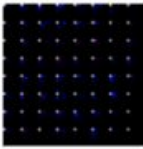

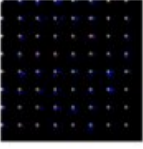



T2CI GAN: A deep learning model that generates compressed images from text

October 26 2022, by Ingrid Fadelli

	DCT image in dataset	After applying IDCT
Before Transformation		
After Transformation		

The significance of applying the simple transformation on JPEG Compressed DCT images. Credit: Rajesh et al.

Generative adversarial networks (GANs), a class of machine learning frameworks that can generate new texts, images, videos, and voice recordings, have been found to be highly valuable for tackling numerous real-world problems. For instance, GANs have been successfully used to generate image datasets to train other deep learning algorithms, to generate videos or animations for specific uses, and to create suitable captions for images.

Researchers at the Computer Vision and Biometrics Lab of IIT Allahabad and Vignan University in India have recently developed a new GAN-based model that can generate compressed images from text-based descriptions. This model, introduced in a paper pre-published on arXiv, could open interesting possibilities for image storage and for the sharing of content between different smart devices.

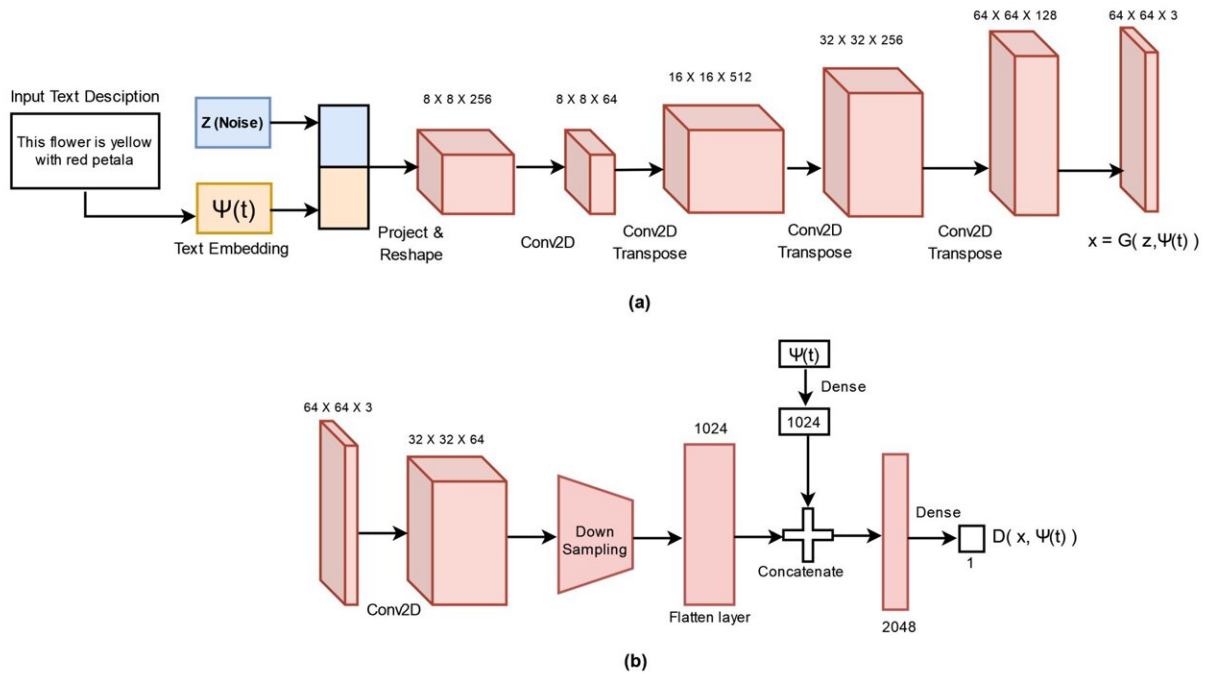
"The idea of T2CI GAN is aligned with the theme of 'direct processing/analytics of data in the compressed domain without full decompression,' on which we have been working on since 2012," Mohammed Javed, one of the researchers who carried out the study, told TechXplore. "However, the idea in T2CI GAN is a bit different, as here we wanted to produce/retrieve images in the compressed form given the text descriptions of the image."

In their past studies, Javed and his colleagues used GANs and other [deep learning models](#) to tackle numerous tasks, including the extraction of features from data, the segmentation of text and [image data](#), spotting words in large text excerpts, and to create compressed JPEG files. The new model they created builds on these previous efforts to address a computational problem that has so far been rarely explored in the literature.

While several other research teams have used deep learning-based methods to generate images based on text descriptions, only a few of these methods produce images in their compressed form. In addition, most existing techniques that generate compressed images approach the task of generating the image and compressing it separately, which increases their computation load and processing time.

"T2CI-GAN is a deep learning-based model that takes text descriptions as an input and produces visual images in the compressed form," Javed explained. "The advantage here is that the conventional methods produce

visual images from text descriptions, and they further subject those images to compression, to produce compressed images. Our model, on the other hand, can directly map/learn the text descriptions and produce compressed images."



The proposed T2CI-GAN Model-1 architecture using backbone networks. (a) Generator network and (b) Discriminator network. Credit: <https://arxiv.org/abs/2210.03734>

Javed and his colleagues developed two distinct GAN-based models for generating compressed images from text descriptions. The first of these models was trained on a dataset containing compressed DCT (discrete cosine transform) images in the JPEG format. After training, this model was able to generate compressed images based on text descriptions.

The researchers' second GAN-based model, on the other hand, was trained on a set of RGB images. This model learned to generate JPEG compressed DCT representations of images, which specifically express a sequence of data points as a mathematical equation.

"T2CI-GAN is the future, because we know that the world is moving towards machine(robot) to machine and man to machine communications," Javed said. "In such a scenario, machines only need data in the compressed form to interpret or understand them. For example, imagine that a person is asking Alexa bot to send her childhood photo to her best friend. Alexa will understand the person's voice message (text description) and try to search for this photo, which would already be stored somewhere in the compressed form, and send it directly to her friend."

Javed and his colleagues evaluated their model in a series of tests, using the renowned Oxtford-102 Flower dataset, which contains several pictures of flowers, categorized into 102 flower types. Their results were highly promising, as their model could generate compressed JPEG versions of images in the flower dataset both quickly and efficiency.

The T2CI-GAN model could be used to improve automated image retrieval systems, particularly when sourced images are meant to be easily shared with smartphones or other smart devices. In addition, it could prove to be a [valuable tool](#) for media and communications professionals, helping them to retrieve lighter versions of specific images to share on online platforms.

"Currently, the T2CI GAN model produces images only in JPEG compressed form," Javed added. "In our future work, we would like to see whether we can have a general model that can produce images in any compressed form, without any constraint of compression algorithm."

More information: Bulla Rajesh, Nandakishore Dusa, Mohammed Javed, Shiv Ram Dubey, P. Nagabhushan, T2CI-GAN: Text to compressed image generation using generative adversarial network. arXiv:2210.03734v1 [cs.CV], arxiv.org/abs/2210.03734

© 2022 Science X Network

Citation: T2CI GAN: A deep learning model that generates compressed images from text (2022, October 26) retrieved 6 May 2024 from <https://techxplore.com/news/2022-10-t2ci-gan-deep-compressed-images.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.