Can a human-like locomotion policy be learned ONLY using biological actuation and rewards, without the need for reference motion clips?

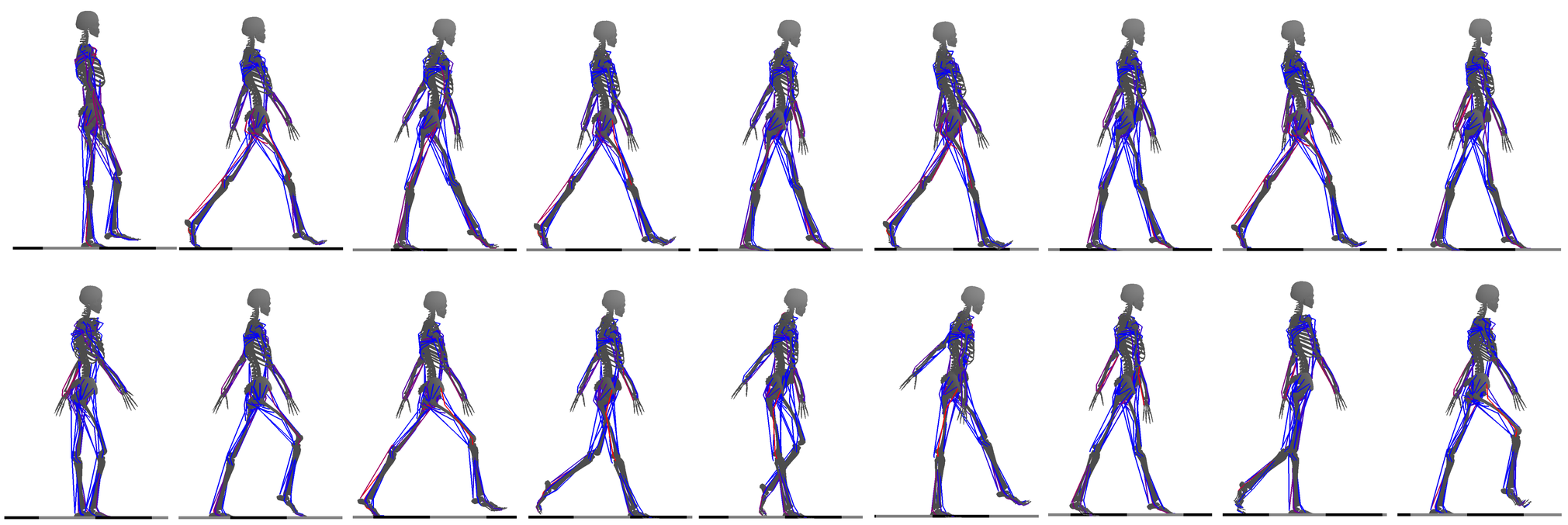
Answer: Yes. We propose a method of learning a policy for human-like locomotion via **DRL** based on a **human anatomical model, muscle actuation, and biologically inspired rewards, without any inherent control rules or reference motions.**

Key Point 1: Two-stage approach for metabolic energy reward
We first train the policy using dense energy reward,
and then fine-tune the policy with sparse energy reward.

Key Point 2: Randomized one leg-lifting starting pose
We introduce a randomized starting position of lifting either the left or right leg at the beginning of each episode.

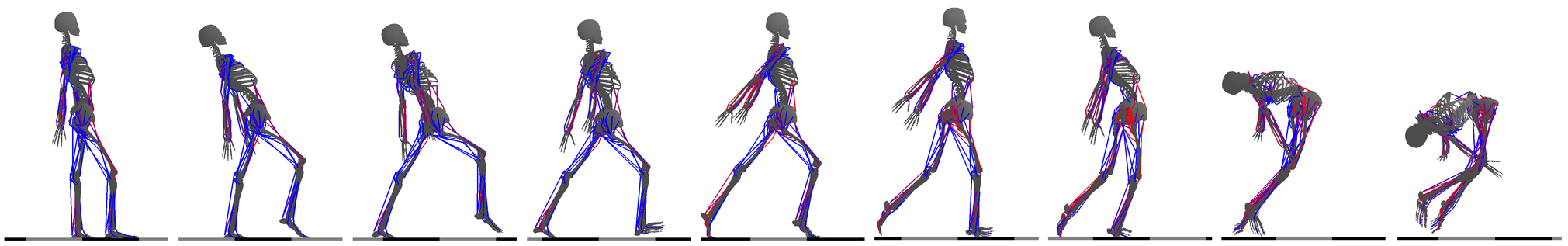
for more details!

(Background: The conventional dominant approach is to control muscle-actuated characters to imitate reference motion clips, but it tends to generate limited resulting motion.)

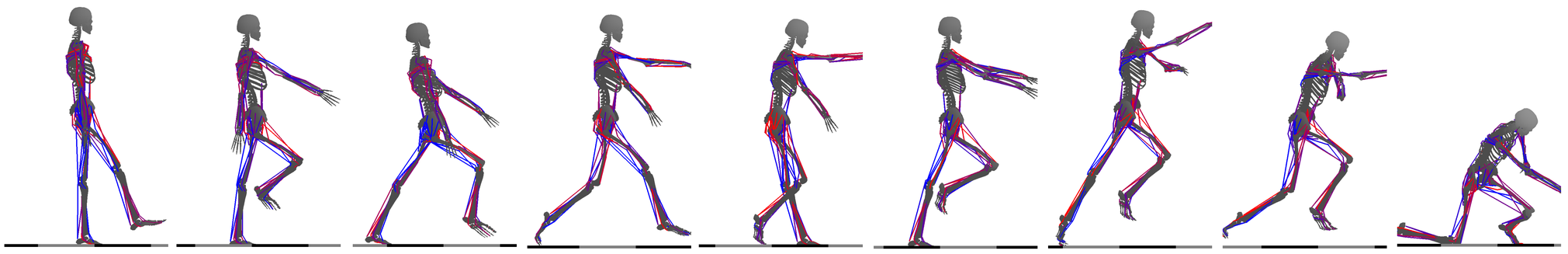


Our best result.
We could train stable, human-like locomotion by sequentially using two types of energy rewards over two stages. (First use MET and then fine-tune with CoT)

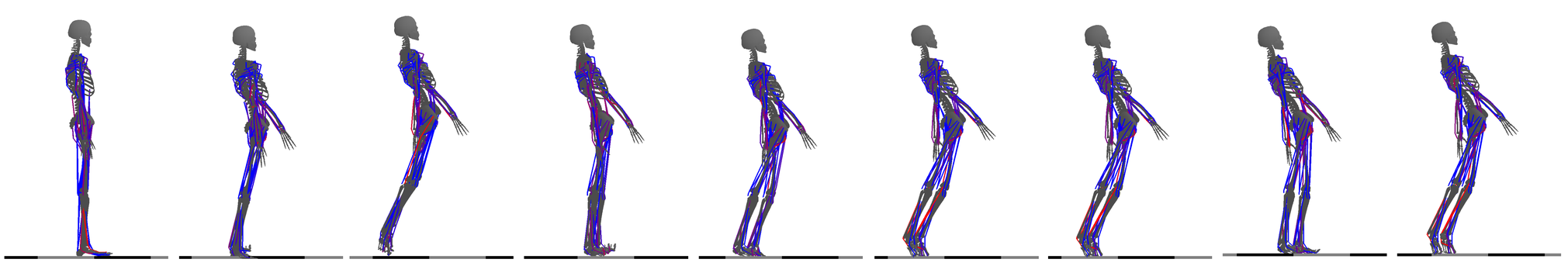
Dense energy (MET) only.
With the dense reward (MET) only, the policy converged after taking only four to five walking steps.



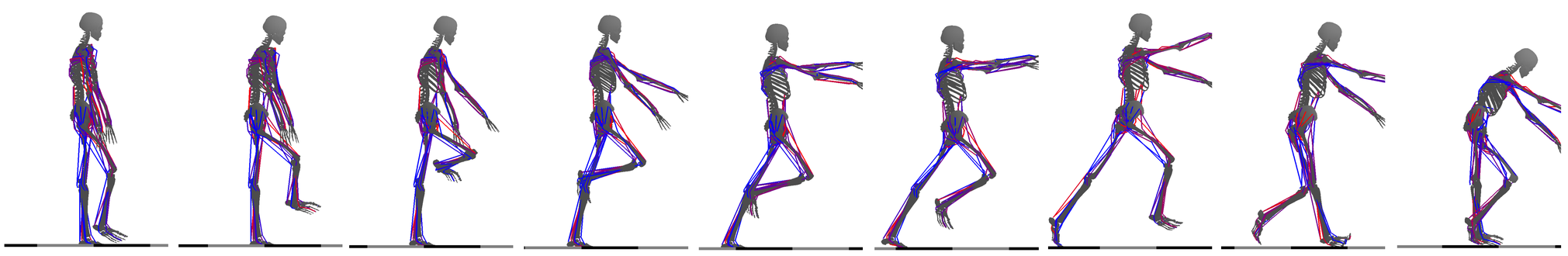
Sparse energy (CoT) only.
With the sparse reward (CoT) only, the policy failed to learn how to move forward and fell easily.



No energy reward.
When no energy reward was used, the policy did not learn how to move at all, resulting in the humanoid falling over immediately.



Double-stance staring pose.
Beginning episodes with both feet in contact led to a learned policy that involved repetitive jumping with both feet.



Activation reward.
The humanoid could not move forward properly by a policy that learned with the mean squared sum of muscle activations as the energy reward.

Learning Human-like Locomotion with Biological Actuation and Rewards

Minkwan Kim, Yoonsang Lee
Hanyang University

SIGGRAPH 2023
LOS ANGELES+ 6-10 AUG