Emerging Information Science and Technology
Vol. 5, No. 2, (2024), pp. 74-79

Classification of Duration in Global Terorism using ResNet

Adinda Nurhayati Adriansya[1,*], Slamet Riyadi[2]

[1,2]*Universitas Muhammadiyah Yogyakarta, Jln.Brawijaya, Tamantirto,Kasihan, Bantul, Yogyakarta 55183*

**Corresponding author: a.nurhayati.ft21@mail.umy.ac.id*

Abstract

Terrorism is a global threat that affects the political, economic, and social stability of many countries. The number of victims killed is based on the duration of the terrorism incident. This study uses the Residual Network (ResNet) model to classify terrorism incidents based on the duration of the incident (less than 24 hours and more than 24 hours) using the Global Terrorism Database (GTD) dataset. The GTD data used covers terrorism incidents from 1970 to 2017, with a total of 181,691. After preprocessing the data by converting categorical features to numeric and removing missing values, the data was divided into training, validation, and test sets with a composition of 70%, 15%, and 15%. The results show that the ResNet model is able to achieve a validation accuracy of 99.61% and a validation loss value of 0.0183. These findings show that the ResNet model is effective in classifying the duration of terrorism incidents and has the potential to be used in the development of better terrorism prevention systems.

Keywords: *Terrorism, Classification, SMOTE, ResNet*

1. Introduction

As an international threat, terrorism has grown to become a significant problem in many industrialized and developing countries. Globally, there have been about 200,000 incidents of terrorism between 1970 and 2019. All countries have challenges as a result of this tragedy. These events not only resulted in huge financial losses, but also caused public anxiety and caused several casualties. The political, economic, and social aspects of the affected countries are all affected by terrorism.

Terrorism has forced governments in several countries to step up prevention and response initiatives. One of the countries that is often the target of terrorist attacks is Indonesia, which has significant obstacles in maintaining national stability and security. A number of bombings, shootings, and other violent incidents have occurred across Indonesia, causing extensive infrastructure damage and high casualties.

Terrorists carry out violence and other acts of terror to gain public support, create fear for institutional governments and collect funds from their supporters [1]. The number of terrorism incidents around the world varies every year, according to data from the World Terrorism Database (GTD)[2]. For example, there were more than 16,000 terrorist attacks in 2014, but only more than 8,000 incidents occurred in 2019[3]. This pattern shows that the threat of terrorism still exists and continues to evolve according to the needs of the world[4].

(Analysis of the classification of terrorist attacks in Indonesia) This study researched the increasing threat of terrorism in various parts of the world, including Indonesia, and analyzed more than 180,000 terrorist attacks that occurred from 1970 to 2017. The classification of terrorist attacks is classified based on their success with the description of successful attacks defined as terrorist attacks that result in casualties. The study used 7 attributes: year, month, type, attack, terrorist name, attack target, city, and type of weapon

Copyright ⓒ 2024 EIST

used [5]. This study uses the naïve bayes algorithm method with an accuracy of 80.45%, C4.5 with an accuracy of 88.82%, and k-NN with an accuracy of 90.79%. This study uses Methodology Validation of several different classification algorithms, namely Naïve Bayes, Decisien Tree (C4.5) and k-Nearest Neighbor (k-NN). The use of these 3 algorithms can be used as a comparison and evaluation of the performance of each method[6].

The study concluded that with an accuracy rate of 90.79%, the k-NN algorithm showed the highest accuracy in the classification of terrorism acts in Indonesia. According to this study, officials in the Republic of Indonesia may be able to prevent terrorism by using appropriate classification algorithms. In addition, the study suggests the use of sentiment analysis to determine patterns of differences in posts and comments on social media. The findings of this study are expected to be integrated with other research findings in the future.

The process of collecting and analyzing data with the aim of extracting significant information from that data is known as data mining[7], [8]. Data mining is used with the aim of explaining, confirming and exploring. One of the main tasks in data mining is classification. Because there is no research that uses the Resnet model in the GTD dataset, this study was made using classification using the ResNet model with tabular data[8], [9], [10]. After we know the significant impact of theorism, this study aims to identify the duration of terrorism incidents using the ResNet classification model. This classification aims to find out about global terrorism incidents that occur less than 24 hours and more than 24 hours and the information that has been obtained after doing this classification can be used in making the development of prevention strategies in global terrorism incidents based on the results of the classification that has been carried out.

2. Related Works

The use of the ResNet model has been proven effective in various previous studies covering various application domains[11]. One study compared the performance of CNN and ResNet algorithms in classifying the degree of malignancy of diabetes based on retinopathy imagery. The study found that ResNet had an accuracy of 81.23% with a loss of 12.59%, while CNN only achieved an accuracy of 68.49% with a loss of 32.57%. These results show that the ResNet algorithm has better performance and accuracy than CNN in this task [12].

Another study used the ResNet method to identify diseases in wheat crops, specifically Septoria and Stripe Rust diseases. This study succeeded in classifying healthy and infected wheat leaves with 98% accuracy. This success shows ResNet's ability to process plant image data to accurately detect diseases [13].

Although many studies have shown the effectiveness of ResNet in various domains, no studies have applied it to the classification of the duration of terrorism incidents. Therefore, this study aims to fill this gap by demonstrating the effectiveness of the ResNet model in classifying the duration of terrorism incidents using tabular data.

3. Methodology

3.1. Data Collection

The dataset utilized in this study is derived from the **Global Terrorism Database (GTD)**, an open-source database maintained by the National Consortium for the Study of Terrorism and Responses to Terrorism (START) at the University of Maryland. The GTD is one of the most comprehensive resources for analyzing global terrorism, systematically documenting incidents from 1970 onward. It provides a detailed account of each recorded event, including information about the date and location of the attack, the type of attack, intended targets, casualties, perpetrator groups (when identified), and weapon types. Additionally, the database offers rich contextual details, such as motives and narratives surrounding the incidents, ensuring a holistic view of terrorism trends worldwide.

The data for this study was retrieved from the official GTD website (https://www.start.umd.edu/gtd/). The dataset version used corresponds to the most recent update available as of [insert date of access]. This format was chosen due to its compatibility with data analysis tools, making it straightforward to process and integrate into the analytical framework. The GTD is curated using an open-source methodology, relying on verified media reports and other credible sources. This ensures the reliability of the data while maintaining a broad coverage of global incidents.

3.2. Data Preprocessing

The process is divided into 2 parts, namely data selection and data cleaning. Selection data is certain data that is selected based on research taken by researchers from the database. Data cleaning is a process of cleaning data so that the data produced does not reduce the accuracy of the data through the process of deleting noun or incomplete data [14]. In this study, the researcher selects all the columns generated after the data cleaning process and does not limit the number of columns after the data cleaning process. These data are materials that will be used to be analyzed with classification algorithms.

3.3. Model Planning

The model used in this study is the ResNet model. ResNet is a specific type of neural network that was introduced in 2015 by Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun under the short name Residual Network.

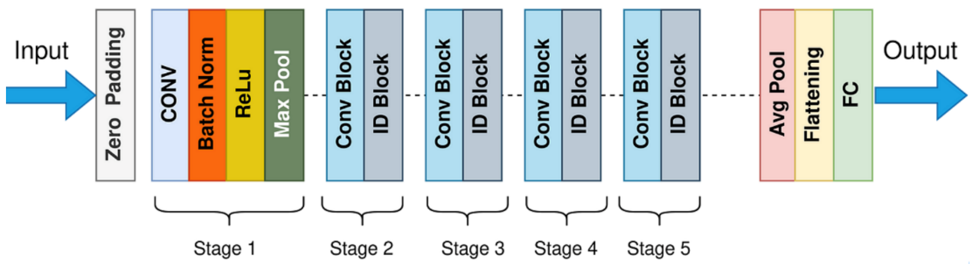

Figure 1. ResNet model

ResNet introduced the concept of "residual learning" to overcome the degradation problem that occurs when the neural network gets deeper. In conventional neural networks, an increase in the number of layers can lead to a decrease in accuracy, not improvement. ResNet solves this problem by using "shortcut connections" or "skip connections" that allow the gradient to flow more smoothly through the network during training [15].

The ResNet used in this study is ResNet in the classification of tabular data. Then the ResNet classification method in classifying tabular data is by applying the concept of residual blocks. Residual blocks are blocks for tabular data consisting of dense layers, batch normalization, and dropouts. The mathematical functions of the residual block are as follows:

$$Output = F(x) + x \quad (1)$$

Remarks: F(x) is the result of two layers of convolution.

The following are the steps taken in this study:

The first step is data preprocessing. At this stage, tabular data is prepared by converting categorical features to numeric using Label Encoder. Next, rows with NA values are removed to ensure the data is clean and ready to use. The second step is data sharing. The data is divided into three parts, namely training, validation, and testing with a test set size of 30% of the total data. This division is done to ensure that the model can be tested and validated with data that has never been seen before. The third step is to address data imbalances. The class imbalance in data training was overcome using the SMOTE (Synthetic Minority Over-sampling Technique) technique.

This technique aims to increase the number of samples in minority classes so that the distribution of data becomes more balanced. The fourth step is model architecture. At this stage, several residual blocks are combined to form a complete ResNet model architecture. These residual blocks help in overcoming the vanishing gradient problem and allow deeper network training. The fifth step is model training. The model is trained using training data and the training process is carried out for 10 epochs. Each epoch involves forward pass and backpropagation to optimize the weight of the model. The sixth step is model evaluation. After training, the model is evaluated using the Confusion Matrix to understand the distribution of true and false predictions. Additionally, metrics such as accuracy and log loss are calculated to measure the overall performance of the model. The steps in implementing the algorithm are expected to produce high accuracy.

4. Result and Discussion

In this chapter, the results of the analysis of the Global Terrorism dataset will be displayed. Previously, researchers had preprocessed data, shared data, and overcome data imbalances in order to improve good performance results.

The researcher conducted a study using the ResNet model to classify terrorism incidents based on extended columns in the Global Terrorism dataset. In the *preprocessing* stage, the researcher converts the categorical features into numeric into labelEncoders and removes the rows with NA values. The results of the preprocessing produce clean data and can be used for data training. Then the data was divided into three sets of training, validation and testing as: 70%:15%:15%.

Figure 2, 3, and 4 show the experiment results. This study produced an accuracy of 99.61% using the ResNet and Validation Loss methods with an accuracy of 0.0183.

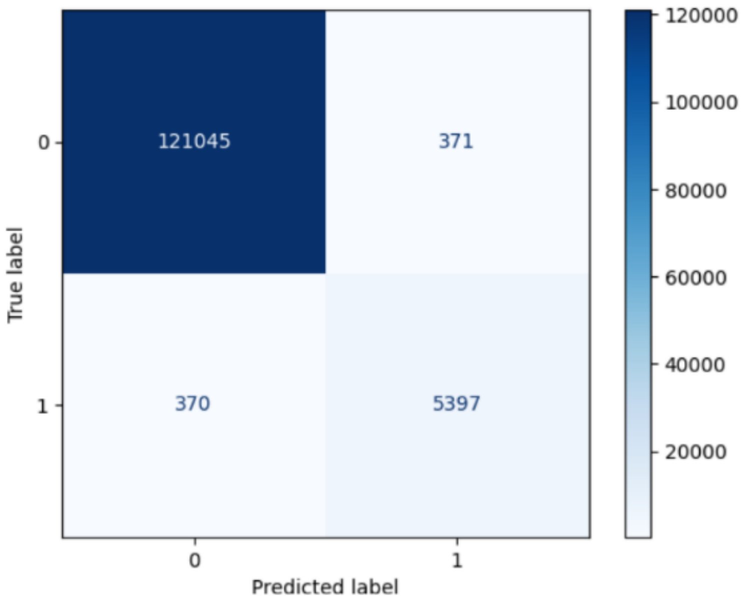

Figure 2. Confusion Marix

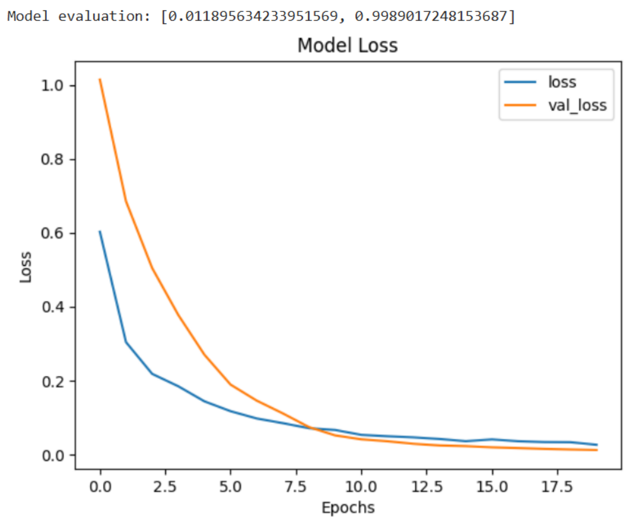

Figure 3. Results of Performance Evaluation of Loss Model

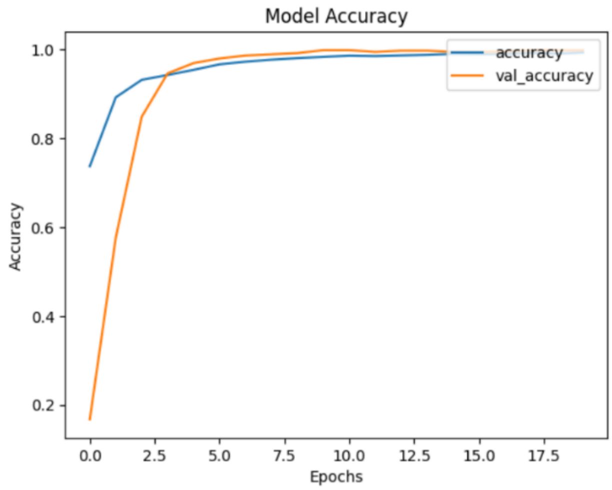

Figure 4. Results of Performance Evaluation of Model Accuracy

4. Conclusions

This study demonstrates the effectiveness of the Residual Network (ResNet) model in classifying the duration of terrorism incidents using the Global Terrorism Database (GTD) dataset. By preprocessing the data and addressing class imbalances with the SMOTE technique, the ResNet model achieved a high validation accuracy of 99.61% and a validation loss of 0.0183. These results indicate that the ResNet model is highly capable of distinguishing between terrorism incidents lasting less than 24 hours and those extending beyond 24 hours.

The successful application of the ResNet model in this context suggests its potential utility in developing advanced terrorism prevention systems. By accurately classifying the duration of terrorism incidents, authorities can better understand the patterns and characteristics of such events, leading to more informed decision-making and strategic planning in counter-terrorism efforts.

Future research could explore the integration of additional features and the application of other deep learning models to further enhance the classification performance. Additionally, expanding the dataset to include more recent incidents and other relevant variables could provide deeper insights into the evolving nature of global terrorism.

References

[1] M. Nuruzzaman, "Terorisme dan Media Sosial Sisi Gelap Berkembangnya Teknologi Informasi Komunikasi," *Syntax Lit. J. Ilm. Indones.*, vol. 3, no. 9, hlm. 61–76, 2018.

[2] M. Dunning dan S. Purohit, "Higher Order Temporal Analysis of Global Terrorism Data.," *arXiv: Social and Information Networks*. 2020. [Daring]. Tersedia pada: http://arxiv.org/pdf/2005.14002.pdf

[3] Z. Li, L. Xiangchun, C. Dong, F. Guo, F. Zhang, dan Q. Zhang, "Quantitative Analysis of Global Terrorist Attacks Based on the Global Terrorism Database," *Sustainability*, vol. 13, no. 14, 2021, doi: 10.3390/SU13147598.

[4] R. Lu, J. Huang, Y. Qu, dan L. Li, "Study on Combined-CNN Model for Classification of Terrorism Text," 2024, doi: 10.1109/icaace61206.2024.10548392.

[5] M. Junaedi, A. Fachrurozi, M. R. Kusumayudha, dan W. Gata, "Analysis of the classification of terrorist attacks in Indonesia," *J. Inform. Telecommun. Eng.*, vol. 4, no. 1, hlm. 57–66, 2020.

[6] K. Singh, A. S. Chaudhary, dan P. Kaur, "A Machine Learning Approach for Enhancing Defence Against Global Terrorism," 2019, hlm. 1–5. doi: 10.1109/IC3.2019.8844947.

[7] J. Zhu, H. Tang, L. Zhang, B. Jin, dan X. P. Wei, "A Global View-Guided Autoregressive Residual Network for Irregular Time Series Classification," dalam *Lecture Notes in Computer Science*, 2023, hlm. 289–300. doi: 10.1007/978-3-031-33383-5_23.

[8] J. Wu, Z. Zhang, Y. Ji, S. Li, dan L. Lin, "A ResNet with GA-based Structure Optimization for Robust Time Series Classification," 2019, hlm. 69–74. doi: 10.1109/SMILE45626.2019.8965287.

[9] A. Ukil, L. Marin, dan A. Jara, "ADV-ResNet: Residual Network with Controlled Adversarial Regularization for Effective Classification of Practical Time Series Under Training Data Scarcity Problem," 2022, hlm. 1–8. doi: 10.1109/IJCNN55064.2022.9892370.

[10] A. Ukil, S. Bandyopadhyay, dan A. Pal, "DyReg-FResNet: Unsupervised Feature Space Amplified Dynamic Regularized Residual Network for Time Series Classification," 2019, hlm. 1–8. doi: 10.1109/IJCNN.2019.8852392.

[11] F. Zhu, H. Wang, dan Y. Zhang, "GRU Deep Residual Network for Time Series Classification," 2023, hlm. 1289–1293. doi: 10.1109/ITNEC56291.2023.10082454.

[12] L. Mutawalli, M. T. A. Zaen, dan Y. Yuliadi, "Komparasi CNN dengan ResNet Untuk Klasifikasi Paling Akurat Tingkat Keganasan Diabetes Berdasarkan Citra Retinopathy," *J. Comput. Syst. Inform. JoSYC*, vol. 4, no. 3, hlm. 522–529, 2023.

[13] A. Ridhovan dan A. Suharso, "Penerapan Metode Residual Network (RESNET) Dalam Klasifikasi Penyakit Pada Daun Gandum," *JIPI J. Ilm. Penelit. Dan Pembelajaran Inform.*, vol. 7, no. 1, hlm. 58–65, 2022.

[14] R. W. Pratiwi dan Y. S. Nugroho, "Prediksi Rating Film Menggunakan Metode Naïve Bayes," *DutaCom*, vol. 12, no. 1, hlm. 91–108, 2016.

[15] K. He, X. Zhang, S. Ren, dan J. Sun, "Deep residual learning for image recognition," dalam *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, hlm. 770–778.