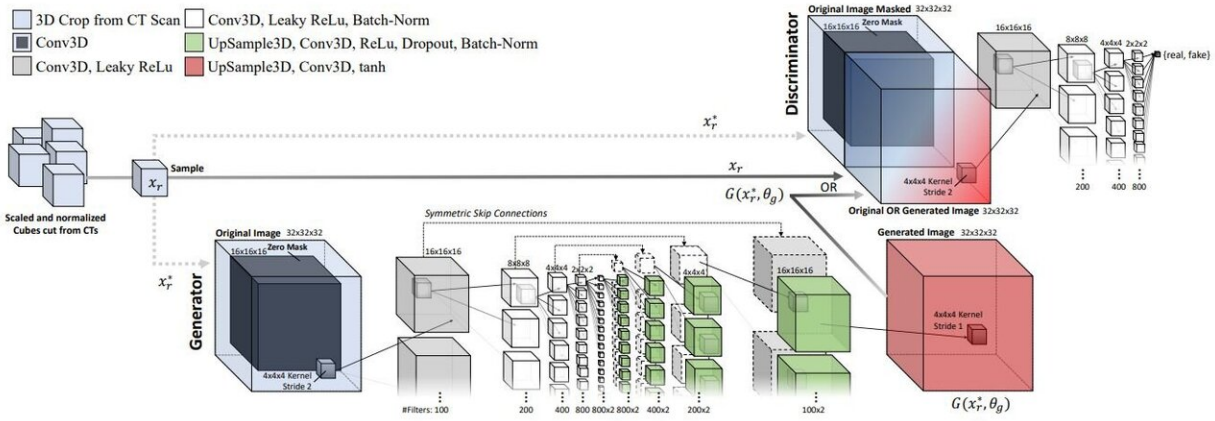
Can attackers inject malice into medical imagery? Fake growths here and there

April 6 2019, by Nancy Cohen



The network architecture, layers, and parameters used for both the injection (GANinj) and removal (GANrem) networks. Credit: arXiv:1901.03597 [cs.CR]

Researchers have found more reason to be concerned over possibilities of tampering within the medical setting. Security watchers are talking about the paper, "CT-GAN: Malicious Tampering of 3-D Medical Imagery using Deep Learning," which is on arXiv. Authors are Yisroel Mirsky, Tom Mahler, Ilan Shelef and Yuval Elovici.

The software, designed by experts at the Ben-Gurion University Cyber Security Research Center, was designed to see if an attacker could tamper with CT and MRI scanning equipment to produce false results about patients with tumors. The findings suggested by all means, yes,

tampering was not difficult to achieve.

The Washington Post carried a much-quoted article on the matter, as it explained their research aims and observations. The news report said that "attackers could target a presidential candidate or other politicians to trick them into believing they have a serious illness and cause them to withdraw from a race to seek treatment."

A growth added here, a growth added there...fake growths through tampering emerged as yet another type of malware that the medical community should know about. Radiologists can be tricked through malware to see fake cancerous nodules, in CT and MRI scanning equipment.

The malware was actually created by researchers in Israel. They wanted to explore security weaknesses both in medical imaging equipment and networks transmitting those images.

The researchers' malware could go both ways—either adding fake growths to the scans or removing real lesions and nodules; the latter maneuver obviously could result in failure to treat patients in critical need of timely attention.

The malware altered 70 images and managed to fool three radiologists into believing patients had cancer. The researchers used lung cancer as the focus. Kim Zetter, *The Washington Post*, described the test. Three radiologists—skilled—were tricked. They misdiagnosed conditions nearly every single time. Look at the numbers.

"In the case of scans with fabricated cancerous nodules, the radiologists diagnosed cancer 99 percent of the time. In cases where the malware removed real cancerous nodules from scans, the radiologists said those patients were healthy 94 percent of the time."

In turn, the study results should lead medical community to consider this about potential impact: Attackers' motives can be general or targeted. They could simply want to introduce chaos and strain the workflow with attention to equipment gone wrong or they could use the malware to target specific patients.

In their paper, the authors offered a grim list of possible goals if an attacker wanted to interfere with the scans. The authors said, "we show how an attacker can use deep-learning to add or remove evidence of medical conditions from volumetric (3-D) medical scans. An attacker may perform this act in order to stop a political candidate, sabotage research, commit insurance fraud, perform an act of terrorism, or even commit murder."

Zetter also brought up the possible scenario where follow-up scans would be messed with to show tumors either spreading or falsely shrinking. Malware could also have adverse effects on drug and medical research trials "to sabotage the results."

The malware's high degree of success makes one wonder how can this happen in hospital settings. Then again, for those who are already familiar with past events, the issue of safety does not surprise.

BBC News refreshed readers' memories. "Hospitals and other healthcare organisations have been a popular target for cyber-attackers and many have been hit by malicious ransomware that encrypts files and only returns the data when victims pay up." The report noted how "The NHS was hit hard in 2017 by the WannaCry ransomware which left many hospitals scrambling to recover data."

Why is the malware able to get past any security gates? *The Washington Post* indicated the problem could be traced to the equipment and networks that transmit and store CT and MRI images.

"These images are sent to radiology workstations and back-end databases through what's known as a picture archiving and communication system (PACS). Mirsky said the attack works because hospitals don't digitally sign the scans to prevent them from being altered without detection and don't use encryption on their PACS networks, allowing an intruder on the network to see the scans and alter them."

PACS networks are generally not encrypted. Another potential problem mentioned in the article rests with those hospitals making do with "20-year-old infrastructure" that does not support newer technologies.

"Although encryption is available for some PACS software now, it's still generally not used for compatibility reasons. It has to communicate with older systems that don't have the ability to decrypt or re-encrypt images," said *The Washington Post*.

Note the title of their research paper has the phrase "GAN." This stands for "a special kind of deep neural network," the generative adversarial network. With GANs, you have two neural networks working against each other: the generator and the discriminator.

"In this paper we introduced the possibility of an attacker modifying 3-D medical imagery using deep learning. We ... presented a manipulation framework (CT-GAN) which can be executed by a malware autonomously."

In addition, "The altered images also managed to trick automated screening systems," said BBC News.

All in all, according to their paper, "both radiologists and AI are highly susceptible to CT-GAN's image tampering attacks," the authors wrote.

More information: CT-GAN: Malicious Tampering of 3D Medical

Imagery using Deep Learning, arXiv:1901.03597 [cs.CR]
arxiv.org/abs/1901.03597