

Research Article

Self-Attention Enhanced Dual BiGRU for Arabic Fake News Detection

Baqer M. Merzah^{1,2,*}, Jafar Razmara², Jaber Karimpour²

¹ Department of Computer Science, Faculty of Education, University of Kufa, Najaf, Iraq

² Department of Computer Science, Faculty of Mathematics, Statistics, and Computer Science, University of Tabriz, Tabriz, Iran

ARTICLEINFO

Article History

Received 18 Jun 2025

Revised 10 Aug 2025

Accepted 27 Aug 2025

Published 2 Sep 2025

Keywords

Fake News Detection

Deep Learning

Natural Language

Processing (NLP)

Bidirectional gated
recurrent unit (BiGRU)

Attention Mechanism



Abstract

The rapid proliferation of social media platforms has greatly amplified the dissemination of fake news, representing significant obstacles to public trust and evidence-based decision-making, particularly for the Arabic-speaking population. Meeting the challenge of Arabic fake news detection is a problem compounded by the complex morphological nature of the language, as well as limited resources. This study presents a hybrid deep learning framework that integrates two Bidirectional Gated Recurrent Units (BiGRUs) along with an attention mechanism for efficiently detecting misinformation in Arabic news. The method leverages FastText word embeddings for disambiguating the intricate semantics of the Arabic language. The model is meticulously crafted to account for the morphological variability and contextual sensitivities of Arabic, with the extensive Arabic fake news dataset (AFND) being used for training and testing. Experimental findings show that our model performs better with an accuracy of 91.92%, outperforming current state-of-the-art approaches. The findings highlight the effectiveness of integrating advanced neural architectures and tailored preprocessing for Arabic, paving the way for more robust and interpretable fake news detection systems in low-resource languages.

1. INTRODUCTION

Social media platforms' explosive growth has completely changed how news is disseminated, allowing information to travel at a previously unheard-of scale and speed. The digital revolution has made it easier for everyone to get information. However, it has also made it easier for false information to spread widely, endangering social cohesiveness, public trust, and the ability to make educated decisions [1]. Misinformation, disinformation, hoaxes, and rumors represent other aspects and types of fake news [2]. What happened in the 2016 US elections was the misleading information published about the US presidential candidates [3][4], and the fabricated information published about the vaccine during the COVID-19 pandemic represents the danger of spreading fake news and its impact on people and public opinion [5]. This phenomenon represents a major challenge to developing effective systems to detect fake news. A number of researchers have attempted to develop various approaches to detecting fake information through machine learning (ML) and deep learning (DL) methods. However, the bulk of prior research has predominantly concentrated on languages rich in resources, particularly English [6][7].

Detecting fake news in Arabic encounters multiple challenges, most notably the linguistic complexity of Arabic itself. It is characterized by an advanced grammatical and morphological structure [8][9], and its texts range from Classical Arabic to local dialects [10]. In addition, the Arabic language suffers from a scarcity of resources, which limits researchers' ability to develop and apply techniques to detect it [11][12]. A good example of the intricacies of the Arabic language is the many meanings of a single word depending on its application [13]. For example, the word "عين" (ayn) can have different meanings depending on its application. It may be a noun and mean "the organ of sight"; in a different context, it may mean "a spring of water". In addition, the term may mean "the thing itself" or "the essence". Lastly, the term "عَيْنَ" (ayn) may be a verb and means "to choose" or "to employ".

This expressiveness of meaning highlights the depth and wealth of the Arabic language. This also presents a great challenge to natural language processing (NLP) systems [14]. To enable these systems to process Arabic texts effectively, advanced techniques need to be employed to effectively sense the context and differentiate between the multiple senses of words. One of the biggest challenges is adapting AI models to the distinctive features of Arabic, underscoring the importance and

*Corresponding author. Email: baqirm.merzah@uokufa.edu.iq

originality of approaches tailored specifically for it [9]. By proposing a DL model with an attention mechanism to detect false news in Arabic, this research tries to fill this gap.

The contributions in this paper are:

- **Utilizing FastText embeddings for Arabic:** FastText word embeddings were utilized to better capture the morphological richness and context-dependent meaning of Arabic, especially for out-of-vocabulary and rare words, and provided stronger feature representations for fake news detection.
- **Proposing a hybrid BiGRU-attention model:** A hybrid architecture combining BiGRU layers and an attention mechanism was developed to capture temporal dependencies and highlight key parts of the text, improving both performance and interpretability.
- **Outperforming state-of-the-art methods:** The model proposed in this paper achieved superior performance to state-of-the-art (SOTA) practices when evaluated on an Arabic fake news dataset (AFND), as demonstrated by higher accuracy and F1 scores.

The rest of the paper is set up like this: Section 2 looks at the important research on fake news detection in Arabic. Section 3 explains how the proposed model was designed. Section 4 shows the experimental results and compares them to the SOTA results. It also talks about an ablation study. Section 5 presents the conclusion and potential avenues for future improvements.

2. RELATED WORKS

This section gives a review of the past research done on Arabic fake news detection. Experiments have explored various methods, from ML and DL models to transformer-based approaches. A method for evaluating the reliability of Arabic news shared on Twitter was created by [15]. The authors applied ML algorithms to a dataset of 800; they found that a decision tree classifier outperformed SVM and Naive Bayes. Their model was that robustness was achieved through the utilization of a combined set of features that comprised content features, source features, external content validation, and sentiment conveyed in user ratings of the news tweets. The authors indicated that their experiment findings demonstrated the superiority of this combination-based method over a number of earlier systems intended to evaluate credibility for specific uses. The author in [16] compared the performance of eight ML algorithms and tested five DL algorithm combinations, such as CNN and LSTM. The results demonstrated that BiLSTM outperformed the other approaches, achieving an accuracy rate of 75% on a dataset of 4561 tweets.

Shifting the focus from model architecture to feature engineering, the researchers in [17] created a unique dataset by crowdsourcing Arabic news articles, identifying the most significant textual features, including part-of-speech tags, syntactic-semantic roles, emotional cues, and contextual polarity. These features were used to train different supervised ML classifiers, with the random forest (RF) model achieving the best result of 78% accuracy.

The authors in [18] introduced the Arabic Fake News Dataset (AFND) due to the absence of comprehensive resources for Arabic fake news detection. This constructive addition comprises over 600,000 news articles. The authors had performed experiments on various DL and ML models and found that DL methods, particularly Capsule Networks (which achieved 79% accuracy in binary classification), showed greater promise.

The authors [19] used the large AFND dataset to study Arabic fake news detection. They proposed a DL model that combines an attention mechanism and a Bi-LSTM, and the final article classification task is performed using an MLP classifier. Their approach showed great performance, achieving an accuracy of 0.8127 on the AFND dataset that outperformed baseline RNN models along with previous work on this specific dataset. In another hybrid Arabic fake news detection model proposed in [20], authors utilized GloVe and FastText for word embedding and subsequently used multiple 2D convolutions and Bi-LSTM for feature extraction. An accuracy of 88% for binary and 78% for multi-class classification was reported on a large Arabic dataset known as AFND.

The common problem of data imbalance in Arabic fake news detection was directly addressed by [21]. The research made use of the AFND and ArCOV-19 datasets to evaluate various balancing techniques (e.g., SMOTE and random under-sampling) on a broad variety of ML and DL frameworks. The findings indicate that the most effective method was SMOTEENN, which significantly enhanced performance metrics such as F1-score and recall, particularly when paired with DL architectures such as BiGRU and BiLSTM.

The author in [22] introduced a new hybrid framework for detecting fake news in Arabic, utilizing pre-trained Arabic bidirectional encoder representations from transformers (APBTMs) like AraBERT, MARBERT, and GigaBERT. These APBTMs are paired with convolutional neural networks (CNNs) featuring both 1D-CNN and 2D-CNN structures to refine feature extraction from the text. Notably, the 2D-CNN architecture processes outputs from each layer of the APBTMs to enhance feature extraction, ultimately improving classification performance. The authors in [23] developed multimodal (text + image) models for Arabic rumor detection on Twitter, but found their text-only model using MARBERTv2 performed better.

3. PROPOSED METHOD

This section presents the architectural framework and components of the suggested DL model, which is especially tailored for Arabic fake news detection. The model is designed to accurately classify news texts by capturing the distinct semantic, morphological, and sequential features of the Arabic language. An input layer for data ingestion and Arabic-specific preprocessing, a word embedding layer, a feature representation layer, and a classification layer for determining the news article's veracity make up the model's structured pipeline, as shown in Fig. 1.

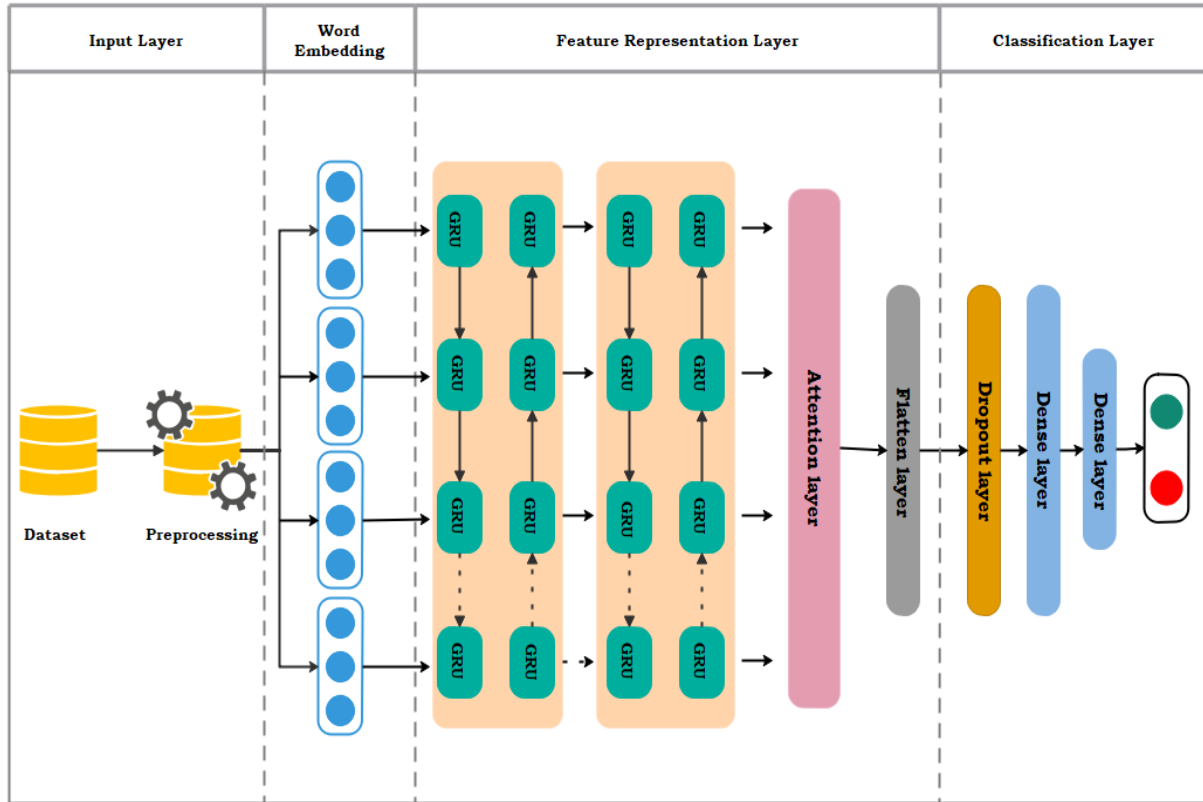


Fig. 1. The proposed model architecture

3.1 Data Collection and Preprocessing

The first step involves preparing the Arabic dataset for analysis. Preprocessing is essential to clean and normalize the data and handle the complexities of Arabic text, which include the following steps:

- Text cleaning: unnecessary characters, punctuation, numbers, and special symbols are removed.
- Normalization: variations in Arabic letters (e.g., "ا", "آ", "إ") are unified, and diacritics are removed.
- Tokenization: The text is split into individual words (tokens).
- Stop word removal: common words that do not add semantic value (e.g., "من", "على", "في") are removed.
- Lemmatization: words are reduced to their root forms to handle different inflections and derivations in Arabic.

3.2 Word Embedding

The second layer of the model involves embedding the output of preprocessing using FastText [24], a pre-trained word embedding model for 157 languages, including Arabic. Unlike traditional word embeddings (Word2Vec [25] and GloVe [26]), FastText represents words as a combination of character n-grams, which makes it particularly effective for morphologically rich languages like Arabic. This subword information enables FastText to generate meaningful embeddings for words not encountered during its training phase (out-of-vocabulary words) and better captures morphological similarities between words. These 300-dimensional vectors, learned from vast Arabic text corpora, provide a rich semantic initialization for our model. The output of this layer is an embedding matrix of shape (sequence length, 300), serving as the input to the feature extraction.

3.3 Feature Extraction Layer

The feature extraction layer serves as a crucial step in our model, bridging the gap between raw word embeddings and the final classification. Its main purpose is to distil the most relevant patterns and contextual cues from the Arabic news texts, ensuring that the model can make accurate predictions. The following sections describe the main components and mechanisms used in this layer.

3.3.1 Bidirectional Gated Recurrent Unit

The bidirectional gated recurrent unit (BiGRU) is a type of recurrent neural network (RNN) that is particularly good at finding sequential dependencies and addressing the vanishing gradient issue in the GRU. The bidirectional capability lets the network read the sequence in both forward (left-to-right) and backward (right-to-left) directions [27][28]. Consequently, the hidden state of every token comprises contextual information from both the words before and after it, so producing a more complete contextual representation.

Stacking two BiGRU layers helps the model learn more complex and abstract hierarchical characteristics from the input sequence. Each token's hidden state thus incorporates context from both the words before and after it. We configured the first BiGRU layer with 64 units; hence, the GRU cell has 64 units for the forward pass and 64 units for the backward pass, producing a concatenated output dimension of 128. The initial word embeddings are processed in this layer. The second BiGRU layer, which is set up with 128 units (128 forward, 128 backward, usually yielding a concatenated output dimension of 256), receives the output sequence from the first BiGRU layer. The model can learn more intricate and abstract hierarchical features from the input sequence by stacking these layers. Rich contextual data gleaned from the complete input sequence in relation to each token position is encoded in the output of the second BiGRU layer, which is a series of hidden states. This output is then used as the input for the attention layer that follows.

3.3.2 Attention Layer

A self-attention mechanism receives the series of hidden states generated by the stacked BiGRUs. When creating the final sequence representation, self-attention enables the model to dynamically balance the significance of various tokens (represented by their BiGRU hidden states) within the same sequence. The relevance of each token to every other token in the sequence is reflected in the attention scores it calculates. Based on these scores, a weighted sum of the hidden states is generated, which allows the model to selectively highlight the most important words or phrases, regardless of their position, in the Arabic news text that are essential for differentiating between fake and real news [29]. This process improves the model's capacity to detect important indicators that are essential for the classification task and to capture long-range dependencies. A sequence of attention-weighted hidden states or a single context vector summarizing the sequence are the usual outputs of the attention layer.

If the output from the self-attention layer is still sequential or multidimensional, a flatten layer transforms it into a single, one-dimensional feature vector. This reshaping is necessary to feed the features into standard fully connected layers.

3.4 Classification Layer

The last section of the model, the classification layer, takes the flattened feature representation produced by the previous layers and processes it to generate the ultimate classification output. The dropout layer is applied to the flattened features. By randomly allocating 0.4 of its input units during the training phase, this layer acts as a regularization method, reducing overfitting. The dropout layer output is then passed through two successive dense (fully connected) 512 and 64 layers. These layers enable the model to learn complex patterns and refine the features for effective classification. Finally, the sigmoid activation function is used for binary classification tasks to determine which news is real or fake.

4 EXPERIMENTAL SETUP AND DISCUSSION

This section describes the results of the proposed model's evaluation conducted on a benchmark dataset. We outline the experimental hyperparameter configurations and the evaluation metrics, followed by an elaborate comparison of the baseline methods against the suggested model.

4.1 Dataset

We trained and tested our model for Arabic fake news detection on the Arabic Fake News Dataset (AFND), a large-scale corpus first presented by Khalil et al. [30]. Consisting of 134 public news websites from 19 Arab countries, this corpus has 606,912 Modern Standard Arabic news articles. Based on figures on the Arabic fact-checking platform, Misbar, each news source's credibility was manually rated as either "credible," "not credible," or "undecided." TABLE I shows the statistical information for the AFND dataset. Articles from sources labelled "undecided" were omitted from our training and testing sets; the proposed method uses binary classification; thus, we mapped the "credible" articles to the "real news" category and the "not credible" articles to the "fake news" category. The AFND is a great tool for training strong models for Arabic language fake news detection because of its great scale, geographic diversity, and natural class imbalance. Fig. 2 shows the percentage of "credible" and "not credible" articles after removing "undecided" articles for the AFND dataset

TABLE I. AFND DATASET STATISTICS

Statistics	Credible	Not credible	Undecided
Sources count	52	51	31
Articles count	207310	167233	232369

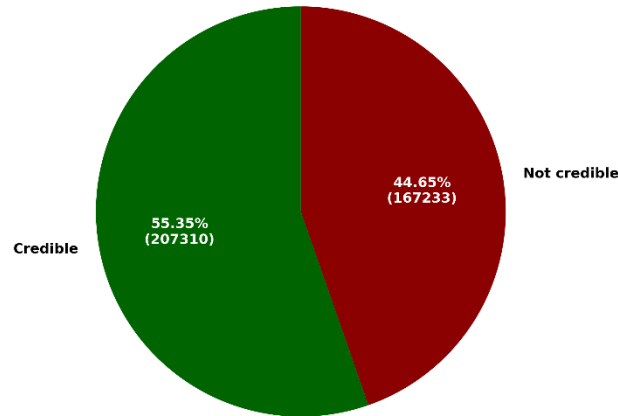


Fig. 2. Class distribution of AFND after removing the undecided class

4.2 Experimental and Hyperparameter Settings

To assess and analyze the efficacy of the suggested approach, we conducted all experiments using Python on Google Colab. The TensorFlow and Keras libraries were used for a deep learning model. Experiments were conducted utilizing Python libraries like Pandas, NLTK, and two prominent Arabic NLP toolkits: Farasa¹ and CAMEL Tools [31]. TABLE II delineates the hyperparameter values employed in the experimental evaluation. Hyperparameter values were determined through a systematic process of iterative experimentations aimed at improving model performance on a validation set. A range of values for key parameters was explored; in particular, batch sizes of 16, 32, and 64 were tried, in addition to different dropout rates from 0.3 to 0.5. The Adam optimizer was chosen for its consistent effectiveness in similar NLP tasks. This systematic parameter tuning process is important to ensure the robustness and replicability of our findings, a principle underscored in structured NLP evaluations [32].

TABLE II. THE MAIN HYPERPARAMETERS

Hyperparameter	Value
Embedding dimension	300
Maximum sequence length	1000
Number of neurons in BiGRUs	64, 128
Dropout	0.4
Dense layers	512, 64
Batch size	32
Output activation	Sigmoid
Optimization algorithm	Adam

¹ <https://farasa.qcri.org/>

4.3 Performance Evaluation Metrics

The effectiveness of the proposed fake news detection model was evaluated using four standard metrics commonly employed in fake news detection tasks: Basis for Calculation: All metrics are derived from the components of the confusion matrix: TP (True Positive), TN (True Negative), FP (False Positive), FN (False Negative).

Precision: Measures the proportion of articles classified as fake that are actually fake. It is calculated by the following equation.

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

Recall: Measures the proportion of actual fake news articles that were correctly identified by the model. The formula of recall is given below.

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

F1-Score: Provides a balance between precision and recall, calculated as their harmonic mean. The F1-score is calculated by the following equation.

$$F1 - score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (3)$$

Accuracy: Measures the overall proportion of correctly classified articles (both real and fake) out of the total number of articles. The accuracy formula is given below.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

4.4 Training Model

The proposed model was trained and assessed using the Arabic AFND dataset. As TABLE III's last row shows, the model attained an overall accuracy of 91.92%, a precision of 92.88%, a recall of 92.45%, and an F1-score of 92.66%. These findings show that the model not only boasts great accuracy but also preserves a strong balance between precision and recall, which is essential in reducing false positives and false negatives.

Fig. 3 presents a graphical representation of the model's performance using a confusion matrix. By showing the distribution of true positives, true negatives, false positives, and false negatives, the confusion matrix offers a deeper understanding of the classification behavior of the model. This enables a more complex evaluation of the strengths and possible areas for development of the model at many levels.

Fig. 4 also shows the training and validation accuracy as well as loss curves. These curves show that the model is well generalized and not overfitting to the training data since the accuracies stabilize at higher values, and both training and validation losses steadily decrease, pointing to effective convergence of the model.

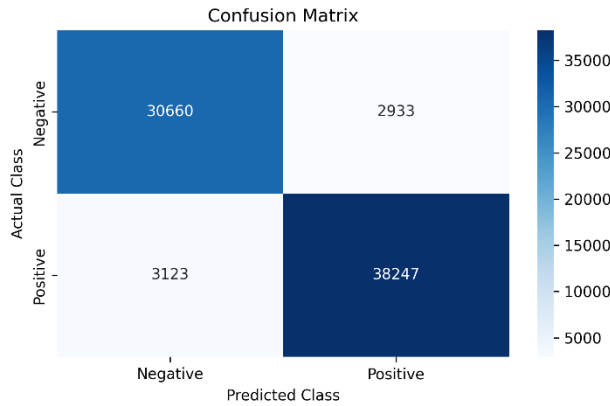


Fig. 3. The proposed model's confusion matrix

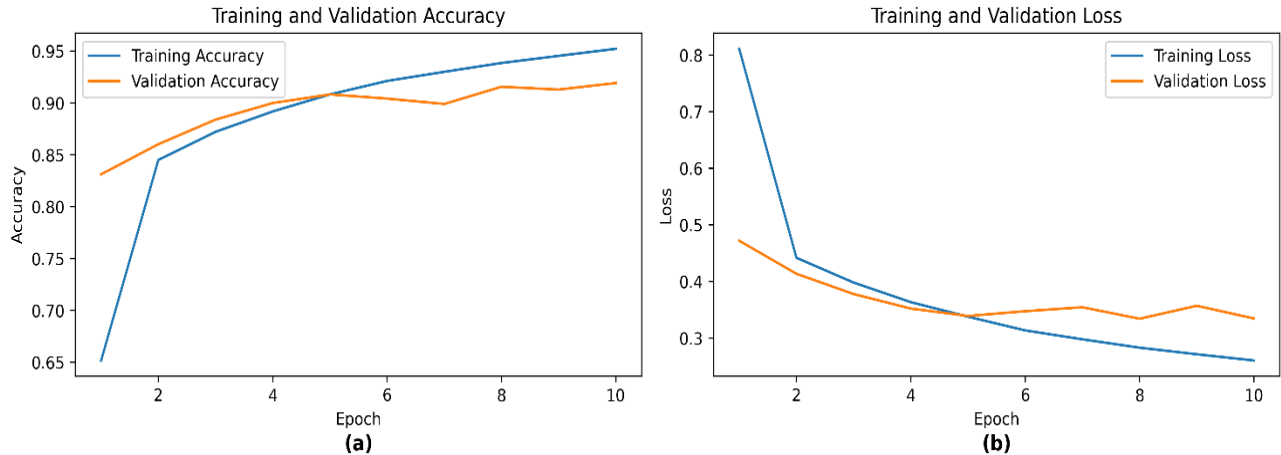


Fig. 4. Training and validation performance for the proposed model: (a) accuracy learning curve, (b) loss learning curve

4.5 Comparison with SOTA Methods

To assess the effectiveness of our proposed hybrid Dual BiGRU-attention model, we compared its performance with several of the most prominent models previously applied to the AFND dataset. The results, as presented in TABLE III, show that our model achieved an accuracy of 91.92%, with precision, recall, and F1-score values of 92.88%, 92.45%, and 92.66%, respectively. These figures are notably higher than those reported for other models, including Capsule Network, CNN-BiLSTM, WaraBERT-V1, BiLSTM-attention, and BiLSTM+SMOTEENN.

For example, WaraBERT-V1, which is considered one of the strongest transformer-based models for Arabic fake news detection, reached an accuracy of 89.91%, falling short of our model by about 2%. The proposed model outperformed the CNN-BiLSTM model, another strong SOTA model, by 4.43%, 4.88%, 5.13%, and 5.66 %, considering accuracy, precision, recall, and F1-score, respectively.

This improvement can be traced back to the hybrid nature of our architecture. By combining stacked Dual BiGRU layers with an attention mechanism, the model can capture both the sequential context and the most relevant features in the news articles. This allows for a deeper and more accurate understanding of the text, especially given the linguistic complexity of Arabic. The use of FastText embeddings further strengthens the model's ability to represent words, particularly those that are rare or morphologically complex, which is a common challenge in Arabic NLP tasks.

Moreover, the training and validation results indicate that our model generalizes well and is not prone to overfitting, as evidenced by the stable accuracy and loss curves. This balance between precision and recall is crucial in fake news detection, where both false positives and false negatives can have significant consequences. Unlike some previous studies that struggled with class imbalance, our model maintained high performance across all metrics, which highlights its robustness and reliability. Fig. 5 shows a graphical accuracy comparison between our model and SOTA approaches.

TABLE III. COMPARISON RESULTS OF DIFFERENT MODELS

Methods	Accuracy	Precision	Recall	F1-score
Capsule network [18]	78.3	-	-	-
BiLSTM-Attention [19]	81.73	-	-	-
BiLSTM+SMOTEENN [21]	85	85	85	85
CNN-BiLSTM [20]	87.49	88	87.32	87
WaraBERT-V1 [33]	89.91	-	-	-
Proposed: Dual BiGRU + Attention	91.92	92.88	92.45	92.66

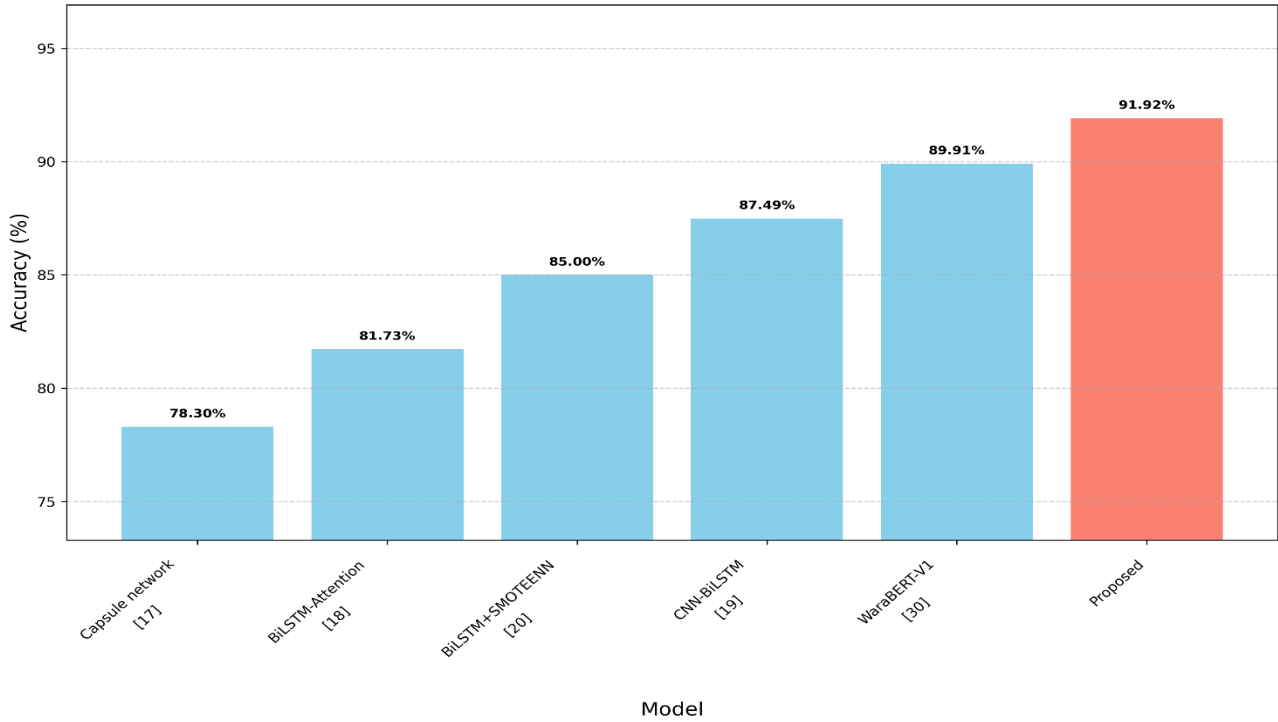


Fig. 5. Comparison results of the proposed model and SOTA models

4.6 Ablation Study

To better understand how each part of our model contributes to its overall performance, we carried out an ablation study, as shown in TABLE IV. This analysis was also a meaningful justification of our choice of using a dual-layer BiGRU architecture. We tested three different configurations to see the effect of removing or changing certain components.

In the first case, we removed the attention layer, so the model was made up of just the dual BiGRU. This setup helped us see how much the model could achieve by relying only on the dual BiGRU for feature extraction. The performance showed a clear drop, which means that even though the dual BiGRU performs well on learning sequential information, it fails to focus on the most significant parts of the input without attention.

For the second case, we kept the attention layer but used only a single BiGRU instead of two. This allowed us to check how much having just one BiGRU would affect the results while still letting the model benefit from the attention mechanism. Experiments show that indeed there was an improvement in performance over the first case, hence the beneficial effect of the attention layer, even in scenarios with less complexity in feature extraction. The results validate our architectural choice, as one BiGRU layer (Case 2) is not sufficient to capture the overall complexity of the data, while the dual-layer structure provides a significant performance boost. Thus, the stacked dual-BiGRU model represents an optimal trade-off between model complexity and performance.

Finally, both the dual BiGRU and the attention layer are incorporated into the model that we propose. From our experiment trials, this arrangement yielded the most optimal results. The attention layer allows for the model to focus on the most relevant features, and the dual BiGRU assists in gaining more comprehensive and diverse information from the input. That the combined application of the two parts is superior to the use of either individually is apparent when one considers. Fig. 6 shows the graphical ablation analysis.

TABLE IV. ABLATION ANALYSIS CONSIDERING FOUR METRICS

Methods	Accuracy	Precision	Recall	F1-score
Dual BiGRU Without Attention	89.73	91.07	90.23	90.65
Single BiGRU + Attention	90.61	91.79	91.16	91.47
Proposed: Dual BiGRU+Attention	91.92	92.88	92.45	92.66

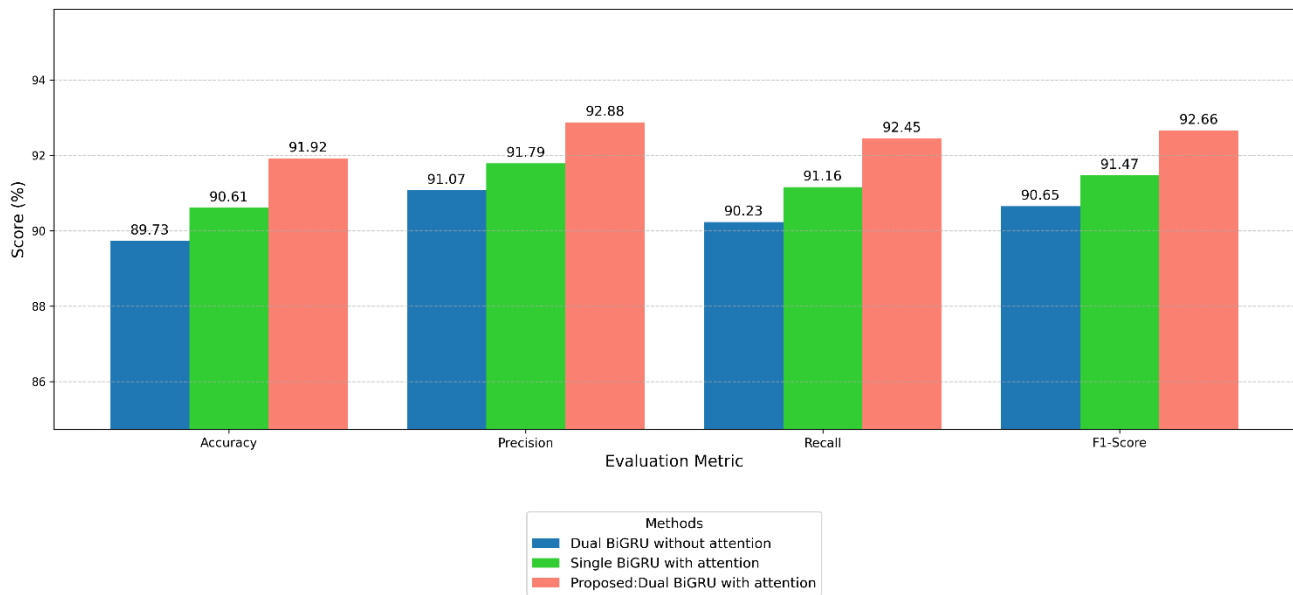


Fig. 6. A graphical ablation analysis

5 CONCLUSION

In this work, we proposed an Arabic fake news detection DL architecture. Through the integration of FastText embeddings and a dual-layer BiGRU along with the attention mechanism, our model can efficiently detect intricate semantic and sequential patterns intrinsic to Arabic news articles. The extensive analysis conducted on the AFND dataset shows that our model not only surpasses existing models in accuracy and reliability, but it also achieves a reasonable trade-off between recall and precision. Moreover, the ablation study emphasizes the need for the dual BiGRU structure and the attention mechanism to attain peak performance.

Such results provide evidence of the efficacy of complex neural architectures for dealing with fake news in multiple morphologically rich but low-resourced languages. These results provide a novel insight into the effectiveness of attention-based architectures in comparison to more complex models, at least when there is an imbalance in labels and limited data, thus fostering the development of Arabic NLP for misinformation detection. Due to the structural similarity between the challenges of low-resource languages, our method can be easily generalized through multilingual transfer, parameter-efficient fine-tuning and linguistically informed data augmentation. In future work we will focus of combining multimodal information and on further adapting the model to work with multiple dialects, thus increasing the robustness and usefulness of systems intended to identify fake news in Arabic.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

There was no outside funding for the research that led to the writing or publishing of this article.

Authors contributions

Conceptualization, B. M.; methodology, B. M., J. R.; software, B. M.; investigation, B. M., J. R.; writing—original draft preparation, B. M.; writing—review and editing, J. R., J. K.; visualization, B. M.; supervision, J. R.; project administration, J. R., advising, J. K.

References

- [1] B. Hu, Z. Mao, and Y. Zhang, “An overview of fake news detection: From a new perspective,” *Fundamental Research*, vol. 5, no. 1, pp. 332–346, Jan. 2025, doi: 10.1016/j.fmre.2024.01.017.
- [2] J. A. Nasir, O. S. Khan, and I. Varlamis, “Fake news detection: A hybrid CNN-RNN based deep learning approach,” *International Journal of Information Management Data Insights*, vol. 1, no. 1, p. 100007, Apr. 2021, doi: 10.1016/j.ijime.2020.100007.
- [3] Z. Jin, J. Cao, H. Guo, Y. Zhang, Y. Wang, and J. Luo, “Detection and Analysis of 2016 US Presidential Election Related Rumors on Twitter,” 2017, pp. 14–24. doi: 10.1007/978-3-319-60240-0_2.
- [4] M. Samadi, M. Mousavian, and S. Momtazi, “Deep contextualized text representation and learning for fake news detection,” *Inf Process Manag*, vol. 58, no. 6, p. 102723, Nov. 2021, doi: 10.1016/j.ipm.2021.102723.
- [5] F. Alam et al., “Fighting the COVID-19 Infodemic in Social Media: A Holistic Perspective and a Call to Arms,” *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 15, pp. 913–922, May 2021, doi: 10.1609/icwsm.v15i1.18114.
- [6] R. S. Matti and S. A. Yousif, “Leveraging Arabic BERT for High-Accuracy Fake News Detection,” *Iraqi Journal of Science*, pp. 751–764, Feb. 2025, doi: 10.24996/ij.s.2025.66.2.18.
- [7] H. Himdi, N. Zamzami, F. Najar, M. Alrehaili, and N. Bouguila, “Arabic Fake News Dataset Development: Humans and AI-Generated Contributions,” *IEEE Access*, vol. 13, pp. 62234–62253, 2025, doi: 10.1109/ACCESS.2025.3556376.
- [8] S. Alyoubi, M. Kalkatawi, and F. Abukhodair, “The Detection of Fake News in Arabic Tweets Using Deep Learning,” *Applied Sciences*, vol. 13, no. 14, p. 8209, Jul. 2023, doi: 10.3390/app13148209.
- [9] A. AlSajri, “Challenges in Translating Arabic Literary Texts Using Artificial Intelligence Techniques,” *EDRAAK*, vol. 2023, pp. 5–10, Feb. 2023, doi: 10.70470/EDRAAK/2023/002.
- [10] I. kh. Alnabrisi and M. kh. Saad, “Detect Arabic fake news through deep learning models and Transformers,” *Expert Syst Appl*, vol. 251, p. 123997, Oct. 2024, doi: 10.1016/j.eswa.2024.123997.
- [11] B. AlEsawi and M. Haqi Al-Tai, “Detecting Arabic Misinformation Using an Attention Mechanism-Based Model,” *Iraqi Journal For Computer Science and Mathematics*, vol. 5, no. 1, pp. 285–298, Feb. 2024, doi: 10.52866/ijcsm.2024.05.01.020.
- [12] E. S. Albtoush, K. H. Gan, and S. A. A. Alrababa, “Fake news detection: state-of-the-art review and advances with attention to Arabic language aspects,” *PeerJ Comput Sci*, vol. 11, p. e2693, Mar. 2025, doi: 10.7717/peerj-cs.2693.
- [13] A. B. Nassif, A. Elnagar, O. Elgendy, and Y. Afadar, “Arabic fake news detection based on deep contextualized embedding models,” *Neural Comput Appl*, vol. 34, no. 18, pp. 16019–16032, Sep. 2022, doi: 10.1007/s00521-022-07206-4.
- [14] N. Abdelhakim Othman, D. S. Elzanfaly, and M. M. M. Elhawary, “Arabic Fake News Detection Using Deep Learning,” *IEEE Access*, vol. 12, pp. 122363–122376, 2024, doi: 10.1109/ACCESS.2024.3451128.
- [15] S. Y. B. SAHAR F. SABBEH, “Arabic news credibility on Twitter: an enhanced model using hybrid features,” *J Theor Appl Inf Technol*, vol. 96, no. 8, pp. 2327–2338, Apr. 2023.
- [16] K. M. Fouad, S. F. Sabbeh, and W. Medhat, “Arabic Fake News Detection Using Deep Learning,” *Computers, Materials & Continua*, vol. 71, no. 2, pp. 3647–3665, 2022, doi: 10.32604/cmc.2022.021449.
- [17] H. Himdi, G. Weir, F. Assiri, and H. Al-Barhamtoshy, “Arabic Fake News Detection Based on Textual Analysis,” *Arab J Sci Eng*, vol. 47, no. 8, pp. 10453–10469, Aug. 2022, doi: 10.1007/s13369-021-06449-y.
- [18] A. Khalil, M. Jarrah, M. Aldwairi, and Y. Jararweh, “Detecting Arabic Fake News Using Machine Learning,” in *2021 Second International Conference on Intelligent Data Science Technologies and Applications (IDSTA)*, IEEE, Nov. 2021, pp. 171–177. doi: 10.1109/IDSTA53674.2021.9660811.
- [19] T. A. Wotaifi and B. N. Dhannoon, “Attention Mechanism Based on a Pre-trained Model for Improving Arabic Fake News Predictions,” *Iraqi Journal of Science*, pp. 6041–6054, Nov. 2023, doi: 10.24996/ij.s.2023.64.11.45.
- [20] A. Khalil, M. Jarrah, and M. Aldwairi, “Hybrid Neural Network Models for Detecting Fake News Articles,” *Human-Centric Intelligent Systems*, vol. 4, no. 1, pp. 136–146, Dec. 2023, doi: 10.1007/s44230-023-00055-x.
- [21] E. Aljohani, “Enhancing Arabic Fake News Detection: Evaluating Data Balancing Techniques Across Multiple Machine Learning Models,” *Engineering, Technology & Applied Science Research*, vol. 14, no. 4, pp. 15947–15956, Aug. 2024, doi: 10.48084/etasr.8019.
- [22] N. Abdelhakim Othman, D. S. Elzanfaly, and M. M. M. Elhawary, “Arabic Fake News Detection Using Deep Learning,” *IEEE Access*, vol. 12, pp. 122363–122376, 2024, doi: 10.1109/ACCESS.2024.3451128.
- [23] R. M. Albalawi, A. T. Jamal, A. O. Khadidos, and A. M. Alhothali, “Multimodal Arabic Rumors Detection,” *IEEE Access*, vol. 11, pp. 9716–9730, 2023, doi: 10.1109/ACCESS.2023.3240373.
- [24] A. Joulin, E. Grave, P. Bojanowski, M. Douze, H. Jégou, and T. Mikolov, “FastText.zip: Compressing text classification models,” Dec. 2016.

- [25] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient Estimation of Word Representations in Vector Space,” Jan. 2013.
- [26] J. Pennington, R. Socher, and C. Manning, “Glove: Global Vectors for Word Representation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2014, pp. 1532–1543. doi: 10.3115/v1/D14-1162.
- [27] 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,” Feb. 2025.
- [28] J. Alghamdi, Y. Lin, and S. Luo, “Enhancing hierarchical attention networks with CNN and stylistic features for fake news detection,” *Expert Syst Appl*, vol. 257, p. 125024, Dec. 2024, doi: 10.1016/j.eswa.2024.125024.
- [29] M. F. C. N. dos S. M. Y. B. X. B. Z. Y. B. Zhouhan Lin, “A Structured Self-attentive Sentence Embedding,” *arXiv preprint arXiv:1703.03130*, 2017.
- [30] A. Khalil, M. Jarrah, M. Aldwairi, and M. Jaradat, “AFND: Arabic fake news dataset for the detection and classification of articles credibility,” *Data Brief*, vol. 42, p. 108141, Jun. 2022, doi: 10.1016/j.dib.2022.108141.
- [31] O. 'Nasser, Z. 'Salam, K. 'Dima, T. 'Mai, O. 'Bashar, A. 'Go, I. 'Fadhl, E. 'Alexander, E. 'Nizar, H. 'Ossama, “CAMEL Tools: An Open Source Python Toolkit for Arabic Natural Language Processing,” in *Proceedings of the Twelfth Language Resources and Evaluation Conference*, Marseille, France: European Language Resources Association, May 2020, pp. 7022–7032.
- [32] M. Aljanabi, “Assessing the Arabic Parsing Capabilities of ChatGPT and Cloude: An Expert-Based Comparative Study,” *Mesopotamian Journal of Arabic Language Studies*, no. 2024, pp. 16–23, Feb. 2024, doi: 10.58496/MJALS/2024/002.
- [33] H. M. Turki et al., “Arabic fake news detection using hybrid contextual features,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 15, no. 1, p. 836, Feb. 2025, doi: 10.11591/ijece.v15i1.pp836-845.