data

MDPI

Data Descriptor

A Multi-Pressure Keyboard Key Sounds Audio Dataset

Diogo Rodrigues [1], Gonçalo Macedo [1], Silvestre Malta [1,2] and Pedro Pinto [3,*]

[1] ESTG, Instituto Politécnico de Viana do Castelo, 4900-347 Viana do Castelo, Portugal; diogo.r@ipvc.pt (D.R.); goncalo.macedo@ipvc.pt (G.M.); smalta@estg.ipvc.pt (S.M.)
[2] Adit-Lab, Instituto Politécnico de Viana do Castelo, 4900-347 Viana do Castelo, Portugal
[3] GECAD, ISEP, Polytechnic of Porto, Rua Dr. António Bernardino de Almeida, 4249-015 Porto, Portugal
* Correspondence: pfp@isep.ipp.pt

Abstract

The acoustic properties of keyboard keystrokes are essential for advancing research in human–computer interaction, security, and accessibility, yet existing datasets often lack force variability, detailed metadata, or controlled recording conditions, limiting their applicability. To address these gaps, we introduce a dataset of keyboard keystroke sounds, systematically recorded under three distinct pressure levels (low, medium, and high) for all alphanumeric keys and the space bar, complemented by high-quality spectrograms and detailed metadata. These datasets enable the development of robust classification models for acoustic side-channel attack detection, behavioral biometrics, assistive sound-based interfaces, and hardware diagnostics while providing a reproducible benchmark for both academic research and industry applications.

Dataset: https://zenodo.org/records/19453177.

Keywords: keystroke; machine learning; acoustic attacks; side-channel attacks; key pressure

1. Introduction

The sound of keyboard keystrokes holds significant potential for applications in cybersecurity, behavioral biometrics, and accessibility—such as detecting user emotions, diagnosing hardware issues, or enabling sound-based interfaces for people with motor impairments. However, existing datasets often lack standardized recordings, detailed metadata, or variability in key pressures, limiting their effectiveness. This paper introduces a new dataset designed to address these challenges by providing high-quality, multi-pressure keystroke audio recordings.

This article proposes a dataset valuable for both academic research and practical industrial applications. Its characteristics—such as multi-press keystroke recordings, high-fidelity spectrograms, and detailed metadata—enable a wide range of uses, including the following:

1. Developing keystroke recognition systems: These datasets can be used to develop keystroke recognition systems that either identify potential attack vectors or serve as the foundation for defensive software. Studies such as [1–3] have demonstrated that attackers can use Deep Learning to identify sensitive data, such as passwords, with high accuracy (in particular, authors in [2] report an accuracy over 90%) by simply listening to keystrokes through nearby microphones or even during online video calls. This

Academic Editor: Giuseppe Ciaburro

Received: 24 April 2026
Revised: 22 June 2026
Accepted: 26 June 2026
Published: 7 July 2026

dataset is a decision-support benchmark, baseline, or seed dataset for exploratory research that enables the study of how these signals are captured and reconstructed [4,5]. Also, it supports the development of countermeasures across various contexts, including mobile banking security [6] and the mitigation of eavesdropping threats in shared workspaces [7]. Our earlier publication [8] used a preliminary version of the current dataset to build a prototype to protect users from acoustic side-channel keyboard attacks. This was accomplished by generating synthetic keystroke sounds to neutralize the acoustic signatures attackers rely on by generating and mixing false, randomized keystroke sounds with real ones to obfuscate keypress identification through sound analysis.

2. Detecting Behavioral Biometrics: Since the proposed dataset introduces three distinct levels of keystroke pressure, it would enable a more granular study of users' behavioral biometrics, detecting their emotional states—such as stress, frustration, or fatigue—based on variations in sound frequency and amplitude. This feature opens new avenues for psychological research related to human–computer interaction and anomaly detection. The work in [9] uses keystroke dynamics to identify or confirm an individual's identity by analysing habitual rhythm patterns as they type on a keyboard and continuously predicts users' emotional states during message-writing sessions. The authors in [10] propose a multi-modal behavioral biometrics framework that fuses keystroke and mouse dynamics using temporal alignment and deep learning, achieving up to 99% accuracy in distinguishing benign from adversarial user behavior while preserving nuanced interaction patterns for robust, non-intrusive authentication and anomaly detection.
3. Enhancing Accessibility Tools: The proposed dataset can be used to leverage sound-based virtual keyboards with acoustic feedback to enable users with motor impairments to interact with computers or mobile devices without relying on physical keypresses. This approach is particularly valuable for individuals with conditions such as cerebral palsy, spinal cord injuries, or severe arthritis, where traditional typing is challenging or impossible. The work in [11] is an example of the use of acoustic sensors to detect voluntary sounds (even if they are not speech) and convert them into inputs for virtual keyboards. This can be combined with the generation of audible feedback to confirm the input, simulating the sound of keys being pressed.
4. Enabling Hardware Diagnostics: The proposed dataset could also aid the development of keystroke recognition systems for hardware monitoring, i.e., acoustic signatures can be used to detect keyboard wear or mechanical failures (e.g., sticky keys) via deviations from the baseline recordings provided. Furthermore, the data can support defensive software that protects users by masking or generating synthetic keystroke sounds to neutralize the acoustic signatures that attackers or diagnostic anomalies rely on [12].
5. Data augmentation techniques: The proposed dataset can be used to artificially expand the dataset for large-scale training using techniques such as time shifting, adding white noise, inserting background noise, reverb, echo, and gain variation (volume) [13,14].

2. Related Work

Rawf et al. (2024) [15] present the Multi-Keyboard Acoustic (MKA) Datasets, a large-scale resource designed to support research on keyboard sound recognition and defenses against acoustic side-channel attacks. The dataset covers six platforms (HP, Lenovo, MSI, Mac, Messenger, and Zoom) with more than 70 key classes and for its structured organization, including raw audio, segmented files, metadata, and MFCC (Mel-frequency Cepstral

Coefficients) matrices. However, potential biases toward specific keyboard models, the absence of noisy real-world conditions, and recording environments that may not fully reflect typical usage scenarios could limit its application. Additionally, some applications may require finer discrimination against the pressure exerted on the keys. These aspects may restrict the generalizability of models trained on the dataset, but overall, it provides an important benchmark for evaluating algorithms and advancing the state of research on acoustic emanation threats.

Dias et al. (2023) [16] introduce the KeyRecs dataset, which focuses on keystroke dynamics and typing pattern recognition for biometric authentication and anomaly detection. The dataset contains fixed-text and free-text samples collected from 99 participants of diverse nationalities, along with demographic data such as age, gender, handedness, and nationality. A key strength of this work is the inclusion of demographic attributes, which allows for broader behavioral analysis and more personalized modeling of typing patterns. Furthermore, the use of both fixed and free text provides valuable variability for training robust authentication systems. However, the dataset records only timing features (inter-key latencies) and not the acoustic properties of keystrokes, limiting its applicability to acoustic side-channel attack research. In addition, data collection was performed online using participants' own keyboards, which introduces heterogeneity but also raises concerns about consistency and environmental control. Despite these limitations, KeyRecs stands out as a significant contribution to keystroke biometrics, offering a valuable benchmark for developing machine learning models in continuous authentication and anomaly detection.

As outlined in Table 1, while recent literature demonstrates significant advancements in Acoustic Side-Channel Attacks (ASCA) and Keystroke Dynamics, existing datasets primarily focus on either hardware heterogeneity or temporal latencies, overlooking the physical nuances of the typing act. For instance, large-scale acoustic datasets like MKA [15] and SKAID [17] provide valuable benchmarks across different platforms and real-world free-text scenarios, but they lack fine-grained mechanical variability, treating all keystrokes of the same key as a uniform acoustic event. Conversely, behavioral biometric studies, such as those utilizing the KeyRecs [16] and BEACON [18] datasets, rely heavily on temporal latencies and hardware events, entirely excluding acoustic properties. Our dataset addresses these gaps by offering standardized force levels, high-fidelity recordings, and comprehensive metadata, enabling more robust and generalizable models.

Table 1. Literature review.

Ref.	Authors	Year	Addressed Problems	Materials	Methods	Participants	Keyboard Devices	Recording Environment(s)	Dataset Size
[1]	Harrison et al.	2023	Deep learning-based acoustic side-channel attack (ASCA) on keyboards.	Laptop keyboards, smartphone, Zoom recordings.	CoAtNet deep learning model, FFT energy peaks for keystroke isolation.	1	1	Quiet room/Zoom	900 keystrokes
[2]	The Guardian	2025	Password identification feasibility via keystroke sound recording.	-	Journalistic review of AI study findings.	n/a	n/a	n/a	n/a
[3]	George & Sagayarajan	2023	Acoustic eavesdropping via AI algorithms.	-	Conceptual review.	n/a	n/a	n/a	n/a
[4]	Venkateswarlu	2013	Fundamental vulnerabilities of keyboard emanations.	-	Broad literature overview of ASCA mechanics.	n/a	n/a	n/a	n/a
[5]	Taheritajar et al.	2025	State-of-the-art overview of ASCA on keyboards.	Academic literature.	Systematic literature review and survey.	n/a	n/a	n/a	n/a
[6]	Savadatti et al.	2023	Keystroke ASCA targeting digital banking on Android platforms.	Android mobile devices.	Elementary review of smartphone acoustic vulnerabilities.	n/a	n/a	n/a	n/a
[7]	Anand & Saxena	2018	Keyboard acoustic leakage during remote voice calls.	Voice call audio streams.	Noise(less) masking defenses and signal obfuscation.	Multiple	Multiple	Remote voice calls	Audio streams
[8]	Rodrigues et al.	2024	Protection against ASCA via acoustic signature neutralization.	Synthetic keystroke software prototype.	Generation and mixing of false randomized keystrokes.	1	1	Controlled	Software logs

Table 1. *Cont.*

Ref.	Authors	Year	Addressed Problems	Materials	Methods	Participants	Keyboard Devices	Recording Environment(s)	Dataset Size
[9]	Marrone & Sansone	2022	Identification of users' emotional states (stress, fatigue).	Keystroke dynamics data.	Behavioral biometric analysis of habitual rhythm patterns.	Multiple	n/a	Uncontrolled	Keystroke dynamics log
[10]	Xiong & Belman	2025	Distinguishing benign from adversarial user behavior.	Keyboard and mouse dynamics data.	Deep learning framework with temporal alignment fusion.	Multiple	Multiple	Controlled lab	Interaction logs
[11]	Prete et al.	2025	Accessibility for motor-impaired and locked-in patients.	Acoustic sensors and virtual keyboards.	Voluntary sound detection to trigger keystroke inputs.	Multiple	n/a	Clinical/Home	Sensor data
[12]	Lobanova et al.	2024	Masking acoustic channels to prevent keystroke inference.	-	Machine-learning-based adaptive sound masking.	n/a	n/a	n/a	n/a
[13]	Tsalera et al.	2025	Improving training robustness for sound classification tasks.	Environmental sound datasets.	Time shifting, noise insertion, and synthetic generation.	n/a	n/a	Various	Augmented sets
[14]	Rzemieniuk et al.	2026	Modeling of ASCA vulnerabilities to interpret acoustic data.	Keyboard input samples.	Convolutional Neural Network (CNN) modeling.	1	Multiple	Controlled	Input samples
[15]	Rawf et al.	2024	Creation of benchmark data for ASCA defense research.	MKA Datasets (6 platforms, >70 key classes).	MFCC matrices extraction and raw audio segmentation.	n/a	6 platforms	Uncontrolled/Various	>70 key classes
[16]	Dias et al.	2023	Keystroke dynamics for continuous biometric authentication.	KeyRecs dataset (fixed/free text, 99 users).	Inter-key latency and temporal feature extraction.	99	99 (own devices)	Online/Home (uncontrolled)	Fixed and free text
[17]	Quattrone & Badr	2025	Lack of realistic and synchronized ASCA datasets for AI inference.	SKAID Dataset (Keyboard audio, transcribed text).	Synchronized audio and key event logging in natural conditions.	Multiple	Multiple	Naturalistic	SKAID audio/logs
[18]	BEACON Authors	2026	Learning behavioral fingerprints under high cognitive load.	Gameplay data (keystrokes and mouse latencies).	Multimodal behavioral biometrics evaluation.	Multiple	Multiple	Gaming environment	Gameplay metrics
[19]	Pipeline Authors	2025	Extracting and snooping keystrokes from mobile phone recordings.	Physical keyboards, smartphone audio.	Wavelet transforms and Temporal Convolutional Networks (TCN).	n/a	Multiple	Controlled	Recording tracks

3. Materials and Methods

The materials used to record the dataset were as follows:

- Samsung Galaxy M52 5G;
  - Manufacturer: Samsung
  - City: Dublin
  - Country: Vietnam
- Mi True Wireless Earphones 2 Pro;
  - Manufacturer: Tiinlab Corporation (a Mi Ecosystem Company)
  - City: Tanglang
  - Country: China
- Victus by Gaming Laptop 16-s0006np.
  - Manufacturer: HP
  - City: Kunshan
  - Country: China

The dataset was recorded by a single person in a low-noise environment with an ambient noise level of 14 dB—an office measuring approximately 20 m^2—using the built-in keyboard. This model is equipped with a full-size membrane keyboard utilizing a scissor-switch mechanism, offering a precise key travel distance of approximately 1.5 mm. The keycaps are constructed from high-grade ABS plastic. The microphone used, rather than a standard single-microphone setup, this device features a 3-mic array system specifically engineered for Environmental Noise Cancellation (ENC) and precise voice/sound isolation.

The microphone was placed in the center of the laptop's touchpad, about 5 cm below the keyboard, for all recording sessions. The recording has been started and stopped for every single key and for each pressure level: Low, Medium, and High.

The Low, Medium, and High ppressure categories correspond to deliberately differentiated, human-like force levels, intended to capture realistic typing variability rather than calibrated physical force values. The RMS Energy values for each of these pressures were obtained using a time-window calculation followed by an averaging aggregation. The extraction of this feature was a two-step process:

1. The digital audio signal is analyzed and split into multiple frames. For each frame, the RMS energy value is calculated, which represents the average magnitude of the signal in that fraction of a second, given by the mathematical formula:

$$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x[n]^2} \quad (1)$$

   where "N" is the number of samples in the frame and "x[n]" is the amplitude of the acoustic signal at sample "n". The result of this step is a vector of RMS values that describes the energy of the keystroke over time.
2. To convert this time series into a one-dimensional tabular dataset, the time vector is collapsed into a single scalar value. This is performed by calculating the arithmetic mean of all frames in the recording.

The obtained average RMS values for each pressure level and for all keys are presented in Table 2. They provide a quantitative acoustic proxy that consistently orders and separates the three categories, offering an objective reference against which future recordings can be aligned.

Table 2. RMS average values for Low, Medium, and High-pressure levels for all keys.

Pressure Level	Average RMS_Energy for All Keys ($\times 10^{-3}$)
Low	1.332
Medium	7.975
High	21.170

When the script loads the audio file, it converts the sound wave into a sequence of decimal numbers. In digital audio, these amplitude values typically fall within a mathematical range between -1.0 and 1.0. Since the mathematical formula for RMS squares these amplitudes, calculates their time-averaged value, and finally extracts the square root, the result will always be a positive value between 0 and 1.0.

To extract the absolute minimum and maximum decibel (dB) levels of the keystrokes, the linear acoustic amplitudes were converted into a logarithmic scale. The transformation was computed using the standard 20-log rule for amplitude magnitudes, mathematically defined as:

$$L_{dB} = 20 \cdot \log_{10}(|A| + \epsilon) \quad (2)$$

where $|A|$ represents the absolute peak amplitude (either maximum or minimum) of the digital audio signal, and ϵ is the machine epsilon (a nearly zero infinitesimal constant provided by the NumPy library), which is systematically added to prevent undefined operations such as the logarithm of zero.

The Short-Time Fourier Transform (STFT) magnitude matrix (S) was mapped to a decibel-scaled spectrogram. The underlying mathematical conversion applied by the library scales the amplitudes relative to a predefined reference value (S_{ref}):

$$S_{dB} = 20 \cdot \log_{10}\left(\frac{S}{S_{ref}}\right) \quad (3)$$

the reference was dynamically set to the maximum amplitude of the specific audio signal. Consequently, the highest energy peak within the spectrogram is normalized and mapped precisely to 0 dB, while all remaining frequency magnitudes are represented as negative decibel values relative to this peak. Additionally, the Librosa library applies a default dynamic range threshold of 80 dB, effectively truncating any underlying background noise weaker than −80 dB, which ensures a clear visual contrast of the mechanical keystroke impact against the ambient silence.

4. Results

These datasets consist of 114 audio files capturing keystroke sounds by an ABNT keyboard (ID 00000816). These files consist of audio of keys "A" to "Z", including the key "Ç" (27 keys), "0" to "9" (10 keys), and the space bar, captured using three different key pressures: low, medium, and high.

Figure 1 presents a diagram illustrating the hierarchical structure of the dataset.

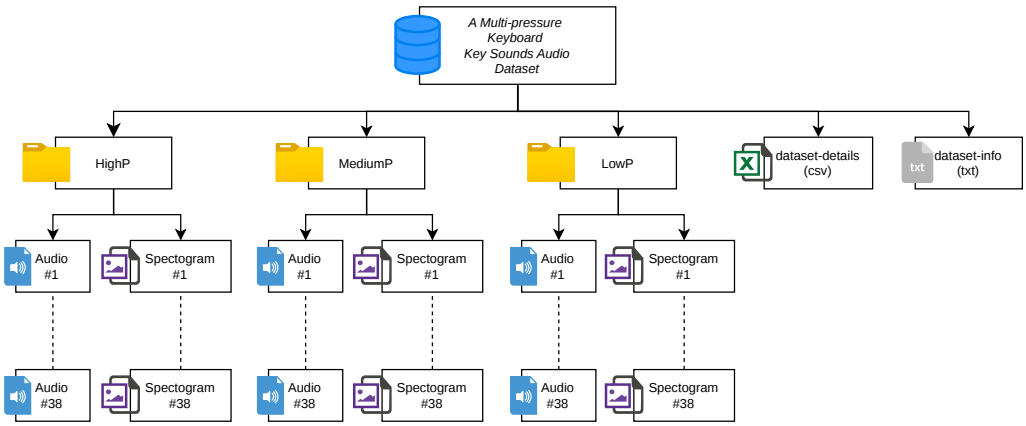

Figure 1. Hierarchical structure of the dataset.

The root directory, "dataset", contains three subdirectories and two files. All three are named with the pressure used on the keyboard to record the data (High, Medium, and Low). Each of these subdirectories contains 38 audio files (WAV) and their corresponding 38 spectrogram images (PNG). The additional 2 files in the root directory are a CSV file with specific information for each audio file and a TXT file with a description for each column presented in the CSV file.

The Main folder "dataset" contains:

- The subfolders: HighP, MediumP, and LowP, with the audio files and spectrograms for the respective High, Medium, and Low pressures.
- A CSV file with detailed information for each audio file, and a TXT file with a description for each column in the CSV file.

Content Overview:

1. Subfolder HighP:
   - 38 audio files for the High pressure (Format: WAV)
   - 38 spectrogram image files of the respective audio files (Format: PNG)
2. Subfolder MediumP:
   - 38 audio files for the Medium pressure (Format: WAV)
   - 38 spectrogram image files of the respective audio files (Format: PNG)
3. Subfolder LowP:
   - 38 audio files for the Low pressure (Format: WAV)

- 38 spectrogram image files of the respective audio files (Format: PNG)

4. File dataset-details (Format: CSV):
   - File with detailed information about the audio files in the HighP, MediumP, and LowP subfolders, namely "Key," "Pressure", "Folder", "FileName", "Spectrum_FileName", "Duration_s", "Max_Amplitude", "Min_Amplitude", "Max_dB", "Min_dB", "RMS_Energy", and "Mean_ZCR".
5. File dataset-info (Format: TXT):
   - File with a description of each column in the dataset-details CSV file.

A comparative visual analysis of the spectrograms reveals distinct acoustic signatures directly correlated with the applied physical force. In the low-pressure keystroke as shown in Figure 2, the acoustic energy is predominantly confined to the lower frequency bands, with a relatively short temporal decay. As the pressure increases to a medium level, the spectrogram illustrates a broader energy distribution, introducing more broadband noise and exciting mid-tier frequencies, as shown in Figure 3. Finally, the high-pressure keystroke, shown in Figure 4, exhibits a highly pronounced temporal impact phase (visible as a dense, vertical acoustic onset) and significantly more energy across the high-frequency spectrum. These visual variations in frequency content and temporal impact align consistently with the quantitative RMS energy findings presented in Table 2, where the energy magnitude scales from 1.332×10^{-3} in low pressure up to 21.170×10^{-3} in high pressure.

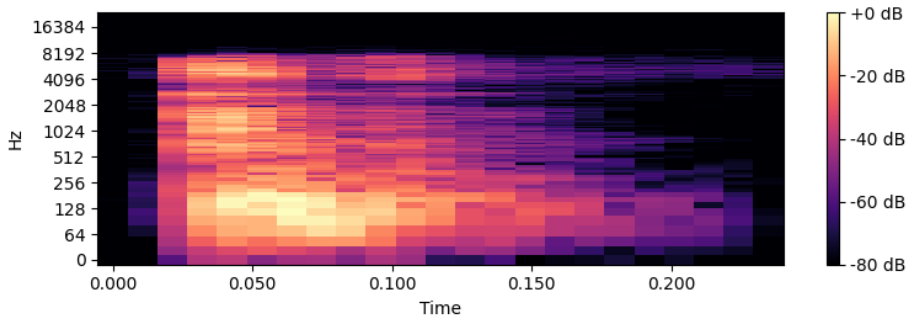

Figure 2. Spectrogram for the audio file recorded with low pressure for the A keystroke, displaying a concentrated energy distribution primarily in the lower frequency bands and a rapid temporal decay.

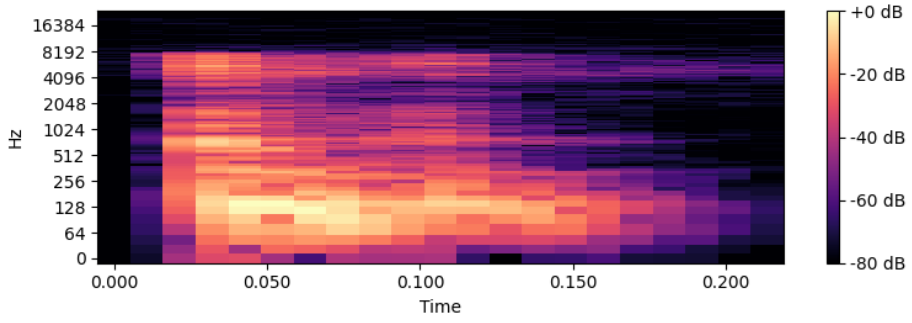

Figure 3. Spectrogram for the audio file recorded with medium pressure for the A keystroke, illustrating an expansion of broadband noise and an increased presence of mid-to-high frequencies.

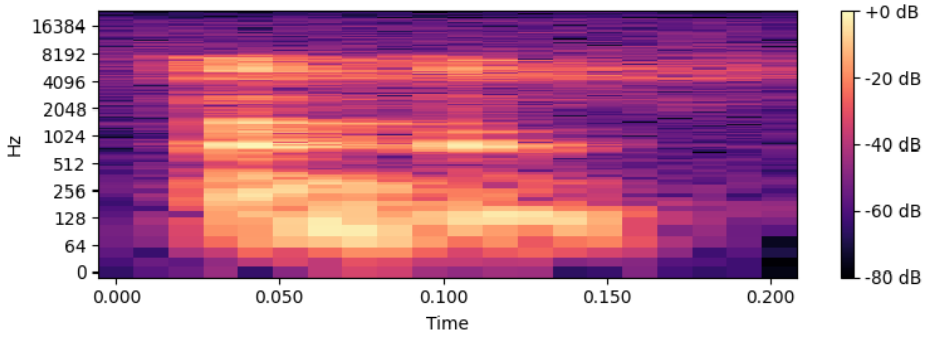

Figure 4. Spectrogram for the audio file recorded with high pressure for the A keystroke, exhibiting prominent high-frequency content, a broader and denser energy spectrum, and a more pronounced temporal impact signature.

Table 3 presents the details on the entire structure of the CSV file, including the column name, data type, and technical description for each feature used in the analysis.

The data in this table were obtained using an automated Algorithm 1, which processes raw audio, extracts acoustic features, and generates spectrograms. The Librosa library (Python) is used to extract acoustic features such as duration, amplitude peaks, decibels, RMS energy, and the Zero-Crossing Rate (ZCR). To generate the spectrograms, the one-dimensional audio is converted into a two-dimensional visual representation using the Short-Time Fourier Transform (STFT) to analyze the frequency spectrum over time. The frequency scale is mapped logarithmically, and the amplitudes are converted to decibels.

Algorithm 1 Script to extract acoustic features and enerate spectrograms

```
Initialize an empty list DatasetRecords
Create output directory SpectrogramFolder
for EACH folder IN ["HighP", "LowP", "MediumP"] do
  pressure_level ← ExtractPressureLabel(folder)
  audio_files ← GetWavFiles(folder)
  for EACH file IN audio_files do
    audio_signal, sample_rate ← LoadAudio(file)
    if loading fails then
      continue
    end if
    // 1. Feature Extraction
    duration ← CalculateDuration(audio_signal, sample_rate)
    max_amp, min_amp ← FindPeakAmplitudes(audio_signal)
    max_dB, min_dB ← ConvertAmplitudesToDecibels(max_amp, min_amp)
    rms_energy ← CalculateMeanRMS(audio_signal)
    mean_zcr ← CalculateMeanZeroCrossingRate(audio_signal)
    // 2. Spectrogram Generation
    stft_matrix ← CalculateSTFT(audio_signal)
    spectrogram_db ← ConvertToLogScale(stft_matrix)
    image_filename ← SaveSpectrogramAsPNG(spectrogram_db, SpectrogramFolder)
    // 3. Data Consolidation
    key_label ← ParseKeyLabelFromFilename(file)
    record ← {Key: key_label, Pressure: pressure_level, FileName: file, ...}
    Append record TO DatasetRecords
  end for
end for
Save DatasetRecords AS "dataset_keystrokes.csv"
```

Table 3. Structure of the CSV file.

Name	Type	Description
#	Integer (int)	Sequential identification number of the record.
Key	String (char)	The key (character) that was pressed (e.g., A, Ç, Z).
Pressure	String (char)	Key strike pressure level (High, Medium, Low).
Folder	String (char)	The original name of the folder where the audio file is (e.g., HighP, LowP).
FileName	String (char)	The full name of the WAV audio file (e.g., Key-A-L.wav).
Spectrum_FileName	String (char)	The full name of the PNG image file (e.g., Key-0-H.png).
Duration_s	Float	Exact duration of the audio file in seconds.
Max_Amplitude	Float	MAXIMUM (positive) peak value of the audio signal waveform (between 0 dB and 1 dB).
Min_Amplitude	Float	MINIMUM (negative) peak value of the audio signal waveform (between −1 dB and 0 dB).
Max_dB	Float	Maximum volume level in decibels.
Min_dB	Float	Minimum volume level in decibels.
RMS_Energy	Float	Root Mean Square of the audio signal energy. This corresponds to the average and total volume of sound energy. It is not a statistically normalized value, but the average energy of the acoustic event.
Mean_ZCR	Float	Mean Zero Crossing Rate. The average number of times per second that the audio signal crosses the zero axis.

5. Discussion

This multi-pressure keystroke sound dataset addresses key gaps in existing research by providing standardized recordings, detailed metadata, and systematic pressure variability. Given its controlled and fully labeled nature, the dataset is suited to a specific set of scenarios. It supports methodological and feasibility studies on press-force acoustic discrimination, where the variation of applied force across three levels enables the investigation of whether, and to what extent, keystroke pressure can be inferred from acoustic emissions. It provides a reliable basis for controlled, reproducible experimentation and pipeline prototyping, allowing researchers to develop, debug, and validate feature-extraction and classification workflows on clean recorded data before transferring them to more heterogeneous settings. And its structured organization, metadata, and accompanying spectrograms serve as a clear educational and exploratory reference resource for teaching and for the initial assessment of keystroke acoustic analysis techniques.

However, its reliance on a single keyboard model and controlled environment may limit real-world generalizability. Applications such as the generalizable acoustic side-channel attack model, robust behavioral biometrics and continuous authentication, and noise-robust real-world systems, require multiple keyboards, environments, users, continuous typing sequences, and recordings under varied ambient conditions that would require additional data collection.

In particular, the following limitations are acknowledged:

1. Hardware and Recording Bias
    - The dataset may be biased toward specific keyboard models, switch types, or construction materials.
    - It uses a single microphone or phone model, which limits generalizability to other recording hardware.
    - Microphone placement, angle, and the acoustics of the recording space may introduce consistent patterns not found in typical usage.

2. Environmental Conditions
    - Recordings were obtained in controlled, low-noise conditions, which differ from real-world environments where background noise, reverberation, or user movement may affect audio.
    - The absence of diverse environments (e.g., office, café, classroom, and home) reduces robustness for deployment-focused applications.
3. User Variability
    - If the dataset includes one or few participants, typing style variability may be limited.
    - Differences in press force or key handling between users may not be fully represented.
4. Data Coverage
    - The dataset focuses on isolated key presses, not continuous typing sequences.
    - Only certain key categories may be included (alphabetic, numeric, and space bar), leaving out modifiers, punctuation, or function keys.
    - Some applications require more granular distinctions, such as precise press-force levels or additional pressure steps.
5. Acoustic Representativeness
    - Controlled recording environments may produce sound characteristics that do not reflect everyday usage, where surfaces, table materials, or keyboard orientation affect acoustics.
    - The recorded volume and dynamics may not match real user behavior.
6. Application-Specific Constraints
    - For security and side-channel research, the lack of real-world noise and typing sequences may reduce external validity.
    - For machine learning tasks, the dataset may not be large enough to train deep models without augmentation.
    - For behavioral biometrics, limited user diversity may constrain generalization.

Nonetheless, the dataset serves as a valuable benchmark for acoustic classification, security defenses, and accessibility tools. Future work should expand the dataset for punctuation keys, modifier keys (Shift, Ctrl), and function keys, and could expand keyboard diversity, include varied typing styles, and incorporate realistic environmental conditions to enhance practical applicability.

6. Conclusions

Keyboard keystroke acoustics are relevant for advancing in areas such as cybersecurity, behavioral biometrics, and accessibility.

This paper introduces a keyboard keystroke sound dataset recorded with three distinct pressure levels. It features multi-press keystroke recordings, high-fidelity spectrograms, and detailed metadata that complement existing datasets. The dataset encompasses 114 audio files per pressure category—covering alphanumeric keys, the "Ç" key, and the space bar—which were carefully captured in a controlled 14 dB low-noise environment.

An automated script was employed to process the raw audio, yielding a multimodal dataset that includes both tabular metadata and high-fidelity spectrograms. The results quantitatively validated the proposed pressure categorizations through the extracted Root Mean Square (RMS) energy metrics. Average RMS values scaled distinctly across the force levels, recording 1.332×10^{-3} for low pressure, 7.975×10^{-3} for medium pressure, and 21.170×10^{-3} for high pressure. Additionally, the pipeline successfully generated two-dimensional spectrograms via Short-Time Fourier Transform (STFT). These visual

representations were normalized using a dynamic range threshold of 80 dB to ensure a clear visual contrast of the mechanical keystroke impact against the ambient silence.

The proposed dataset provides a basis for future research on advancements in areas such as acoustic side-channel attack detection, behavioral biometrics, assistive technologies, and hardware diagnostics, for both academic research and industry applications.

While the current dataset is limited to isolated key presses on a single keyboard model, it establishes a highly reliable baseline. Future work may include expanding the hardware diversity to other keyboard models, incorporating continuous user typing sequences, and introducing real-world environmental noise to further enhance practical applicability. Ultimately, this work contributes to the growing body of research on keystroke acoustics, providing a foundational resource for developing more secure, accessible, and adaptive human–computer interaction systems.

Author Contributions: D.R.: Methodology, Software, Formal Analysis, Investigation, and Writing—Original Draft; G.M.: Methodology, Software, Formal Analysis, Investigation, and Writing—Original Draft; S.M.: Writing—Review and Editing; P.P.: Conceptualization, Supervision, Project administration, Writing—Review and Editing, Resources, and Validation. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported and received funding from "CyberPRAISE—Cybersecurity research for PRivAte, Intelligent and truStablE solutions"—NORTE2030-FEDER-01820300.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset presented in this study is publicly available in Zenodo at https://zenodo.org/records/19453177 (accessed on 7 April 2026).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CVS	Comma-Separated Values
MKA	Multi-Keyboard Acoustic
MFCC	Mel-frequency Cepstral Coefficients
PNG	Portable Network Graphics
RMS	Root Mean Square
TXT	Plain Text File
WAV	Waveform Audio File Format
STFT	Short-Time Fourier Transform
ENC	Environmental Noise Cancellation

References

1. Harrison, J.; Toreini, E.; Mehrnezhad, M. A Practical Deep Learning-Based Acoustic Side Channel Attack on Keyboards. In *2023 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2023; pp. 270–280. [CrossRef]
2. AI Can Identify Passwords by Sound of Keys Being Pressed, Study Suggests. AI (Artificial Intelligence). *The Guardian*, 2025. Available online: https://www.theguardian.com/technology/2023/aug/08/ai-could-identify-passwords-by-sound-of-keys-being-pressed-study-suggests (accessed on 21 April 2026).
3. George, A.; Sagayarajan, S. Acoustic Eavesdropping: How AIs Can Steal Your Secrets by Listening to Your Typing. *Partners Univers. Int. Innov. J.* **2023**, *1*, 1–14. [CrossRef]
4. Venkateswarlu, S. An Overview of Acoustic Side-Channel Attack. 2013. Available online: https://www.academia.edu/91617557/An_Overview_of_Acoustic_Side_Channel_Attack (accessed on 18 April 2026).

5. Taheritajar, A.; Harris, Z.M.; Rahaeimehr, R. A Survey on Acoustic Side Channel Attacks on Keyboards. In *Information and Communications Security*; Katsikas, S., Xenakis, C., Kalloniatis, C., Lambrinoudakis, C., Eds.; Springer: Singapore, 2025; pp. 99–121. [CrossRef]
6. Savadatti, M.B.; Kulkarni, N.J.; Bansod, G.D.; Sarkar, P.; Kumar, P.; Sargam, K.; Gulati, P.; Bale, A.S. Keyboard Acoustic Side Channel Attacks on Android Phones in the AI Era of Digital Banking: An Elementary Review. *Int. J. Intell. Syst. Appl. Eng.* **2023**, *12*, 292–300.
7. Anand, S.A.; Saxena, N. Keyboard Emanations in Remote Voice Calls: Password Leakage and Noise(less) Masking Defenses. In *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy*; CODASPY '18; Association for Computing Machinery: New York, NY, USA, 2018; pp. 103–110. [CrossRef]
8. Rodrigues, D.; Macedo, G.; Conti, M.; Pinto, P. A Prototype for Generating Random Key Sounds to Prevent Keyboard Acoustic Side-Channel Attacks. In *2024 IEEE 22nd Mediterranean Electrotechnical Conference (MELECON)*; IEEE: Piscataway, NJ, USA, 2024; pp. 1287–1292. [CrossRef]
9. Marrone, S.; Sansone, C. Identifying Users' Emotional States through Keystroke Dynamics. In *Proceedings of the 3rd International Conference on Deep Learning Theory and Applications DeLTA-Volume 1*; SciTePress: Setúbal, Portugal, 2022; pp. 207–214. [CrossRef]
10. Xiong, J.; Belman, A.K. Multi-Modal Adversarial Activity Detection Using Keyboard and Mouse Dynamics. In *2025 IEEE World AI IoT Congress (AIIoT)*; IEEE: Piscataway, NJ, USA, 2025; pp. 0340–0345. [CrossRef]
11. Prete, A.L.; Andreini, P.; Bonechi, S.; Bianchini, M. A smart virtual keyboard to improve communication of locked-in patients. *Comput. Stand. Interfaces* **2025**, *93*, 103963. [CrossRef]
12. Lobanova, A.; He, J.; Zhu, N.; Nazir, A.; Ma, X. Machine-learning-based adaptive keyboard sound masking against acoustic channel attacks. In Proceedings of the Fifth International Conference on Computer Communication and Network Security (CCNS 2024), Guangzhou, China, 3–5 May 2024; p. 51. [CrossRef]
13. Tsalera, E.; Papadakis, A.; Pagiatakis, G.; Samarakou, M. Impact Evaluation of Sound Dataset Augmentation and Synthetic Generation upon Classification Accuracy. *J. Sens. Actuator Netw.* **2025**, *14*, 91. [CrossRef]
14. Rzemieniuk, M.; Niewiarowski, A.; Książek, W. Acoustic Side-Channel Vulnerabilities in Keyboard Input Explored Through Convolutional Neural Network Modeling: A Pilot Study. *Appl. Sci.* **2026**, *16*, 563. [CrossRef]
15. Rawf, K.M.H.; Abdulrahman, A.O.; Kamel, H.O.; Hassan, L.M.; Ali, A.O. Multi-Datasets for Different Keyboard Key Sound Recognition. *Data Brief* **2024**, *19*, 110949. Available online: https://pubmed.ncbi.nlm.nih.gov/39391001/ (accessed on 25 June 2026). [CrossRef]
16. Dias, T.; Vitorino, J.; Maia, E.; Sousa, O.; Praça, I. KeyRecs: A Keystroke Dynamics and Typing Pattern Recognition Dataset. *Data Brief* **2023**, *50*, 109509. [CrossRef] [PubMed]
17. Quattrone, B.; Badr, Y. SKAID: A Realistic Dataset for Acoustic Side-Channel Attacks with Synchronized Keyboard Audio and Keystroke Logs for AI-based Inference. (V 1.0) [Data Set]. Zenodo. Available online: https://zenodo.org/records/17282184 (accessed on 7 October 2025). [CrossRef]
18. Singh, I.; Kaur, G.; Atwal, U.P.S.; Singh, G.; Singh, G.; Singh, M. BEACON: A Multimodal Dataset for Learning Behavioral Fingerprints from Gameplay Data. *arXiv* **2026**, arXiv:2605.10867. [CrossRef]
19. Spata, M.O.; Maria Russo, V.; Ortis, A.; Battiato, S. A New Pipeline for Snooping Keystroke Based on Deep Learning Algorithm. *IEEE Access* **2025**, *13*, 24498–24514. [CrossRef]