable, easily bored, and hot-tempered traits. *TaeEum* has a developed waist and weak neck; exhibits gentle, endurable, and humorous traits. Finally, *TaeYang* has a thin waist and developed neck; exhibits creative, positive, and heroic traits.

Based on the physical descriptions of Sasang typology,[14] we selected 22 clips, each with 13 to 25 frames. Of these 22 clips, six were of type *SoEum*, five were of type *SoYang*, six were *TaeEum*, and the remaining five clips were of type *TaeYang*. Figure 5 shows some labeled samples from the final dataset.

After labeling the selected samples and extracting their poses with DLC, we obtained 22 motion files. Before correlation analysis, we make further adjustments to this data by removing rows with a small likelihood from each file, representing data with a high probability of falsehood. We use 0.7 as the threshold. By averaging the remaining rows for

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Sasang type correlation of the features.

	SoEum	SoYang	TaeEum	TaeYang
F1	0.823	**−0.913**	**0.824**	−0.712
F2	**−0.949**	0.895	−0.802	**0.865**
F3	0.901	−0.677	−0.695	0.686
F4	0.821	−0.769	−0.428	0.655
F5	0.858	−0.749	−0.539	0.663
F6	0.799	−0.555	−0.525	0.635

Note: For each Sasang type, the highest correlated feature is given in bold.

each file, we have 22 rows in the end, with an unequal number of instances for each Sasang type. We select multiple subsets from each file using consecutive frames to extend our data and balance the sample count for all Sasang types. For calculating correlation, we consider multiple motion features. We report *Point Biserial Correlation* since motion features are continuous and the Sasang types are categorical.

We consider all possible pairs of joints to find the correlation for each feature. For each pair, we need to determine whether it has a positive impact on correlation or not. This approach results in a binary string, 1 showing every tuple that improves the correlation and 0 indicating otherwise. More detail on our motion analysis is available in Appendix E. We use the following features: *F1*: arctan angle of joint pairs, *F2*: arctan2 angle of joint pairs, *F3*: x-axis distances of joint pairs, *F4*: y-axis distances of joint pairs, *F5*: Euclidean distances of joint pairs, and *F6*: the area of the convex hull of a subset of joints.

Table 1 shows the correlation calculated from the best combination of pairs for each Sasang type. Angle features have the best correlation for all types. The *SoEum* type shows the most correlation in distance-related features. In this type, the best correlation happens when pairs containing higher body parts (such as ears, eyes, or nose) are included. Aside from that, the attributes of being short and small for *SoEum* and *SoYang* types make the distance of their body parts more impactful on distance correlations.

We summarize the joint pair combinations yielding the best correlations in Figure 6. Rows and columns indicate the related joints and the cells show the features the pair included. To simplify the figures, we combine the left and right joints for ears, eyes, elbows, hands, knees, and ankles, as well as the tail tip and tail base. Consequently, a joint paired with itself means its left and right components contribute to a feature. We use a darker color to indicate the pairs used in multiple features. The similar number of experiments between two cells does not guarantee the parity of their colors. For instance, the pairs (`ear_r`, `eye_r`) and (`ear_l`, `eye_l`) are both included in the best combination of Feature 1 (`arctan angle`) for the *TaeYang* type, meaning that the equivalent pair of (`ear`, `eye`) has to be counted twice for the relevant cell.

Figure 6a summarizes the pairs used in motion features of the *SoEum* type. We observe that the relationship between the tail and knees and the positioning of the head are essential factors for this type. We also find the relation of the elbows to the joints of the lower body to be highly influential. We observe that the joints close to each other in the skeleton hierarchy are often used as pairs for the *SoEum*, which could be due to the small body shape of this type. Figure 6b depicts the joint pairs used for the *SoYang* type. This type mainly utilizes the relationship between the hands, elbows, and head joints. Compared to *SoEum*, the role of the hands and elbows is more prominent for *SoYang*, probably due to the developed chest area of this type. The tail joints play an active role in distance-based features for *SoYang*; the small hips of animals of this type could be the reason for the highly prominent tail movements. We illustrate the features of the *TaeEum* type in Figure 6c. This type is influenced mainly by the relation of the hands to the other joints. Especially the hand-ankle relationship is very prominent in this category. Due to the wide waist area of this type, our system rarely used the lower body joints for *TaeEum*. The head joints are also very dominant for this type. Finally, we show the combinations yielding the best results for the *TaeYang* type in Figure 6d. This type is mainly related to the head and hand joints, with the lower portion of the body being less influential. The most pronounced joint pair is the eye-nose, an indicator of the head tilt. The developed neck area of the *TaeYang* can explain this type's prominent nature of head movements. The results indicate that our approach can find interpretable motion features for animal characters of different Sasang categories.

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

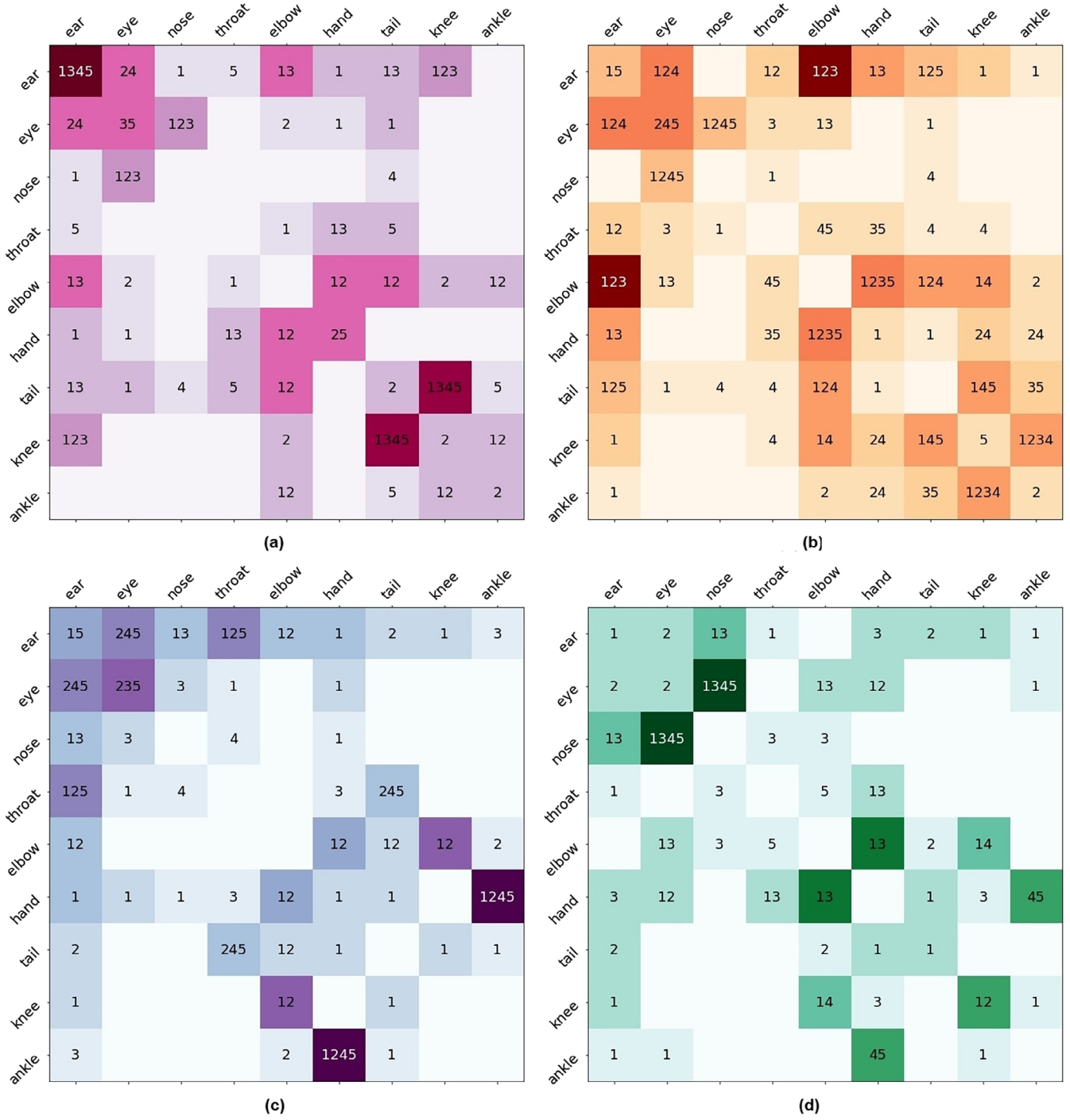

FIGURE 6 Joint pairs used in calculating motion features (F1–F5) for each Sasang type. F6 is excluded as it does not use pair data. Left-right components of the related joint are combined for ease of visualization. Numbers in each cell show the features that use the corresponding joint pair. Cells with darker colors are used in many features; these pairs are highly descriptive of the corresponding Sasang type. Each motion feature is calculated as the sum of the corresponding measurements (angle or distance) of the joint pairs it uses. The chosen pairs per feature and Sasang type are optimal for our data. (a) SoEum. (b) SoYang. (c) TaeEum. (d) TaeYang.

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

5 | CONCLUSION

We examine a well-known pose estimation system's performance on animal characters using the extracted joints in motion personality analysis. We categorize the samples of an animated walk cycle dataset following the physical descriptions of Sasang types. We report the point bi-serial correlation of different Sasang types with our motion parameters inspired by LMA. Our results show a high correlation between the angle-based motion features and the Sasang types. Although existing theories relate human motion to apparent personality, similar work for animal characters is yet to be done. Our results show that the Sasang typology can form a basis for animal characters by relating to personality and physiology. The physiological structure of the different body types results in different movement styles that we can capture using LMA-inspired motion parameters.

Different Sasang types' relation to the physiological body properties simplifies the categorization process. In human personality analysis, ground truth personality labels of the samples are usually based on the results of standard questionnaires; for animals, the lack of such standardized tests creates difficulties. Sasang typology can be an intermediary element for relating animals to personality traits based on physiological factors. Since Sasang types are strongly related to the apparent features, their automatic classification is possible. Our motion analysis can work with different categorizations to compare the performance of using Sasang types with any future theory on animal personality.

A natural extension to this work would be an automatic personality classification system that utilizes the proposed motion features for animal characters. Modifying the movements to influence certain motion features can also help with applications that generate expressive animal animation, helping animators to add an adjustable personality layer to their designs.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Animated Character Personality repository at https://github.com/hamilamailee/AnimatedCharacterPersonality.

ORCID

Hamila Mailee https://orcid.org/0000-0001-8762-055X
Sinan Sonlu https://orcid.org/0000-0002-9743-6833
Uğur Güdükbay https://orcid.org/0000-0003-2462-6959

REFERENCES

1. Yang C, Su X, Liu D, Guo Z, Wang F, Lu Y. A New method of aquatic animal personality analysis based on ML (PAML): taking swimming crab *Portunus trituberculatus* as an example. Front Mar Sci. 2020;7:32.
2. Fritz RG, Zimmermann E, Meier M, Mestre-Francés N, Radespiel U, Schmidtke D. Neurobiological substrates of animal personality and cognition in a nonhuman primate *(Microcebus murinus)*. Brain Behav. 2020;10(9):e01752.
3. Delp SL, Loan JP. A computational framework for simulating and analyzing human and animal movement. Comput Sci Eng. 2000;2(5):46–55.
4. Tang W, Bennett DA. Agent-based modeling of animal movement: a review. Geogr Compass. 2010;4(7):682–700.
5. Erkoç Z, Demirci S, Sonlu S, Güdükbay U. Skeleton-based personality recognition using Laban movement analysis. In: Palmero C, Jacques Junior JCS, Clapés A, Guyon I, Tu W-W, Moeslund TB, et al., editors. Understanding social behavior in dyadic and small group interactions. Volume 173. Maastricht, The Netherlands: ML Research Press; 2022. p. 74–87.
6. Chen Y, Tian Y, He M. Monocular human pose estimation: a survey of deep learning-based methods. Comput Vis Image Underst. 2020;192:102897.
7. Nath T, Mathis A, Chen AC, Patel A, Bethge M, Mathis MW. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. Nat Protoc. 2019;14(7):2152–76.
8. Durupinar F, Kapadia M, Deutsch S, Neff M, Badler NI. PERFORM: perceptual approach for adding OCEAN personality to human motion using Laban movement analysis. ACM Trans Graph. 2016;36(1):6.
9. Foroud A, Pellis SM. The development of "roughness" in the play fighting of rats: a Laban movement analysis perspective. Dev Psychobiol. 2003;42(1):35–43.
10. Siarohin A, Lathuilière S, Tulyakov S, Ricci E, Sebe N. Animating arbitrary objects via deep motion transfer. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE; 2019. p. 2372-81.
11. Hogan R, Sherman RA. Personality theory and the nature of human nature. Pers Individ Dif. 2020;152:109561.
12. Weiss A. Personality traits: a view from the animal kingdom. J Pers. 2018;86(1):12–22.

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

13. Spiegel O, Leu S, Bull C, Sih A. What's your move? Movement as a link between personality and spatial dynamics in animal populations. Ecol Lett. 2017;20:3–18. https://doi.org/10.1111/ele.12708
14. Chae H, Lyoo IK, Lee SJ, Cho S, Bae H, Hong M, et al. An alternative way to individualized medicine: psychological and physical traits of Sasang typology. J Altern Complement Med. 2003;9(4):519–28. https://doi.org/10.1089/107555303322284811
15. Chae H, Park SH, Lee SJ, Kim MG, Wedding D, Kwon YK. Psychological profile of *Sasang* typology: a systematic review. Evid Based Complement Alternat Med. 2009;6(s1):21–9. https://doi.org/10.1093/ecam/nep079
16. Yoon YJ, Hwang BK, Lee SJ, Lee JO, Chae H. Analysis of seven animation characters in Pororo the Little Penguin with Sasang typology. Integr Med Res. 2017;6(2):156–64. https://doi.org/10.1016/j.imr.2017.02.002
17. Mathis A, Schneider S, Lauer J, Mathis MW. A primer on motion capture with deep learning: principles, pitfalls, and perspectives. Neuron. 2020;108(1):44–65.
18. Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, et al. DeepLabCut: a software package for animal pose estimation. http://www.mackenziemathislab.org/deeplabcut (2021). Accessed 10 April 2023.
19. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Washington, DC: IEEE Computer Society. 2016. p. 770–8.
20. Jung AB, Wada K, Crall J, Tanaka S, Graving J, Reinders C, et al. Imgaug: Python library for augmenting images in ML projects. https://github.com/aleju/imgaug (2020). Accessed 10 April 2023.
21. Wu Y et al. Tensorpack. https://github.com/tensorpack/ (2016). Accessed 10 April 2023.
22. Jiang L, Lee C, Teotia D, Ostadabbas S. Animal pose estimation: a closer look at the state-of-the-art, existing gaps and opportunities. Comput Vis Image Underst. 2022;222:103483.
23. Augereau O, Iwata M, Kise K. A survey of comics research in computer science. J Imaging. 2018;4(7):87.
24. Yiannakides A, Aristidou A, Chrysanthou Y. Real-time 3D human pose and motion reconstruction from monocular RGB videos. Comput Animat Virtual Worlds. 2019;30(3-4):e1887. https://doi.org/10.1002/cav.1887
25. Ferres K, Schloesser T, Gloor PA. Predicting dog emotions based on posture analysis using DeepLabCut. Future Internet. 2022;14(4):97.
26. Aslan S, Güdükbay U, Dibeklioğlu H. Multimodal assessment of apparent personality using feature attention and error consistency constraint. Image Vis Comput. 2021;110:104163.
27. Smith HJ, Neff M. Understanding the impact of animated gesture performance on personality perceptions. ACM Trans Graph. 2017;36(4):49.
28. Sung WO, Song MJ, Chung K-w. Applying Sasang typology theory to robot appearance design. Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication. Piscataway, NJ: IEEE. 2007. p. 1066–71.
29. Aristidou A, Stavrakis E, Charalambous P, Chrysanthou Y, Himona SL. Folk dance evaluation using Laban movement analysis. J Comput Cult Herit. 2015;8(4):20.
30. Pereira TD, Tabris N, Matsliah A, Turner DM, Li J, Ravindranath S, et al. SLEAP: a deep learning system for multi-animal pose tracking. Nat Methods. 2022;19(4):486–95.
31. Pereira T, Aldarondo DE, Willmore L, Kislin M, Wang SS-H, Murthy M, et al. Fast animal pose estimation using deep neural networks. Nat Methods. 2019;16:117–25.
32. Graving JM, Chae D, Naik H, Li L, Koger B, Costelloe BR, et al. DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. eLife. 2019;8:e47994. https://doi.org/10.7554/eLife.47994
33. Liu X, Sy Y, Flierman N, Loyola S, Kamermans M, Hoogland TM, et al. OptiFlex: video-based animal pose estimation using deep learning enhanced by optical flow. Front Cell Neurosci. 2021;15:621252.
34. Xiao B, Wu H, Wei Y. Simple baselines for human pose estimation and tracking. Proceedings of the Computer Vision, Lecture Notes on Computer Science (ECCV 2018), vol 11210. Cham, Switzerland: Springer. 2018. p. 472-87.
35. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. MobileNets: efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861v1, 2017.
36. Tan M, Le Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In: Chaudhuri K, Salakhutdinov R, editors. Proceedings of the 36th International Conference on Machine Learning. Volume 97. Maastricht, The Netherlands: ML Research Press; 2019. p. 6105–14.
37. Willett NS, Shin HV, Jin Z, Li W, Finkelstein A. Pose2Pose: pose selection and transfer for 2D character animation. Proceedings of the 25th International Conference on Intelligent User Interfaces. New York, NY: ACM. 2020. p. 88–99.
38. Khungurn P, Chou D. Pose estimation of anime/manga characters: a case for synthetic data. Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding. New York: ACM. 2016. p. 1–6.
39. Chu WT, Li WW. Manga FaceNet: face detection in manga based on deep neural network. Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval. New York: ACM. 2017. p. 412-5.
40. Aberman K, Wu R, Lischinski D, Chen B, Cohen-Or D. Learning character-agnostic motion for motion retargeting in 2D. ACM Trans Graph. 2019;38(4):75.
41. Cao J, Tang H, Fang HS, Shen X, Lu C, Tai YW. Cross-domain adaptation for animal pose estimation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE. 2019. p. 9497-9506.

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mailee H, Sonlu S, Güdükbay U. Personality expression in cartoon animal characters using Sasang typology. Comput Anim Virtual Worlds. 2023;e2164. https://doi.org/10.1002/cav.2164

APPENDIX A. MOTION CAPTURE SYSTEMS

SLEAP[30] is a successor of *LEAP*,[31] using a class of fully convolutional network architectures alongside different training strategies for multi-animal projects (top-down and bottom-up). The encoder–decoder meta-architecture has considerably increased the inference speed of *SLEAP* and enables it to work competently in real-time applications.

DeepPoseKit[32] is the first to use a Stacked DenseNet model with augmented data for tracking animals. The model is generalized by data augmentation, using multiple spatial and noise transformations. One strength of *DeepPoseKit* is its proposition of a GPU-based convolutional layer called *subpixel*, which optimizes the timing of the process without sacrificing accuracy.

Unlike other models, *OptiFlex*[33] considers a sequence of frames for labeling instead of random selection. The architecture used by this model is known as *FlexibleBaseline* and is inspired by the *Simple Baselines*.[34] Another difference between this model and its previous counterparts is its evaluation metric, which uses the percentage of correct key points instead of root mean square error.

DeepLabCut[7] supports various network options (*ResNets*,[19] *MobileNets*,[35] and *EfficientNet*[36]) and augmenters (*tensorpack*,[21] *scalecrop*,[7] *imgaug*,[20] and *deterministic*[18]). We use *DeepLabCut* in our research for the following reasons:

- Most other packages adapted DLC's methods as it was the first model working on animal detection and tracking.
- Well-documented and user-friendly GUI of DLC makes labeling easier. The package also supports scripting and training via online platforms like Google Collaboratory.
- Regular maintenance and feature updates of DLC prevent possible incompatibility issues. Dependency on rapidly changing deep learning packages can cause problems if not maintained, resulting in difficulties for further studies.
- Accepting a user-defined skeleton is the feature that empowers animal models for our task. Many animal characters exhibit human-like motion but have distinct bodies, so we need to specify the different skeletons of each character.

APPENDIX B. CUSTOM SKELETON

Most studies use pose-estimation algorithms to generate motion for 2D or 3D human characters,[24,37] while a few focus on extracting the pose of animated characters. Khungurn and Chou[38] used pose estimation on Anime and Manga characters, and work focuses on face detection in Manga.[39] Current research focuses on human characters, while the animation industry comprises diverse characters. This variety encourages us to present methods that would apply to several domains. The first step in achieving this goal is using a skeleton that would fit a variety of characters. Using a standard skeleton for all animal types has the following benefits:

- We aim to analyze motion personality in animal characters using the extracted poses. Using a standard skeleton facilitates calculating the motion parameters using the same joints for each character.
- The results can then be retargeted to each other using various approaches.[40]
- Camera angle and coloring do not drastically change throughout real-world videos used in studies. In contrast, finding animation segments with consistent angles and coloring is challenging. Excluding the concern for the animal type can compensate for the difficulty of video collection.

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

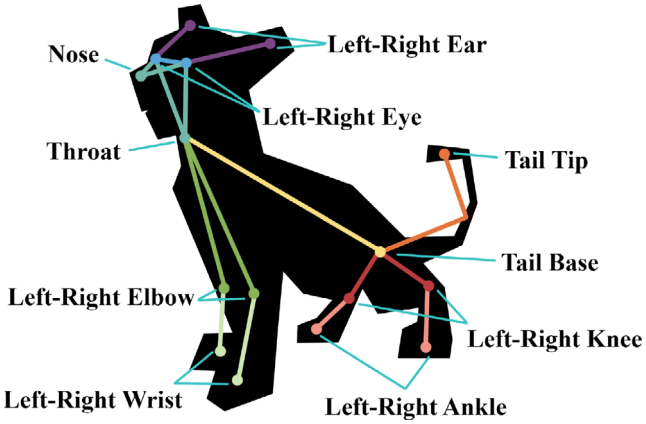

FIGURE B1 The proposed skeleton and the corresponding joint labels. This skeleton can fit different character types with minor adjustments. For instance, we can label the middle section of the wings as elbows and the tip of the wings as wrists for bird characters.

Consequently, we propose our version of the shared skeleton, shown in Figure B1, adapted from the study by Cao et al.[41] This modified skeleton can fit different characters with minor changes.

APPENDIX C. ITERATION COUNT EXPERIMENT

The last part of the experiments was determining the number of iterations needed to train our network. To check the performance, we continued the "*Shuffle3* Modified" task until the 150k point, getting the plots depicted in Figure C1. The horizontal line shows the errors in iteration 30k, our termination point for training. The arrow shows the point with the minimum error. We prefer our stopping point to be 30k instead of 27k (in "Test Error") or 21k

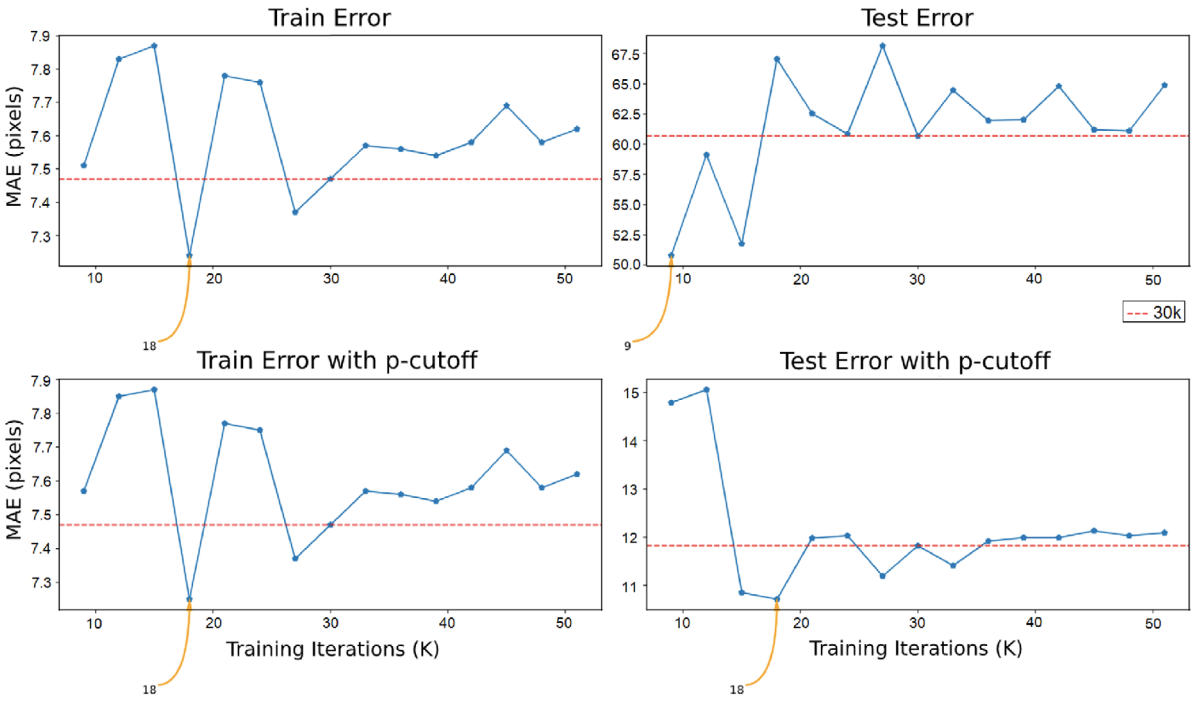

FIGURE C1 Results of continuing the training for 150k iterations. The horizontal line shows the error in the 30k iteration, our chosen termination point. Arrows point to the minimum error.

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE D1 The results of the additional experiments.

	Properties			Error	
Project name	**Frames**	**Extracted**	**p-cutoff**	**Train**	**Test**
Zootopia	153	15	0.6	5.70	6.03
Madagascar	261	30	0.6	7.45	10.43
Finding Nemo	152	20	0.6	0.9	7.51
The Secret Life of Pets	478	48	0.6	1.41	11.97

Note: We determine the number of extracted frames based on the total frame count of the input. Errors respect the p-cutoff.

(in "Test Error with p-cutoff") since the error value at these points may be undependable. Furthermore, 30k point marks the start of divergence in three charts (excluding the "Test Error"). The difference between the error of minimum checkpoints and the 30k does not exceed 10 pixels in all tables, which is an acceptable trade-off for achieving stability.

APPENDIX D. ADDITIONAL EXPERIMENTS

After finding the best network settings and the termination point, we tested three more animation cuts to evaluate the performance of our network. The chosen videos are from recent animations, "Zootopia," "Madagascar," "Finding Nemo," and "The Secret Life of Pets." Table D1 shows the settings and results of these additional experiments. The error value depends highly on the character's movements in the original clips. The error in the *Zootopia* and *Finding Nemo* experiments is relatively low since the videos sketch simple walking and swimming animations accordingly. The increased error in the remaining experiments results from complicated movements; *Alex* dancing in *Madagascar* and a dog rolling on his back in *The Secret Life of Pets*.

APPENDIX E. MOTION ANALYSIS

We use Algorithm 1 to prepare the DLC output data in our motion analysis. Then, in our motion analysis experiment, we consider all possible pairs of joints to find the correlation between our motion features and the corresponding Sasang types. We summarize this process in Algorithm 2.

Algorithm 1. The algorithm used for data preparation

files ← all *csv* files produced from DLC
final_file ← empty *csv* file
for all $f \in$ *files* **do**
  Remove rows with low likelihood from f
  Add "Sasang type" column to f
  expand ← number of needed samples from f
  for $i \leftarrow 0$ to *expand* **do**
    $temp \leftarrow f$ from row $i$to row $(len(f) - (expand - i) + 1)$
    mean ← average of all values in *temp*
    append *mean* to *final_file*
  end for
end for

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Algorithm 2. The algorithm for motion analysis. The function *calculate* operates on all samples based on the current motion feature to calculate the corresponding high-level movement parameters

```
bests ← empty array [motion features, Sasang types]
for all feature ∈ motion features and type ∈ Sasang types do
    bests[feature, type] ← empty combination set
    correlation ← calculate(bests[feature, type])
    while correlation can be improved do
        for all joint pair ∉ bests[feature, type] do
            if correlation improve by adding joint pair then
                bests[feature, type].add(joint pair)
                correlation ← calculate(bests[feature, type])
            end if
        end for
        for all joint pair ∈ bests[feature, type] do
            if correlation improve by removing joint pair then
                bests[feature, type].remove(joint pair)
                correlation ← calculate(bests[feature, type])
            end if
        end for
    end while
    report correlation and bests[feature,type]
end for
```

AUTHOR BIOGRAPHIES

Hamila Mailee is pursuing her B.S. in Computer Engineering at the Sharif University of Technology, Tehran, Iran. As a bachelor's student, she wrote her thesis on Computer Graphics and Animation at the Sharif University of Technology. Her research interests include computer vision and computer graphics, specifically the animation of virtual characters.

Sinan Sonlu received B.S. and M.S. degrees in Computer Engineering from Bilkent University, Ankara, Turkey, in 2014 and 2019, respectively. He is working toward a Ph.D. in Computer Engineering at Bilkent University. His research interests include computer vision, deep learning, computer graphics, specifically conversational agents, agent personality, virtual and augmented reality, and crowd simulation.

Uğur Güdükbay received a B.S. degree in Computer Engineering from the Middle East Technical University, Ankara, Turkey, in 1987 and the M.S. and Ph.D. degrees in Computer Engineering and Information Science from Bilkent University, Ankara, Turkey, in 1989 and 1994, respectively. He conducted research as a Postdoctoral Fellow at the Human Modeling and Simulation Laboratory at the University of Pennsylvania. Currently, he is a professor in the Department of Computer Engineering at Bilkent University. His research interests are different aspects of computer graphics, including human modeling and animation, crowd simulation, physically-based modeling, rendering, and visualization.

1546427x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cav.2164 by Bilkent University, Wiley Online Library on [18/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License