Indones. J. Math. Nat. Sci. 46(2) 2023
Indonesian Journal of Mathematics and Natural Sciences
http://journal.unnes.ac.id/nju/index.php/JM

Penerapan Model Hibrida CNN-GRU-BiLSTM-PCA Untuk Meningkatkan Akurasi Deteksi Serangan Jaringan Pada *Intrusion Detection System*

Muhammad Iqbal Yoshanda*, Alamsyah

Program Studi Teknik Informatika, FMIPA, Universitas Negeri Semarang, Indonesia
Gedung D5 Lt.2, Kampus Sekaran Gunungpati, Semarang 50229
E-mail: 1yoshanda95@students.unnes.ac.id, alamsyah@mail.unnes.ac.id

Diterima 12 Februari 2023 Disetujui 16 September 2023 Dipublikasikan 16 Oktober 2023

Abstrak

Intrusion detection system merupakan teknik pertahanan yang populer terhadap serangan siber. Berdasarkan metode pendeteksiannya dapat dibedakan menjadi *signatured-based* dan *anomaly-based*. Kedua metode tersebut memiliki kelemahan masing-masing, *signatured-based* memerlukan pemeliharaan database yang mencatat perilaku abnormal yang mahal, memakan waktu dan tidak dapat secara efektif menemukan serangan yang muncul untuk pertama kalinya. Sementara *anomaly-based* sulit mendapatkan data benchmark yang akurat, tingkat false alarm yang lebih tinggi, dan keterlambatan data benchmark. Pada penelitian ini, dilakukan penerapan model hibrida untuk meningkatkan akurasi deteksi yaitu metode Convolutional Neural Network (CNN) untuk mengekstrak fitur dari dataset, Gated Recurrent Unit (GRU) dan Bidirectional Long Short-Term Memory (BiLSTM) untuk memproses dan memahami informasi lalu lintas jaringan dalam dataset, serta metode Principal Component Analysis (PCA) untuk mereduksi dimensi data. Model hibrida CNN-GRU-BILSTM-PCA berhasil mendapatkan nilai akurasi sebesar 98,34% dan loss sebesar 0,048%. Model hibrida ini merupakan model terbaik di antara model-model penelitian sebelumnya.

Kata kunci: *Intrusion Detection System*, CNN, GRU, BiLSTM, PCA

Abstract

Intrusion detection system is a popular defense technique against cyber attacks. Based on the detection method, it can be divided into signatured-based and anomaly-based. Both methods have their own weaknesses, signatured-based requires maintaining a database that records abnormal behavior which is expensive and time-consuming and cannot effectively find attacks that appear for the first time. While anomaly-based is difficult to get accurate benchmark data, higher false alarm rates, and delays in benchmark data. In this research, a hybrid model is applied with the aim of improving detection accuracy where Convolutional Neural Network (CNN) method is used to extract features from the dataset, Gated Recurrent Unit (GRU) and Bidirectional Long Short-Term Memory (BiLSTM) are used to process and understand network traffic information in the dataset, and Principal Component Analysis (PCA) method is used to reduce data dimension. The CNN-GRU-BILSTM-PCA hybrid model succeeded in getting an accuracy value of 98.34% and a loss of 0.048%. This hybrid model is the best model among previous research models.

Keywords: *Intrusion Detection System*, CNN, GRU, BiLSTM, PCA

How to cite:
Yoshanda M. I., Alamsyah. (2023). Penerapan model hibrida CNN-GRU-BiLSTM-PCA untuk meningkatkan akurasi deteksi serangan jaringan pada intrusion detection system. *Indonesian Journal of Mathematics and Natural Sciences*, 46(2), 61-67.

PENDAHULUAN

Penggunaan teknologi dan komputer terus berkembang dengan cepat karena perkembangan informasi yang pesat. Banyak organisasi dan individu menggunakan sistem server dan komputer untuk menyimpan data penting dan sensitif. Perkembangan teknologi membuat organisasi memiliki

informasi yang sangat besar dalam berbagai aspek, termasuk komputasi awan (Gartner, 2022), teknologi 5G (Admaja, 2018), *internet of things*, dan *smart manufacturing* (Lee *et al*., 2022). Oleh karena itu, kebanyakan perangkat komputer harus terhubung ke internet. Namun, hal ini dapat menimbulkan resiko bagi keamanan jaringan, seperti serangan *phishing*, *ransomware*, dan *distributed denial of service* (DDoS).

Menurut laporan monitoring keamanan siber yang dipublikasikan oleh Badan Siber dan Sandi Negara (BSSN) Republik Indonesia tahun 2021, terdapat 1.637.973.022 perilaku abnormal nasional di Indonesia. Statistik monitoring menunjukkan bahwa perilaku abnormal tertinggi terjadi pada bulan Desember, sebanyak 242.006.168. Ini menunjukkan bahwa serangan siber meningkat selama masa pandemi Covid-19. Hal ini disebabkan karena orang-orang bekerja dari rumah untuk menghindari kerumunan yang membahayakan, dan layanan dan sistem internal organisasi bisa diakses publik oleh karyawan melalui akses internet. Tanpa perlindungan keamanan informasi yang memadai, penyerang dapat dengan mudah masuk dan mencuri informasi penting dari sistem server organisasi. Pandemi ini juga dapat dimanfaatkan oleh penyerang untuk melakukan serangan siber melalui pengiriman *spam* atau *phishing*.

Serangan siber menimbulkan kerugian yang sangat besar bagi organisasi maupun individu. Menurut Lezzi *et al*. (2018), serangan siber pada sistem manufaktur berdampak negatif pada bisnis, termasuk sabotase terhadap infrastruktur penting, gangguan pada jaringan dan komputer, pencurian rahasia dagang dan kekayaan intelektual, pelanggaran regulasi keselamatan dan lingkungan, sampai ancaman terhadap keselamatan pekerja. Dalam situasi ini, perusahaan akan mengalami kerugian ekonomi yang signifikan untuk memulihkan produktivitas dan bersaing di pasar yang relevan.

Untuk mengatasi permasalahan keamanan jaringan diperlukan sebuah *Intrusion Detection System* (IDS) agar dapat mendeteksi potensi pola serangan. IDS merupakan teknik pertahanan yang populer terhadap serangan siber. Berdasarkan metode pendeteksiannya dapat dibedakan menjadi *signature-based* dan *anomaly-based* (Gamage & Samarabandu, 2020). IDS *signature-based* mengumpulkan berbagai karakteristik serangan terlebih dahulu, mengidentifikasi lalu lintas abnormal dengan mencocokkan input dengan berbagai pola lalu lintas yang diketahui pada jaringan. Hal ini memungkinkan sistem untuk mendeteksi dan menganalisis perilaku serangan siber. IDS *signature-based* memiliki kelemahan, karena memerlukan pemeliharaan database yang mencatat perilaku abnormal yang mahal dan memakan waktu (Lin *et al*., 2022) serta tidak dapat secara efektif menemukan serangan yang muncul untuk pertama kalinya (Lee *et al*., 2022). Hal ini dikarenakan IDS *signature-based* hanya dapat mendeteksi serangan yang diketahui.

IDS *anomaly-based* memiliki kemampuan menemukan lalu lintas jaringan yang abnormal berdasarkan perbedaan antara lalu lintas target dan lalu lintas normal. Bila ditemukan lalu lintas yang sangat berbeda dari lalu lintas normal, akan dikategorikan sebagai lalu lintas berbahaya. Meskipun bisa mendeteksi lalu lintas abnormal, IDS *anomaly-based* memiliki beberapa keterbatasan seperti sulit mendapatkan data benchmark yang akurat, tingkat false alarm yang lebih tinggi, dan keterlambatan data benchmark (Lee *et al*., 2022). Kelemahan dalam mendeteksi metode serangan dan perilaku serangan menjadi masalah serius yang dapat menghambat keamanan *information technology* (IT).

Penggunaan *machine learning* untuk meningkatkan akurasi deteksi serangan jaringan pada IDS banyak diterapkan dalam penelitian sebelumnya. Hosseini dan Zade (2020) dalam risetnya mengusulkan metode hibrida baru untuk mendeteksi serangan dengan menggabungkan algoritma evolusi, *support vector machine* (SVM) dan *artificial neural network* (ANN). Pooja dan Shrinivasacharya (2021) dalam papernya telah mengevaluasi model jaringan saraf berbasis Bidirectional LSTM (BiLSTM) dalam penerapan metode *deep learning* dan model IDS. Sementara Ansari *et al*. (2022) mengusulkan metode pendekatan *deep learning* menggunakan *gated recurrent unit* (GRU) untuk memprediksi peringatan intrusi jaringan. Dalam risetnya, Yao *et al*. (2022) membangun model pendekatan hibrida CNN-transformer untuk mendeteksi serangan pada infratruktur pengukuran tingkat lanjut. Riset Kim *et al*. (2016) menggunakan *long short-term memory* (LSTM) yang dikombinasikan dengan *recurrent neural network* (RNN) untuk melatih model deteksi serangan secara efektif. Penelitian Yin *et al*. (2017) menunjukkan bahwa RNN-IDS dapat meningkatkan akurasi deteksi serangan melebihi performa mesin tradisional dengan mempelajari model *binary classification* dan *multiclass classification*. Penelitian Xiao *et al*. (2019) menggunakan model CNN dan metode PCA untuk meningkatkan kinerja deteksi dan klasifikasi lalu lintas intrusi jaringan, serta mengurangi waktu klasifikasi secara signifikan.

Penelitian ini menerapkan model hibrida CNN-GRU-BiLSTM-PCA untuk meningkatkan akurasi deteksi serangan jaringan pada IDS. Kombinasi model ini diharapkan dapat memisahkan

berbagai serangan yang tidak diketahui untuk mencapai akurasi deteksi serangan yang lebih tinggi, dan mampu membedakan serangan dengan kategori tertentu sesuai dengan karakteristik input.

.

METODE

Dataset

Dataset yang digunakan dalam penelitian ini adalah NSL-KDD yang diperoleh dari laman *University of New Brunswick* (UNB). Dalam dataset tersebut, terdiri dari KDD *train set* berisi 3.925.650 *attack records* dan 972.781 *normal records* (Tabel 1) dan KDD *test set* berisi 250.436 *attack records* dan 60.591 *normal records* (Tabel 2). Pada tahap pertama dimensi data akan direduksi menggunakan metode PCA menjadi 8 komponen utama, kemudian dilatih mengunakan model hibrida CNN-GRU-BiLSTM. CNN digunakan untuk mengekstrak fitur dari dataset, sementara GRU dan BiLSTM digunakan untuk memproses dan memahami informasi lalu lintas jaringan dalam dataset.

Tabel 1. Statistik *redundant records* pada *KDD train set*

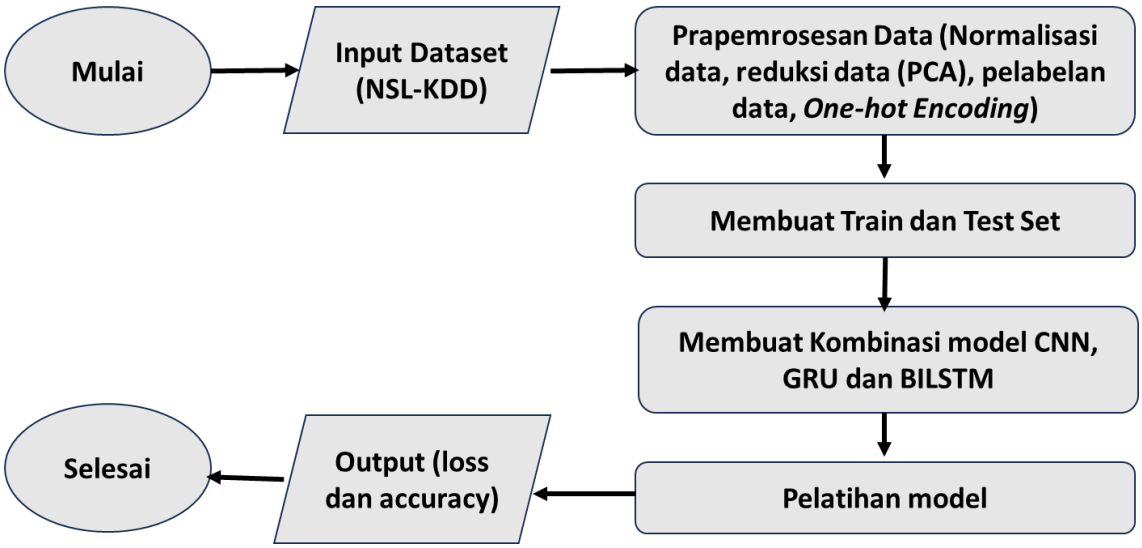
Type	Original Records	Distinct Records	Reduction Rate
Attacks	3.925.650	262.178	93,32%
Normal	972.781	812.814	16,44%
Total	4.898.431	1.074.992	78,05%

Tabel 2. Statistik *redundant records* pada *KDD test set*

Type	Original Records	Distinct Records	Reduction Rate
Attacks	250.436	29.378	88,26%
Normal	60.591	47.911	20,92%
Total	311.027	77.289	75,15%

Tahap Eksperimen

Pada tahap ini digunakan *principal component analysis* (PCA) (Alamsyah & Fadila, 2021) sebagai metode untuk mereduksi dimensi data dan *one-hot encoding* sebagai pelabelan data, serta penggunaan metode normalisasi dan standarisasi data pada pra-pemrosesan data. Selanjutnya digunakan *convolutional neural network* (CNN) (Prasetiyo *et al*., 2023) untuk mengesktrak fitur dalam dataset yang dikombinasikan dengan *gated recurrent unit* (GRU) dan *bidirectional long short-term memory* (BiLSTM) untuk memproses dan memahami informasi lalu lintas jaringan dalam dataset. Model dilatih dan dievaluasi menggunakan metrik *accuracy* dan *loss*. *Flowchart* yang digunakan pada metode ini ditunjukkan pada Gambar 1.



Gambar 1. Langkah-langkah penelitian

HASIL DAN PEMBAHASAN

Model diimplementasikan menggunakan Google Colaboratory yang berjalan pada perangkat Windows 10, dengan spesifikasi prosesor Intel Core i5-7200U 2.5GHz, NVIDIA GeForce 940MX dengan 2GB VRAM. Hasil penelitian yang didapatkan adalah peningkatan akurasi deteksi serangan jaringan pada *intrusion detection system* dengan model hibrida CNN-GRU-BiLSTM-PCA.

Sebelum model dilatih dan dievaluasi, dilakukan pembentukan model dengan menggabungkan CNN dengan GRU dan BiLSTM sehingga didapatkan ringkasan arsitektur yang disajikan pada Tabel 3.

Tabel 3. Ringkasan arsitektur

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 91, 32)	128
max_pooling1d	(None, 22, 32)	0
(MaxPooling1D)	(None, 22, 64)	18.816
gru (GRU)	(None, 22, 64)	0
dropout (Dropout)	(None, 128)	66.048
bidirectional (Bidirectional)	(None, 128)	0
dropout_1 (Dropout)	(None, 128)	0
flatten (Flatten)	(None, 50)	6.450
dense (Dense)	(None, 5)	255
dense_1 (Dense)		
Total params: 91.697		
Trainable params: 91.697		
Non-trainable params: 0		

Pada input *layer* model CNN menggunakan aktivasi relu dengan metode MaxPool dengan *pooling layer* berukuran 4 dan dengan *dropout* bernilai 0,2. Model GRU dan BiLSTM ditambahkan agar model dapat memproses dan memahami informasi lalu lintas jaringan dalam dataset. Untuk proses klasifikasi pada input *layer* menggunakan *fully connection* sebanyak 50 unit. Penggunaan MaxPool dan *dropout* digunakan agar model pada saat dilatih dapat mengurangi masalah *overfitting*. Pada output *layer* menggunakan aktivasi softmax karena fokus penelitian ini adalah klasifikasi multikelas.

Evaluasi Model

Sebelum model dievaluasi, dilakukan pelatihan terhadap model dengan parameter yang disajikan pada Tabel 4.

Tabel 4. Parameter latih dan uji

Parameter	Dimensi
Data latih	80%
Data uji	20%
epoch	10
batch_size	32
validation_split	0,2

Model dievaluasi menggunakan fungsi *evaluate* mengambil dataset sebagai argumen dan menghitung *loss* dan *accuracy* model pada dataset. Fungsi ini dimulai dengan mengevaluasi model pada dataset latih dan uji menghasilkan metrik *accuracy* dan *loss*. *Compiling* model dilakukan dengan menggunakan *categorical_crossentropy* sebagai fungsi *loss*, adam sebagai *optimizer* dan *library sklearn metrics*. Penggunaan fungsi *categorical_crossentropy* digunakan untuk mengukur perbedaan antara distribusi probabilitas prediksi yang dihasilkan oleh model dengan distribusi probabilitas target yang diharapkan. Penggunaan fungsi *loss* sesuai dengan konteks klasifikasi multikelas dengan pelabelan target menggunakan representasi *one-hot encoding*. *Optimizer* adam digunakan untuk menentukan algoritma optimasi dalam memperbarui bobot (*weights*) dan bias (*biases*) dalam proses pelatihan model. *Optimizer* adam dapat meningkatkan efisiensi komputasi, pengaturan laju pembelajaran adaptif, dan stabil dalam ruang parameter yang besar dan kompleks. Penggunaan *library sklearn metrics* digunakan untuk mengukur akurasi model atau algoritma pembelajaran mesin.

Evaluasi Metrik

Salah satu metrik evaluasi yang umum digunakan untuk mengukur performa model klasifikasi adalah akurasi. Akurasi dihitung dengan membagi jumlah prediksi yang benar dengan total jumlah pengamatan yang dievaluasi (Udas *et al.*, 2022). Dengan menggunakan persamaan (1), dapat ditentukann sejauh mana model klasifikasi mampu memberikan prediksi yang tepat.

$$\text{Accuracy} = (TP+TN)/(TP+FP+TN+FN) \quad (1)$$

Dalam konteks ini, *true positive* (TP) merupakan jumlah data positif yang diklasifikasikan dengan benar oleh model. Dengan kata lain, model memprediksi data sebagai positif, dan kelas sebenarnya dari data tersebut juga positif. *False positive* (FP) merupakan jumlah data negatif yang salah diklasifikasikan sebagai positif oleh model. Dengan kata lain, model memprediksi data sebagai positif, tetapi kelas sebenarnya dari data tersebut adalah negatif. *True negative* (TN) merupakan jumlah data negatif yang diklasifikasikan dengan benar oleh model. Dengan kata lain, model memprediksi data sebagai negatif, dan kelas sebenarnya dari data tersebut juga negatif. *False negative* (FN) merupakan jumlah data positif yang salah diklasifikasikan sebagai negatif oleh model. Dengan kata lain, model memprediksi data sebagai negatif, tetapi kelas sebenarnya dari data tersebut adalah positif.

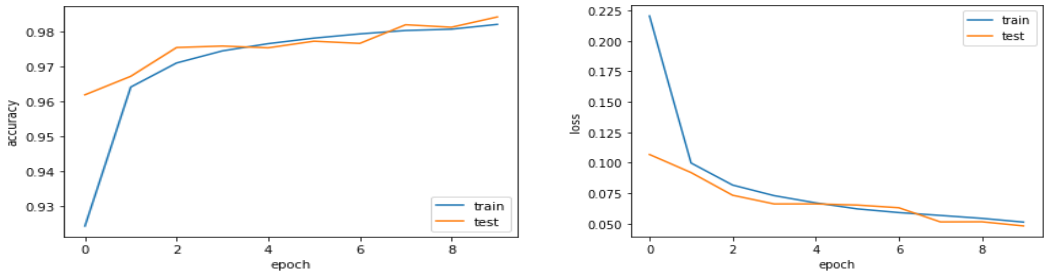
Selain akurasi, terdapat fungsi log *loss categorical crossentropy* sebagai metrik evaluasi. Fungsi ini merupakan fungsi matematis yang digunakan sebagai fungsi *loss* dalam pelatihan model klasifikasi multikelas menggunakan metode *gradient descent*. Fungsi ini memberikan ukuran sejauh mana probabilitas prediksi model cocok dengan label sebenarnya dan menghitung penalti untuk setiap prediksi yang salah, yaitu ketika probabilitas prediksi yang dihasilkan oleh model berbeda dengan label sebenarnya. Semakin tinggi perbedaan antara prediksi dan label, semakin tinggi nilai *loss*. Begitu juga sebaliknya, semakin rendah perbedaan antara prediksi dan label, maka semakin rendah nilai *loss*. Secara matematis persamaan log *loss categorical crossentropy* didefinisikan pada Persamaan (2).

$$loss = -\sum (y\_true \times log\ log\ (y\_pred)) \quad (2)$$

Dimana:

y_true adalah distribusi probabilitas target yang diharapkan dalam bentuk one-hot encoding
y_pred adalah distribusi probabilitas prediksi yang dihasilkan oleh model

Melalui kombinasi hibrida CNN-GRU-BILSTM-PCA, didapatkan hasil akurasi terbaik untuk mendeteksi serangan dalam jaringan pada *intrusion detection system*. Dengan menerapkan metode *principal component analysis* waktu pelatihan model dapat direduksi dengan baik. Gambar 5 menyajikan grafik plot *accuracy* dan *loss versus epoch* untuk dataset latih dan uji.



Gambar 5. Plot akurasi vs epoch (a) dan plot loss vs epoch (b) untuk dataset latih dan uji

Hasil akurasi analisis multikelas dari setiap model yang diteliti pada penelitian sebelumnya ditunjukkan pada Tabel 5.

Tabel 5. Perbandingan akurasi model dengan penelitian terdahulu

Penulis	Model	Tahun	Akurasi Multikelas (%)
Udas *et al.* (2022)	SPIDER	2022	82,91%
Li *et al.* (2020)	Multi-CNN Fusion	2020	81,33%
Hu *et al.* (2020)	AS-CNN	2020	84,08%
Liu *et al.* (2021)	CNN-RF	2021	85,24%
Jiang *et al.* (2020)	CNN-OSS-BiLSTM	2020	82,74%
Metode yang diusulkan	**CNN-GRU-BiLSTM-PCA**	**2023**	**98,34%**

Seperti ditunjukkan pada Tabel 5, metode yang diusulkan mendapatkan hasil yang lebih baik dibandingkan dengan model pada penelitian terdahulu. Metode yang diusulkan dengan model hibrida CNN-GRU-BiLSTM-PCA berhasil mendapatkan tingkat akurasi sebesar 98,34% dan *loss* sebesar 0,048%.

SIMPULAN

Model hibrida CNN-GRU-BILSTM-PCA berhasil mendapatkan nilai akurasi sebesar 98,34% dan loss sebesar 0,048%. Model hibrida ini merupakan model terbaik di antara model-model penelitian sebelumnya. Perlu menggunakan dataset realtime agar hasil akurasi dapat lebih optimal.

DAFTAR PUSTAKA

Admaja, A.F.S. (2018). Pemetaan riset teknologi 5G. *Buletin Pos dan Telekomunikasi*, *16*(1), 27-40. https://doi.org/10.17933/bpostel.2018.160103.

Ansari, M. S., Bartoš, V., & Lee, B. (2022). GRU-based deep learning approach for network intrusion alert prediction. *Future Generation Computer Systems*, *128*, 235-247. https://doi.org/10.1016/ j.future.2021.09.040

Gamage, S., & Samarabandu, J. (2020). Deep learning methods in network intrusion detection: A survey and an objective comparison. *Journal of Network and Computer Applications*, *169*, 102767. https://doi.org/10.1016/j.jnca.2020.102767

Gartner. (2022). *Gartner Top Strategic Technology Trends for 2022*. https://content.intland. com/blog/gartner-top-strategic-technology-trends-for-2022. 1-19.

Hosseini, S., & Zade, B. M. H. (2020). New hybrid method for attack detection using combination of evolutionary algorithms, SVM, and ANN. *Computer Networks*, *173*, 107168. https://doi.org/10.1016/j.comnet.2020.107168

Hu, Z., Wang, L., Qi, L., Li, Y., & Yang, W. (2020). A novel wireless network intrusion detection method based on adaptive synthetic sampling and an improved convolutional neural network. *IEEE Access*, *8*, 195741-195751. https://doi.org/10.1109/ACCESS.2020.3034015

Jiang, K., Wang, W., Wang, A., & Wu, H. (2020). Network intrusion detection combined hybrid sampling with deep hierarchical network. *IEEE Access*, *8*(3), 32464-32476. https://doi.org/ 10.1109/ACCESS.2020.2973730

Kim, J., Kim, J., Thu, H. L. T., & Kim, H. (2016). Long short term memory recurrent neural network classifier for intrusion detection. *2016 International Conference on Platform Technology and Service, PlatCon 2016 - Proceedings*. https://doi.org/10.1109/PlatCon.2016.7456805

Lee, J. S., Chen, Y. C., Chew, C. J., Chen, C. L., Huynh, T. N., & Kuo, C. W. (2022). CoNN-IDS: Intrusion detection system based on collaborative neural networks and agile training. *Computers and Security*, *122*, 102908. https://doi.org/10.1016/j.cose.2022.102908

Lezzi, M., Lazoi, M., & Corallo, A. (2018). Cybersecurity for industry 4.0 in the current literature: A reference framework. *Computers in Industry*, *103*, 97-110. https://doi.org/10.1016/ j.compind. 2018.09.004

Li, Y., Xu, Y., Liu, Z., Hou, H., Zheng, Y., Xin, Y., Zhao, Y., & Cui, L. (2020). Robust detection for network intrusion of industrial IoT based on multi-CNN fusion. *Measurement: Journal of the International Measurement Confederation*, *154*, 107450. https://doi.org/10.1016/j.measurement. 2019.107450

Lin, K., Xu, X., & Xiao, F. (2022). MFFusion: A multi-level features fusion model for malicious traffic detection based on deep learning. *Computer Networks*, *202*, 108658. https://doi.org/10.1016/j.comnet.2021.108658

Liu, C., Gu, Z., & Wang, J. (2021). A hybrid intrusion detection system based on scalable k-means+ random forest and deep learning. *IEEE Access*, *9*, 75729-75740. https://doi.org/10.1109/ ACCESS.2021.3082147

Pooja, T.S., & Shrinivasacharya, P. (2021). Evaluating neural networks using Bi-Directional LSTM for network IDS (intrusion detection systems) in cyber security. *Global Transitions Proceedings*, *2*(2), 448-454. https://doi.org/10.1016/j.gltp.2021.08.017

Prasetiyo, B., Alamsyah, A., Al Hakim, M.F., Jumanto, J., & Adi, M. H. (2023). Differential augmentation data for vehicle classification using convolutional neural network. *AIP Conference Proceedings*, *2614*, 040001. https://doi.org/10.1063/5.0126720

Udas, P. B., Karim, M. E., & Roy, K. S. (2022). SPIDER: A shallow PCA based network intrusion detection system with enhanced recurrent neural networks. *Journal of King Saud University -*

Computer and Information Sciences, *34*(10), 10246-1027. https://doi.org/10.1016/j.jksuci.2022.10.019

Xiao, Y., Xing, C., Zhang, T., & Zhao, Z. (2019). An intrusion detection model based on feature reduction and convolutional neural networks. *IEEE Access*, *7*, 42210-42219. https://doi.org/10.1109/ACCESS.2019.2904620

Yao, R., Wang, N., Chen, P., Ma, D., & Sheng, X. (2022). A CNN-transformer hybrid approach for an intrusion detection system in advanced metering infrastructure. *Multimedia Tools and Applications*, *82*(6), 19463-19486. https://doi.org/10.1007/s11042-022-14121-2