25 December 2024 · CC-BY 4.0

Peer Review

Review of: "Continuous Video Process: Modeling Videos as Continuous Multi-Dimensional Processes for Video Prediction"

Tasin Islam[1]

[1]. Brunel University Uxbridge, United Kingdom

The author introduces a model that synthesises video from a continuous multi-dimensional process rather than from a series of discrete images.

The paper argues that "transitions between consecutive frames in a video do not uniformly contain the same amount of motion," and this variation depends on the frame rate. Is this issue relevant when a video is set to 60 frames per second, where consecutive frames tend to exhibit similar motion?

References are needed when mentioning that temporal attention misses the natural flow of motion in the video.

In Figures 1 and 2, can you explain how the intermediate time step equation works? Does it contain merged images of x and y of different strengths depending on t + white Gaussian noise? Please confirm.

You have repeated a bullet point in the summary contribution section of the introduction.

Please provide a reference to: "Majorly, methods falling under this category enforce temporal consistency using artificial external constraints such as the introduction of temporal attention blocks. This might be effective but comes at a cost of significant computing power." Which papers state that using temporal attention requires high computational power?

In section 2, briefly explain how InDI and Cold diffusion are similar to your approach.

In section 5, you only used the last four frames of the KTH dataset as context frames to predict upcoming frames. Is there any performance difference if you were to use more frames?

doi.org/10.32388/BAWHBR

Table 1 indicates that your model produces only one frame at a time, and it will be reused to generate longer video sequences. I was wondering, how long does it take your model, as well as previous work, to synthesise 40 frames?

Declarations

Potential competing interests: No potential competing interests to declare.

doi.org/10.32388/BAWHBR