

Intrusion Detection Based on Keystroke Biometrics and Siamese Neural Networks

ID I-7



Arnoldas Budžys, Viktor Medvedev, Olga Kurasova
Vilnius University, Institute of Data Science and Digital Technologies

arnoldas.budzys@mif.stud.vu.lt



Introduction

Cyber security is becoming one of the most important topics in today's critical infrastructure to ensure a secure connection between the administrator and the session. In recent years, the development of intrusion detection systems to protect against insiders has been a relevant and complex challenge in critical infrastructures.

If the system administrator's password is compromised or otherwise misappropriated, it could cause significant damage to the critical infrastructure. An insider threat is a harmful activity against an organisation by users with legitimate access to the organisation's infrastructure, software, or databases. These users may be current or former employees with access to the organisation's data. A methodology for user authentication of critical infrastructure systems based on a deep learning network is proposed.

Keystroke Biometrics

Let's say you are typing

"Welcome to 13th Data Analyses Methods For Software Engineering Conference"

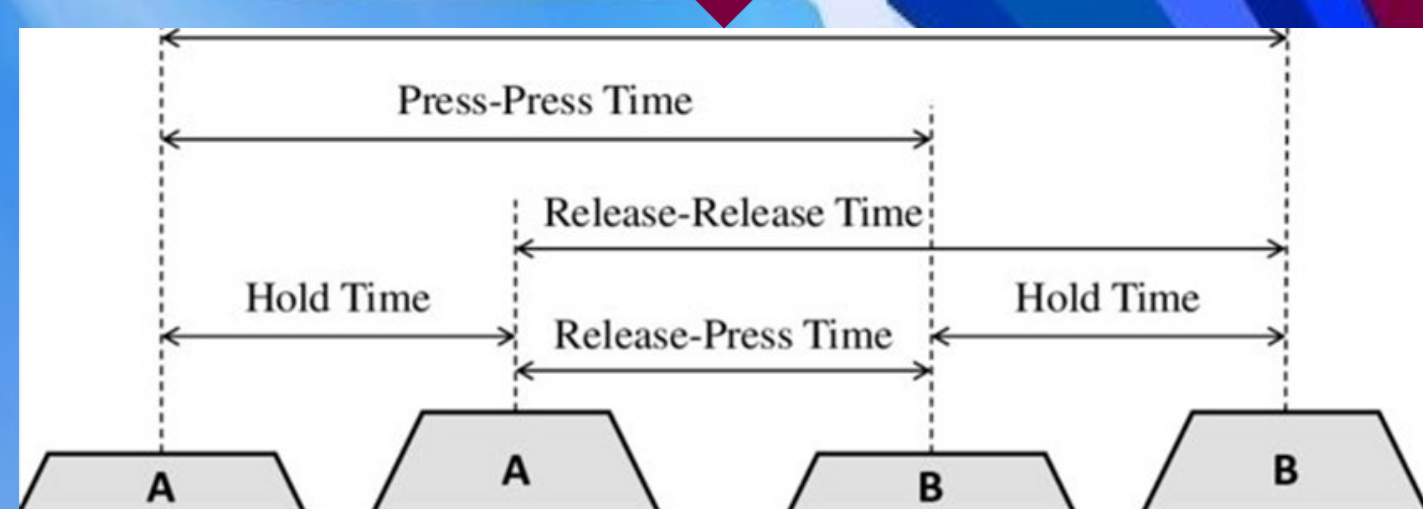
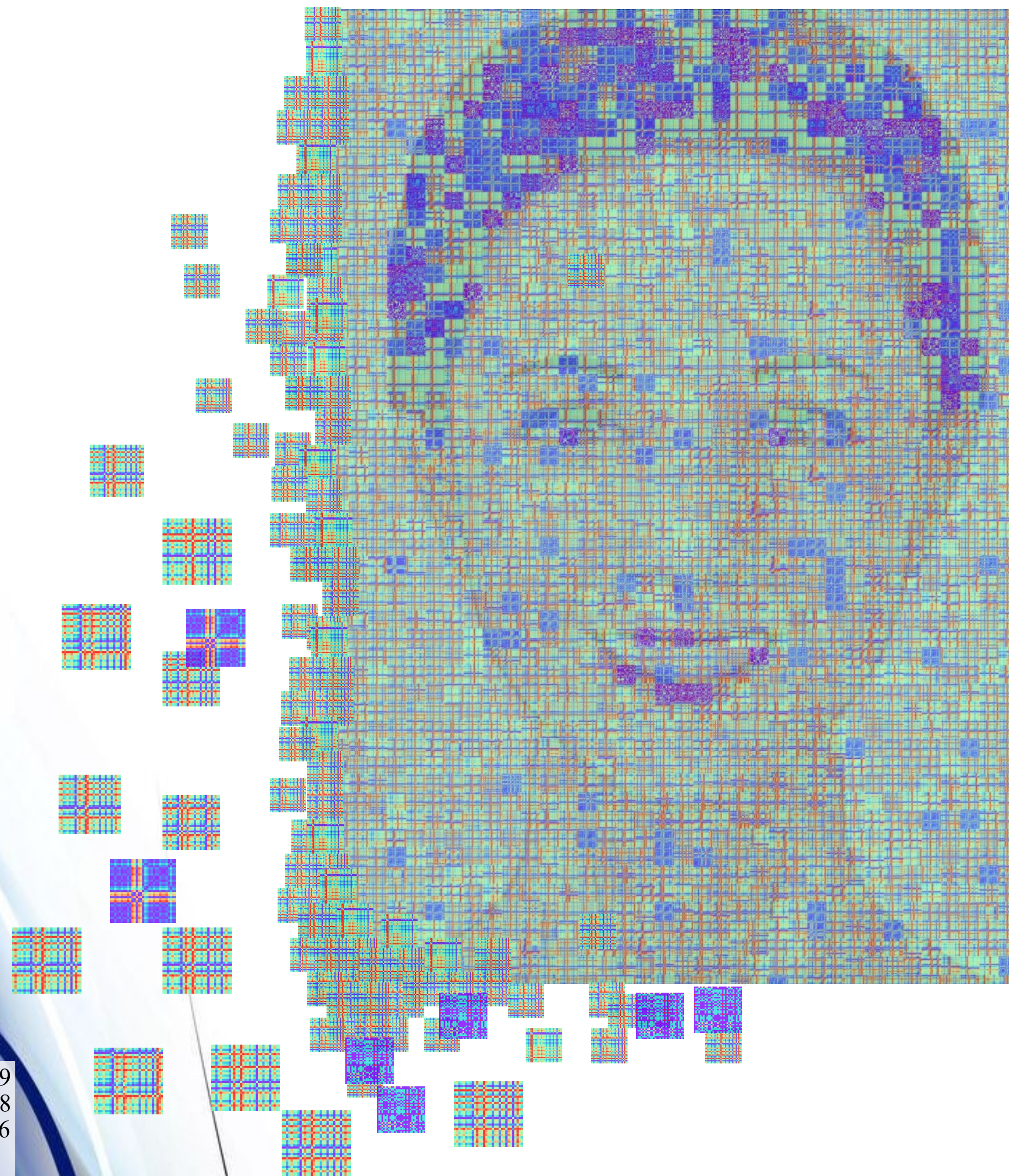


Figure 1. Keystroke Dynamics¹

Keystroke Data to Image



0.1491 0.3979 0.2488 0.1069 0.1674 0.0605 0.1169 0.2212 0.104 0.1417 1.1885 1.0468 0.1146 1.6055 1.4909 0.1067 0.759 0.652 0.1016 0.2136 0.112 0.1349 0.1484 0.0135 0.0932 0.3515 0.258 0.1338 0.3509 0.2171 0.0742 0.1491 0.3979 0.2488 0.1069 0.1674 0.0605 0.1169 0.2212 0.104 0.1417 1.1885 1.0468 0.1146 1.6055 1.4909 0.1067 0.759 0.652 0.1016 0.2136 0.112 0.1349 0.1484 0.0135 0.0932 0.3515 0.258 0.1338 0.3509 0.2171 0.0742 0.0137 0.0932 0.3515 0.258 0.1338 0.652

Methodology

Behavioural biometric data or user behavioural characteristics are converted into an image (Non-Image to Image) and further used in the proposed methodology for authentication. Converted password into image is more acceptable to convolutional neural networks.

The CMU dataset² was used for this experiment: 51 users from the Carnegie-Mellon University. All users typed the same password (.tie5Roanl) 400 times.

For future work other datasets will be used: Aalto University dataset³, GREYC dataset⁴.

- Markov Transition Field (MTF)
- Reccurence Plots (RPs)
- Gramian Angular Difference Field (GADF)
- Gramian Angular Summation Field (GASF)



Experiments and Results

Siamese neural networks can be employed for image similarity detection to distinguish a legal user from an insider. The results are promising, demonstrating that using a deep learning-based approach to analyse images obtained from user keystroke data can improve intrusion detection accuracy and perform user authentication more efficiently.

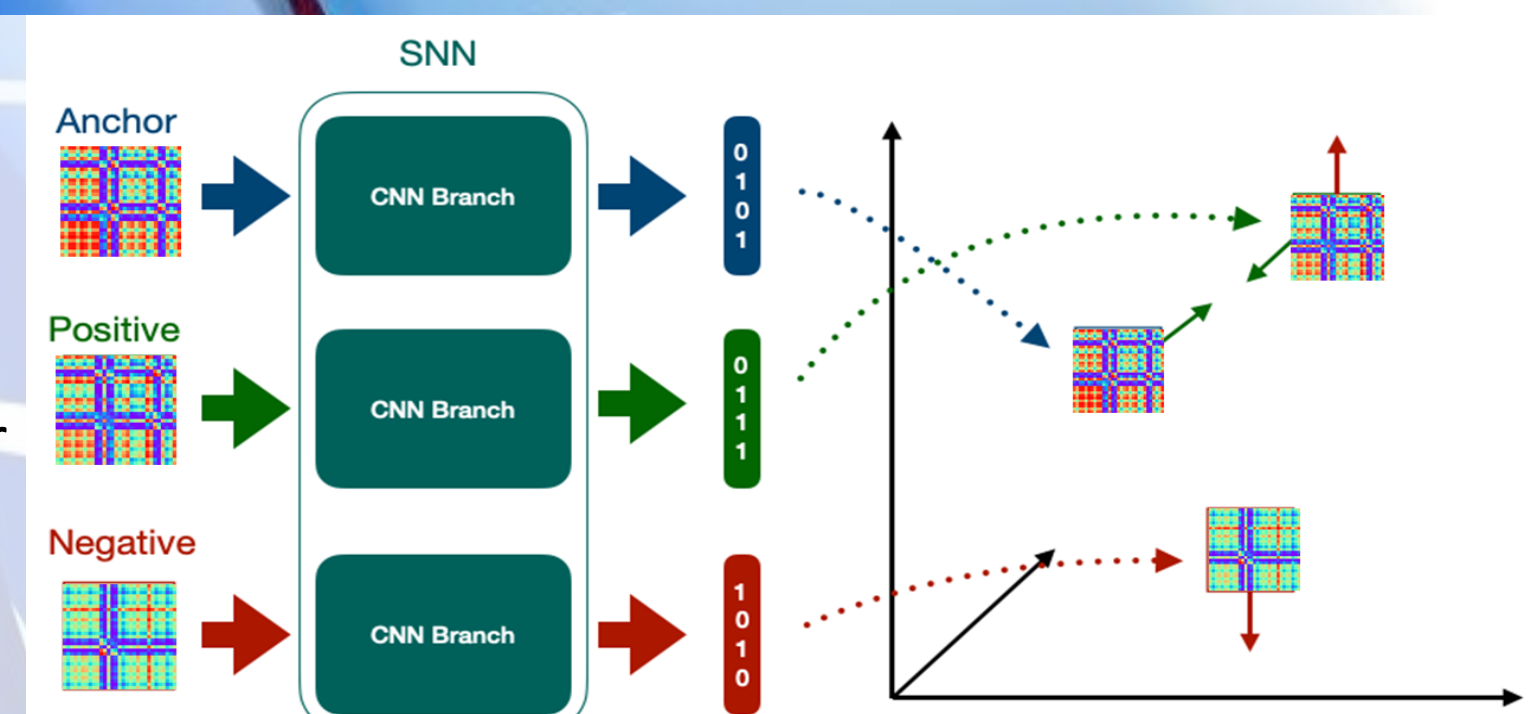


Figure 3. Siamese Neural Network using Triple Loss Function

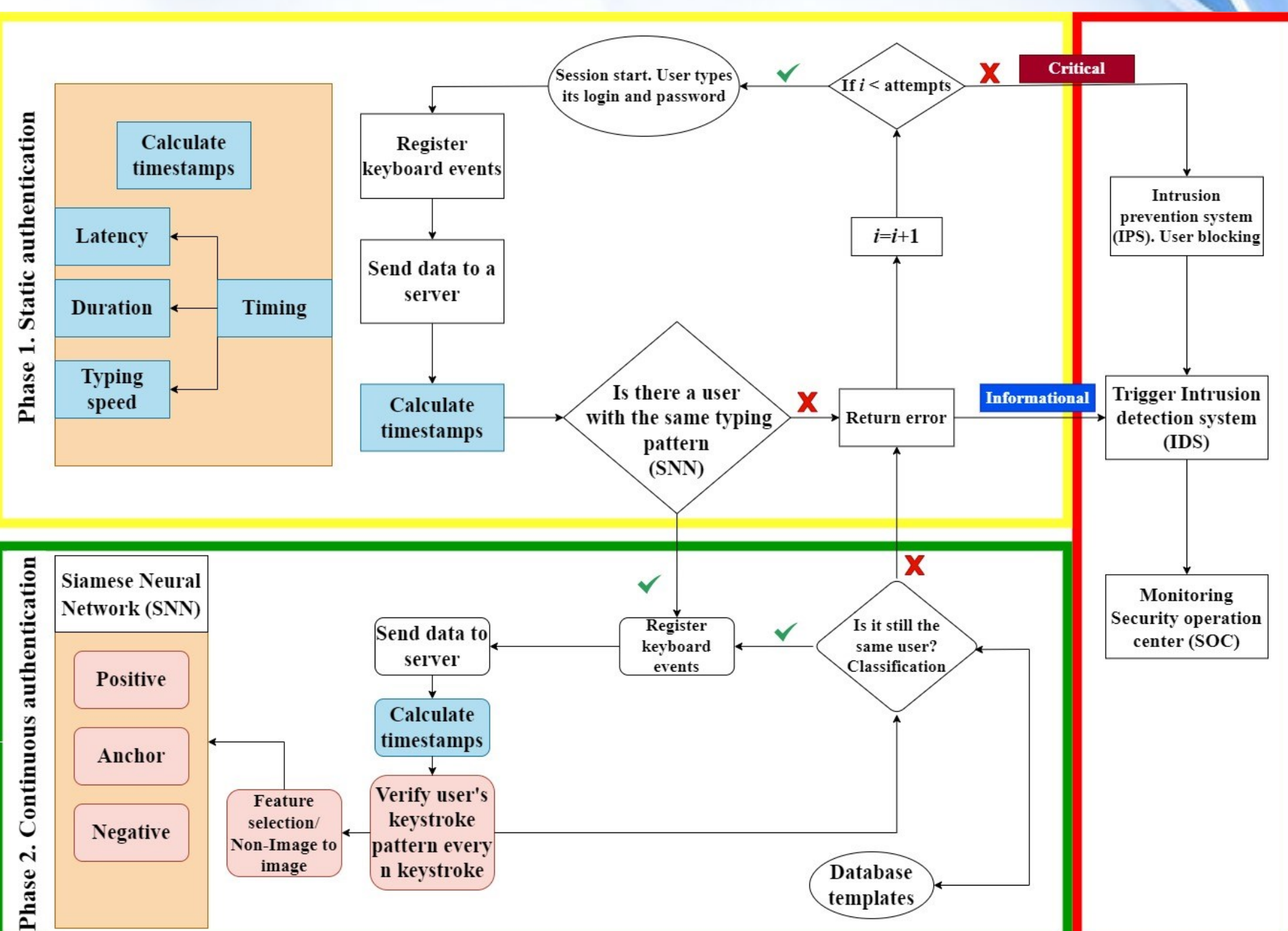


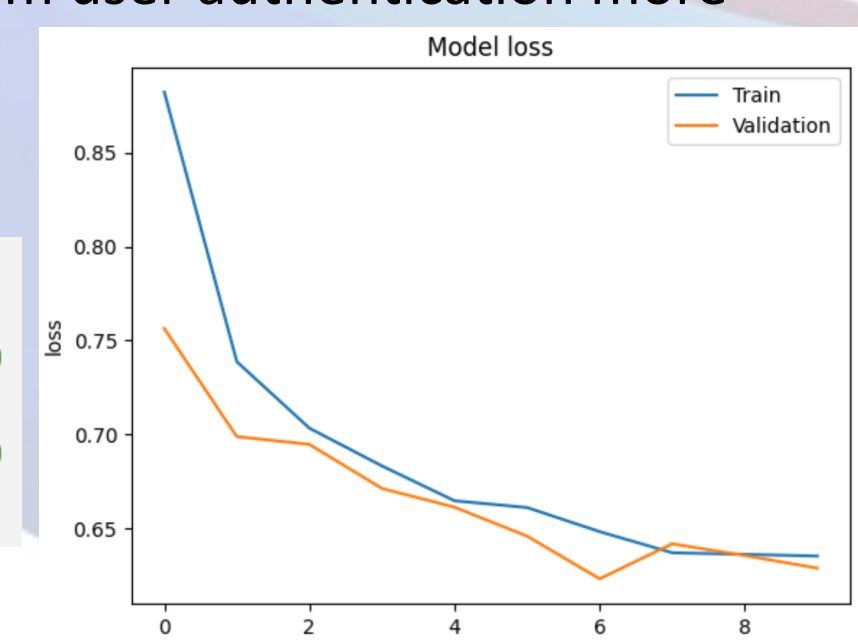
Figure 2. Methodology of User Authentication

```

1 cosine_similarity = metrics.CosineSimilarity()
2 sinus_similarity = metrics.Accuracy()
3 positive_similarity = cosine_similarity(anchor_embed, positive_embed)
4 print("Positive similarity:", positive_similarity.numpy())
5 negative_similarity = cosine_similarity(anchor_embed, negative_embed)
6 print("Negative similarity", negative_similarity.numpy())

```

Positive similarity: 0.96270293
Negative similarity 0.78226334



References: [1] Harilal, A., et al. (2017). Twos: A dataset of malicious insider threat behavior based on a gamified competition. In *Proceedings of the 2017 International Workshop on Managing Insider Security Threats* (pp. 45-56). [2] Killourhy, K. S., & Maxion, R. A. (2009). Comparing anomaly-detection algorithms for keystroke dynamics. In *2009 IEEE/IFIP International Conference on Dependable Systems & Networks* (pp. 125-134). IEEE. [3] Dhakal, V., Feit, A. M., Kristensson, P. O., & Oulasvirta, A. (2018, April). Observations on typing from 136 million keystrokes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-12). [4] Giot, R., El-Abed, M., & Rosenberger, C. (2009, September). Greyc keystroke: a benchmark for keystroke dynamics biometric systems. In *2009 IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems* (pp. 1-6). IEEE.