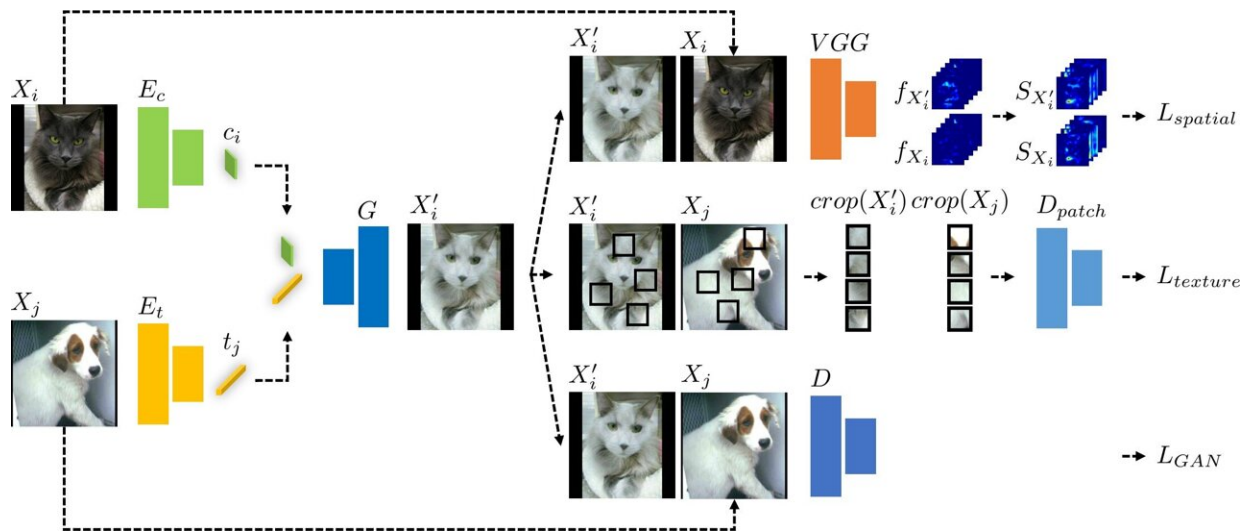


Research team develops an AI model for effectively removing biases in a dataset

October 31 2023



Model structure devised by DGIST Professor Park Sang-hyun's research team.
Credit: DGIST Professor Park Sang-hyun's research team

The research team of Professor Sang-hyun Park, at Daegu Gyeongbuk Institute of Science and Technology's (DGIST) Department of Robotics and Mechatronics Engineering, has developed a new image translation model that could effectively reduce biases in data.

In the process of developing an artificial intelligence (AI) model using images collected from different sources, contrary to the user's intention, data biases may occur because of various factors. The developed model

can remove data biases despite the absence of information on such factors, thereby providing a high image-analysis performance. This solution is expected to facilitate innovations in the fields of self-driving, content creation, and medicine.

The datasets used to train deep learning models tend to exhibit biases. For example, when creating a dataset to distinguish bacterial pneumonia from coronavirus disease 2019 (COVID-19), image collection conditions may vary because of the risk of COVID-19 infection. Consequently, these variations result in subtle differences in the images, causing existing deep learning models to discern diseases based on features stemming from differences in image protocols rather than the critical characteristics for practical disease identification.

In this case, these models exhibit [high performance](#) based on data used for their training process. However, they show limited performance on data obtained from [different places](#) owing to their inability to effectively generalize, which can lead to over-fitting issues. Particularly, existing deep learning techniques tend to use differences in textures as crucial data, which may lead to inaccurate predictions.

To address these challenges, Prof. Park's research team developed an image translation model that could generate a dataset applying texture debiasing and perform the learning process based on the generated dataset.

Existing image translation models are often limited by the issue of texture changes leading to unintended content alterations, as textures and contents are intertwined. To address this issue, Prof. Park's research team developed a new model that simultaneously uses error functions for both textures and contents. The work is [published](#) in the journal *Neural Networks*.

The new image translation model proposed by this research team operates by extracting information on the contents of an input image and on textures from a different domain and combining them.

To simultaneously maintain information on not only the contents of input images but also the texture of the new domain, the developed model is trained using both error functions for spatial self-similarity and texture co-occurrence. Through these processes, the model can generate an image that has the texture of a different domain while maintaining information on the contents of the input image.

As the developed deep learning model generates a dataset applying texture debiasing and uses the generated dataset for training, it exhibits a better performance than the existing models.

It achieved a superior performance compared to existing debiasing and image translation techniques when tested on datasets with texture biases, such as a classification dataset for distinguishing numbers, classification dataset for distinguishing dogs and cats with different hair colors, and classification dataset applying different image protocols for distinguishing COVID-19 from bacterial pneumonia. Moreover, it outperformed existing methods when applied to datasets with various biases, such as a classification dataset for distinguishing multi-label numbers and that for distinguishing photos, images, animations, and sketches.

Furthermore, the image translation technology proposed by Prof. Park's research team can be implemented in image manipulation. The research team found that the developed method altered only the textures of an image while preserving its original contents. This analytic result confirmed the superior performance of the developed method compared to existing image manipulation methods.

Additionally, this solution can be effectively used in other environments. The research team compared the performance of the developed method with that of existing image translation methods based on various domains, such as medical and self-driving images. Based on the analytical results, the developed method demonstrated a better performance than existing methods.

Prof. Park said, "The technology developed in this research offers a significant performance boost in situations where biased [datasets](#) are inevitably used to train [deep learning models](#) in industrial and medical fields."

He added, "It is expected that this solution will make a substantial contribution to enhance the robustness of AI models commercially used or distributed in diverse environments for commercial purposes."

More information: Myeongkyun Kang et al, Content preserving image translation with texture co-occurrence and spatial self-similarity for texture debiasing and domain adaptation, *Neural Networks* (2023). [DOI: 10.1016/j.neunet.2023.07.049](#)

Provided by Daegu Gyeongbuk Institute of Science and Technology

Citation: Research team develops an AI model for effectively removing biases in a dataset (2023, October 31) retrieved 28 April 2024 from <https://techxplore.com/news/2023-10-team-ai-effectively-biases-dataset.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.