A New Method for the Exploitation of Speech Recognition Systems

Submitted by akarns on Thu, 09/27/2018 - 8:35am. Contributors:
Suha HusseinZahra GhodsiRamesh Karri

BIO

Suha Hussain is a senior at Queens High School for the Sciences, working as a research intern at the NYU Center for Cybersecurity. She has been recognized for her work in machine learning and cybersecurity, notably discovering and exploiting a vulnerability within speech recognition systems. For her work, Suha received multiple awards, including the 2018 Intel ISEF Second Award in Systems Software. At her school, she runs and is heavily involved in several organizations, including the robotics team and research program. Previously, she completed a hardware engineering internship and studied at Google's igniteCS Bootcamp at Columbia University.

ABSTRACT

The rapid proliferation and adoption of speech recognition systems in our day-to-day lives result in greater consequences for possible vulnerabilities. Previous research has proven that host hardware and preprocessing can be leveraged to successfully deceive speech recognition systems. Additionally, neural networks, algorithms within modern systems, can be effectively fooled by generating adversarial noise. However, a method to exploit speech recognition systems by leveraging neural networks was notably absent. An algorithm was developed that crafts universal, transformable adversarial noise for the inputs of a speech recognition system that would result in deliberate misclassification. To evaluate this algorithm, adversarial noises for five randomly chosen target classes were produced using a substitute neural network.The noises were then added to the inputs of a victim system in a black-box setting. On average, the crafted adversarial noises led to deliberate misclassification 60.42% of the time. The universality of the generated noises increased the inconspicuousness, aided by limitations set on the noises. The feasibility and practicality of the attack was increased by the fact that the adversarial noises were transformable. Thus, the neural networks in speech recognition systems are a significant vulnerability. It is imperative that attacks such as these are mitigated for speech recognition systems to be considered safe. Future research can improve upon the proposed attack for the purpose of finding more vulnerabilities or focus upon building an optimal defense strategy.

Suha Hussein | Zahra Ghodsi | Ramesh Karri

Other available formats:

A New Method for the Exploitation of Speech Recognition Systems
Switch to normal viewerSwitch to experimental viewer

machine learning Speech recognition Adversarial Machine Learning adversarial attack Deep Neural Network Poster Posters