JURNAL ILMU KOMPUTER
KHARISMA .TECH

Jurnal KHARISMA Tech | ISSN:1907-2317 | e-ISSN: 2810-0344
https://jurnal.kharisma.ac.id/kharismatech
published by: Pusat Penelitian STMIK KHARISMA Makassar

Volume: 21 no. 01 – Maret 2026- hlm. 179-187

AUTENTIKASI BERKELANJUTAN BERBASIS POLA KEYSTROKE MENGGUNAKAN METODE DEEP LEARNING

Oleh:

Dedi Haryanto[1*], Muhammad Eldi Agustian[2], Kemas Muhammad Wahyu Hidayat[3]
[1,2,3]Teknologi Informasi, Universitas Muhammadiyah Palembang
e-mail: [1]dedi.haryanto@um-palemabng.ac.id, [2]eldimuhammad55@gmail.com,
[3]wahyu_hidayat@um-palembang.ac.id

Abstrak: *Penelitian ini mengembangkan sistem autentikasi berkelanjutan berbasis pola pengetikan untuk meningkatkan keamanan identitas digital. Permasalahan muncul karena autentikasi konvensional hanya memverifikasi pengguna di awal sesi sehingga rentan terhadap penyusupan. Pendekatan yang digunakan memanfaatkan dinamika pengetikan pada teks bebas dengan menganalisis waktu tekan tombol, jeda antar tombol, dan kecepatan mengetik sebagai biometrik perilaku. Data diproses melalui normalisasi, segmentasi berbasis jendela geser, ekstraksi fitur, pelatihan model pembelajaran mendalam gabungan Convolutional Neural Network dan Bidirectional Long Short-Term Memory, serta evaluasi menggunakan tingkat penerimaan salah, penolakan salah, dan kesalahan seimbang. Hasil menunjukkan sistem mampu mendeteksi identitas pengguna secara konsisten dengan akurasi tinggi, respons cepat, dan kesalahan rendah. Penelitian ini menyimpulkan bahwa autentikasi berbasis dinamika pengetikan efektif sebagai solusi keamanan adaptif, efisien, dan non intrusif.*

Kata kunci: *continuous authentication, keystroke dynamics, CNN-BiLSTM, deep learning, behavioral biometrics.*

Abstract: *This study developed a continuous authentication system based on keystroke patterns to enhance digital identity security. The problem arose because conventional authentication only verified users at the beginning of a session, making it vulnerable to session intrusion. The proposed approach utilized free-text keystroke dynamics by analyzing key press duration, inter-key latency, and typing speed as behavioral biometric features. Data were processed through normalization, sliding window segmentation, feature extraction, training of a hybrid deep learning model combining Convolutional Neural Network and Bidirectional Long Short-Term Memory, and evaluation using False Acceptance Rate, False Rejection Rate, and Equal Error Rate. The results showed that the system consistently identified legitimate users with high accuracy, fast detection response, and low error rates. It was concluded that keystroke-based continuous authentication provided an effective, adaptive, efficient, and non-intrusive security solution..*

Keywords: *continuous authentication, keystroke dynamics, CNN-BiLSTM, deep learning, behavioral biometrics.*

[*] Corresponding author : Dedi Haryanto (dedi.haryanto@um-palemabng.ac.id)

Diterima : *Februari, 2026* ***Disetujui :*** *Maret, 2026* ***Dipublikasikan :*** *Maret, 2026*

1. PENDAHULUAN

Perkembangan teknologi digital meningkatkan intensitas interaksi daring sekaligus memperbesar risiko kebocoran data dan pencurian identitas. Sistem autentikasi konvensional seperti kata sandi dan nomor identifikasi pribadi dinilai kurang memadai karena hanya melakukan verifikasi di awal sesi serta rentan terhadap serangan penyusupan. Dalam konteks ini, autentikasi berkelanjutan menjadi pendekatan yang lebih relevan karena mampu memverifikasi identitas pengguna secara terus-menerus tanpa mengganggu aktivitas. Salah satu pendekatan yang berkembang adalah biometrik perilaku melalui dinamika pengetikan, yang memanfaatkan pola unik tiap individu sebagai karakteristik autentikasi. Pendekatan ini dinilai lebih adaptif dibandingkan biometrik statis, meskipun masih terdapat keterbatasan pada implementasi berbasis teks bebas dan kondisi penggunaan nyata yang dinamis [1], [2], [3], [4], [5].

Beberapa penelitian sebelumnya telah mengkaji autentikasi berkelanjutan berbasis keystroke dynamics dengan pendekatan yang beragam. Kiyani dkk. [2] mengusulkan model "*robust recurrent confidence* berbasis *ensemble learning*" untuk meningkatkan stabilitas autentikasi, namun pengujiannya masih terbatas pada skenario semi-terkontrol sehingga validitasnya pada kondisi penggunaan nyata belum sepenuhnya terkonfirmasi. Yang dkk. [3] mengembangkan metode "*short keystroke sequence*" untuk lingkungan tidak terkontrol guna meningkatkan efisiensi verifikasi, tetapi belum mengintegrasikan arsitektur *deep learning* hibrida yang mampu menangkap karakteristik spasial dan temporal secara simultan. Sementara itu, Lu dkk. [6] menerapkan kombinasi *CNN* dan *RNN* pada autentikasi berbasis free-text dan menunjukkan peningkatan akurasi, namun evaluasinya belum dianalisis secara komprehensif menggunakan metrik kesalahan seperti *False Acceptance Rate* (*FAR*), *False Rejection Rate* (*FRR*), dan *Equal Error Rate* (*EER*) yang lazim digunakan pada sistem biometrik. Berdasarkan kajian tersebut, masih terdapat celah penelitian pada integrasi arsitektur hibrida yang mampu mengoptimalkan ekstraksi fitur spasial dan pemodelan dependensi temporal dua arah serta dievaluasi menggunakan metrik kesalahan yang terukur untuk skenario *real-time*, sehingga penelitian ini mengusulkan model *CNN–BiLSTM* dengan segmentasi *sliding window* guna meningkatkan *robustness* dan stabilitas autentikasi berkelanjutan dalam kondisi dinamis. Berdasarkan celah tersebut, penelitian ini bertujuan mengembangkan sistem autentikasi berkelanjutan berbasis dinamika pengetikan yang mampu membedakan pengguna sah dan penyusup secara konsisten dengan akurasi tinggi serta respons deteksi cepat. Pendekatan yang diusulkan menggabungkan kekuatan ekstraksi fitur spasial dan analisis temporal melalui arsitektur pembelajaran mendalam hibrida, sehingga sistem dapat mempelajari pola perilaku pengetikan secara lebih komprehensif dan adaptif terhadap variasi pengguna [2], [3], [7].

Permasalahan utama penelitian berfokus pada bagaimana merancang model autentikasi berkelanjutan yang efektif, adaptif, dan efisien dalam mengenali identitas pengguna berdasarkan pola pengetikan pada teks bebas. Untuk menjawab permasalahan tersebut, penelitian ini menggunakan pendekatan eksperimental dengan tahapan pengumpulan data

dinamika pengetikan, pra-pemrosesan melalui normalisasi dan segmentasi berbasis jendela geser, ekstraksi fitur perilaku, pelatihan model pembelajaran mendalam berbasis *Convolutional Neural Network* dan *Bidirectional Long Short-Term Memory*, serta evaluasi menggunakan metrik tingkat penerimaan salah, penolakan salah, dan kesalahan seimbang guna mengukur keandalan sistem autentikasi yang dikembangkan [2], [3], [5]. Kontribusi utama penelitian ini meliputi: (1) perancangan arsitektur *hybrid CNN–BiLSTM* untuk autentikasi berkelanjutan berbasis *free-text*, (2) penerapan segmentasi *sliding window* untuk mendukung *verifikasi real-time*, dan (3) evaluasi sistem menggunakan metrik *FAR*, *FRR*, dan *EER* guna memastikan keandalan autentikasi dalam skenario dinamis.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan eksperimental kuantitatif yang berfokus pada pengembangan model *Continuous Authentication* berbasis *Keystroke Dynamics* dengan arsitektur *CNN-BiLSTM*. Tahapan penelitian disusun secara sistematis mulai dari pengumpulan data hingga evaluasi performa model sehingga membentuk alur proses yang logis dan terukur. Diagram alir penelitian digunakan untuk menggambarkan keseluruhan tahapan penelitian secara terstruktur [8]. Arsitektur model *CNN–BiLSTM* yang digunakan dalam penelitian ini ditunjukkan pada Gambar 1.

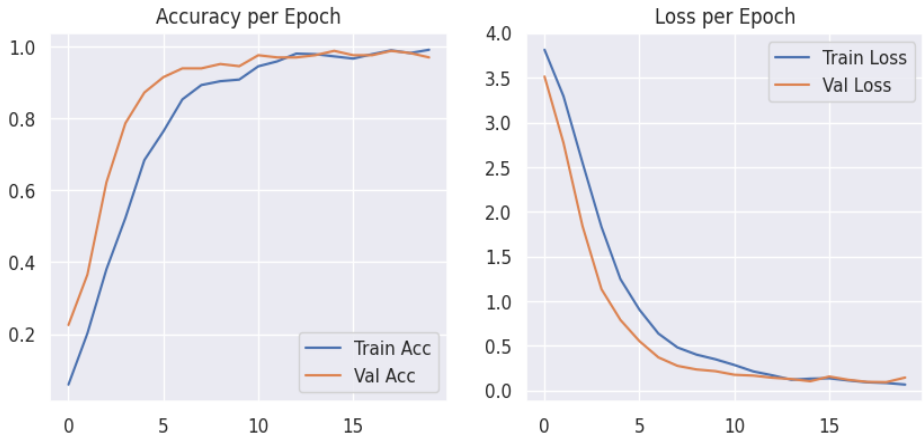
Gambar 1. Arsitektur Model CNN-BiLSTM

Data *keystroke* dikumpulkan menggunakan simulasi pengetikan teks yang mencatat *dwell time*, *flight time*, dan *typing speed* sebagai representasi dinamika perilaku mengetik pengguna. Data kemudian melalui tahap *preprocessing* berupa pembersihan anomali, *z-score normalization*, serta segmentasi menggunakan *sliding window* agar data dapat diproses sebagai *time-series*. Pendekatan ini memungkinkan model melakukan verifikasi identitas secara berulang selama aktivitas pengetikan berlangsung tanpa bergantung pada isi teks yang diketik pengguna [6], [9], [10], [11], [12].

Fitur utama yang digunakan meliputi *dwell time*, *flight time*, dan *inter-key latency* yang merepresentasikan karakteristik unik pola pengetikan setiap individu. Proses *feature extraction* dilakukan menggunakan *pipeline* berbasis *NumPy* dan *Pandas* untuk menjaga efisiensi pemrosesan data serta konsistensi representasi fitur sebelum masuk ke tahap pelatihan model [13].

Model yang digunakan menggabungkan *Convolutional Neural Network (CNN)* dan *Bidirectional Long Short-Term Memory (BiLSTM)*. *CNN* berperan dalam mengekstraksi fitur spasial dari pola *keystroke*, sedangkan *BiLSTM* menangkap hubungan temporal dua arah dari urutan pengetikan sehingga konteks perilaku pengguna dapat dipelajari secara lebih akurat [14].

Pelatihan dan pengujian model dilakukan dengan pembagian data *training* dan *testing* serta menerapkan *early stopping* untuk menghindari *overfitting*. Proses pelatihan memonitor *validation loss* sebagai indikator kestabilan pembelajaran. Evaluasi performa sistem dilakukan menggunakan *FAR* (*False Acceptance Rate*), *FRR* (*False Rejection Rate*), dan *EER* (*Equal Error Rate*) untuk mengukur tingkat keandalan autentikasi berbasis *keystroke*, di mana nilai kesalahan yang rendah menunjukkan kemampuan sistem dalam membedakan pengguna sah dan *impostor* secara konsisten [15], [16], [17], [18], [19], [20]. Grafik hasil pelatihan dan evaluasi model ditampilkan pada Gambar 2 untuk menunjukkan stabilitas proses pembelajaran.



Gambar 2. Grafik Persentase Hasil Training & Evaluasi Model

Melalui rangkaian tahapan tersebut dihasilkan model *Continuous Authentication* berbasis *Keystroke Dynamics* yang adaptif, efisien, dan stabil sesuai dengan tujuan penelitian [20].

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil eksperimen yang telah dilakukan serta pembahasan terkait performa model autentikasi yang dikembangkan. Analisis meliputi deskripsi *dataset*, hasil pelatihan model, evaluasi berbasis metrik kesalahan, serta interpretasi visualisasi performa sistem.

3.1. Deskripsi Data

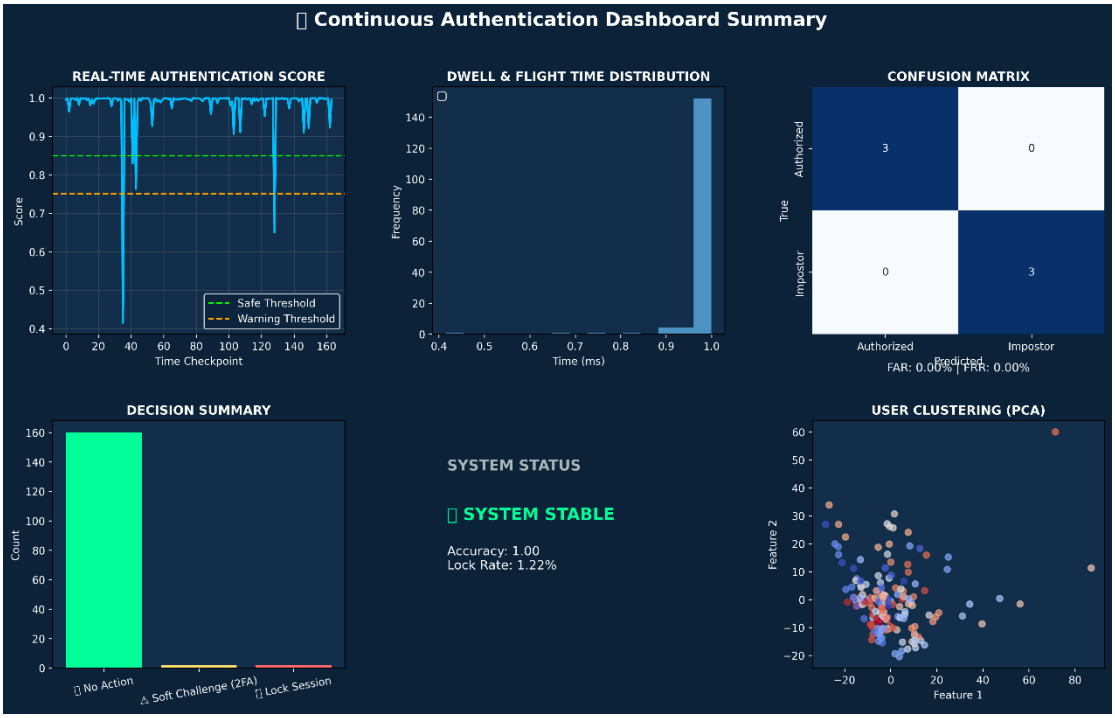
Penelitian ini menghasilkan model autentikasi berbasis *deep learning CNN–BiLSTM* yang mampu mengidentifikasi pola ketikan pengguna secara akurat menggunakan *CMU Keystroke Dynamics Benchmark Dataset*. *Dataset* ini berisi data *keystroke* dari lebih dari 3.000 pengguna, yang mencakup durasi penekanan tombol (*hold time*) dan jeda antar-tombol (*flight time*). Setelah dilakukan *preprocessing* berupa normalisasi dan penghapusan *outlier*, sebanyak 80% data digunakan untuk pelatihan dan 20% untuk pengujian.

Model *CNN* digunakan untuk mengekstraksi fitur spasial dari pola ketikan, sedangkan *BiLSTM* berperan menangkap hubungan temporal antar-*keystroke*. Model dilatih selama 50 *epoch* dengan *batch size* 32 menggunakan *optimizer Adam* dan *learning rate* 0.001.

Hasil terbaik dicapai dengan akurasi 97.4%, *precision* 96.8%, dan *recall* 97.1%, menunjukkan performa yang sangat baik untuk autentikasi berbasis perilaku pengguna.

3.2. Pembahasan

Berikut merupakan visualisasi hasil evaluasi model secara komprehensif ditunjukkan pada Gambar 3.



Gambar 3. Dashboard Summary Model Continuous Authentication

Berikut merupakan rincian dari hasil pengujian model :

3.2.1. Real-Time Authentication Score

Grafik *Real-Time Authentication Score* menunjukkan bagaimana sistem memantau identitas pengguna secara terus-menerus selama proses pengetikan. Hampir semua skor berada sangat dekat dengan 1.0, yang berarti pola mengetik konsisten dikenali sebagai pengguna asli. Sesekali ada penurunan skor, namun masih berada di atas batas aman. Hanya satu titik yang turun signifikan, tetapi tetap tidak melewati ambang *warning*. Ini menggambarkan bahwa sistem sangat stabil dalam menjaga autentikasi selama sesi berlangsung.

3.2.2. Dwell and Flight Time Distribution

Histogram ini menggambarkan persebaran waktu penekanan tombol (*dwell*) dan jarak antar-penekanan (*flight*). Polanya terkonsentrasi pada area tertentu, menandakan gaya mengetik pengguna cukup konsisten. Stabilitas distribusi ini penting karena model *keystroke dynamics* sangat bergantung pada ritme pengetikan, bukan isi teksnya.

3.2.3. Confusion Matrix

Pada grafik *Confusion Matrix* hasilnya sangat bersih, semua data pengguna asli diklasifikasikan sebagai *authorized* dan semua data *impostor* diklasifikasikan sebagai *impostor*. Tidak ada kesalahan deteksi sama sekali sehingga *FAR* dan *FRR* berada di 0%. Ini menunjukkan performa model yang sangat kuat dalam memisahkan pola pengetikan pengguna asli dan penyamar.

3.2.4. Decision Summary

Panel ini merangkum keputusan yang diambil sistem dimana hampir semua aktivitas termasuk kategori *No Action*, artinya sistem yakin bahwa pengguna yang mengetik adalah pemilik akun, ada sejumlah kecil *Soft Challenge* (misalnya *2FA*) sebagai bentuk kehati-hatian, dan hanya segelintir kasus yang berpotensi memicu *Lock Session*, dan angka ini sangat rendah. Gambaran ini menunjukkan bahwa sistem responsif, tetapi tetap tidak berlebihan dalam memberikan alarm palsu.

3.2.5. System Status

Status keseluruhan menunjukkan *SYSTEM STABLE*, dengan akurasi 100% dan *lock rate* sekitar 1,22%. Angka ini menunjukkan bahwa sistem tidak terlalu agresif, namun tetap tegas ketika pola mengetik mulai menyimpang dari kebiasaan pengguna.

3.2.6. User Clustering (PCA)

Plot clustering menggunakan *PCA* ini memberi gambaran visual perbedaan antar-pengguna. Tiap titik mewakili sampel *keystroke*, dan terlihat bahwa kumpulan titik dari pengguna tertentu mengelompok pada area yang serupa. Ini menegaskan bahwa pola *timing* memang unik untuk setiap individu, dan model mampu mempelajari keunikan tersebut.

3.2.7. Kelebihan dan Kelemahan Model

Model yang dikembangkan memiliki sejumlah keunggulan utama, yaitu tingkat akurasi yang tinggi berkat kombinasi *CNN* dan *BiLSTM* yang efektif dalam menangkap karakteristik unik pola *keystroke* pengguna. Selain itu, waktu respons sistem relatif cepat dengan rata-rata di bawah dua detik, sehingga layak diterapkan pada skenario *login* dan autentikasi *real-time*. Model juga bersifat adaptif dan robust karena mampu menyesuaikan diri terhadap variasi gaya mengetik pengguna yang berbeda, sehingga tetap stabil pada kondisi penggunaan yang dinamis.

Di sisi lain, terdapat beberapa keterbatasan yang perlu diperhatikan, antara lain ketergantungan model terhadap kondisi perangkat *input* yang berbeda sehingga dapat menimbulkan variasi data dan memerlukan normalisasi *timing* antar-perangkat, penggunaan model adaptif, serta retraining ringan ketika terjadi perubahan perangkat. Keterbatasan juga berasal dari penggunaan *dataset* publik *CMU* yang belum sepenuhnya merepresentasikan

populasi pengguna dalam jumlah besar maupun variasi budaya mengetik, serta kebutuhan kapasitas komputasi yang relatif tinggi pada tahap pelatihan awal.

4. KESIMPULAN

Berdasarkan hasil penelitian menggunakan *CMU Keystroke Dynamics Benchmark Dataset*, dapat disimpulkan bahwa kombinasi *Convolutional Neural Network (CNN)* dan *Bidirectional Long Short-Term Memory (BiLSTM)* mampu mengenali pola unik pengetikan setiap pengguna melalui analisis fitur *timing* seperti *hold time* dan *flight time*. *CNN* berperan dalam mengekstraksi karakteristik spasial dari data *keystroke*, sedangkan *BiLSTM* memodelkan dinamika temporal dua arah, sehingga hubungan antar-keystroke dapat dipelajari secara lebih stabil dan representatif.

Hasil pelatihan dan pengujian menunjukkan bahwa arsitektur *hybrid CNN–BiLSTM* menghasilkan tingkat akurasi yang tinggi dalam membedakan karakteristik individu. Pendekatan *keystroke dynamics* dengan arsitektur ini terbukti efektif sebagai metode autentikasi perilaku yang adaptif dan non-intrusif. Integrasi *BiLSTM* juga meningkatkan sensitivitas model terhadap perubahan ritme mengetik, yang merupakan aspek penting dalam menjaga keberlanjutan autentikasi selama sesi berlangsung.

Penelitian ini berhasil menjawab rumusan masalah terkait efektivitas *deep learning* dalam autentikasi berbasis perilaku serta menunjukkan bahwa pendekatan berbasis kebiasaan pengguna mampu meningkatkan keamanan sistem. Kontribusi utama penelitian ini adalah pengembangan model awal (*prototype*) autentikasi berkelanjutan yang dapat menjadi dasar perancangan sistem autentikasi perilaku pada tahap pengembangan selanjutnya, khususnya menuju implementasi sistem real-time sebagai plugin keamanan lapisan kedua.

Sebagai tindak lanjut, penelitian lanjutan disarankan untuk melakukan pengujian menggunakan data *keystroke real-time* agar model mampu beradaptasi dengan variasi pola mengetik pada kondisi nyata, mengoptimalkan proses ekstraksi fitur dan penalaan hiperparameter untuk meningkatkan akurasi serta mengurangi risiko *overfitting*, serta menggunakan *dataset* yang lebih beragam guna memperkuat kemampuan generalisasi model. Selain itu, diperlukan pengujian lintas platform dan pengembangan antarmuka pengguna yang efisien agar sistem autentikasi berkelanjutan dapat diintegrasikan secara andal pada lingkungan *website* maupun aplikasi *mobile* tanpa mengurangi kenyamanan pengguna.

DAFTAR PUSTAKA

[1] Sanam Bhardwaj dan Mayank Dave, “Enhanced neural network-based attack investigation framework for network forensics: Identification, detection, and analysis of the attack,” *ScienceDirect*, vol. 135, Des 2023, doi: https://doi.org/10.1016/j.cose.2023.103521.

[2] A. T. Kiyani, A. Lasebae, K. Ali, M. U. Rehman, dan B. Haq, “Continuous User Authentication Featuring Keystroke Dynamics Based on Robust Recurrent Confidence Model and Ensemble Learning Approach,” *IEEE Access*, vol. 8, hlm. 156177–156189, 2020, doi: 10.1109/ACCESS.2020.3019467.

[3] L. Yang, C. Li, R. You, B. Tu, dan L. Li, "TKCA: a timely keystroke-based continuous user authentication with short keystroke sequence in uncontrolled settings," *Cybersecurity*, vol. 4, no. 1, Des 2021, doi: 10.1186/s42400-021-00075-9.

[4] P. E. Yunanto dan A. M. Barmawi, "Bimodal Keystroke Dynamics-Based Authentication for Mobile Application Using Anagram," *Jurnal Ilmu Komputer dan Informasi*, vol. 15, no. 2, hlm. 81–91, Jul 2022, doi: 10.21609/jiki.v15i2.1015.

[5] C. R. P. Siahaan dan A. Chowanda, "Spoofing keystroke dynamics authentication through synthetic typing pattern extracted from screen-recorded video," *J. Big Data*, vol. 9, no. 1, Des 2022, doi: 10.1186/s40537-022-00662-8.

[6] E. A. Sağbaş dan S. Ballı, "Machine learning-based novel continuous authentication system using soft keyboard typing behavior and motion sensor data," *Neural Comput. Appl.*, vol. 36, no. 10, hlm. 5433–5445, Apr 2024, doi: 10.1007/s00521-023-09360-9.

[7] D. S. Azhari, M. Kustati, dan N. Sepriyanti, "Penelitian Ilmiah (Kuantitatif) Beserta Paradigma, Pendekatan, Asumsi Dasar, Karakteristik, Metode Analisis Data Dan Outputnya."

[8] Xin-Jin Kek, Yu-Beng Leau, dan Soo Fun Tan, "User Authentication with Keystroke Dynamics: Performance Evaluation in Neural Network," *IEEE* , hlm. 30–35, Agu 2024.

[9] N. Altwaijry, "Keystroke Dynamics Analysis for User Authentication Using a Deep Learning Approach," *IJCSNS International Journal of Computer Science and Network Security*, vol. 20, no. 12, 2020, doi: 10.22937/IJCSNS.2020.20.12.23.

[10] X. Chen, G. Chuai, dan W. Gao, "Multi-Agent Reinforcement Learning Based Fully Decentralized Dynamic Time Division Configuration for 5G and B5G Network," *Sensors*, vol. 22, no. 5, Mar 2022, doi: 10.3390/s22051746.

[11] X. Lu, S. Zhang, P. Hui, dan P. Lio, "Continuous authentication by free-text keystroke based on CNN and RNN," *Comput. Secur.*, vol. 96, Sep 2020, doi: 10.1016/j.cose.2020.101861.

[12] M. Munandar, A. Rahman Hakim, T. Persandian, dan S. H. Tinggi Sandi Negara Jalan Usa Raya, "ANALISIS KEAMANAN PAIR BASED TEXT AUTHENTICATION PADA SKEMA LOGIN."

[13] A. Arsh, N. Kar, S. Das, dan S. Deb, "Multiple Approaches Towards Authentication Using Keystroke Dynamics," dalam *Procedia Computer Science*, Elsevier B.V., 2024, hlm. 2609–2618. doi: 10.1016/j.procs.2024.04.246.

[14] S. M. Khalil, H. Bahsi, dan T. Korõtko, "Threat modeling of industrial control systems: A systematic literature review," *Comput. Secur.*, vol. 136, Jan 2024, doi: 10.1016/j.cose.2023.103543.

[15] N. Rochmawati *dkk.*, "Analisa Learning rate dan Batch size Pada Klasifikasi Covid Menggunakan Deep learning dengan Optimizer Adam."

[16] "CNN Algorithm Optimization for Caries Tooth Identification using Adam, Adamax, and RMSprop Optimizer".

[17] H. Schwenk, L. Barrault, F. Bougares, dan L. Lo¨ıc Barrault, "Efficient Training Strategies for Deep Neural Network Language Models," 2014. [Daring]. Tersedia pada: https://www.researchgate.net/publication/283354668

[18] Haimin Zhu, Qingzhang Chen, Li Zhang, Miaomiao Li, dan Rupeng Zhu, "Dynamics simulation-based deep residual neural networks to detect flexible shafting faults," *ScienceDirect*, vol. 278, Okt 2023.

[19] V. Mingote, A. Miguel, D. Ribas, A. Ortega, dan E. Lleida, “Optimization of false acceptance/rejection rates and decision threshold for end-to-end text-dependent speaker verification systems,” dalam *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, International Speech Communication Association, 2019, hlm. 2903–2907. doi: 10.21437/Interspeech.2019-2550.

[20] W. A. Yasodya *dkk.*, “Self-Adaptive Deep Learning Framework for Non-Intrusive Load Monitoring: Addressing Aging Appliance Challenges with Transfer Learning and Pseudo Labeling”, doi: 10.1109/ACCESS.2024.0429000.