**Name:** Adrian Wong

**SID:** 520452275

**GitHub:** https://github.com/womba/elec5305-project-520452275

---

**Project Tile**

Keystroke Classification using Deep Image Classifiers

**Project Overview**

The aim of this project is to develop an acoustic side-channel attack for typical home keyboards using a deep image classifier. The goal is to be able to discreetly extract the intended typed text from typing audio alone.

**Background and Motivation**

This project explores the concept of an acoustic side-channel attack – a class of security vulnerabilities that exploit the sounds emitted by computers and input devices in order to discreetly obtain sensitive information. This works under the pretence that both hardware components (e.g. power supply coil whine) and human-interaction-devices (e.g. keyboards, printers) emit distinctive acoustic signatures that vary with usage. For example, previous research has demonstrated the extraction of encryption keys from coil whine and the reconstruction of printed text from the sounds of dot-matrix printers.

Keyboards are of particular interest because each keystroke generates a unique acoustic signal influenced by factors such as key position, switch design, and resonance within the frame. By recording these sounds with a nearby microphone – or even via video conferencing software – an attacker can infer typed text, including sensitive information such as passwords and private communications.

Historically, acoustic side-channel attacks have relied on statistical language models, combined with features such as timing patterns, frequency components, and amplitude variations, to distinguish between keys. While these approaches have achieved moderate success, accuracy has typically been limited to ~75%.

Recent advances in deep learning models – particularly transformer-based image classifiers – has shown promise in their use for audio classification through spectrogram representations. This project is motivated by the paper *A Practical Deep Learning-Based Acoustic Side Channel Attack on Keyboards.* Where the authors demonstrated the feasibility of using deep learning to classify keystrokes from Mel-spectrograms of recorded audio, achieving 95% accuracy with a nearby smartphone microphone and 93% accuracy using recordings obtained through Zoom.

**Proposed Methodology**

In current literature, acoustic side-channel attacks on keyboards generally consist of two primary stages: (1) keystroke segmentation and (2) classification of the segmented audio samples. This project will adopt and extend this framework, combining classical signal processing methods with deep learning techniques.

***Keystroke Segmentation***

The first step is isolating individual keystroke events from a continuous audio stream. Current literature primarily applies classical signal processing techniques for this task, many of which have been covered in ELEC5305. Examples include:

- Energy Thresholding: Performing an STFFT to compute time-varying energy and identifying peaks above a threshold.
- Matched Filtering: Convolving a template waveform to identify candidate keystrokes.
- Onset Detection: Using spectral flux to identify keystroke boundaries.

This project will implement and evaluate these methods in MATLAB, in an attempt to find a robust approach that generalises across different recording environments.

***Feature Extraction and Classification***

Following segmentation, the isolated keystrokes will be converted into representations suitable for classification using a deep image classifier. Listed options in the referenced paper include an FFT, MFCCs, or a Mel-Spectrogram. This project aims to train an image classifier model in PyTorch.

***Dataset***

The dataset from the referenced paper is publicly available on GitHub. However, it is limited to readings from a MacBook Pro equipped with scissor switches.

Alternatively, a custom dataset can be obtained through recording keystrokes from another common keyboard model.

**Expected Outcomes**

I will first attempt to re-create the deep learning model seen in the referenced paper, before attempting to iterate certain aspects and produce a working prototype which can be used for a live (albeit not real-time) demonstration. I aim to achieve a F1 score of at least ~80%.

All accompanying source-code and documentation for the project will be hosted on GitHub.

**Timeline**

| Week | Task |
|---|---|
| 6-7 | Literature Review, Dataset Collection, Keystroke Isolation |
| 8-9 | ML Model Initial Implementation and Training |
| 10-11 | Testing and Optimisations |
| 12-13 | Final Report and Documentations |

**References**

Harrison, J., Toreini, E., & Mehrnezhad, M. (2023, July). A practical deep learning-based acoustic side channel attack on keyboards. In *2023 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)* (pp. 270-280). IEEE.

Halevi, T., & Saxena, N. (2015). Keyboard acoustic side channel attacks: exploring realistic and security-sensitive scenarios. *International Journal of Information Security*, *14*(5), (pp. 443-456).

Berger, Y., Wool, A., & Yeredor, A. (2006, October). Dictionary attacks using keyboard acoustic emanations. In *Proceedings of the 13th ACM conference on Computer and communications security* (pp. 245-254).