Robot Behavior Learning of Animal Locomotion with Imitation Learning

Arthur Tabary, Alexander Dittrich, Dario Floreano

Abstract—**Learning complex behaviors in high-dimensional action and state spaces, such as animal locomotion, remains a challenging research question. This paper provides an overview of recent approaches of Imitation Learning (IL) methods that leverage online videos of animals as expert demonstrations to train animal-shaped robots in the MuJoCo simulator. Drawing on existing methods, this paper proposes a modular 3-stage cohesive pipeline integrating state-of-the-art techniques in 3D Pose Estimation, Motion Retargeting, and Adversarial Imitation Learning.**

Among the studied approaches, the most reliable combination involves first performing pose estimation using DeepLabCut (DLC) [1]. This is followed by 3D skeleton reconstruction via triangulation with the AcinoSet framework [2]. Subsequently, a model-based technique such as spatio-temporal motion retargeting (STMR) [3] is employed to adapt the animal skeleton's shape to match the morphology of the target robot. This adapted skeleton then serves as expert demonstrations for training the robot in the MuJoCo simulator, utilizing the Variational Adversarial Imitation Learning from Observation (VAILO) [4] algorithm. This work also points out the main bottleneck for such a pipeline is the motion retargeting part, and it must account for the natural physical constraints of robots to ensure that the adapted motion is feasible for the target morphology.

Index Terms—**Imitation learning, Legged robots, Animal locomotion, Learning from Observation.**

I. Introduction

DEEP Reinforcement Learning (DRL) has emerged as a powerful tool for training robots to perform complex behaviors by learning from interactions with their environment. DRL leverages deep neural networks to approximate policies and value functions, enabling robots to solve intricate tasks. However, one significant drawback of DRL is its sample inefficiency, often requiring a vast amount of training data to achieve proficient performance. This inefficiency poses challenges in practical applications, such as mimicking complex behaviors of animal locomotion, where data collection can be particularly costly and time-consuming. Imitation Learning (IL) offers a more sample-efficient alternative. Inspired by the learning processes of living beings, this approach relies on expert demonstrations to guide policy search.

This paper proposes an overview of recent approaches that capitalize on online videos of animals as a source of demonstrations, facilitating the training of animal-shaped robots within the MuJoCo simulator. While IL from videos has been extensively explored for different applications [5], its implementation to legged robot remains relatively new [3], [6]–[8] in scientific literature. Drawing inspiration from existing methods, this paper proposes a unified pipeline that integrates state-of-the-art techniques including 3D pose estimation, motion retargeting and Adversarial Imitation Learning. Each of these components are supported by extensive academic research, encompassing numerous techniques. The aim of this paper is to extract from all these initiatives the most promising ones, with the goal of integrating them into a cohesive framework.

The authors are with the Laboratory of Intelligent Systems, Ecole Polytechnique Federale de Lausanne (EPFL), CH1015 Lausanne, Switzerland.

II. Related work

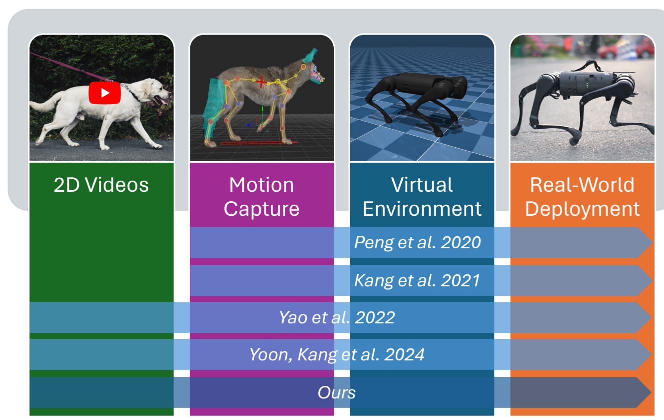

Fig. 1. Papers providing a complete pipeline for training legged robots (real or virtual) by imitating animal motion observed on a record (motion capture or 2D video) [3], [6]–[8].

Legged robots are increasingly integrated into human environments due to their ability to walk alongside people. There's growing interest in creating natural and nuanced movements beyond standard walking. IL is a successful approach for generating natural motion by mimicking prerecorded animal motions [9]–[11].

In recent advancements, *Peng et al.* [6] presents an IL framework enabling robots to acquire skills through imitation of animal Motion Capture (MOCAP) data and real-world robot control. First, the motion clip is mapped from the original subject's morphology to the robot's morphology via inverse-kinematics. Next, the retargeted reference motion is used to train a control policy via DRL on a simulated robot model. Finally, the policy is transferred to a real robot via a domain adaptation process, which adapts the policy's behavior using a learned latent dynamics representation. Nevertheless, this framework demonstrates proficiency in walking behaviors but falls short in replicating more dynamic motions like running or jumping.

Afterward, *Kang et al.* [7] explores an adaptive imitation

framework that involves analyzing animal motion data to extract key points, utilizing motion adaptation techniques, and employing DRL to train the robot to mimic the observed actions. Its effectiveness is constrained by the availability and scope of the existing dataset, potentially limiting the diversity and applicability of learned behaviors. Additionally, the framework has issues such as foot sliding effects, which can detract from the realism and effectiveness of the learned motions. Furthermore, the lack of availability of the code impedes the reproducibility and further development of the proposed framework.
Yao et al. [8] introduces a methodology based on a motion adapter, distinguishing between periodic and aperiodic motion patterns observed in the animal motion and a consistency reward integrated into the training process to ensure alignment between learned behaviors and the observed animal movements.
Finally, a recent contribution [3] presents a comprehensive pipeline for training a legged robot (A1 and go1) using only one expert video of cat from internet. The proposed methodology encompasses the entire process from animal video to real-world deployment, offering a promising avenue for leveraging vast datasets of animal movements available online. The proposed method revolves around model-based motion retargeting to ensure the generation of physically feasible motion for legged robots. This approach aims to transfer whole-body motion from baseless motion by leveraging the contact sequence, a key aspect of animal movements. Accurate contact estimation is crucial for this method, but it's challenging to achieve. If contact phases, especially determined by foot velocity thresholds, are inaccurately estimated, it can cause erratic movements, making it hard for the robot to mimic animal behaviors faithfully.

III. Methods

Based on these various works, this paper presents a new pipeline that leverages the sample efficiency of IL by using internet videos as experts (see Fig. 1). This pipeline aims to be more modular than previous ones by separating it into three distinct blocks, thereby making it more adaptable to future research results and facilitating fine-tuning.

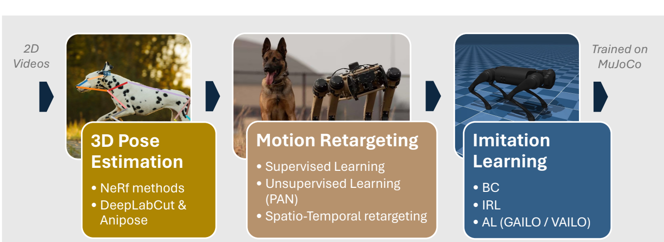

Fig. 2. Summary of state-of-the-art methods for the various stages of our pipeline [3], [4], [10], [12]–[17].

The pipeline presented in this paper consists of 3 main stages: pose estimation, motion retargeting and imitation learning. These 3 components are intended to be independent and versatile. Unlike other papers, such an interchangeable structure has the advantage of being updatable, as it is more open to future novelties. In this way, each step can be fine-tuned and the bottleneck easily identified.

A. Pose Estimation

The aim of this first step is to reconstruct in 3D the skeleton motion of an animal from a 2D video. Key points must be extracted and tracked. For example, to reproduce the motion observed on a dog video, the position and rotations of different joints of its body (hip, legs, paws, head, etc.) must be monitored. The main difficulty at this stage lies in the fact that the animal's skeleton has to be reconstructed in 3D.
There exist 2 main approaches for achieving the complete stage:

- **An unsupervised approach**: Employing a deep learning algorithm like Neural Radiance Fields (NeRFs) involves creating a 3D representation of an animal by interpreting 2D video data [10], [12], [13]. NeRF models the scene as a continuous 3D volume, using volumetric rendering to synthesize novel views and generate a detailed 3D reconstruction. This method is advantageous because it is label-free and automates the process, reducing the need for human intervention. However, NeRF requires pre-training on a large dataset and inferences are computationally intensive.
- **A supervised approach**: Using DeepLabCut (DLC) [1], key points of an animal's body are extracted and tracked in 2D based on initial human labelling. These points are then triangulated to obtain a 3D skeleton using multiple camera angles, synchronized through the wand calibration technique. This approach, although labor-intensive and requiring multiple records of the same scene, provides a highly accurate representation of the skeleton and its movements. This accuracy is crucial for the subsequent stages of motion retargeting and IL, as it directly influences the quality of the robot's learned behaviors.

In the following, these two approaches will be compared and analyzed through the results claimed by their authors.

B. Motion Retargeting

In motion imitation, overcoming the morphological differences between the source and target systems is crucial, and motion retargeting plays a key role in this process. This stage presents a significant challenge in all research efforts attempting to implement IL with an agent and an expert of differing morphologies (e.g., using a video of a dog to train an UnitreeA1 robot). Among the many approaches available in the literature, this selection is a good representation of the best methods currently available in this context.

- **Supervised learning with paired data**: Techniques such as those used by *Yamane et al.* (2010) [18], *Seol et al.* (2013) [19], and *Kim et al.* (2022) [9] directly map human motions to characters. *Yamane et al.* utilized a Gaussian process latent model, while *Seol et al.* blended

retargeted motion with the nearest motion data points. *Kim et al.* employed a mixture of experts trained on paired datasets to generate real-time quadruped motions. While these methods are straightforward and efficient in learning the mapping from source motion to retargeted motion, they require a laborious data collection process, posing challenges in scaling for numerous configurations. This method is presented here for the sake of completeness, but will not be further explored in the rest of this report.
- **Unsupervised learning**: While *Li et al.* (2023b) used key point-wise feature loss and adversarial loss, *Aberman et al.* (2020) [20] introduced a common skeleton to create an intermediate latent space shared among different kinematic structures. By combining these two techniques, *Hu et al.* (2023) employs a Pose-Aware Attention Network (PAN) [14] to dynamically extract spatial features of motion and implement body-part retargeting strategies. It enhances the process through latent space creation, adversarial training, and automatic retargeting.
- **Model-based optimization methods** At the kinematic level, retargeting consists in transferring movement by matching certain key points between the initial morphology and the final morphology, then applying inverse kinematics to calculate the parameters (position, rotation) of the joints in the targeted morphologies [21], [22]. While the kinematic methods can generate visually convincing motions, incorporating dynamic properties can significantly aid in achieving physically feasible motions and deploying robots in the real world [23], [24]. By incorporating both kinematic and dynamic aspects, Spatio-Temporal Motion Retargeting (STMR) claims to produce physically feasible movements, leveraging established methods like the Unit Vector method and Differential Dynamic Programming to ensure realistic and effective motion transfer [3].

These last two techniques are discussed in greater detail later in the paper.

C. Imitation Learning from Observation

The prevailing paradigm in IL assumes that the learner has access to both states and actions demonstrated by an expert. However, this often necessitates collecting data explicitly for IL purposes. These limiting factors have motivated recent efforts in Imitation from Observation (IfO) [25], where the expert's actions are unknown. In contrast to previous Reinforcement Learning (RL) methods, imitation from observation is a more natural way to learn from experts and is more in tune with how humans and animals approach imitation in general. Note that IfO is the core of the pipeline, whereas the previous steps of estimation and motion retargeting are just a way of feeding the algorithm with expert demonstrations from internet videos.
Here are the main approaches when it comes to IfO:

- **Behavioral Cloning (BC)** [15] involves training a model to mimic an expert's behavior by learning to map environment states to corresponding expert actions using demonstrations. While BC is computationally efficient and requires no knowledge of environment dynamics, it suffers from the covariate shift problem. This issue arises because during training, the learner is trained on states generated by the expert policy, but during testing, the learner is tested on states induced by its action. As a result, the state distribution observed during testing can differ from that observed during training. The problem with BC supervised approach is that the agent does not know how to return to the demonstrated states when it drifts and encounters out-of-distribution states [26].
- **Inverse Reinforcement Learning (IRL)** [16] involves an agent inferring the reward function from expert demonstrations and then optimizing it through RL. Unlike behavioral cloning, RL agents interact with the environment to maximize long-term rewards, making IRL less sensitive to covariate shift. However, IRL poses challenges in computational complexity, resource-intensiveness, and safety in high-risk applications due to the need for repeated environment interactions [26].
- **Adversarial Learning (AL)** [17] methods involve two competing networks: a generator and a discriminator. The generator aims to produce behavior indistinguishable from the expert's, while the discriminator differentiates between the expert's and the agent's behavior. This adversarial process helps improve policy learning, handling diverse and unstructured environments effectively, though it requires careful tuning and significant computational resources.

The latter approach exhibits promising results for Imitation from Observation (IfO) with noisy expert data. Among the existing AL algorithms, this work focuses on two of them: Generative Adversarial Imitation Learning from Observation (GAILO) [17], which is widely used and serves as a baseline, and Variational Adversarial Imitation Learning from Observation (VAILO) [4], which is an improved version. VAILO enhances GAILO by imposing a constraint on the mutual information between the observations and the discriminator's internal representation. This modification is intended to improve performance by ensuring the discriminator retains relevant information from the observations, thereby increasing robustness to noise and variations in the expert demonstrations. In summary, this study aims to evaluate and compare several existing methods from the literature across different stages of the proposed pipeline. The objective of this paper is to assess the performance and potential combinations of existing techniques.

IV. Results and Discussions

A. Pose Estimation

The results illustrating the supervised technique via DeepLabCut are obtained via the AcinoSet dataset, which tunes the raw results from DLC using an extended Kalman filter and a trajectory optimization-based method [2]. For each video, around 200 human-labelled frames (around 1 hour of work) are necessary to train the network.
The 3D NeRF reconstruction is achieved with the Lab4D framework based on the work of *Yang et al.* [10], [12], [13].

Training was done on 8 RTX 3090 GPUs for 24 hours. No human labelling is required.

The skeletons obtained through the supervised method with

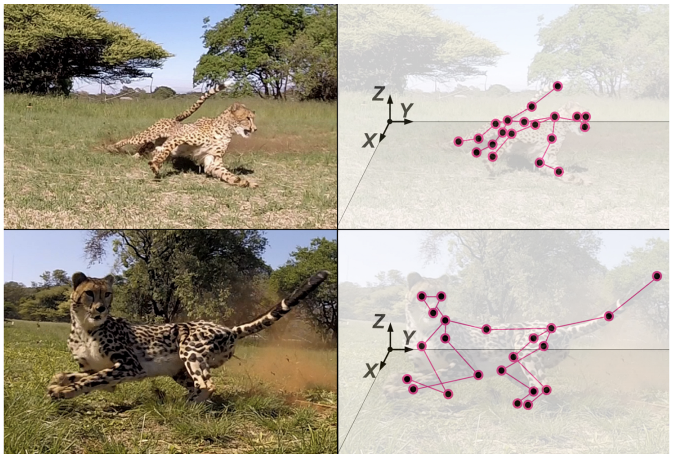

Fig. 3. 3D mesh of a running cheetah obtained by triangulating via 6 synchronized cameras the pose estimator calculated by DLC and applying an Extended Kalman Filter and a trajectory optimization-based method [2].

Fig. 4. 3D skeleton of a dog, without human intervention, by applying the NeRF approach presented in the work of Yang and al. [10], [12], [13].

DeepLabCut (DLC) more accurately reproduce the contours of the animal. Specifically, the head and various joints are qualitatively accurate, even though the mesh sometimes lacks detail, not extending to the ends of the cheetah's paw due to grass occlusion (see Fig.4). This imprecision is significant because poor estimation of the contact schedule can severely impact the performance of motion retargeting algorithms such as STMR.

On the other hand, the results obtained using the NeRFs method show visually less accurate dog skeletons, with joints not perfectly aligned with the animal's joints (see Fig.4). This is the trade-off faced when using a label-free method compared to a supervised method like DLC. Furthermore, the NeRFs models are computationally intensive and requires a substantial amount of data to accurately a category of animal (e.g. cat, dogs, etc.) [12].

B. Motion Retargeting

For this stage, only the unsupervised learning method has been tested in our configuration. Among the existing algorithms, PAN was chosen for its promising results [14]. The retargeting was performed between a dog from a public MOCAP dataset [27] and the UnitreeA1 robot, whose MOCAP data were generated from a MuJoCo simulation. The network was trained for 800 epochs on an NVIDIA V100 GPU, and the training process took 20 hours.

Despite training the model for the same number of epochs

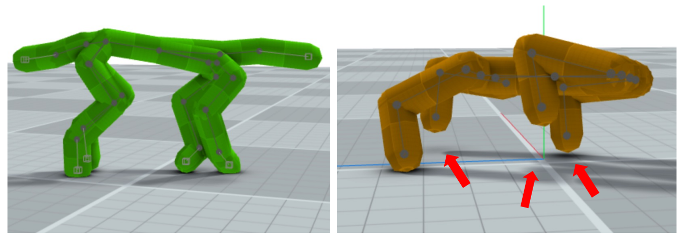

Fig. 5. Source (left) and result (right) of motion retargeting between a dog walking (from a Motion Capture dataset [27]) and the UnitreeA1 robot [28]. This result was obtained using the PAN algorithm [14] after 20h of training on the Izar GPU cluster. The red arrows highlight the non-physical behavior of the imitated movement.

as in the original paper, the performance remained suboptimal. The retargeted skeleton exhibited slow, inaccurate, and nonphysical movements, such as single-point ground contact (see Fig. 5) and foot penetration. This issue could stem from incorrect parameterization of PAN model before training, but other factors are also at play. This algorithm uses MOCAP data, which only captures the skeleton's position and joint rotations per frame. Consequently, the algorithm lacks awareness of the specific constraints of the target skeletons, such as the degrees of freedom of its joints. Additionally, the algorithm does not account for contact forces or gravity, explaining the nonphysical behavior observed in the target skeleton. These limitations highlight significant issues with the PAN algorithm for our configuration. To address these, this work suggests integrating the algorithm with knowledge of the physical constraints inherent to the target shape, such as the degree of freedom for each joint.

Yang et al. recently developed the STMR algorithm, in which such constraints are implemented [3]. By integrating techniques such as the unit vector method (UV) for kinematic motion retargeting and Differential Dynamic Programming (DDP) for dynamic motion retargeting, STMR ensures that the resulting movements not only match the desired poses, but are also physically possible, preventing foot sliding or foot penetration for example. However, these strengths come with certain limitations. Firstly, the sequential refinement process and the utilization of DDP can impose a significant computational burden, potentially requiring substantial computational resources for efficient implementation. Moreover, the complexity of the algorithm requires careful tuning and a deep understanding of both kinematic and dynamic retargeting techniques, which may pose challenges in practical application and deployment. Due to the recency of this article, this method has not yet been tested and verified in our configuration.

C. Imitation Learning

The results presented are obtained after training 400 over epochs, using 1 V100 GPU, and the LocoMuJoCo framework. For the two GAN algorithms compared, the reward function

	3D Pose Estimation		Motion Retargeting		Adversarial Imitation Learning	
Methods	**Supervised (DLC)**	**Unsupervised (NeRF)**	**Model-Based (STMR)**	**Unsupervised (PAN)**	**GAILO (baseline)**	**VAILO**
Benefits	- High accuracy - Robustness	- Label free and automation - No need for multiple camera angles	- Ensure physical feasibility - Comprehensive	- No manual intervention	- On-policy - GAIL has extensive evidences of its performance	- Achieved imitation within a reasonable timeframe
Limitations	- Labour intensive - Need calibrated cameras	- Computationally intensive for both training and inferences - Lack of precision	- Computationally intensive because of DDP - Hard to implement - Requires precise contact estimation	- No physical constraints - Heavily relies on mesh quality	- Fail in our case	

TABLE I
SUMMARY TABLE OF THE RESULTS OBTAINED IN THIS WORK IN TERMS OF BENEFITS AND LIMITATIONS FOR EACH OF THE TECHNIQUES STUDIED.

is defined in terms of the difference between the expert's and agent's center-of-mass velocities $R = \exp(-5 \cdot \|\mathbf{v}_{\text{expert}} - \mathbf{v}_{\text{agent}}\|)$.

In the context of simple walk imitation, where the agent has

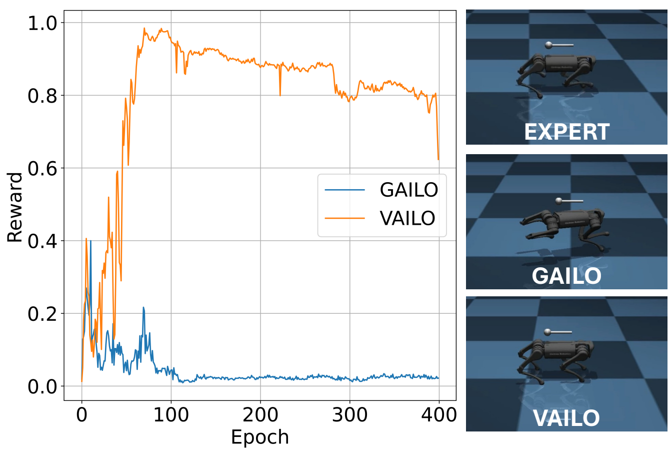

Fig. 6. Performance of the GAILO (baseline) [17] and VAILO [4] algorithms after training 400 epochs. On the left, the reward function during the training and on the right a qualitative inspection of the trained robot's motion.

access only to the expert's state at any given time, VAILO outperforms GAILO for training the same number of samples. After 400 epochs, GAILO fails to reproduce the expert robot's walking motion and merely hops in place (see Fig. 6). As the information sent to the discriminator is not constrained, GAILO might require longer training than VAILO.

Conversely, VAILO demonstrates promising results in mimicking the expert's movement. This suggests that incorporating an information bottleneck in the discriminator significantly enhances imitation, at least in our use cases. However, there are some limitations: for instance, the VAILO-trained robot lifts one leg at a time, whereas the expert moves by simultaneously lifting both opposite legs, similar to a real quadruped. This issue could potentially be addressed by adding a regularization term that accounts for contact schedules.

V. Conclusion

Imitation learning from observation using internet videos is growing fast, and there are promising results in this area. Currently, the major bottleneck is the motion retargeting step. Although this research topic has been present in the literature for decades, there is no commonly-approved method. One key insight from the experiments conducted in this paper is that motion retargeting must account for the inherent physical constraints of the target shape, such as the degrees of freedom of joints, to ensure that the adapted motion remains feasible for the retarget skeleton.

For future work, it would be interesting to test the NeRFs method and the STMR method in our configuration, and to obtain performance results when considering the complete pipeline. Moreover, the paper published in April 2024 by *Yoon et al.* [3] presents results of a complete pipeline from the motion estimation with a single 2D animal video into a legged robot deployed in the real world. The paper extensively details their method for the motion retargeting, and the publicly available code makes it an ideal candidate to extend the tests carried out in this report.

References

[1] A. Mathis, P. Mamidanna, K. M. Cury, T. Abe, V. N. Murthy, M. W. Mathis, and M. Bethge, "Deeplabcut: markerless pose estimation of user-defined body parts with deep learning," *Nature Neuroscience*, 2018. [Online]. Available: https://www.nature.com/articles/s41593-018-0209-y

[2] D. Joska, L. Clark, N. Muramatsu, R. Jericevich, F. Nicolls, A. Mathis, M. W. Mathis, and A. Patel, "Acinoset: A 3d pose estimation dataset and baseline models for cheetahs in the wild," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2021. [Online]. Available: http://dx.doi.org/10.1109/ICRA48506.2021.9561338

[3] T. Yoon, D. Kang, S. Kim, M. Ahn, S. Coros, and S. Choi, "Spatio-temporal motion retargeting for quadruped robots," 2024.

[4] X. B. Peng, A. Kanazawa, S. Toyer, P. Abbeel, and S. Levine, "Variational discriminator bottleneck: Improving imitation learning, inverse rl, and gans by constraining information flow," 2020.

[5] R. McCarthy, D. C. H. Tan, D. Schmidt, F. Acero, N. Herr, Y. Du, T. G. Thuruthel, and Z. Li, "Towards generalist robot learning from internet video: A survey," 2024.

[6] X. B. Peng, E. Coumans, T. Zhang, T.-W. Lee, J. Tan, and S. Levine, "Learning agile robotic locomotion skills by imitating animals," 2020.

[7] D. Kang, S. Zimmermann, and S. Coros, "Animal gaits on quadrupedal robots using motion matching and model-based control," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 8500–8507.

[8] Q. Yao, J. Wang, S. Yang, C. Wang, H. Zhang, Q. Zhang, and D. Wang, "Imitation and adaptation based on consistency: A quadruped robot imitates animals from videos using deep reinforcement learning," in *2022 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2022, pp. 1414–1419.

[9] S. Kim, M. Sorokin, J. Lee, and S. Ha, "Humanconquad: Human motion control of quadrupedal robots using deep reinforcement learning," in *SIGGRAPH Asia 2022 Emerging Technologies*, ser. SA '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3550471.3564762

[10] G. Yang, M. Vo, N. Neverova, D. Ramanan, A. Vedaldi, and H. Joo, "Banmo: Building animatable 3d neural models from many casual videos," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 2853–2863.

[11] X. B. Peng, P. Abbeel, S. Levine, and M. van de Panne, "Deepmimic: example-guided deep reinforcement learning of physics-based character skills," *ACM Trans. Graph.*, vol. 37, no. 4, jul 2018. [Online]. Available: https://doi.org/10.1145/3197517.3201311

[12] G. Yang, C. Wang, N. D. Reddy, and D. Ramanan, "Reconstructing animatable categories from videos," in *CVPR*, 2023.

[13] C. Song, G. Yang, K. Deng, J.-Y. Zhu, and D. Ramanan, "Total-recon: Deformable scene reconstruction for embodied view synthesis," 2023.

[14] L. Hu, Z. Zhang, C. Zhong, B. Jiang, and S. Xia, "Pose-aware attention network for flexible motion retargeting by body part," *IEEE Transactions on Visualization and Computer Graphics*, 2023.

[15] F. Torabi, G. Warnell, and P. Stone, "Behavioral cloning from observation," 2018.

[16] A. Jaegle, Y. Sulsky, A. Ahuja, J. Bruce, R. Fergus, and G. Wayne, "Imitation by predicting observations," in *Proceedings of the 38th International Conference on Machine Learning, ICML 2021*, ser. Proceedings of Machine Learning Research. ML Research Press, 2021, pp. 4665–4676, publisher Copyright: Copyright © 2021 by the author(s); 38th International Conference on Machine Learning, ICML 2021 ; Conference date: 18-07-2021 Through 24-07-2021.

[17] F. Torabi, G. Warnell, and P. Stone, "Generative adversarial imitation from observation," 2019.

[18] K. Yamane, Y. Ariki, and J. Hodgins, "Animating non-humanoid characters with human motion data," in *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, ser. SCA '10. Goslar, DEU: Eurographics Association, 2010, p. 169–178.

[19] Y. Seol, C. O'Sullivan, and J. Lee, "Creature features: online motion puppetry for non-human characters," in *Proceedings of the 12th ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, ser. SCA '13. New York, NY, USA: Association for Computing Machinery, 2013, p. 213–221. [Online]. Available: https://doi.org/10.1145/2485895.2485903

[20] K. Aberman, P. Li, D. Lischinski, O. Sorkine-Hornung, D. Cohen-Or, and B. Chen, "Skeleton-aware networks for deep motion retargeting," *ACM Transactions on Graphics*, vol. 39, no. 4, Aug. 2020. [Online]. Available: http://dx.doi.org/10.1145/3386569.3392462

[21] K.-J. Choi and H.-S. Ko, "On-line motion retargeting," *Journal of Visualization and Computer Animation*, vol. 11, pp. 223–235, 12 2000.

[22] S. Choi and J. Kim, "Towards a natural motion generator: a pipeline to control a humanoid based on motion data," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 4373–4380.

[23] R. Grandia, F. Farshidian, E. Knoop, C. Schumacher, M. Hutter, and M. Bächer, "Doc: Differentiable optimal control for retargeting motions onto legged robots," *ACM Trans. Graph.*, vol. 42, no. 4, jul 2023. [Online]. Available: https://doi.org/10.1145/3592454

[24] M. Al Borno, L. Righetti, M. J. Black, S. L. Delp, E. Fiume, and J. Romero, "Robust physics-based motion retargeting with realistic body shapes," *Computer Graphics Forum*, vol. 37, no. 8, pp. 81–92, 2018. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.13514

[25] Y. Liu, A. Gupta, P. Abbeel, and S. Levine, "Imitation from observation: Learning to imitate behaviors from raw video via context translation," 2018.

[26] M. Zare, P. M. Kebria, A. Khosravi, and S. Nahavandi, "A survey of imitation learning: Algorithms, recent developments, and challenges," 2023.

[27] H. Zhang, S. Starke, T. Komura, and J. Saito, "Mode-adaptive neural networks for quadruped motion control," *ACM Transactions on Graphics (TOG)*, vol. 37, pp. 1 – 11, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:51692385

[28] [Online]. Available: https://www.unitree.com/a1/