

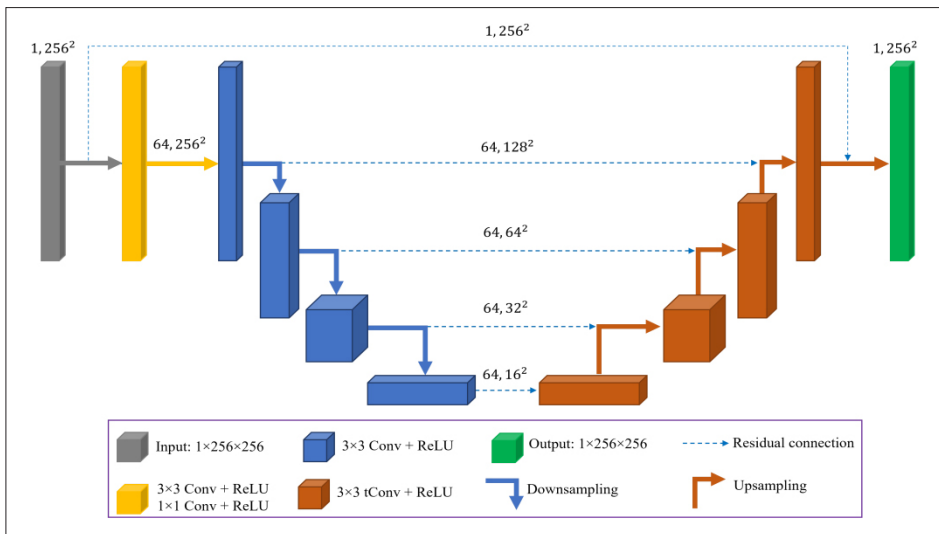
영상의 적대적 공격에 대응가능한 강인한 오픈셋 학습 방법

Kutub Uddin / 한국항공대학교 AI 미디어 연구실

In recent times, digital image forensics is gaining increased attention in multimedia forensics owing to the widespread scam alertness. Anti-forensic (AF) attacks on manipulated images, particularly generative adversarial network (GAN), have been successfully applied to delude forensic methods.

Consequently, a robust counter-AF (CAF) method is required to secure the integrity of digital images.

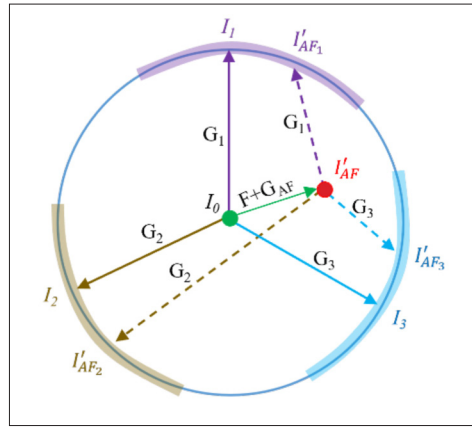
In this study, we propose a robust open-set multi-instance learning by introducing additional GAN-based operations. In details, we introduce double GAN-based operations to detect AF images in which additional



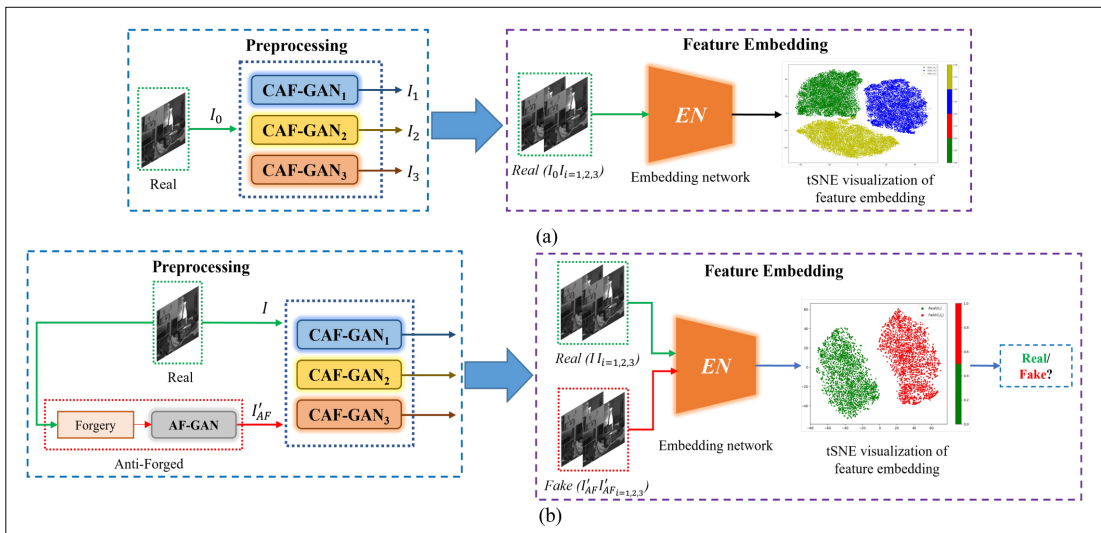
<Fig. 1> The architecture of the proposed CAF-GAN

GAN-based operations were purposely applied using multiple CAF-GAN models in <Fig. 1>. The generator models vary based on the number of parameters and produce three different real-generated images that are projected onto three different generated image spaces marked in blue, yellow, and purple in <Fig. 2>. Finally, the input and CAF-GAN generated images are concatenated for the final decision.

The proposed CAF was divided into two phases: training and testing. In the training phase, the embedding network learns only real images as shown in <Fig. 3> (a). In testing phase, both the real and AF images were processed to distinguish between them as shown in <Fig. 3> (b). The proposed method exhibited promising performance and robust against transferable updating AF attacks.



<Fig. 2> Working principle of the proposed open-set CAF



<Fig. 3> The architecture of the proposed open-set CAF method: (a) training scenario of the embedding network to learn multiple instances of real images in open-set fashion and (b) testing scenario of real and AF images which do not appear in the training



Kutub Uddin

- 2017 : *Bachelor of Science, University of Chittagong*
- 2020 : *Master of Science, Korea Aerospace University*
- 2023 : *Ph. D., Korea Aerospace University*
- *Research interests : Image processing, Image forensic*