International Journal of Advances in Applied Sciences (IJAAS)
Vol. 15, No. 1, March 2026, pp. 322~332
ISSN: 2252-8814, DOI: 10.11591/ijaas.v15.i1.pp322-332

Hybrid deep learning approach for Indonesian hoax detection: a comparative evaluation with IndoBERT

Siti Mujilahwati, Moh. Rosidi Zamroni, Miftahus Sholihin
Department of Informatics Engineering, Faculty of Science and Technology, Universitas Islam Lamongan, Lamongan, Indonesia

Article Info

Article history:

Received May 26, 2025
Revised Sep 23, 2025
Accepted Jan 22, 2026

Keywords:

Bidirectional gated recurrent unit
Bidirectional LSTM
Deep learning
Hoax detection
Hybrid model

ABSTRACT

The spread of hoaxes in Indonesia has escalated significantly, with over 12,547 cases recorded between 2018 and 2023. Low public literacy and uncontrolled information flow contribute to the rapid dissemination of false content that fuels disinformation and social unrest. Previous studies have utilized artificial intelligence (AI) approaches such as Indonesia bidirectional encoder representations from Transformers (IndoBERT) and deep learning models like long short-term memory (LSTM), bidirectional LSTM (BiLSTM), convolutional neural network (CNN), and Transformer-based methods. However, most relied on a single modeling paradigm and did not address the trade-offs between classification performance and computational efficiency. This study proposes a hybrid architecture that integrates IndoBERT with bidirectional gated recurrent unit (BiGRU) and BiLSTM to enhance Indonesian hoax detection. Using 4,312 news articles and 10-fold cross-validation, we compare the performance of IndoBERT–BiGRU, IndoBERT–BiLSTM, and the proposed hybrid IndoBERT–BiGRU–BiLSTM model. Evaluation metrics include accuracy, precision, recall, F1-score, and training time. The hybrid model achieved the best performance with 98.73% accuracy, 99.01% recall, 98.04% precision, and 98.98% F1-score, while also reducing training time compared to single models. These findings demonstrate that combining BiGRU and BiLSTM within the IndoBERT framework effectively balances performance and efficiency, making it a robust solution for Indonesian text classification.

Corresponding Author:

Siti Mujilahwati
Department of Informatics Engineering, Faculty of Science and Technology, Universitas Islam Lamongan
Lamongan, East Java, Indonesia
Email: moedjee@unisla.ac.id

1. INTRODUCTION

The spread of fake news (hoaxes) has become a global issue that demands increasing attention, particularly in developing countries where digital literacy levels vary significantly among the population. The rapid dissemination of unverified information through digital platforms has contributed to widespread disinformation, public anxiety, ultimately, and social division [1]. This phenomenon has escalated sharply in Indonesia, with 12,547 hoax cases reported online between 2018 and 2023. A key contributing factor to this trend is the low level of information literacy, which hampers the public's ability to effectively distinguish between credible sources and misleading content [2].

To address this challenge, many researchers have devoted themselves to utilizing artificial intelligence (AI), particularly deep learning and natural language processing (NLP), to detect fake news. Transformer-based models like bidirectional encoder representations from Transformers (BERT) and its variants in various languages, such as Indonesia BERT (IndoBERT), have demonstrated outstanding performance in Transformer-based models such as BERT and its variants in various languages, such as IndoBERT, have shown excellent performance in summarization approaches for text-based depression detection [3]. Rahmawati *et al*. [4] their research compared several of the best models in hoax news classification and concluded that IndoBERT was the best model out of support vector machine (SVM) and naive Bayes. Similarly, Ridho and Yulianti [5] integrated IndoBERT with classical machine learning models (e.g., SVM and naïve Bayes), showing that this combination improved classification accuracy by over 10% compared to models without contextual embeddings. Nugraha and Fudholi [6] applied BERT to detect COVID-19 misinformation on Indonesian Twitter and found that BERT effectively handled informal and noisy text, achieving strong F1-scores with minimal preprocessing. Yusuf and Suyanto [7] employed long short-term memory (LSTM) for hoax classification in Indonesian news, demonstrating the model's capability to learn sequential dependencies, although constrained by training time and gradient issues. Fakhruzzaman and Gunawan [8] used convolutional neural networks (CNN) with data augmentation techniques, which proved effective in enhancing model generalization on small datasets.

Although these studies demonstrate the potential of deep learning for hoax detection, most rely on a single modeling paradigm, either recurrent neural networks (RNN)-based or Transformer-based, without a comprehensive analysis of trade-offs between classification performance and computational efficiency. For instance, bidirectional LSTM (BiLSTM) captures long-range dependencies effectively but suffers from high computational cost [9]. In contrast, bidirectional gated recurrent unit (BiGRU) offers a more lightweight architecture with faster convergence, though it may be less expressive for modeling complex linguistic patterns. Transformer-based models like IndoBERT provide rich contextual representations, yet they inherently lack sequential modeling capabilities. These limitations are further amplified in the Indonesian language, which features complex morphology, inflection, and highly flexible word order. Therefore, a model that combines the strengths of both recurrent and Transformer-based approaches, while addressing their respective limitations, is essential.

While previous studies have explored the use of either Transformer-based models or RNNs for Indonesian hoax detection, they have not explicitly addressed the trade-offs between classification effectiveness and computational efficiency, nor proposed a unified hybrid framework that leverages the strengths of both. To address this gap, this paper presents a novel hybrid deep learning model that integrates BiGRU and BiLSTM layers with IndoBERT embeddings for Indonesian fake news detection. The hybrid architecture is designed to leverage the complementary strengths of each component: BiGRU offers efficient training and faster convergence, while BiLSTM captures complex and long-range dependencies in sequential data. Notably, BiGRU has demonstrated superior validation accuracy compared to BiLSTM (0.969 vs. 0.947) and exhibits faster convergence during training [10]. When combined with contextualized embeddings from IndoBERT, the hybrid configuration is expected to outperform models that rely solely on Transformer-based or non-contextual architectures. For example, Transformer-based hybrid models incorporating BiGRU have achieved up to 99.7% accuracy in fake news detection before the 10th epoch [11]. This combination offers a balanced trade-off between accuracy, convergence, and computational efficiency, making it a practical solution for NLP tasks in Indonesian.

The proposed model is evaluated using a dataset of 4,312 news articles, with performance validated through 10-fold cross-validation. Evaluation metrics include accuracy, precision, recall, F1-score, and training time. Despite the promising progress of previous studies, two critical gaps remain unresolved. First, most approaches have focused on a single modeling paradigm, either Transformer-based or RNN, without explicitly addressing the trade-offs between classification performance and computational efficiency. Second, the complementary strengths of BiGRU and BiLSTM have not been systematically integrated within the IndoBERT framework for Indonesian hoax detection. To bridge these gaps, this study introduces a hybrid IndoBERT-BiGRU-BiLSTM architecture and makes the following contributions:

i) Novel hybrid framework: we design and implement a hybrid architecture that integrates contextualized embeddings from IndoBERT with BiGRU and BiLSTM layers, enabling efficient convergence while preserving the ability to capture long-term dependencies.
ii) Comprehensive evaluation: we systematically evaluate three model variants (IndoBERT-BiGRU, IndoBERT-BiLSTM, and the proposed hybrid IndoBERT-BiGRU-BiLSTM) on a dataset of 4,312 Indonesian news articles using 10-fold cross-validation.
iii) Performance and efficiency: we demonstrate that the hybrid model achieves superior accuracy (98.73%) and F1-score (98.98%) while reducing training time compared to standalone models, thereby balancing effectiveness and efficiency.
iv) Generalisability and practical implications: beyond hoax detection, the proposed hybrid framework can be extended to other NLP tasks in low-resource and morphologically rich languages, supporting real-world applications such as misinformation filtering, media verification, and digital literacy initiatives.

Although the experimental evaluation in this study focuses on Indonesian texts, the proposed hybrid IndoBERT-BiGRU-BiLSTM framework is conceptually generalizable. Its design is particularly relevant for multilingual hoax detection in low-resource and morphologically rich languages, where similar challenges of data scarcity, noisy text, and complex linguistic structures occur.

The remainder of this paper is organized as follows: section 2 presents the methodology and architecture of the proposed model. Section 3 describes the datasets and experimental setup. Section 4 discusses the results and performance evaluation. Finally, section 5 concludes the paper and suggests directions for future research.

2. METHOD

This study compares the performance of three IndoBERT-based models, BiGRU, BiLSTM, and a hybrid BiGRU-BiLSTM for Indonesian hoax detection. The methodological workflow covers five stages: data collection, preprocessing, feature extraction with IndoBERT, model training, and performance evaluation. The hybrid design is motivated by prior work showing the effectiveness of combining Transformer embeddings with recurrent networks (e.g., BERT-BiGRU for sentiment analysis and BERT-CNN-LSTM for hoax detection). Accordingly, this study integrates IndoBERT with BiGRU and BiLSTM layers to exploit their complementary strengths in capturing contextual and sequential information.

Figure 1 illustrates the architecture of the research pipeline. The process begins with data preparation, including preprocessing steps such as cleaning and tokenization. The cleaned data is then passed through the IndoBERT model to obtain contextual embeddings, which are subsequently fed into three types of deep learning classifiers: BiLSTM, BiGRU, and a hybrid BiGRU-BiLSTM model. The outputs are evaluated using classification metrics including accuracy, precision, recall, F1-score, and training time.

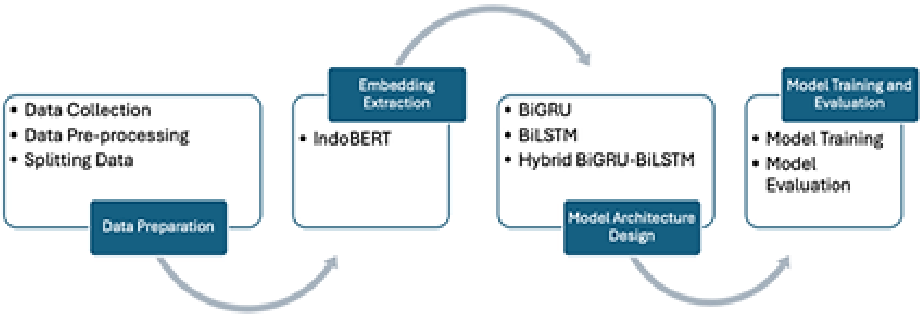

Figure 1. Architecture of the research

2.1. Dataset

The dataset was compiled from two reliable sources: TurnBackHoax.id (labeled as hoax) and Detik.com (labeled as non-hoax). The Detik.com subset covers three categories: politics, sports, and technology. In total, 4,312 news articles were collected. For model development, the dataset was divided into training (70%), validation (20%), and testing (10%) sets to ensure robust and reliable evaluation (see Table 1).

Table 1. Dataset splitting overview

Data	Count
Train 70%	3,019
Test 10%	432
Validation 20%	861

2.2. Preprocessing

In this stage, the dataset undergoes cleaning and labeling to prepare for model training. Cleaning includes lowercasing, removal of non-alphabetic punctuation, normalization, stopword elimination, and tokenization [12]. Labeling is source-based: news from TurnBackHoax.id is assigned label 0 (hoax), while news from Detik.com is assigned label 1 (non-hoax). The texts are then tokenized with the IndoBERT WordPiece tokenizer, padded or truncated to a maximum sequence length of 128 tokens, and converted into input_ids and attention_mask tensors. Finally, the 4,312 news articles are split into training (70%), validation (20%), and test (10%) sets to ensure reliable evaluation.

2.3. Embedding extraction with IndoBERT

The IndoBERT model employed in this study follows the standard BERT Transformer encoder architecture, as adapted for the Indonesian language [13]. Each input text is tokenized using the WordPiece tokenizer and enclosed by special tokens [CLS] and [SEP]. The tokens are transformed into vector representations through a combination of token, segment, and positional embeddings, which are then processed by multiple bidirectional Transformer encoder layers to capture contextual relationships within the sentence. The resulting contextual embeddings (including the [CLS] token representation) serve as the input features for the subsequent deep learning models. Unlike typical IndoBERT classification setups that directly apply a Softmax layer, in this study, the embeddings are passed into BiGRU, BiLSTM, and the proposed BiGRU-BiLSTM hybrid architectures to enhance sequential modeling (see Figure 2).

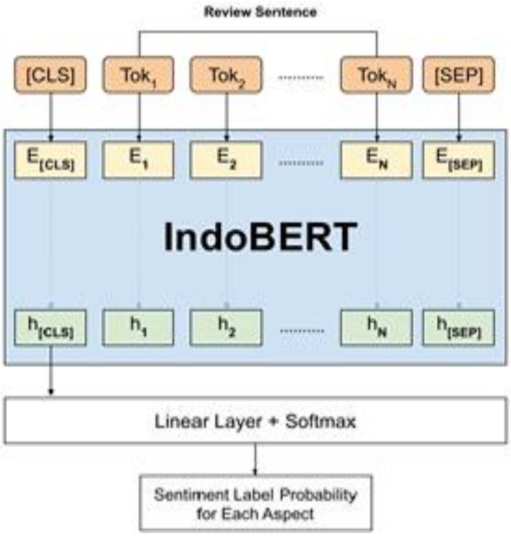

Figure 2. IndoBERT architecture

2.4. Design model

2.4.1. Bidirectional long short-term memory architecture

The BiLSTM model extends the conventional LSTM by processing the input sequence in both forward and backward directions. This bidirectional structure allows the model to capture past and future context simultaneously, producing richer sequence representations [14]. The outputs from the two directions are concatenated at each time step and then passed to a dense or classification layer. Compared to standard LSTM, BiLSTM is more effective in modeling long-term dependencies in sequential data, making it suitable for text classification tasks. The overall design is illustrated in Figure 3. This architecture has been widely applied in NLP tasks such as sentiment analysis, machine translation, and fake news detection, demonstrating its effectiveness in capturing sequential dependencies.

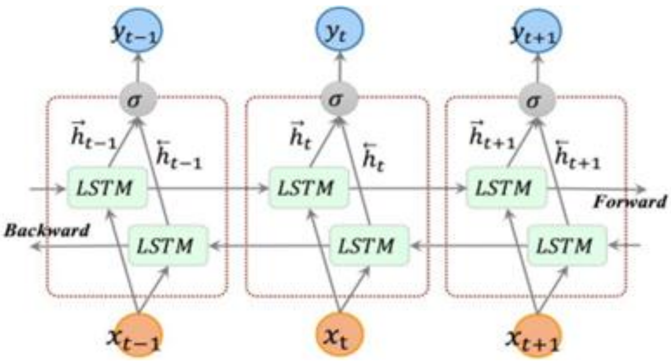

Figure 3. BiLSTM architecture

2.4.2. Bidirectional gated recurrent unit architecture

BiGRU functions similarly to BiLSTM in that both processes data in two directions, forward and backward, simultaneously. This allows for the recognition of dependency relationships in both directions and the handling of complex temporal dynamics [15], [16]. In the BiGRU architecture, the input sequence is divided into two paths: one following the original sequence (forward path) and one following the reverse sequence (backward path) [17]. Each pathway comprises GRU units that utilize gating mechanisms to regulate the flow of information (see Figure 4). The forward path GRU captures information from the previous time to the current time, while the backward path GRU captures information from the later time to the current time. The results of these two paths are then combined to form the final representation at each point in time. Compared to BiLSTM, BiGRU uses fewer parameters, which reduces computational cost while maintaining competitive performance in sequential modeling tasks.

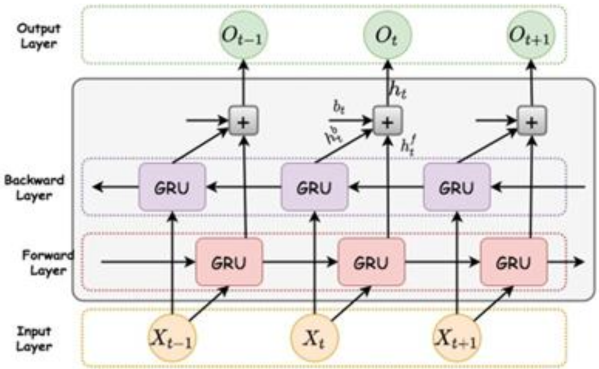

Figure 4. BiGRU architecture [18]

2.4.3. Proposed hybrid

In addition to evaluating BiGRU and BiLSTM separately, this study proposes a hybrid architecture that integrates both on top of IndoBERT embeddings. The aim is to combine BiGRU's training efficiency with BiLSTM's ability to capture long-term sequential dependencies, thereby addressing the trade-off between computational speed and contextual accuracy in Indonesian hoax detection. The overall design of the proposed model is illustrated in Figure 5.

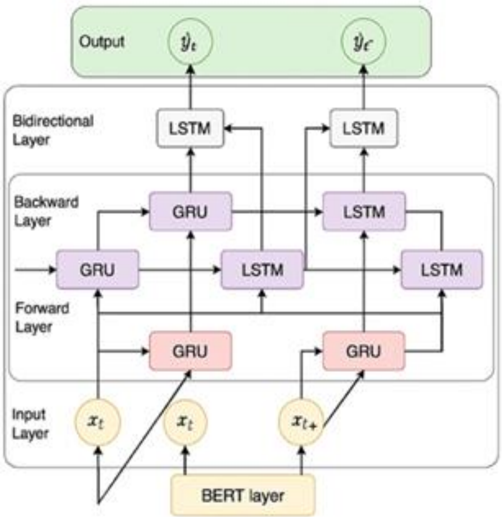

Figure 5. Hybrid model architecture

We stack a bidirectional GRU followed by a BiLSTM on top of IndoBERT embeddings to combine fast convergence (BiGRU) with long-range dependency modeling (BiLSTM). The sentence embedding is then pooled and fed to a lightweight classifier.

Pipeline.
1) Input text → IndoBERT tokenizer → input_ids, attention_mask (max seq len e.g., 128).
2) IndoBERT encoder → contextual token embeddings (hidden size per token per step).
3) BiGRU (bidirectional) → captures efficient bidirectional patterns.
4) BiLSTM (bidirectional) → enriches long-term sequential dependencies.
5) Global average pooling → BatchNorm → dropout.
6) Dense (ReLU) → Dense (Softmax, 2 classes: hoax/non-hoax).

Design rationale.
– BiGRU → BiLSTM order: GRU's lighter gating accelerates early convergence and stabilizes gradients; LSTM then captures longer dependencies that GRU may miss.
– Pooling and regularization: global pooling, batch normalization, and dropout reduce dimensionality and mitigate overfitting while keeping the classifier compact.
– Compatibility with IndoBERT: leverages contextual embeddings without heavy task-specific heads, preserving Transformer semantics.

As summarized in Table 2, tokenized inputs are encoded by IndoBERT into (128, 768) contextual embeddings, passed through a bidirectional GRU-LSTM sequence layer, pooled to a fixed-length vector, regularized, and finally classified via a two-layer dense head with Softmax.

Table 2. Hybrid model layer configuration

Layer type	Name/description	Output shape	Parameters
InputLayer	input_ids, attention_mask	(None, 128)	0
Lambda	-	(None, 128, 768)	0
Bidirectional (GRU+LSTM)	-	(None, 128, 256)	394,240
GlobalAveragePooling1D	-	(None, 256)	0
Dropout	-	(None, 128)	0
Dropout	-	(None, 256)	0
Dense	-	(None, 128)	32,896
Dense	Softmax output (2 classes)	(None, 2)	258

2.5. Model evaluation

The evaluation of model performance is conducted using conventional classification metrics, including accuracy, precision, recall, and the F1-score. In this scenario, T_t (true true) denotes the quantity of data that is accurately identified as true (true positive), whereas T_f (true false) signifies the amount of data that is erroneously classified as true (false negative). Furthermore, F_t (false true) denotes the amount of data that is incorrectly classified as false (false positive), and F_f (false false) denotes the amount of data that is correctly classified as false (true negative).

$$Accuracy = \frac{T_t + T_f}{T_t + T_n + F_t + F_f} \tag{1}$$

$$Precision = \frac{T_t}{T_t + F_t} \tag{2}$$

$$Recall = \frac{T_t}{T_t + T_f} \tag{3}$$

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

2.6. Model validation

To prevent overfitting and ensure the generalizability of the model, a k-fold cross-validation technique was employed with 10 folds. Each model was trained for 20 epochs per fold, resulting in a total of 200 training iterations across the folds. Early stopping was applied during training to monitor validation loss and halt the training process if no improvement was observed over a predefined patience threshold. This approach not only minimizes the risk of overfitting but also ensures more stable and reliable evaluation results across different data partitions.

3. RESULTS AND DISCUSSION

The IndoBERT-based BiGRU-BiLSTM hybrid model was trained and tested using a cross-validation technique for 10 folds (k-fold =10) to ensure the stability and generalization of the model to the data used. The training was conducted for 20 epochs with a batch size of 32 in each fold. The evaluation results demonstrate the model's capacity to generate consistent performance in the training, testing, and validation processes, with accuracy values and other evaluation metrics exhibiting relative stability across each fold (see Figure 6). This finding suggests that the hybrid model possesses effective generalization capabilities and can efficiently handle data variations. The employment of k-fold validation offers a more comprehensive depiction of the model's performance and assists in mitigating the potential for bias arising from random data distribution.

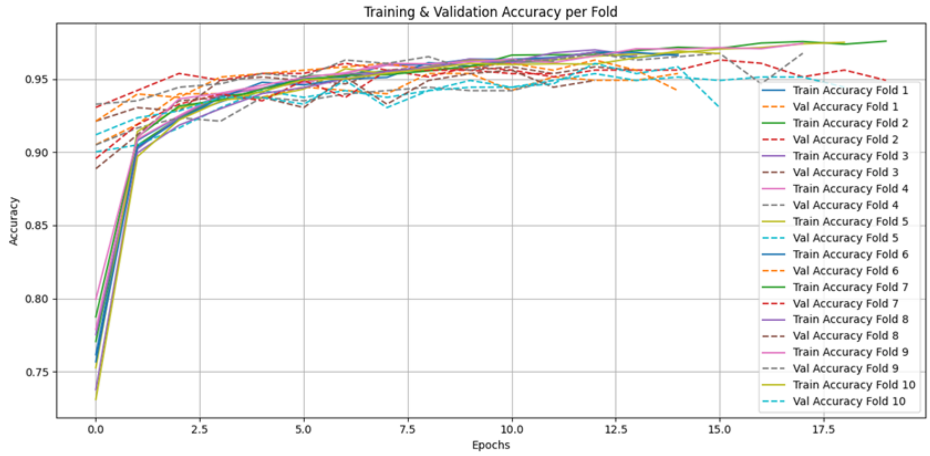

Figure 6. Graphic training and validation accuracy for 10-fold

In the proposed hybrid model, BiGRU and BiLSTM are arranged sequentially to combine their complementary strengths. BiGRU, placed first, accelerates convergence with fewer parameters while effectively modeling bidirectional dependencies. BiLSTM then enhances the ability to capture long-term dependencies and complex contextual patterns. This configuration enables richer sequence representations and leads to improved classification performance, as summarized in Table 3. Figure 7 shows a graph of the results of the IndoBERT-based BiGRU-BiLSTM hybrid model training process, which includes changes in accuracy and loss values over 20 epochs. This graph provides an overview of the stability and convergence of the model during the training and validation process.

Training accuracy increases gradually from the beginning to the end of the epoch, indicating that the model can learn patterns from the training data effectively. The validation accuracy demonstrates stability and high value from the outset, exhibiting slight variations but maintaining a consistent upward trend. The absence of any discernible indication of overfitting is evidenced by the lack of a substantial decline in validation accuracy, despite the sustained increase in training accuracy. Meanwhile, the second Figure 8 is a confusion matrix that illustrates the distribution of model predictions for each of the classes and identifies the type of misclassification that occurred.

To evaluate performance improvements, the proposed hybrid model was compared with standalone BiGRU and BiLSTM baselines. While BiGRU excelled in training efficiency and BiLSTM in modeling long-term context, the IndoBERT-based BiGRU-BiLSTM hybrid achieved superior accuracy, precision, and stability by combining their strengths. The comparative results are summarized in Table 4.

The evaluation results show that the IndoBERT-based BiGRU–BiLSTM hybrid model provides the most superior performance across all evaluation metrics. Specifically, it achieved an accuracy of 98.73%, a recall of 99.01%, a precision of 98.04%, and an F1-score of 98.98%. In comparison, BiGRU and BiLSTM models achieved accuracies of 95.54% and 95.94%, respectively, with lower values across other metrics. This advantage was consistently observed across all 10 folds, indicating strong generalization and stability. Category-wise results (politics, sports, and technology) also remain consistently high with only minor variations (see Table 5), indicating robustness across heterogeneous content types. Interestingly, despite its higher architectural complexity, the hybrid model also achieved the fastest training time (172.72s), outperforming BiGRU (231.04s) and BiLSTM (280.94s). These findings suggest that the integration of BiGRU and BiLSTM effectively leverages the sequential modeling strengths of both architectures, where BiGRU contributes to faster convergence, and BiLSTM captures longer-term dependencies. Unlike prior studies that utilized BiLSTM [20] or BiGRU [19] in isolation, this study demonstrates that combining both within the IndoBERT framework offers a more balanced and efficient solution for hoax detection in Indonesian language.

Table 3. Result evaluation performance of the hybrid model

Evaluation	Value %
Accuracy	98.73
Recall	99.01
Precision	98.04
F1-score	98.98

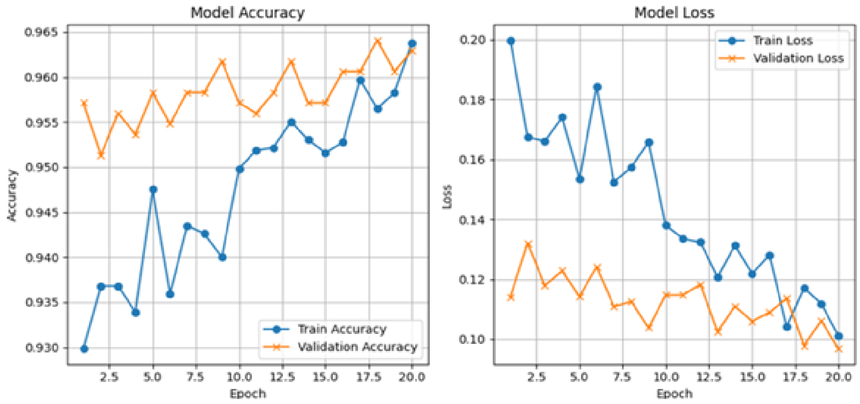

Figure 7. Graphic training accuracy, validation, and loss

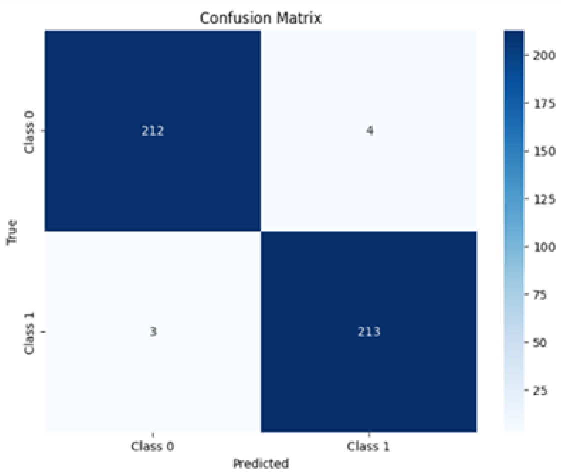

Figure 8. Confusion matrix hoax detection

Table 4. Comparison of the hybrid model with the baseline

Model-based IndoBERT	Accuracy %	Recall %	Precision %	F1-score%	Time/second
BiGRU [19]	95.54	96.15	95.01	95.58	231.04
BiLSTM [20]	95.94	96.66	95.30	95.97	280.94
BiGRU-BiLSTM hybrid	98.73	99.01	98.04	98.98	172.72

Table 5. Performance of the hybrid model across categories

Category	Accuracy %	Recall %	Precision %	F1-score %
Politics	99.16	99.16	98.89	99.02
Technology	98.98	99.44	98.48	98.96
Sports	99.04	99.86	98.62	99.24

Overall, these results prove that the IndoBERT-based hybrid BiGRU-BiLSTM approach is an effective strategy to improve the performance of text classification systems, especially in the context of news or information detection based on the Indonesian language. The superiority in all aspects of evaluation metrics and time efficiency makes this model very feasible to be applied in real scenarios with large and complex data volumes. Despite its high performance, the proposed hybrid model has several limitations. Its dual-layer sequential architecture increases the number of parameters and memory usage compared to single models. Moreover, the model has only been evaluated on a binary-class dataset, and its effectiveness on multi-class or multi-label hoax detection remains to be investigated. Future work should also evaluate inference speed and deployability on low-resource or edge devices to ensure real-time feasibility.

The findings of this study have several important implications. First, they validate the synergistic effect of combining BiGRU and BiLSTM on top of IndoBERT, resulting in improved classification accuracy and computational efficiency. This directly addresses a key trade-off in prior studies, where either model effectiveness or training efficiency was often prioritized, rarely both. Unlike previous works that relied solely on either Transformer-based or RNN-based models, the proposed hybrid framework leverages the strengths of both, making it a practical and high-performing solution. Residual errors are concentrated on borderline cases such as satire, clickbait, or partially true claims, suggesting the value of incorporating stylistic/pragmatic cues in future work.

The model's ability to achieve faster convergence and higher accuracy demonstrates its suitability for real-world deployments, especially on mobile or web platforms where computational resources are limited. This is particularly significant in the Indonesian context, where the language's rich morphology and complex syntax pose challenges that this hybrid approach can effectively address. Beyond technical performance, this suitability also extends to real-world applications such as social media moderation, media verification, and digital literacy initiatives, reinforcing the role of AI in strengthening public trust in information systems. The hybrid model architecture presents promising avenues for expansion into diverse NLP applications, including rumor detection [21], cyberbullying, and offensive content detection [22]–[25], and mental health classification from social media [26], [27]. These tasks have seen growing interest due to their social impact and data complexity. Further work could explore model compression and quantization to enhance real-time deployment in constrained environments.

4. CONCLUSION

This study introduced a hybrid deep learning model integrating IndoBERT with BiGRU and BiLSTM for Indonesian hoax detection. The proposed approach achieved superior performance (accuracy 98.73% and F1-score 98.98%) compared to standalone models, demonstrating the benefit of combining contextual and sequential learning. The synergy between BiGRU's efficiency and BiLSTM's ability to capture long-term dependencies offers a balanced solution for text classification in morphologically rich languages such as Indonesian language. Despite these strengths, limitations remain regarding dataset diversity, model complexity, and reliance on a single pre-trained embedding. Future research should address these aspects and explore broader applications such as rumor detection, offensive content filtering, and real-time analytics. Overall, this work provides a practical and effective framework with potential to support media monitoring, fact-checking, and digital literacy initiatives. These findings highlight not only the model's technical merits but also its potential for real-time filtering, fact-checking support, and civic education, thereby contributing to broader social impact.

FUNDING INFORMATION

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. All research expenses were fully supported by the authors through personal funding.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Siti Mujilahwati	✓	✓		✓		✓	✓		✓			✓	✓	✓
Moh. Rosidi Zamroni	✓	✓	✓			✓	✓	✓		✓	✓			
Miftahus Sholihin			✓	✓	✓					✓	✓			

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known conflicts of interest, whether financial or non-financial (including political, personal, religious, ideological, academic, or intellectual), that could have influenced or been perceived to influence the results and interpretation of the research reported in this paper.

INFORMED CONSENT

The data analyzed in this study were obtained from publicly available sources. No personally identifiable information was collected or disclosed; therefore, informed consent was not required.

ETHICAL APPROVAL

The data used in this study were obtained from publicly available sources and did not involve direct interaction with human subjects or the collection of personally identifiable information. Therefore, ethical approval was not required.

DATA AVAILABILITY

The data that support the findings of this study are openly available on Kaggle at the following repository: hoax news Indonesia dataset (https://www.kaggle.com/datasets/vijayandika/hoax-news-indonesia). The dataset was accessed and used in accordance with Kaggle's data usage policies.

REFERENCES

[1] N. Rai, D. Kumar, N. Kaushik, C. Raj, and A. Ali, "Fake news classification using transformer based enhanced LSTM and BERT," *International Journal of Cognitive Computing in Engineering*, vol. 3, pp. 98–105, 2022, doi: 10.1016/j.ijcce.2022.03.003.

[2] A. Z. M. Noor, R. Gernowo, and O. D. Nurhayati, "Data augmentation for hoax detection through the method of convolutional neural network in Indonesian news," *Jurnal Penelitian Pendidikan IPA*, vol. 9, no. 7, pp. 5078–5084, 2023, doi: 10.29303/jppipa.v9i7.4214.

[3] H. S. Gavalan, M. N. Rastgoo, and B. Nakisa, "A BERT-based summarization approach for depression detection," 2024, *arXiv:2409.08483*.

[4] A. Rahmawati, A. Alamsyah, and A. Romadhony, "Hoax news detection analysis using IndoBERT deep learning methodology," in *2022 10th International Conference on Information and Communication Technology (ICoICT)*, Aug. 2022, pp. 368–373, doi: 10.1109/ICoICT55009.2022.9914902.

[5] M. Y. Ridho and E. Yulianti, "From text to truth: leveraging IndoBERT and machine learning models for hoax detection in Indonesian news," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 10, no. 3, pp. 544–555, 2024, doi: 10.26555/jiteki.v10i3.29450.

[6] F. A. Nugraha and D. H. Fudholi, "COVID-19 misinformation detection in Indonesian tweets using BERT," *Jurnal Media Informatika Budidarma*, vol. 7, no. 2, pp. 600–608, 2023, doi: 10.30865/mib.v7i2.5668.

[7] R. Yusuf and S. Suyanto, "Hoax detection on Indonesian text using long short-term memory," in *ICOIACT 2022 - 5th International Conference on Information and Communications Technology: A New Way to Make AI Useful for Everyone in the New Normal Era, Proceeding*, 2022, pp. 268–271, doi: 10.1109/ICOIACT55506.2022.9972086.

[8] M. N. Fakhruzzaman and S. W. Gunawan, "Web-based application for detecting Indonesian clickbait headlines using IndoBERT," 2021, *arXiv:2102.10601*.

[9] J. Dong, "Improving the BiLSTM model performance for tweet sentiment analysis," *Applied and Computational Engineering*, vol. 6, no. 1, pp. 1106–1117, Jun. 2023, doi: 10.54254/2755-2721/6/20230415.

[10] A. Abdullah, "Classifying innovation with biLSTM and biGRU: a study of labor market data in the Netherlands," *M.S. Thesis*, Department of Cognitive Science and Artificial Intelligence, Tilburg University, Tilburg, The Netherlands, 2021.

[11] T. Huang, Z. Xu, P. Yu, J. Yi, and X. Xu, "A hybrid transformer model for fake news detection: leveraging Bayesian optimization and bidirectional recurrent unit," in *2025 8th International Symposium on Big Data and Applied Statistics (ISBDAS)*, Feb. 2025, pp. 696–701, doi: 10.1109/ISBDAS64762.2025.11117058.

[12] N. Noor, N. Z. M. Safar, and C. Supriyanto, "Segmentation word to improve performance sentiment analysis for Indonesian language," *Fusion: Practice and Applications*, vol. 15, no. 2, pp. 145–154, 2024, doi: 10.54216/FPA.150213.

[13] F. Koto, J. H. Lau, and T. Baldwin, "INDOBERTWEET: a pretrained language model for Indonesian Twitter with effective domain-specific vocabulary initialization," in *EMNLP 2021 - Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 2021, pp. 10660–10668, doi: 10.18653/v1/2021.emnlp-main.833.

[14] F. C. Zegarra, J. V.-Machuca, and A. M. Coronado, "A comparative study of CNN, LSTM, BiLSTM, and GRU architectures for tool wear prediction in milling processes," *Journal of Machine Engineering*, vol. 23, no. 4, pp. 122–136, 2023, doi: 10.36897/jme/174019.

[15] B. Khaldi, F. Harrou, A. Dairi, and Y. Sun, "A deep recurrent neural network framework for swarm motion speed prediction," *Journal of Electrical Engineering and Technology*, vol. 18, no. 5, pp. 3811–3825, 2023, doi: 10.1007/s42835-023-01446-7.

[16] X. Lin, Z. Quan, Z. J. Wang, H. Huang, and X. Zeng, "A novel molecular representation with BiGRU neural networks for learning atom," *Briefings in Bioinformatics*, vol. 21, no. 6, pp. 2099–2111, 2020, doi: 10.1093/bib/bbz125.

[17] A. B. Prasetijo, R. R. Isnanto, D. Eridani, Y. A. A. Soetrisno, M. Arfan, and A. Sofwan, "Hoax detection system on Indonesian news sites based on text classification using SVM and SGD," in *Proceedings - 2017 4th International Conference on Information Technology, Computer, and Electrical Engineering, ICITACEE 2017*, 2017, pp. 45–49, doi: 10.1109/ICITACEE.2017.8257673.

[18] F. Harrou, A. Dairi, A. Dorbane, and Y. Sun, "Energy consumption prediction in water treatment plants using deep learning with data augmentation," *Results in Engineering*, vol. 20, 2023, doi: 10.1016/j.rineng.2023.101428.

[19] X. Gu, "Research on the design of news text classification algorithm based on bidirectional GRU neural network," *OALib*, vol. 12, no. 4, pp. 1–16, 2025, doi: 10.4236/oalib.1113226.

[20] S. G. Arkaan, A. R. Atmadja, and M. D. Firdaus, "Fake news detection in the 2024 Indonesian general election using bidirectional long short-term memory (BI-LSTM) algorithm," *Komputasi: Jurnal Ilmiah Ilmu Komputer dan Matematika*, vol. 21, no. 2, pp. 22–30, 2024, doi: 10.33751/komputasi.v21i2.9987.

[21] A. Bondielli and F. Marcelloni, "A survey on fake news and rumour detection techniques," *Information Sciences*, vol. 497, pp. 38–55, 2019, doi: 10.1016/j.ins.2019.05.035.

[22] R. Zhao, A. Zhou, and K. Mao, "Automatic detection of cyberbullying on social networks based on bullying features," in *Proceedings of the 17th International Conference on Distributed Computing and Networking*, Jan. 2016, pp. 1–6, doi: 10.1145/2833312.2849567.

[23] M. Zampieri, S. Malmasi, P. Nakov, S. Rosenthal, N. Farra, and R. Kumar, "Predicting the type and target of offensive posts in social media," in *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, 2019, pp. 1415–1420, doi: 10.18653/v1/n19-1144.

[24] M. R. Zamroni, R. A. Hamid, S. Mujilahwati, M. Sholihin, and D. M. Leksana, "Detection of bullying content in online news using a combination of RoBERTa-BiLSTM," *Jurnal Teknik Informatika*, vol. 6, no. 1, pp. 41–50, 2025, doi: 10.52436/1.jutif.2025.6.1.4140.

[25] K. Chandra, K. A. Prasetya, R. D. Saputra, and M. F. Hasani, "Leveraging IndoBERT for cyberbullying classification on social media," in *2024 4th International Conference of Science and Information Technology in Smart Administration (ICSINTESA)*, Jul. 2024, pp. 407–411, doi: 10.1109/ICSINTESA62455.2024.10747874.

[26] R. A. Ramadhan, D. Pramesti, and H. Fakhrurroja, "Multi-label classification of depression and anxiety in Indonesian social media: a transformer approach," in *Proceedings of Data Analytics and Management*, 2026, pp. 177–190, doi: 10.1007/978-3-032-03740-4_16.

[27] A. S. Rizky and E. Y. Hidayat, "Emotion classification in Indonesian text using IndoBERT," *Computer Engineering and Applications Journal*, vol. 14, no. 1, pp. 1–11, Feb. 2025, doi: 10.18495/comengapp.v14i1.494.

BIOGRAPHIES OF AUTHORS

Siti Mujilahwati is an assistant professor in natural language processing at the Department of Informatics Engineering, Faculty of Science and Technology, Universitas Islam Lamongan. She is currently a Ph.D. candidate in computer science at Universiti Tun Hussein Onn Malaysia (UTHM). She can be contacted at email: moedjee@unisla.ac.id.

Moh. Rosidi Zamroni is a member of the Faculty of Science and Technology, Computer Science Program. Currently, he is pursuing a Ph.D. at FSKM Tun Hussein Onn Malaysia, with a focus on natural language processing. He can be contacted at email: rosidizamroni@unisla.ac.id.

Miftahus Sholihin is a Ph.D. candidate at the Faculty of Computer Sciences and Information Technology (FSKTM), Universiti Tun Hussein Onn Malaysia. His research interest includes digital image processing, pattern recognition, data mining, and artificial intelligence. He can be contacted at email: miftahus.sholihin@unisla.ac.id.