SPRINGER LINK

Log in

Menu | Search   Cart

VTM-GAN: video-text matcher based generative adversarial network for generating videos from textual description

Original Research   Published: 16 September 2023

(2023)   Cite this article

International Journal of Information Technology

Aims and scope

Submit manuscript

Rayeesa Mehmood, Rumaan Bashir ✉ & Kaiser J. Giri

Abstract

Text-to-video synthesis has garnered significant attention as a challenging task in the domain of vision computing. With the advent of unsupervised learning techniques, text-to-video synthesis has become more feasible. In this context, Generative Adversarial Network (GAN)–based training networks have emerged as the leading unsupervised deep learning methods, exhibiting promising results. However, achieving visual quality, temporal coherence, and semantic consistency between the generated video and textual descriptions remains a considerable challenge. In this paper, we propose a novel approach called Video-Text Matcher (VTM) based GAN for text-to-video synthesis. The proposed VTM is based on Contrastive

Language-Image Pre-training (CLIP) but with modifications. It incorporates both global sentence-level and fine-grained word-level information to calculate the similarity between the generated video and the provided textual descriptions. Unlike CLIP, which focuses on matching losses at the global sentence-image level only, our VTM includes a word-region level loss to enhance the fine granularity consistency between the text and video. We evaluate our proposed approach using the Single Digit Bouncing MNIST GIFs (SBMG) dataset and conduct both qualitative and quantitative analyses. The results demonstrate that our proposed method generates appealing videos that align well with the given textual descriptions, showcasing the effectiveness of our approach for text-to-video synthesis.

ℹ️ This is a preview of subscription content, <u>log in via an institution</u> ↗ to check access.

Access this article

Log in via an institution

Buy article PDF 39,95 €

Price includes VAT (India)

Instant access to the full article PDF.

Rent this article via <u>DeepDyve</u> ↗

Institutional subscriptions →

Data availability

Not applicable.

References