IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 24 June 2024, accepted 20 July 2024, date of publication 29 July 2024, date of current version 27 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3435497

SURVEY

Fake News Detection Using Deep Learning: A Systematic Literature Review

MOHAMMAD Q. ALNABHAN AND PAULA BRANCO
Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada
Corresponding author: Mohammad Q. Alnabhan (e-mail: malna035@uottawa.ca).

ABSTRACT Nowadays, we witness rapid technological advancements in online communication platforms, with increasing volumes of people using a vast range of communication solutions. The fast flow of information and the enormous number of users opens the door to the publication of non-truthful news, which has the potential to reach many people. Disseminating this news through low- or no-cost channels resulted in a flood of fake news that is difficult to detect by humans. Social media networks are one of these channels that are used to quickly spread this fake news by manipulating it in ways that influence readers in many aspects. That influence appears in a recent example amid the COVID-19 pandemic and various political events such as the recent US presidential elections. Given how this phenomenon impacts society, it is crucial to understand it well and study mechanisms that allow its timely detection. Deep learning (DL) has proven its potential for multiple complex tasks in the last few years with outstanding results. In particular, multiple specialized solutions have been put forward for natural language processing (NLP) tasks. In this paper, we systematically review existing fake news detection (FND) strategies that use DL techniques. We systematically surveyed the existing research articles by investigating the DL algorithms used in the detection process. Our focus then shifts to the datasets utilized in previous research and the effectiveness of the different DL solutions. Special attention was given to the application of strategies for transfer learning and dealing with the class imbalance problem. The effect of these solutions on the detection accuracy is also discussed. Finally, our survey provides an overview of key challenges that remain unsolved in the context of FND.

INDEX TERMS Classification, deep learning, fake news, misinformation, systematic literature review.

I. INTRODUCTION

Due to a greater interest in the use of the internet, the spread of fake news has become more common than ever before. Before the popularity of social media platforms, fake news was less common and much more difficult to spread to a vast amount of people, as it was achieved either through word of mouth or through printed media. Fake news can be defined as the phenomenon that occurs when incorrect information is purposefully spread throughout social media outlets with a significant ability to convince the reader of the content written [1]. Nowadays, anyone can publish content without regulation or scrutiny. Several social media platforms, such as Facebook and Twitter, serve as means for disseminating fake news, as people and influencers utilize them to share their opinions, videos, and various activities [2], [3].

Fake news greatly increased in 2016 during the period preceding the United States (US) presidential election [4]. As such, fake news on social media networks has captured the attention of many researchers. Recently, detecting fake news has become an emerging area of interest for many researchers, such as [4] and [5]. However, fake news detection is a complicated task requiring the use of complex models to compare related or unrelated information with known truthful information [6]. Furthermore, fake news is perceived in several ways by researchers, leading to multiple ways of addressing and solving this issue. Some terms related to misinformation are used interchangeably in multiple cases. These terms include fake news, rumors, spam, and disinformation which usually contain numerical, categorical, textual,

The associate editor coordinating the review of this manuscript and approving it for publication was Mohamad Afendee Mohamed.

and image contents [7], [8], [9]. Unfortunately, many people have the urge to spread false information on social media, backed with professionally written, long, and referenced comments that allow the reader to more easily agree with the misinformation provided (e.g., [10], [11]). Researchers aim to eliminate the increased spread of misinformation by detecting the varied manners in which misinformation can be spread. As such, researchers have resorted to the use of deep learning (DL) algorithms to detect fake news before it spreads (e.g., [12]). This is accomplished by collecting or creating a dataset containing both true and false information within articles. Then, a pattern is determined, creating a model that can predict whether a given article contains true or false information.

There are noticeable gaps in the existing studies on fake news detection that our research highlights. This includes (i) a lack of clear distinction between the definitions of misinformation, disinformation, and false information; (ii) a lack of DL-based systematic reviews on varying types of misinformation problems; (iii) a lack of generalizable DL models that allow achieving a base acceptable detection accuracy on different datasets, which introduces the scarce use of transfer learning in this context; and (iv) a lack of models that deal with different levels of imbalance datasets in a fake news detection environment.

As technology progresses, the ability to detect misinformation becomes more complicated and thus more difficult to detect using standard machine learning (ML) techniques. This motivates our focus on DL techniques for the problem of fake news detection.

In this systematic literature review (SLR), we investigate existing fake news detection (FND) strategies that use deep learning. We focus on publicly available datasets used in FND and their NLP approaches. We aim to gather information about the transfer learning techniques applied and the methods used for addressing class imbalance, to examine their effect on detection accuracy. Our survey aims to identify open issues and research gaps in current studies. To the best of our knowledge, we are the first to provide a comprehensive SLR that investigates the effects of transfer learning and class imbalance treatment in the fake news detection domain.

Key Contributions:

The main contributions of this paper are as follows:

- We provide a detailed discussion of the main deep learning-based algorithms used to detect fake news, including their effectiveness.
- We discuss the main datasets available for fake news detection as well as their respective characteristics, advantages, and disadvantages.
- We study transfer learning techniques and strategies for dealing with class imbalance in this application domain. We also investigate their effects on the detection of fake news and the challenges associated with implementing these strategies.

Paper Organization:

This paper is organized as follows. Section II, presents the research methodology, including the search strategy, research questions, source databases, search query, inclusion and extraction criteria, and data collection summary. In Section III, we investigate the deep learning (DL) algorithms used for detecting fake news. Section IV describes the publicly available datasets in the fake news domain and the associated challenges. SectionV, discusses transfer learning strategies and open challenges in the FND context. Section VI analyzes the class imbalance problem in fake news detection. Section VII provides a summary of the data collected in this SLR and answers to our research questions. Section VIII addresses the research threats to validity, and Section IX discusses the main gaps and open issues that still exist in fake news detection. Lastly, Section X concludes our paper.

II. RESEARCH METHODOLOGY

A. SEARCH STRATEGY OVERVIEW

Our SLR is generated based on a set of detailed steps described in [13]. We begin by defining our research questions, after which we build the keywords for the search query to obtain the relevant papers for our study. Then, we select the most relevant databases to query and establish the inclusion and exclusion criteria. Finally, we define the fields to be extracted from the retrieved documents.

B. RESEARCH QUESTIONS

The key focus of our SLR is on understanding how the DL techniques have been used to address the FND problem. We are also interested in how TL has been applied in this field and how the class imbalance problem has been tackled.

- **RQ1:** Which deep learning algorithms have been used for fake news detection throughout time?
- **RQ2:** Which datasets are used in the fake news detection domain?
- **RQ3:** How effective are deep learning methods for fake news detection?
- **RQ4:** Which solutions incorporate transfer learning mechanisms, if any?
- **RQ5:** Which solutions deal with different levels of imbalanced datasets (if any)?

C. SOURCE DATABASES AND SEARCH QUERY

For the purpose of collecting research articles, we selected four digital databases that are renowned for their comprehensive coverage and relevance to our field of study. These databases include:

- Google Scholar (we selected the articles that appeared in the first thirteen retrieved pages);
- Association for Computing Machinery (ACM) Digital Library database;
- IEEE Xplore database; and
- Scopus.

Based on the research questions established in Section II-B, we collected a set of precise concepts that can cover the topic we are studying. We, therefore, formulated the search query as follows:

((fake ***OR*** *misinformation* ***OR*** *false* ***OR*** *unverified* ***OR*** *inaccurate* ***OR*** *rumor** ***OR*** *misleading)*
AND
(information ***OR*** *news* ***OR*** *article** ***OR*** *media)*
AND
*(detect** ***OR*** *classification)*
AND
("deep learning" ***OR*** *"machine learning"* ***OR*** *"neural"* ***OR*** *"artificial intelligence"))*

FIGURE 1. Search query used in our SLR.

The above search statement addresses the research questions by focusing on the 4 key concepts in the studied topic: "fake", "information", "detect", and "deep learning".

We searched both the title and the abstract for articles published between January 2018 and December 2023 inclusive. Limiting the search on this date range is motivated by the fact that FND has become more popular throughout the last years, especially during the COVID-19 pandemic that started around the beginning of 2020.

We defined the following set of restrictions on the results that limit the selection among the returned articles.

- The selected articles must be published in peer-reviewed journals or conferences. Thus, we excluded patents and any articles that did not conform to this condition.
- The language of the surveyed papers must be English. Any papers retrieved that were not written in English were excluded.
- Articles containing classification models that do not mention the performance evaluation of the methods (e.g., accuracy, precision, recall, F1-score, etc.) were excluded.
- We excluded the articles that have not mentioned the classifier/model used in the detection task in their methodology.
- We excluded the articles that only applied standard ML algorithms instead of DL ones.
- We excluded older articles when extensions and more recent editions were found.
- We excluded articles that were published in domains outside of Computer Science such as art, business, or other domains.

The total number of articles obtained from the search query, adhering to the extraction, duplicate removal, inclusion, and exclusion criteria, is 176. This includes 88 journal articles and 88 conference articles. The process of obtaining the articles is detailed in Section II-F.

D. INCLUSION CRITERIA

We considered the following inclusion criteria for our systematic literature review:

- Peer-reviewed journals and conference articles retrieved from the search query defined in Figure 1.
- Articles from the Computer Science domain.
- Research articles that focus on detecting or classifying fake news.

We applied the backward snowballing technique [14] to gather relevant articles that might have been missed in our search by inspecting the reference sections of the retrieved papers. We identified two articles that were not picked up through our search query and were added to the set of manuscripts to analyze.

E. DATA EXTRACTION

We used Covidence [15], a special web-based software for supporting the data aggregation and extraction of SLRs. The extracted data was organized in a spreadsheet that was exported from Covidence. The data that was extracted from the retrieved and selected articles is the following:

- Date: date of the publication;
- Publication Type: where the article has been published (conference/journal).
- Classifier/Model: algorithms used for FND in the article.
- Network Structure: the architecture of the network (details including the number/types of layers and any special setup in the network.
- Dataset: name of the fake news corpus or dataset(s) used.
- TL Techniques: the TL mechanism(s) used in the proposed solution.
- Imbalance Techniques: shows whether the imbalanced issue was treated in the proposed solution and how they dealt with it.
- Effectiveness: depicts the performance of the model in terms of accuracy, precision, recall, F-measure, and other evaluation metrics.

F. DATA COLLECTION SUMMARY

Overall, our search query retrieved 1642 articles. We found 436 duplicate articles that were removed. After the first screening of the titles and abstracts, we ended up with 393 research papers, matching our research keywords and the inclusion and exclusion criteria. After a second full text screening, we excluded 217 articles obtaining 176 research papers for analysis. Figure 2 shows the PRISMA chart demonstrating the retrieved papers' selection strategy

III. DEEP LEARNING ALGORITHMS USED FOR FAKE NEWS DETECTION

The thorough examination of various models and techniques pointed out the significant role that DL plays in different classification tasks including detecting fake news. Building and improving such algorithms became a pressing necessity, especially during the COVID-19 pandemic when a large volume of fake news and rumours were being disseminated widely. Figure 3 demonstrates a clear increase in the use of DL models over the years.

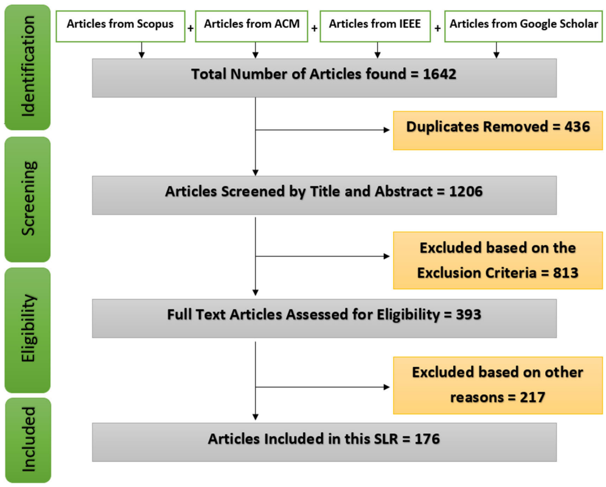

FIGURE 2. PRISMA Chart of selecting and retrieving the articles.

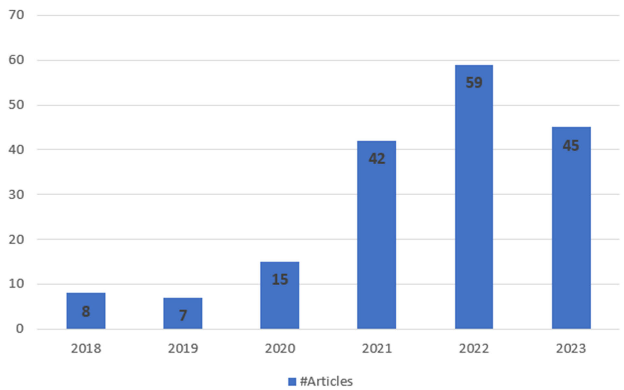

FIGURE 3. DL Models used for FND between the years of 2018 and 2023.

The extracted data shows that the FND task usually follows a generic framework as is shown in Figure 4. Initially, the process involved acquiring or generating a dataset. The majority of studies have utilized news articles that were gathered from openly accessible datasets. After collecting the dataset, preprocessing techniques were employed to prepare the data for input into a neural network. Prior investigations have mainly employed Word2vec and GloVe word embedding methods to transform words into vectors [16]. Finally, the neural network model is trained and the predictions are obtained.

Neural networks for FND can be categorized into different types based on their architecture and how they process data. The first type is feedforward neural networks, including single-layer and multi-layer perceptrons. Convolutional neural networks (CNNs) are another type, which are designed to process data with a grid-like topology, such as images. They include traditional convolutional neural networks, residual networks, and dense networks.

Recurrent neural networks (RNNs) are designed to handle sequential data, such as time series or language, and include basic recurrent neural networks and bi-directional RNNs, long short-term memory networks (LSTM), gated recurrent units (GRUs) and bi-directional GRUs, and bi-directional

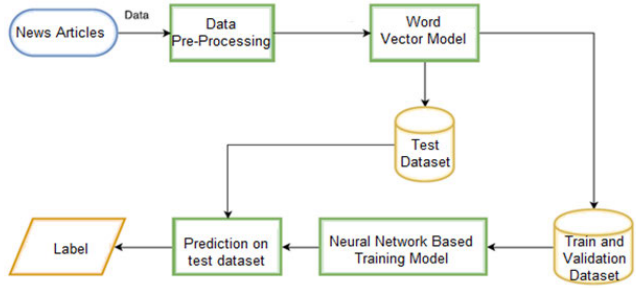

FIGURE 4. The General DL Framework that Used for FND.

long short-term memory networks (BiLSTM). we will refer to the model and its bi-directional version collectively as (Bi)X, where X is the model. Graph neural networks (GNNs), a newer type of neural network, are designed to operate on graph-structured data, such as social networks, chemical molecules, or protein structures.

Recently, attention-based models have gained popularity due to their ability to focus on certain parts of the input data selectively. They include self-attention networks and multi-head attention networks. Hybrid models, which combine different types of neural networks, have also become popular. For example, convolutional recurrent neural networks combine the spatial processing capabilities of convolutional neural networks with the temporal modeling capabilities of recurrent neural networks. Transformer networks, like BERT, combine self-attention mechanisms with feedforward neural networks to process sequences of data. Figure 5 shows a taxonomy of the various types of neural networks that are used for FND.

Based on the data that we gathered from the surveyed articles, it is evident that researchers extensively explored several DL algorithms for the detection task. Figure 6 shows the usage of different DL detection models for fake news. More precisely, this figure displays the percentage of papers where a particular model was used. We observe that the (Bi)LSTM was the most frequently included model used in 72% of articles and the CNN model was the second most used model utilized in 61% of the articles reviewed. The third architecture used is the hybrid architecture, which combines different types of neural networks in the detection process. Since multiple models may be used in the same research paper, summing up the percentages in Figure 6 exceeds a total of 100%. The following sections provide a detailed discussion of the main architecture used for FND.

A. ARCHITECTURES BASED ON CONVOLUTIONAL NEURAL NETWORKS

Our findings also show that 61% of the previous works used Convolutional Neural Networks (CNNs) to handle the detection issues, attempting to boost the performance of the FND process through the use of this DL algorithm [12], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50],

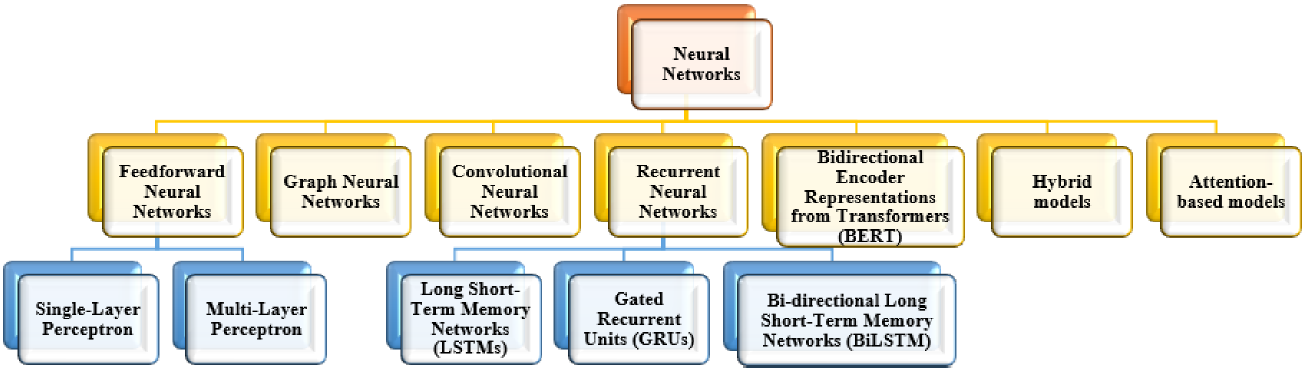

FIGURE 5. Taxonomy of the main neural network categories used for FND.

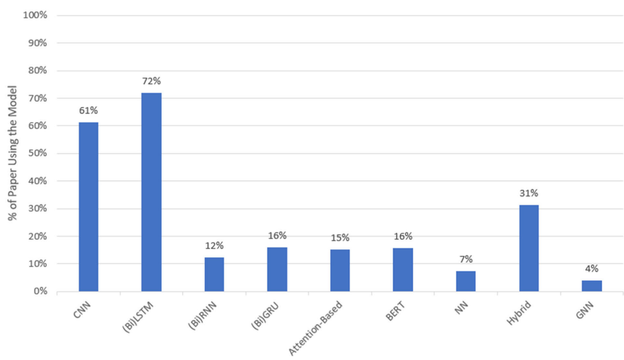

FIGURE 6. Deep Learning Models for FND.

[51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86], [87], [88], [89], [90], [91], [92], [93], [94], [95], [96], [97], [98], [99], [100], [101], [102], [103], [104], [105].

The detection effectiveness is the result of CNN's ability to carry out feature extraction [106]. It is worth noting here that CNNs were trained on different fake news datasets. The CNN achieved notable effectiveness, with accuracy ranging between 95% and 98%, depending on whether it was used individually [78] or in conjunction with another model, such as the Gated Recurrent Unit (GRU) [41], respectively. It is also worth mentioning that the CNN has fallen in some cases to about 47% detection accuracy [30] which leads to the conclusion that some key points may affect the effectiveness of the CNNs in FND tasks. The first point is the degree of the deepness of the network being used. A deeper CNN is considered an advantage for solving the overfitting issue [62]. This is what we discovered using the data collected which contains a case of building a deeper CNN, called FNDNet, which solved the overfitting problem by learning the discriminatory features for FND using multiple hidden layers [12]. The second point affecting the CNNs' effectiveness concerns the selected dataset that will be used in the detection task and its readability and cleanness before being fed into the model [28]. Finally, the overall architecture that will be used for the detection task may adopt either the CNN itself or a CNN in a hybrid approach as we can see in our extracted results [41].

Figure 7 illustrates an example of a CNN architecture used for FND as proposed in [100]. The CNN architecture used in this study is composed of an input layer, an embedding layer, and three sets of convolutional and max pooling layers. The input layer resizes the input data to a uniform size of 1000, while the embedding layer reduces the size to 100 by embedding the data. The convolutional and max pooling pairs extract features from the input. To perform this task, filters are applied to each convolutional layer, each of which consists of 128 filters with a kernel size of 5 and a ReLU activation function. Additionally, the fully connected network includes both a flat and a dense layer. Lastly, the feature maps are classified using a dense layer with a softmax activation function.

B. ARCHITECTURES BASED ON RECURRENT NEURAL NETWORKS

Another popular FND algorithm examined in the previous studies is the Recurrent Neural Network (RNN) and its variations. Authors have investigated various RNN models to detect fake news in sequential data. They have proposed Long Short-Term Memory (LSTM), GRU, unidirectional LSTM-RNN, vanilla RNN, and Bi-directional LSTM ((Bi)LSTM).

Our findings show that researchers' focus is highly shifted toward RNNs and their variations in fake news detection. Figure 8 shows the utilization of RNNs in the previous studies.

It is noticeable from our findings that researchers examined FND using classic RNNs in only 12% of the total number of the surveyed articles [16], [31], [46], [50], [59], [67], [72], [85], [89], [95], [98], [107], [108], [109], [110], [111], [112], [113], [114]. Despite the importance of the RNN in such domains, research authors discussed the RNN vanishing problem [115]. One solution to solve vanishing in RNNs is to use other architectures such as LSTM and (Bi)LSTM. The percentage of the articles that examined both LSTM and

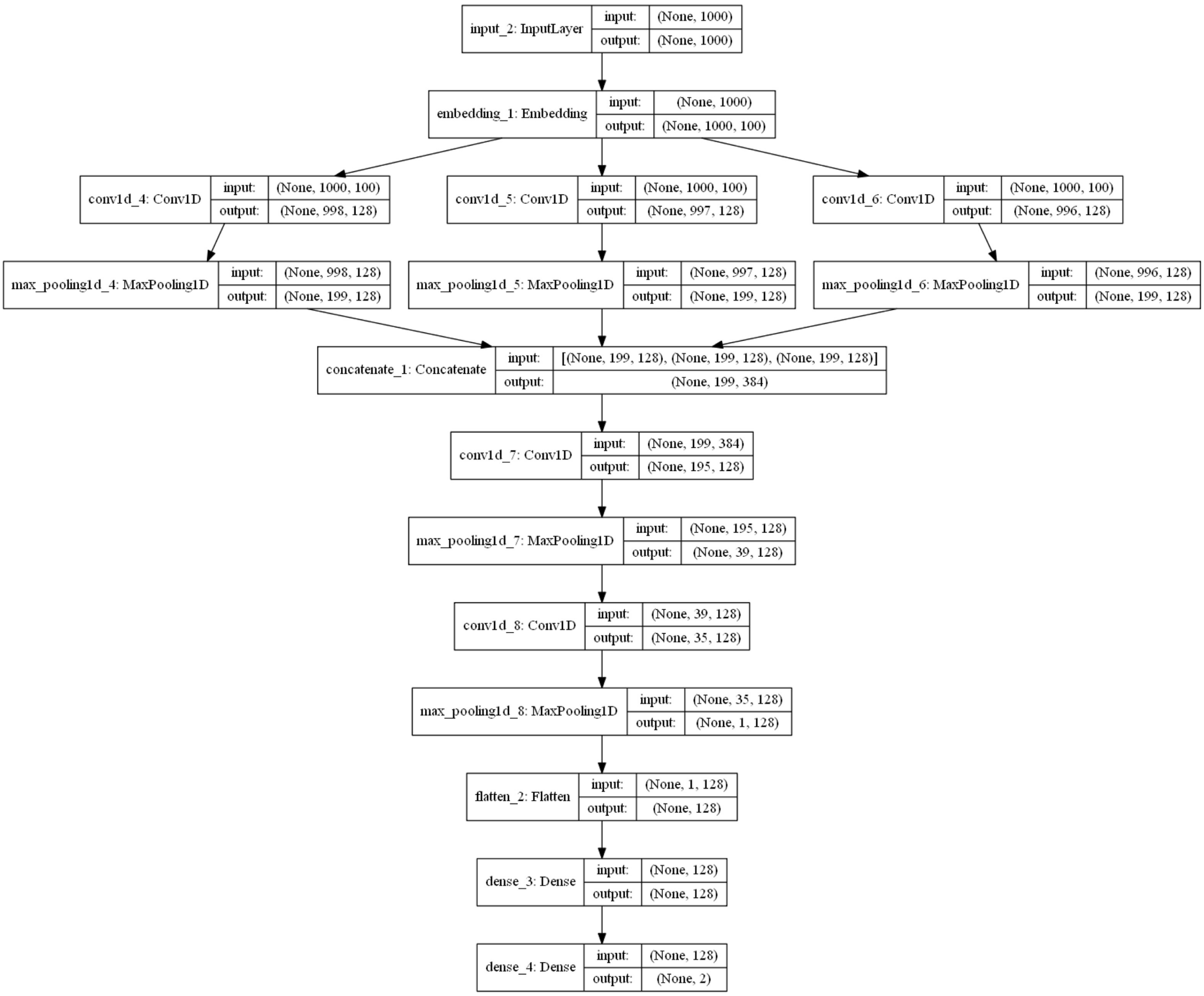

FIGURE 7. An example of CNN architecture used in FND.

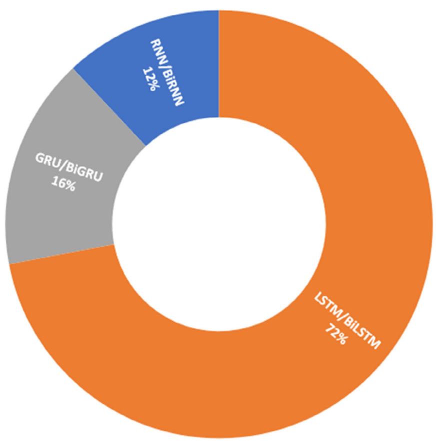

FIGURE 8. RNNs utilization in FND.

(Bi)LSTM was around 72% of the total articles [16], [17], [18], [19], [20], [21], [23], [24], [25], [26], [27], [29], [30], [31], [32], [33], [34], [35], [36], [43], [44], [49], [50], [52], [53], [56], [57], [59], [60], [62], [63], [64], [67], [69], [70], [72], [73], [76], [77], [79], [80], [82], [83], [85], [86], [87], [89], [90], [91], [95], [96], [99], [100], [102], [103], [104], [107], [108], [109], [110], [111], [112], [116], [117], [118], [119], [120], [121], [122], [123], [124], [125], [126], [127], [128], [129], [130], [131], [132], [133], [134], [135], [136], [137], [138], [139], [140], [141], [142], [143], [144], [145], [146], [147], [148], [149], [150], [151], [152], [153], [154], [155], [156], [157], [158], [159].

Other solutions were also adopted in the previous studies which include using (Bi)GRU as a detection architecture. (Bi)GRU has been examined in 16% of the total surveyed articles [16], [24], [25], [31], [41], [56], [72], [75], [79], [84], [89], [93], [100], [109], [119], [121], [136], [143], [153], [160], [161], [162], [163].

Figure 9 shows the RNN GRU-based architecture for FND that was presented in [100]. In this proposed solution, the use of GRU RNNs for FND is explored. The model proposed includes an input layer and an embedding layer with data sizes of 1000 and 100, respectively. The GRU layer is then implemented with identical hyperparameters as the LSTM layer to facilitate a reliable comparison between the two. Finally, fully connected networks are used, along with a batch normalization layer, and a dense layer with a softmax activation function is applied for classification.

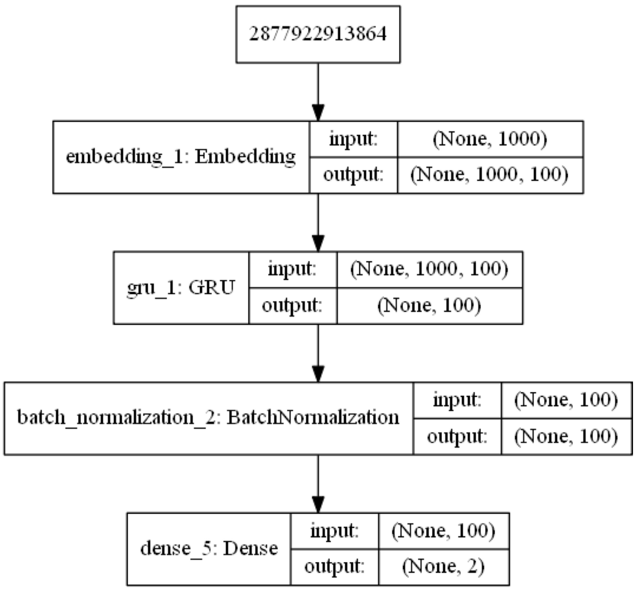

FIGURE 9. An example of GRU architecture used in FND.

The findings from our survey also show that RNNs and their variations had a remarkable detection accuracy in the fake news domain when compared against other detection models and taking into consideration the usage of different datasets. The RNN detection effectiveness ranged from 48% [50] to around 92% [109] to around 99% [96] detection accuracy.

Using another architecture with the RNN does not seem to increase the accuracy of the detection results as shown in the works of Ilie et al. [31] and Nasir et al. [46]. It is also noticeable in our findings that the GRU model had also participated in detecting fake news with an accuracy ranging between about 76% [161] and 97% [24]. These satisfactory results are not the case when using the BiGRU instead of the standard GRU architecture. In the latter case, the detection accuracy decreases to a range from 28% and 71% [136]. Finally, it was clear that the detection was more accurate when it was done by a second model besides GRU in a hybrid mechanism. This was obvious when the researchers used the GRU with a CNN in [41] and [161] and when a GRU was used with (Bi)LSTM in [119].

Despite the effectiveness of the above-mentioned architectures in fake news detection, previous studies showed that the LSTM and the (Bi)LSTM are the future key players in enhancing fake news detection. The average accuracy of detecting fake news using LSTM architectures was ranging

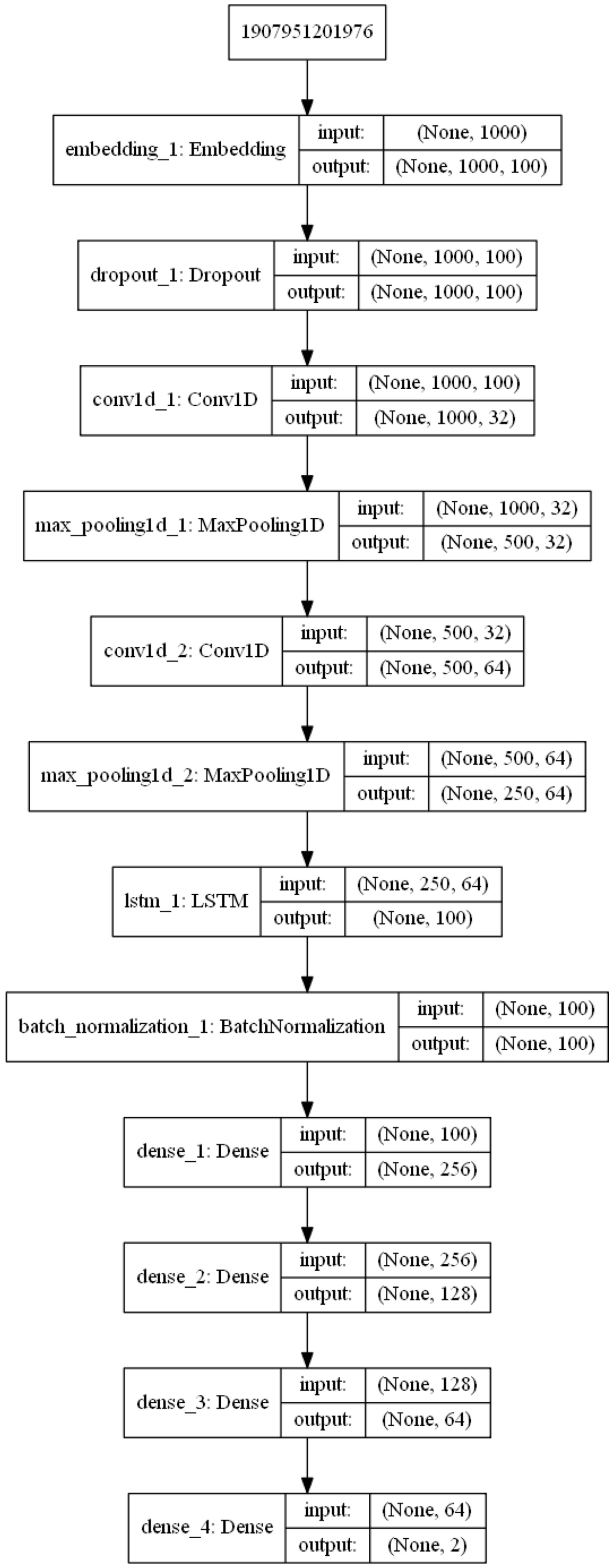

FIGURE 10. An example of CLSTM architecture used in FND.

between 79.03% and 81.21%. These models also reported a maximum of 99.9% detection accuracy in [25] and a minimum of 11% in [43].

In addition, the findings show that LSTM was used in a hybrid fashion with one or more architectures to determine the optimal FND system among the proposed systems. Figure 10 shows the architecture of the CNNs-LSTM model proposed in [100]. This model utilizes both hybrid and recurrent models on collected news data. The proposed hybrid model incorporates both CNNs and LSTM models. The algorithm includes an input layer that resizes the input data frames to 1000 and an embedding layer that embeds the input tensor size from 1000 to 100. The embedded tensors are then processed through two sets of convolutional and max-pooling layers for feature extraction. The convolutional layers have 32 and 64 filters, respectively, and a kernel size of 3. The feature extraction process is then performed by the LSTM layer with 100 units, a dropout rate of 0.2, and a recurrent dropout rate of 0.2. Additionally, the fully connected network is designed with a batch normalization layer, followed by three dense layers with a ReLU activation function and several filters of 256, 128, and 64, respectively. The classification task is carried out using a dense layer with a softmax activation function.

Researchers focus more on testing the effects of developing hybrid models that adopt LSTM in the detection process. They tested the importance of (Bi)LSTMs over LSTM and reported that the (Bi)LSTM+CNN achieved considerably higher accuracy than when they attempted to use the LSTM with the CNNs. They reported a detection accuracy of about 99% detection accuracy when they attempted to use the (Bi)LSTM instead of the LSTM [96].

When LSTM is combined with CNN, studies also reported an accuracy ranging between 97.8% in [129] and 47.06% in [30], with an average accuracy of 82.3%. In the case where LSTM is combined with a DNN architecture, we observed an accuracy of 91.16% [133], while when it is combined with BERT the accuracy achieved was 84.10% [134].

Bi(LSTM) is also getting popular in fake news detection as our survey findings show. It recorded the highest detection accuracy of 99.52% [126] and the lowest of 28% [136] with an average of 75.22%. It also appeared connected to other detection architectures such as CNN and GRU. Bi(LSTM) with CNN recorded the highest accuracy of 98.65% in [132] and the lowest accuracy of 35.13% in [56]. The average detection accuracy in such cases was about 77.6%. Bi(LSTM) with GRU reached 89.8% detection accuracy [119].

C. ARCHITECTURES BASED ON GRAPH NEURAL NETWORKS

Another popular model in fake news detection is the Graph Neural Network (GNN) and its variants such as Sequence Graph Transform (SGT) [164], Graph Attention Networks (GAT) [165], GraphSAGE [166], and Graph Convolutional Networks(GCN) [167]. GNN is a neural network that directly operates on the graph structure. One of its popular applications is node classification in which every node in the network has a label. This network predicts the label of the node without the ground truth [168]. In FND using GNNs, news articles, and related information are represented as a graph. The nodes of the graph represent the individual entities, such as news articles, users, or social media posts, and the edges represent the relationships or interactions between them. To create the graph, the news articles are typically preprocessed to extract features such as the article content, metadata, and social media interactions. These features are used to construct the nodes and edges of the graph, with the nodes representing the articles and the edges representing the relationships between articles, users, or other entities. For example, edges could represent similarities between articles or social media interactions such as retweets or mentions. Once the graph is constructed, Graph Neural Networks are used to analyze the graph structure and extract useful features for fake news detection. The GNNs use graph algorithms to propagate information across the graph and learn representations of the nodes and edges that capture their relationships and interactions. These learned representations can then be used to classify the news articles as fake or real based on their similarity to other articles and the overall structure of the graph.

Our findings show that only 4% of the selected research articles adopted GNN architectures for fake news detection [93], [169], [170], [171], [172], [173], [174], [175]. The claimed detection accuracy was incredibly low compared to the other deep-learning models on different datasets used. The highest detection accuracy obtained when using the GraphSAGE was 89.7% accuracy without mentioning whether this was on the training or the testing dataset [171]. The accuracy went deeply down to 61.5% when they adopted the GNN. It also recorded a 73.12% [169] with GCN with a maximum of 88.6% [171]. The other variants such as SGT, GCN, and GAT had reached an average accuracy of about 83.1%.

D. ATTENTION-BASED AND BERT-BASED ARCHITECTURES

Another notable advancement happened in fake news detection with the use of attention-based approaches using different datasets. Our findings show that their use has been increasing since the year of 2018 and has reached the maximum in the year of 2022. In addition, this approach appeared in 15% of the surveyed articles mostly in the year 2022. Authors have applied it to the other detection models including RNNs [31], GRU [31], [75], [160], [163], [176], LSTM, and (Bi)LSTM [19], [49], [57], [77], [123], [132], [140], [154], [174], [177], BERT [57], and CNN alone [49], [77] or with other models [19], [49], [140]. The detection accuracy ranged between 54% [49] and 98.65% [132].

Another deep learning model present in our surveyed works that shows cutting-edge detection is the BERT [178] model. It is a sophisticated pre-trained word-embedding model built on a transformer-encoded architecture. The findings show that 16% of the surveyed studies adopted the BERT as a detection mechanism [43], [51], [57], [66], [74], [79], [81], [83], [88], [89], [91], [99], [110], [111], [134], [144], [156], [157], [179], [180], [181], [182], [183], [184].

The findings also show that authors started using the BERT as a detection model for fake news in 2021 which makes it still a novel tool for the detection model and a future direction in the fake news detection field. Our findings show that this model has reached a remarkable detection accuracy with the highest recorded accuracy of 98.5% [181] and an average accuracy of around 90%. It is also clear in our findings that researchers experimented with the effectiveness of applying the BERT with other models such as LSTM [134] and CNN [51] for the detection of fake news using different datasets. An example of using BERT in the fake news detection process, FakeBERT has been proposed in [185] which outperforms all other models with an accuracy of 98.9%. Figure 11 illustrates the proposed FakeBERT.

As Figure 11 shows, this design employs three parallel blocks of 1D-CNN with 128 filters, with each block having one convolutional layer. The first layer has a kernel size of 3 and 128 filters, reducing the input embedding vector from 1000 to 998. The second layer has a kernel size of 4 and 128 filters, reducing the input vector from 1000 to 997. The third layer has a kernel size of 5 and 128 filters, decreasing the input vector from 1000 to 996. Max-pooling layers are also included after each convolutional layer to further reduce the dimension. A max-pooling layer with a kernel size of 5 reduces the vector to 1/5th of 996, which is 199. After concatenating the three convolutional layers, another convolution layer with a kernel size of 5 and 128 filters is applied. This is followed by two hidden layers with 384 and 128 nodes respectively. The number of trainable parameters for each layer is also provided in the "Param number" column for further details.

A recent study has conducted a thorough comparison between different deep learning models in fake news detection using various datasets [186]. The authors studied the effect of deploying (Bi)LSTM, CNN-RNN, C-LSTM, CNN, and BERT in the detection of fake news. They used seven fake news detection datasets with each model to be able to draw a generalized conclusion. They figured out that the (Bi)LSTM and BERT detection models achieved the best detection accuracies and F-scores. The authors have also concluded that BERT performs better than the (Bi)LSTM when the model aims at detecting fake news in different contexts from the one it was trained on [186].

E. ARCHITECTURE BASED ON FEEDFORWARD NEURAL NETWORKS

Finally, other deep learning models have been used in the fake news detection field with basic and standard feedforward neural network (FFN) settings. Authors categorized these under simple neural networks (NN; ANN, DNN, and FNN). Although these models are referred to as simple detection techniques and were used in 7% of the total surveyed articles, they still reached a noticeable accuracy in detecting fake news. Our findings show that FFN has reached a detection accuracy of 89.8% [44] to about 95% when it was provided by solid support from a strong embedding technique [148] and a lower accuracy of 83.35% [107]. On the other hand, DNN reached an accuracy of 94.68% [187] while it was less accurate when applying it with an LSTM by 2.8% [133]. The findings also show that using a multichannel ANN [187] has increased the detection accuracy by approximately 13% of the basic ANN which was 80.9% accurate in detecting fake news [188].

In conclusion, we observe a clear growing trend in the solutions proposed using (Bi)LSTM, CNN, BERT, etc. throughout the years, as Figure 12 shows.

F. CHALLENGES RELATED TO DEEP LEARNING METHODS FOR FAKE NEWS DETECTION

Despite the promising results of deep learning methods for fake news detection, several challenges remain to be addressed. These include issues related to dataset quality, model performance on imbalanced datasets, and the generalizability of models across different datasets. In this article, we will explore these challenges in more detail and discuss potential solutions.

It is important to recognize that there are several challenges when it comes to achieving effective fake news detection using deep learning methods. One major issue is the potential for overfitting, where models achieve high accuracy on the training data but perform poorly on new, unseen data [60]. Some previous research has reported extremely high accuracy results, but these were obtained by evaluating the model on the same data that was used for training. The performance of their models achieved a high accuracy of 99.9% [25], [53], [60], [120], [124], [126], [131], and [132]. This raises questions about the model's ability to generalize to new data. Another challenge is the use of accuracy as the sole evaluation measure for imbalanced datasets, where the number of fake news samples vastly outweighs the number of real news samples [17], [18], [19], [20], [36], [57], [60], [61], [62], [117], [120], [123], [131], [137], [138], [139]. Accuracy can be misleading in these cases, as it can be skewed by the dominance of the majority class. A more appropriate measure, such as precision or recall, would provide a better understanding of the model's performance. Additionally, different datasets can have varying characteristics and biases, and models that perform well on one dataset may not generalize to other datasets. This is noted from our findings in [43], [50], [125], [135], and [136]. This was also proved by the thorough experiments that were made in a recent study of cross-domain fake news detection [186]. Finally, the quality and diversity of the training data can greatly impact the performance of the model [189], [190]. In some cases, models have been trained on datasets that are not representative of the full range of fake news content, leading to poor detection performance [190], [191].

IV. DATASETS USED FOR FAKE NEWS DETECTION

In this section, we first discuss the main characteristics of the datasets used in the surveyed works. Then, we discuss

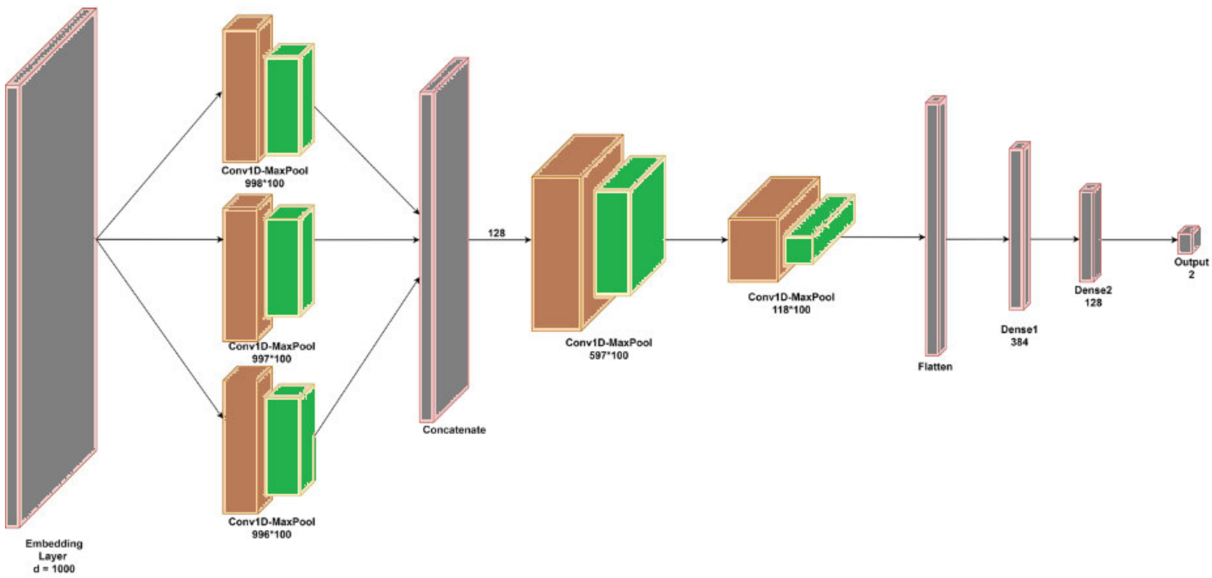

FIGURE 11. An example of BERT architecture used in FND.

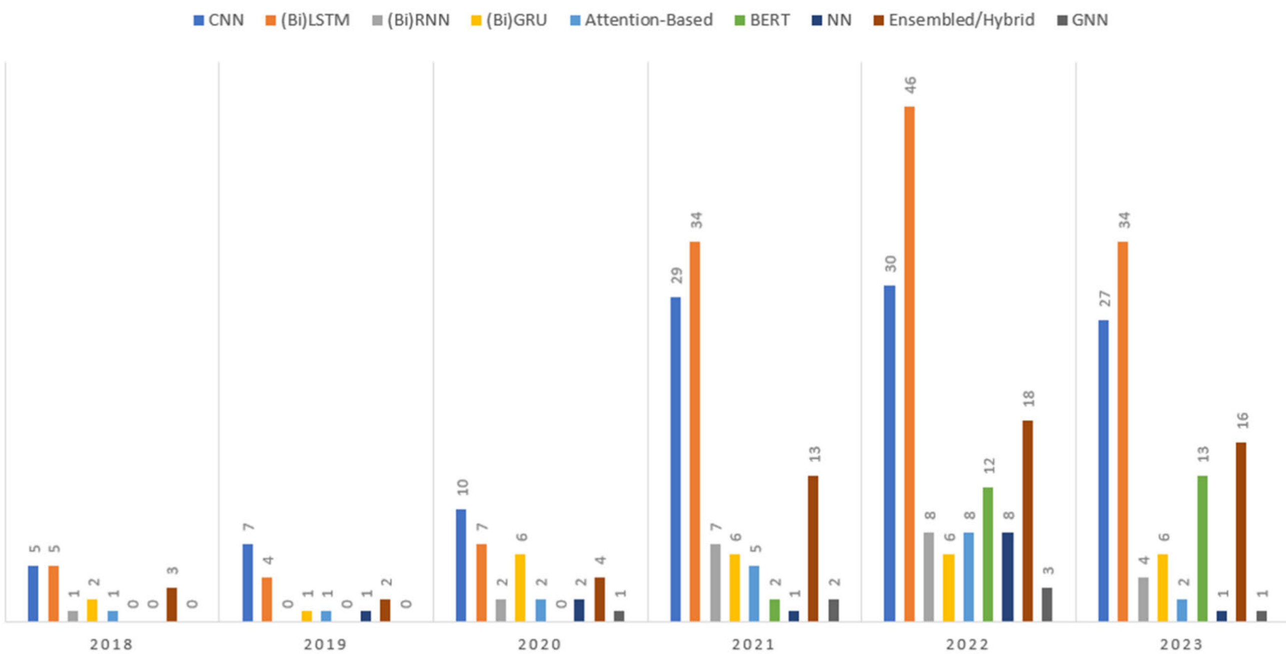

FIGURE 12. DL models in fake news detection throughout the time.

some of the open challenges related to the datasets in this application domain.

A. MAIN DATASETS USED FOR FAKE NEWS DETECTION

Researchers have used several datasets in the context of fake news detection. However, we found that only a small part of these datasets is publicly available, while a considerable percentage is created by the researchers and/or is not disclosed publicly. A pie chart of the used datasets in the surveyed studies is presented in Figure 13.

We observe that ISOT [192], PHEME [193], Liar [194], and FakeNewsNet [195], with its three sub-datasets, GossipCop, PolitiFact, and BuzzFeedNews, are examples of publicly available fake news datasets. These are among the most popular and frequently used datasets.

The LIAR dataset includes short statements obtained from the Politifact fact-checking website. This dataset includes a total of 12.8 K labelled short statements. The annotation task has been done by the Politifact site, and the statements are classified into 6 classes: pants-fire, false, barely-true,

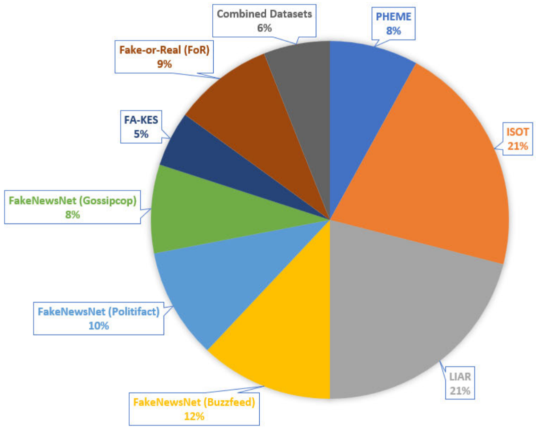

FIGURE 13. Datasets used in the surveyed studies.

half-true, mostly-true, and true. In addition, another fake news dataset was collected from real-world news articles called ISOT. The real news cases were collected by crawling news articles from Reuters.com, and the fake news examples were obtained from unreliable websites, which were annotated by the Politifact website. The PHEME dataset was collected from Twitter based on 9 newsworthy events classified by journalists. The annotation process was conducted by journalists (human annotators) and each tweet was annotated with one of the following labels: "proven to be false", "confirmed as true" or "unverified". The FakeNewsNet dataset consists of three subdatasets, which are GossipCop, PolitiFact, and BuzzFeedNews. In total, the FakeNewsNet dataset contains approximately 19,838 news articles labelled as either "fake" or "real". The news articles in the FakeNewsNet dataset were annotated by a team of human annotators. The annotators were given guidelines for identifying fake news and were trained to identify various characteristics of fake news, such as misleading headlines, fabricated content, and misleading images. Table 1 summarizes the main characteristics related to these datasets.

TABLE 1. Main characteristics of the publicly available FND datasets.

Dataset	Size	Num. of Labels	Type
LIAR	12.8K	6	Political Statements
PHEME	5800	2	Social media (tweets)
ISOT	45k	2	News articles
BuzzFeedNews	5,835k	2	News articles
PolitiFact	12,835k	2	News articles
GossipCop	1168k	2	News articles

From our findings, LIAR achieved a maximum of 98.95% when a Bi(LSTM) model was used for training [125]. The same dataset was an option for training the (Bi)GRU in [136] which recorded a low detection accuracy of 28.12%. ISOT recorded high detection effectiveness in many cases, especially when the trained model was a Bi(LSTM) [53] with an accuracy of 99.95%. Still, the accuracy decreased when the models used for training were RNN and CNN, with the lowest performance recorded at 82.5% [46]. PHEME also exhibited high performance when the CLSTM model was used, achieving an accuracy of 91.88% and recording a minimum accuracy of 65.5% with the training of a CNN [55]. Lastly, FakeNewsNet sub-datasets used in training different models such as CNNs [50], various RNNs [50], [110], [111], [134], [136], GNNs [170], and BERT [110], [111], [134], [180]. The best detection accuracy that achieved when training the GAT with a 96.42% accuracy while it recorded a 71.16% accuracy when it was used to train a (Bi)GRU.

B. CHALLENGES RELATED TO THE DATASETS USED FOR FAKE NEWS DETECTION

One of the main difficulties in the fake news detection field is the scarcity of labelled cases [196], [197]. Even though multiple datasets with a massive amount of records exist they are mostly unlabeled or have only a few records labeled. Researchers have collected datasets over the last few years for use with DL models in different contexts associated with fake news detection. Datasets are massively diverse from one another due to having different research goals inside the fake news detection application domain [198]. For example, some datasets contain exclusively political statements, while other datasets only include news articles or social media posts [186].

To collect appropriate datasets to serve in fake news detection, we need fake articles and non-fake articles. Fake articles are gathered from deceitful websites that are designed on purpose to disseminate misinformation and fake news. The fake news published on these websites will eventually be shared on social media to be read and circulated by innocent people who do not check the news source.

It is also clear from our findings that the datasets used in fake news detection are insufficient for training models due to their characteristics, such as language features or size [199]. That leads us to the question of creating a dataset to serve as a benchmark in the detection process. However, this can be challenging due to several reasons, some of which are:

- Sources of fake and non-fake news: Identifying reliable sources of fake and non-fake news can be difficult, especially in today's world where there are numerous sources of information and not all of them are trustworthy [200]. It is crucial to ensure that the dataset contains a diverse range of sources to ensure that the model is trained to detect fake news from a variety of sources.
- Bias in the data: Bias can be introduced in the data due to various reasons such as the sources of the data, the labelling process, or the selection criteria for the dataset. Bias can affect the accuracy of the model and can also lead to unfair predictions [201].
- Labeling issues: Labeling data for fake news detection can be challenging, as there can be discrepancies in the definition of what constitutes fake news. Human labels

may be subjective, and there may be inconsistencies in the labelling process. Automatic labels generated using machine learning techniques can also have their limitations [197], [202].
- Bots involvement: Bots can be used to generate large volumes of fake news and spread it rapidly across the internet, making it difficult to detect and remove. Bots can also be used to manipulate the labelling process by providing biased labels, leading to inaccuracies in the dataset [200].
- Rapid evolving nature of fake news: The nature of fake news is constantly evolving, and new techniques for creating and spreading it are being developed all the time. This makes it difficult to create a comprehensive dataset and up-to-date [198], [203].

To address these issues, it is crucial to have a well-designed and diverse dataset that is regularly updated to reflect the changing nature of fake news. It is also important to have robust labelling procedures in place, using a combination of human and machine labels, to ensure that the dataset is unbiased and accurate. Additionally, researchers should consider incorporating techniques such as adversarial training to improve the robustness of the model to adversarial attacks.

V. STRATEGIES FOR TRANSFER LEARNING

A. TRANSFER LEARNING STRATEGIES APPLIED TO FAKE NEWS DETECTION

Numerous real-world applications have made use of the machine and deep learning techniques. These learning methodologies assume that the input feature space and data distribution properties are maintained across the experiments carried out because the training data and testing data are drawn from the same domain [204]. This assumption, however, may not be accurate in some real-world machine-learning situations. In fact, in some circumstances, gathering training data can be costly and/or challenging. As a result, the research community has been considering the development of high-performance learners who are trained using data that could be more easily obtained from other various domains instead of the deployment domain.

Transfer learning is a technique used to advance a learner in one domain by transferring knowledge from a related domain. Real-world, non-technical experiences can help us comprehend why transfer learning is feasible. Take the case of two individuals who wish to learn how to play the piano. One person has no prior musical training, whereas the other plays the guitar and has a wealth of musical expertise. By applying previously acquired musical information to the goal of learning to play the piano, a person with a strong musical background will be able to learn the piano more quickly and effectively [205]. One can employ knowledge from a task they have already mastered to help them learn a new one that is related.

The essence and necessity of transfer learning appear when there is a dearth of target training data [204]. This can be the result of the data being rare, expensive to gather and label, or inaccessible. The use of other existing datasets that are related to, but not precisely the same as a given target domain of interest makes transfer learning solutions an alluring strategy since big data repositories become more widespread. Transfer learning has been successfully used in many machine and deep learning applications, including text sentiment classification [206], image classification [207], [208], [209], classification of human activity [210], classification of software defects [211], and classification of multi-language text [212].

Different techniques can be utilized in transfer learning to accomplish tasks as the following [213]:

- Training models in similar domains: This transfer learning method trains models that belong to similar domains. For instance, if there is insufficient data to complete task X, but task Y is similar and has adequate data, a model can be trained on task Y and then used to create a new model for task X [214].
- Feature extraction: Feature extraction is another transfer learning approach where deep neural networks are trained to extract features automatically. After training them on pre-existing models, the representations are exported to new models. This technique is commonly employed by data scientists [215].
- Utilizing pre-trained models: This approach involves developing pre-trained models that take transfer learning variables into account. Companies experienced in model development often have access to a library of models that can be used to create future models. This means that when dealing with a new problem, a pre-trained model can be selected, optimized for the problem at hand, and then reused to train another model [214].

The first transfer learning technique involves training models in similar domains by using a pre-trained model from a source domain that is similar or related to the target domain. The idea is that the knowledge learned from the source domain can be leveraged to improve model performance on the target domain, even if the target domain has limited labelled data.

Training models in similar domains typically involve the following steps:

1) Selecting a source domain: The source domain should be chosen based on its similarity or relevance to the target domain. Ideally, the source domain should have similar data distribution, task, or domain characteristics as the target domain, so the knowledge learned from the source domain can be effectively transferred to the target domain.
2) Acquiring or creating a labelled dataset in the source domain: A labelled dataset in the source domain is needed for training the pre-trained model. This dataset should be representative of the data in the source domain and should cover the task or tasks of interest.

3) Pre-training the model on the source domain: The pre-trained model is trained on the labelled dataset in the source domain. This involves training the model using standard machine learning or deep learning techniques, such as supervised learning or unsupervised learning, depending on the availability of labelled data in the source domain.
4) Fine-tuning or adapting the pre-trained model to the target domain: After pre-training on the source domain, the pre-trained model is fine-tuned or adapted to the target domain. This typically involves further training the model using the limited labelled data available in the target domain, while retaining the knowledge learned from the source domain. Fine-tuning can be done by updating the weights of some or all of the layers of the pre-trained model, depending on the specific task and data.
5) Evaluating and validating the model performance: The fine-tuned model is evaluated and validated on the target domain dataset to assess its performance. This may involve measuring metrics such as accuracy, precision, recall, F1 score, or other relevant performance indicators to determine the effectiveness of the transfer learning approach.

Transfer learning by training models in similar domains can be useful when the target domain has limited labelled data, but related or similar domains have abundant labelled data. By leveraging the knowledge learned from the related source domain, the model can benefit from the additional data and potentially achieve better performance on the target domain task. However, it is important to carefully consider the similarity and relevance between the source and target domains to ensure that the knowledge transfer is effective and results in improved performance.

For the second transfer learning technique, feature extraction is one of the common techniques used in transfer learning, where a pre-trained model is used to extract features from data in one domain and these features are then used to train a new model for a different task or domain [216].

In transfer learning with feature extraction, the pre-trained model is typically a deep neural network trained on a large dataset from a source domain. This model has learned to extract relevant features from the source domain data, which can be representations or embeddings of the input data at different layers of the network. These learned features are then used as inputs to a new model, often referred to as the target model, which is trained on the limited labelled data available in the target domain.

The process of using feature extraction in transfer learning typically involves the following steps:

1) Selecting a pre-trained model: The pre-trained model should be chosen based on its relevance to the target task or domain. Ideally, the pre-trained model should have been trained on a large dataset from a source domain that is similar or related to the target domain, so that the learned features are relevant to the target task.
2) Removing the last layers of the pre-trained model: The last layers of the pre-trained model, which are often responsible for task-specific predictions, are removed to retain the feature extraction capability of the model. These last layers are replaced with new layers that are specific to the target task.
3) Extracting features from the source data: The pre-trained model is used to extract features from the data in the source domain. This typically involves passing the data through the layers of the pre-trained model up to a certain layer and using the outputs of that layer as the learned features.
4) Training a new model on top of the extracted features: The extracted features are then used as inputs to a new model, which is trained on the limited labelled data available in the target domain. This new model, often referred to as the target model, is trained using standard machine learning or deep learning techniques, such as supervised learning or fine-tuning, depending on the availability of data in the target domain.
5) Evaluating and validating the target model performance: The trained target model is evaluated and validated on the target domain dataset to assess its performance. This may involve measuring metrics such as accuracy, precision, recall, F1 score, or other relevant performance indicators to determine the effectiveness of the transfer learning approach.

Feature extraction in transfer learning allows leveraging the knowledge learned from the source domain to extract relevant features from the data in the target domain, even if the target domain has limited labelled data. By using the learned features as inputs to a new model, the target model can potentially benefit from the representations or embeddings learned from the source domain. This can help improve the performance of the target model on the target domain task. However, it is important to carefully consider the similarity and relevance between the source and target domains to ensure that the features extracted from the source domain are relevant to the target task.

For the third transfer learning technique, utilizing pre-trained models is a common approach in transfer learning where a pre-trained model, typically trained on a large dataset, is used as a starting point for training a new model on a smaller target dataset. The idea is that the knowledge learned from the source domain can be transferred to the target domain, even if the two domains are different, to improve the performance of the target model [217].

Here are some key steps involved in utilizing pre-trained models for transfer learning:

1) Select a pre-trained model: Choose a pre-trained model that is trained on a large dataset and is relevant to your target task. For example, suppose you are working on an image classification task. In that case, you can

choose a pre-trained Convolutional Neural Network (CNN) such as VGG, ResNet, or Inception, which have been trained on large image datasets like ImageNet.

2) Remove or freeze some layers: Depending on the architecture of the pre-trained model, it may be necessary to remove or freeze some layers. For example, you can remove the output layer(s) of the pre-trained model and replace them with new layers that are suitable for your target task. Alternatively, you can freeze the weights of some of the layers in the pre-trained model and only fine-tune the remaining layers during the training process.
3) Add new layers: Add new layers on top of the pre-trained model to adapt it to your target task. These new layers are typically randomly initialized and are trained using the target dataset. The output of these new layers serves as the final prediction layer for your target task.
4) Fine-tune the model: Train the entire model, including the pre-trained layers and the newly added layers, on your target dataset. During the fine-tuning process, the weights of the pre-trained layers and the new layers are updated using the gradients computed from the target dataset. Fine-tuning allows the model to learn task-specific representations while leveraging the knowledge from the pre-trained model.
5) Evaluate and tune: After training, evaluate the performance of the transferred model on your target task. You may need to tune the hyperparameters and architecture of the transferred model to optimize its performance.

Utilizing pre-trained models can be an effective transfer learning approach as it allows leveraging the knowledge learned from large datasets, reducing the need for extensive training data in the target domain, and potentially improving the performance of the target model. However, it's important to carefully choose the pre-trained model, architecture, and fine-tuning strategy to ensure that the transferred knowledge is relevant and beneficial for the target task.

There are various pre-trained machine learning models available in the market, such as Google's Inception model [218], Microsoft's MicrosoftML R package [219] and Microsoftml Python package [220], and others like AlexNet [221], Oxford's VGG Model [222], and Microsoft's ResNet [223]. In addition, some of the well-known pre-trained models used for NLP-related data problems are Google's word2vec Model [224], Stanford's GloVe Model [225] and BERT [178].

BERT is a pre-trained language model that was initially introduced by Google in 2018. The model is trained on a large corpus of unlabeled text data to learn the underlying structure of the language. It utilizes a transformer-based architecture that allows it to capture long-term dependencies and contextual relationships between words. After pre-training, the model is fine-tuned on a specific downstream NLP task, such as sentiment analysis, question answering, or named entity recognition. This fine-tuning step enables the model to adapt to the specific requirements of the downstream task. In summary, BERT is a transfer learning technique that leverages pre-training on unlabeled text data and fine-tuning on specific NLP tasks to achieve state-of-the-art performance on a variety of NLP benchmarks [178].

In particular, the most common transfer learning strategy in fake news detection is fine-tuning pre-trained models. Models like BERT, Llama, and GPT (Generative Pre-trained Transformer) have been pre-trained on extensive text corpora and can be fine-tuned for fake news detection [91]. By adjusting the weights of these models on a specific fake news dataset, researchers can achieve high detection accuracy with relatively low computational resources.

Other transfer learning strategies used for fake news detection include the adaptation of Convolutional Neural Networks (CNNs), traditionally used for image recognition, to text classification tasks, including fake news detection [22]. Models like VGG16, which were previously trained on large image datasets, may be reused by replacing the last layers and retraining on textual data [226]. This strategy takes advantage of CNNs' hierarchical feature extraction capabilities, which enable them to detect detailed patterns in textual data that indicate fake news. In addition, pre-training hybrid models on large datasets and then fine-tuning them on specific fake news datasets used in fake news to exploit the strengths of both architectures.

Recently, the researchers shifted the whole focus to transformer-based models, particularly those like BERT, GPT-3, and Llama [227], [228]. These models are pre-trained on massive datasets using self-supervised learning techniques, which enable them to understand and generate human-like text. For fake news detection, these models can be fine-tuned on labelled datasets specific to fake news, enabling them to distinguish between fake and real news with high precision.

Based on the collected data from the surveyed articles, along with their corresponding fake news detection effectiveness, the following conclusions can be drawn related to the use of transfer learning techniques in this domain:

1) CNN with AlexNet as a transfer learning technique achieved an accuracy of 93.2%. In comparison, not applying transfer learning recorded an accuracy of 70.1% in [22].
2) In [179], pre-trained BERT as a transfer learning technique achieved an accuracy of 94.66%. Similarly, a pre-trained BERT has also helped in the detection of fake news using ISOT dataset [91]. In another case of BERT variations, RoBERTa achieved an accuracy of 92.77% and 91.7% on Politifact and Gossipcop respectively [227] which outperform the state-of-the-art, without transfer learning, techniques by achieving an average accuracy of 10.49% and 14.53% improvements on Politifact and Gossipcop, respectively.
3) CNN with various transfer learning techniques, such as AlexNet, ResNet50, MobileNet, DenseNet,

XceptionNet, InceptionV3, VGG16, and VGG19, achieved high accuracy on the EMERGENT dataset [226]. The detection accuracy was ranging between 91.22% and 97.68% in [48]. VGG16 was also used as a pre-trained model with freezing some layers and trained on a self-created dataset, achieving about 98% detection accuracy [71].
4) The Universal Language Model Fine-tuning transfer learning technique has achieved over 80% for all the evaluation metrics (Accuracy, Precision, Recall, F1) on PHEME dataset [55].

B. TRANSFER LEARNING CHALLENGES FOR FAKE NEWS DETECTION

Transfer learning has been used in various natural language processing (NLP) applications, including fake news detection. However, there are several challenges associated with applying transfer learning in this domain.

One initial aspect that we must highlight is that transfer learning has not been extensively explored in fake news detection. we consider that this is partly due to the complexity of the task, which requires identifying subtle linguistic cues and context-specific information.

Another challenge concerns the difficulty in finding related domains and publicly available datasets that can be useful for training the models. The success of transfer learning relies on the availability of large and diverse datasets that share some commonality with the target task. However, in the case of fake news detection, relevant datasets are often limited, and it can be challenging to find related domains that can be used for transfer learning.

The rarity of data is another significant challenge in fake news detection. Since the detection of fake news is a relatively new area of research, there are limited annotated datasets available for training and testing models. This scarcity of data makes it difficult to apply transfer learning techniques, which rely on large amounts of labelled data for pre-training.

In addition to these challenges, other issues need to be addressed to apply transfer learning effectively in fake news detection. For instance, the choice of pre-trained models and their adaptation to specific tasks can significantly impact the performance of the models. Furthermore, the transferability of pre-trained models across different languages, domains, and cultures is still an active area of research. Considering transfer learning strategies is a relevant area for further research that can lead to improved solutions for FND.

VI. STRATEGIES FOR DEALING WITH IMBALANCE

Deep learning and machine learning algorithms presuppose that the target classes of the training data have similar prior probabilities. This assumption, however, is flagrantly violated in a variety of real-world applications, including fake news detection. In this section, we start by summarizing the main techniques used to deal with the class imbalance problem and describe our main findings from the surveyed articles in the context of fake news detection. Then we summarize the main open challenges related to the class imbalance problem that are still open in this domain.

A. THE CLASS IMBALANCE PROBLEM IN THE CONTEXT OF FAKE NEWS DETECTION

In many real-world domains, the majority of the available examples belong to one class (the majority or negative class) while a much smaller number belongs to the other class (the minority or positive class), which is typically the most important class [229]. This situation is known as the class imbalance problem. The dominant class tends to overpower classifiers in this situation, causing them to overlook the minority class. The significance of the imbalance problem grew as more researchers discovered that it leads to inadequate classification performance and that most algorithms perform poorly when datasets are highly imbalanced [230]. From the standpoint of applications, the nature of the imbalance can be divided into two categories: data that is naturally imbalanced (e.g., credit card frauds, earthquakes, shuttle failure and rare diseases) or data for which it is too expensive to obtain data on the minority class for learning such as natural disasters prediction, or uncommon events prediction such as volcanic eruptions or tsunamis, may require historical data or expert knowledge, which could be sparse or expensive to obtain [230]. This is also the case for fake news detection where the number of fake news available is much less represented in the available data.

Several techniques have been proposed to address the issues associated with class imbalance. The three main types of techniques that can be applied are resampling techniques (or data pre-processing), algorithmic level techniques, and data post-processing techniques [231]. The solutions most commonly used are the data pre-processing or algorithm-level techniques.

In data preprocessing techniques, sampling is applied to the training data to add new samples or remove existing ones. These techniques aim to change the training data distribution to force the learning algorithm to focus on the most relevant class. This change in the training data can be accomplished through over- and/or under-sampling. Over-sampling is the process of adding new samples to the training data while under-sampling is the process of removing samples. Figure 14 and Figure 15 illustrate the random under-sampling and random over-sampling techniques. These techniques act by randomly removing cases or adding copies of existing cases.

In Random Under-sampling, examples from the majority class are randomly removed from the training dataset until the class distribution becomes more balanced. This can be achieved by randomly selecting examples from the majority class and removing them from the training dataset. Random under-sampling can be a simple and quick technique to address class imbalance, but it may result in the loss of valuable information from the majority class, leading to a potential loss of predictive performance.

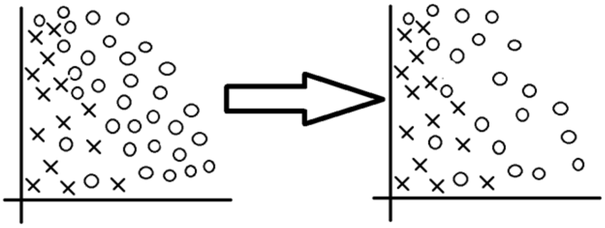

FIGURE 14. Random under-sampling.

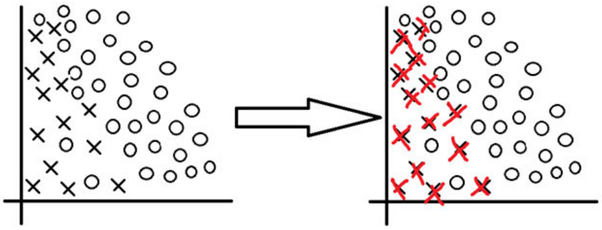

FIGURE 15. Random over-sampling.

In Random over-sampling, examples from the minority class are randomly duplicated or synthetically generated to increase their representation in the training dataset. This can be achieved by randomly selecting examples from the minority class and duplicating them or generating synthetic examples using techniques such as SMOTE (Synthetic Minority Over-sampling Technique) [232] or ADASYN (Adaptive Synthetic Sampling) [233]. Random over-sampling can help in increasing the representation of the minority class, but it may also result in overfitting or amplification of noise if not done carefully.

The second method for resolving class imbalance is to create or modify an existing algorithm. Instead of changing the distribution of the training data, the change is applied to the learning and the decision process by increasing the importance of the positive class. The cost-sensitive method and recognition-based approaches, kernel-based learning, such as support vector machine (SVM) and radial basis function [234], are among the algorithms that have been adapted to address the class imbalance problem. Typically, specially developed algorithms for dealing with the class imbalance issue will work very well for a specific domain for which they were thought. However, they will fail under other domains and they require a thorough understanding of the algorithm to implement the modifications [231].

Our findings show that the use of various imbalance techniques, such as oversampling and downsampling, has shown promising results in improving the performance of different classifiers, including RNN variations, CNN, and hybrid models like CNN+LSTM. The results indicate that oversampling has been effective in improving the accuracy of LSTM and CNN models in [21], achieving accuracies of 95.51% and 98.96%, respectively. Similarly, oversampling has also been beneficial for BERT, achieving an accuracy of 94.66% in [179].

Moreover, the use of SMOTE oversampling has demonstrated effectiveness in dealing with class imbalance. For instance, in [235], SMOTE was used to improve the performance of a DNN model, achieving an accuracy of 98% on the Politifact dataset.

Additionally, another study [177] utilized the focal loss function to prevent classification bias towards the majority class, which significantly improved the performance of their models on imbalanced datasets.

Downsampling has shown effectiveness in improving the training accuracy of hybrid models like CNN+LSTM on PHEME and FN-COV datasets in [34], achieving accuracy rates of 91.88% and 98.62% respectively. Additionally, downsampling has improved the accuracy of CNN and LSTM models in [52], achieving accuracies of 92.38% and 93.56% respectively.

It should be noted that despite the effectiveness of class imbalance techniques in improving the accuracy of fake news detection models, a significant portion of the literature has not thoroughly investigated or addressed this issue. From our findings, only five research articles investigated the class imbalance effect on fake news detection. This highlights the need for further research and exploration of various imbalance techniques to better understand their impact on model performance and generalizability in the context of fake news detection.

B. CHALLENGES RELATED TO THE CLASS IMBALANCE PROBLEM IN FAKE NEW DETECTION

Class imbalance is a common problem in multiple application domains, and fake news detection is not an exception. However, it has not received as much attention as it deserves in the context of fake news detection, which we consider a big challenge to be addressed. The imbalance between real and fake news samples in the dataset can lead to biased classification, where the model performs well on the majority class but poorly on the minority class [231]. Even when considering the usage of deep learning models, it was shown that the class imbalance problem will still affect the performance of the models [236].

One of the challenges related to the class imbalance problem in fake news detection is the issue of using adequate performance assessment metrics to evaluate the model's performance. Traditional metrics such as accuracy can be misleading, as the model may perform well on the majority class but miss out on correctly identifying the minority class. This issue emphasizes the need for specialized metrics such as F1 score, precision, and recall [237].

In the FND domain, there is a lack of systematic studies that evaluate the impact of known techniques for dealing with class imbalance. Techniques such as oversampling, undersampling, and ensemble methods have been widely used in many other domains. However, their effectiveness in fake news detection remains understudied. Therefore, more research is needed to explore the effectiveness of these techniques in the FND domain. An important challenge with the application of these techniques for fake news detection is related to the generation of fake news texts. In this case, it is necessary to generate complete texts that look like real news, but it is also necessary to generate texts that correspond to fake news. This leads to another challenge connected to the need to carefully craft the synthetic text generation so that it corresponds to either fake news or real news. In particular, the generated fake news articles should be realistic and representative of the actual fake news articles to ensure the effectiveness of the model.

The context is also a challenge when considering the generation of fake news. Since fake news is often generated in response to specific events or situations, it can be difficult to apply generic techniques for dealing with a class imbalance that does not consider the specific context in which the fake news was generated.

Lastly, special-purpose algorithms that can deal with the class imbalance problem have not been explored or evaluated for FND. These algorithms include cost-sensitive learning, manipulating the loss functions, or building ensembles that are specially developed to address the class imbalance problem [231]. These techniques have shown promising results in multiple domains, and their effectiveness in FND requires further investigation.

In conclusion, addressing the class imbalance problem in fake news detection is crucial for developing accurate and reliable models. Still, not much research has been done to address this problem. Researchers and practitioners need to pay more attention to this problem and explore various techniques to overcome it. This is a possible area where future researchers should focus on that may lead to improved solutions for FND.

VII. ANSWERS TO RESEARCH QUESTIONS

In this section, we attempt to answer the research questions presented in Section II-B based on our findings. The detailed answers are described below.

RQ1: Which algorithms are used for fake news detection throughout time?

Given our findings, deep learning models are considered effective models in fake news detection. There is a notable increase in the number of articles that address the different models and architectures for this task. We also noticed that the research focus shifted towards deep learning models for FND during the global COVID-19 Pandemic in 2021 which forms about 83% of the research effort that was conducted on FND. The remaining 17% of the FND research was conducted before this year.

Our findings also show that fake news can be detected by CNNs, RNNs, GRUs, LSTMs, and BERTs models in many variations and with different architectures. We noticed that LSTM/(Bi)LSTM were the models that appeared more frequently in the surveyed articles. The detection was also examined using hybrid models which increased the detection effectiveness at some points. It is also noticeable that using the BERT model in the detection of fake news exhibits a huge positive impact on the detection effectiveness.

RQ2: Which datasets are used in the fake news detection domain?

The most difficult part of detecting fake news is the absence of a labelled dataset with trustworthy ground truth labels with an accepted size [195]. For several usages in DL, researchers attempted to collect datasets over the last few years. The collected datasets are massively varied from one another due to the purpose of the study. For instance, some of these datasets are political and consist of political statements as is the case in PolitiFact. Other datasets are built with news articles collected in a specific time frame, while other datasets include social media posts such as Twitter. Moreover, fake news is frequently collected from duplicitous websites intended to disseminate misinformation. This fake news will end up being shared on social media platforms by its creator. This fake news will also be shared by other individuals unintentionally without checking the news source or by other malicious users and bots.

Our findings show that Liar, ISOT, PHEM, and FakeNewsNet (with their three variations) are the most popular datasets being used in fake news detection. These six datasets have been used in about 80% of the surveyed articles. We also noticed that researchers frequently attempted to create their own dataset to reach the required size and the domain which is obvious in about 45% of the surveyed studies. Other researchers combined two or more datasets to have an acceptable-sized dataset.

It is also worth mentioning that selecting a proper dataset is a crucial task in fake news detection since it will impact the detection effectiveness. It is noticeable from our findings that applying the same detection model in different datasets has an enormous difference in the detection accuracy [28], [30], [41], [42], [43], [46], [50], [122], [125], [135], [136], [163], [170], [172], [186], [188].

RQ3: How effective are deep learning methods for fake news detection?

Researchers studied various DL algorithms in the detection and classification of fake news as we mentioned previously. These algorithms include CNN, RNN (with it is variations), GNN, BERT and Attention-based mechanisms, and hybrid approaches. The detection effectiveness of these algorithms is influenced by the datasets used and the combination of different architectures for detection.

CNN and (Bi)LSTM have been the most used detection models and achieved the highest detection accuracy when compared against other approaches. RNNs, including their variations such as LSTM/(Bi)LSTM and GRU, are utilized with considerable effectiveness in about 70%. Their ability to maintain information over sequences allows them to understand context better, which is essential for identifying fake news. CNNs on the other hand have proven to be effective for fake news detection tasks, appearing in 61% of the research articles we surveyed. BERT and hybrid detection models have also made a noticeable detection effectiveness appearing in about 47% of the surveyed articles. Feedforward

Neural Networks and Graph Neural Networks were also used in the detection process even though not in many studies.

It is worth mentioning that in one research article, many deep learning models were developed to draw comprehensive conclusions. Hence, the total percentage of all the models that appeared in the surveyed articles is more than 100%.

The detailed effectiveness of DL detection models in the fake news field is in Section III.

RQ4: Which solutions incorporate transfer learning mechanisms, if any?

Transfer learning is the process of exploiting what has been learned in one task to improve the generalization in another task [204]. The goal of transfer learning is to improve learning in the target task by leveraging knowledge from the source task.

Transfer learning is not applied in many fake news detection studies as our findings show. There are only seven research articles that examined the effect of transfer learning on detection accuracy [22], [71], [91], [98], [179], [183], [227]. However, utilizing transfer learning strategies increased the detection accuracy. The transfer learning that was utilized in the FND domain may be categorized under fine-tuning pre-trained models, using CNN-based architectures, employing pre-trained hybrid models, and leveraging transformer-based models. The highest improvement presented by utilizing transfer learning was by reaching an accuracy of 93.2% when applying the Alexnet pre-trained model which represents an improvement of 23.1% compared to the baseline case which is done without applying transfer learning.

It is worth mentioning that applying the same detection model on different datasets recorded enormous differences in the detection accuracy [28], [29], [30], [41], [43], [46], [50], [122], [125], [135], [136], [163], [170], [172], [186], [188]. This issue might be tackled by including a transfer learning approach so the detection model can report an approximate accuracy.

RQ5: Which solutions deal with different levels of an imbalanced dataset?

A dataset with a skewed class distribution where the end-user preferences are biased towards the least represented class(es) suffers from a class imbalance problem. A model learned under these conditions will focus on the majority class and will not learn correctly the minority and important classes [231].

Most of the available fake news datasets are imbalanced. From the articles we surveyed, only seven papers specifically treated class imbalance and studied its effect on fake news detection by utilizing various strategies to handle this issue. These strategies were: random and advance oversampling in four articles, random undersampling in two articles, and utilizing a different loss function in one article. The oversampling has been deployed by increasing the number of instances in the minority class to match the majority class which improved the detection effectiveness [21], [31], [179]. In addition, advanced oversampling techniques such as the Synthetic Minority Over-sampling Technique (SMOTE), generate synthetic examples of the minority class by interpolating between existing instances rather than duplicating existing ones. SMOTE helped the trained model get about 95% detection accuracy with a noticeable improvement compared to the baseline case without treating the class imbalance [235]. This helped to mitigate the risk of overfitting and enhanced the model's generalization ability.

Another strategy presented to balance the dataset was the random undersampling which involves reducing the number of instances in the majority class to match the minority class [34], [52].

Finally, the focal loss function is designed to address the class imbalance by down-weighting the loss assigned to well-classified examples and focusing more on hard-to-classify instances [177]. This approach helps to prevent the model from becoming biased towards the majority class and ensures that the minority class instances are given appropriate attention during training.

Handling the imbalanced dataset achieved a better accuracy result compared to the baseline cases that do not deal with the class imbalance. Thus, this is a relevant area for further research that can lead to improved solutions for FND.

VIII. THREATS TO VALIDITY

SLRs are prone to several threats to validity that may lead to a bias in the review outcomes. These threats are publication bias and errors in data collection, study exclusion, and data extraction. Regarding publication bias, studies with positive results are more expected to be selected over negative studies. This issue is alleviated by attempting to determine whether the studies discuss their results and limitations. Moreover, the sole purpose of this SLR is to report the effectiveness of DL models rather than present new results. In addition, there is no motivation from our SLR to select studies reporting only positive results.

Regarding filtering out studies based on the search criteria, we aimed to have a broad search query as we mentioned in Section II-C to alleviate this threat. We could also expand the survey date range to contain the studies that were published before 2018. However, fake news became more popular from 2018 onward, and we aimed to provide an updated review of the most recent trends in this application domain. This motivation is supported by Figure 3 which demonstrates the remarkable increase in fake news detection publications over time.

Regarding the issue of incorrectly excluded articles and extracting the data, we alleviated this issue by asking another researcher to review some random studies. There is no rule for determining the number of articles for the random check task, but about half of the surveyed articles were selected for this special check.

IX. MAIN GAPS AND OPEN ISSUES

From our investigation, we gathered a list of the main gaps and open challenges that still deserve the attention of the research community for the fake news detection problem. We must highlight that this is a challenging task, involving several difficulties which we describe to allow future researchers to focus on the most important open issues.

- ***Lack of labelled data:*** One of the major challenges in training deep learning models for fake news detection is the limited availability of labelled data [197], [202]. Fake news datasets are often small, and obtaining accurate and comprehensive annotations for training can be challenging. This can impact the performance and generalization of deep learning models, as they heavily rely on large amounts of labelled data for effective training.
- ***Potentially biased datasets:*** Another issue in fake news detection is the potential bias in the datasets used for training and evaluation [201]. Fake news datasets may contain inherent biases, such as political or cultural biases, that can affect the performance and fairness of deep learning models. It is essential to carefully curate and preprocess datasets to mitigate these biases and ensure the reliability and generalizability of the models.
- ***Lack of benchmarks:*** There is a lack of standardized benchmarks for evaluating the performance of deep learning models in fake news detection. The absence of benchmark datasets, evaluation metrics, and protocols makes it challenging to compare the performance of different models and assess their effectiveness [186], [238]. The development of standardized benchmarks can facilitate fair and rigorous comparisons and foster advancements in the field.
- ***Transfer learning solutions not sufficiently explored:*** Transfer learning, which leverages pre-trained models for feature extraction or model initialization, has shown promise in improving the performance of deep learning models for various tasks [239], [240]. However, in the context of fake news detection, the exploration of transfer learning solutions is still limited. There is a need to further investigate and optimize transfer learning approaches for fake news detection to leverage knowledge from related tasks and domains.
- ***Class imbalance not adequately addressed:*** Class imbalance, where the number of samples in different classes is significantly imbalanced, is a common issue in fake news detection. Deep learning models trained on imbalanced datasets may result in biased and inaccurate predictions, as they tend to be biased towards the majority class [241], [242]. Although some studies have explored imbalance techniques such as oversampling or undersampling, the effectiveness of these techniques in deep learning for fake news detection needs further investigation.
- ***Limited understanding of fake news dynamics:*** Despite extensive research on fake news, there is still a limited understanding of the complex dynamics and mechanisms underlying the spread and impact of misinformation [243]. Deep learning models for fake news detection may be limited by the lack of a comprehensive understanding of how fake news is created, disseminated, and received, which can impact the models' accuracy and effectiveness. Further research is needed to better understand the underlying dynamics of fake news and inform the development of more effective solutions.
- ***Real-world applicability:*** While deep learning models for fake news detection show promising results in controlled research settings, their real-world applicability, and effectiveness in detecting fake news in diverse and dynamic environments, such as social media or online news platforms, is still a challenge. Real-world factors, such as varying levels of information quality, diverse sources of misinformation, and rapid information spread, can impact the performance and reliability of deep learning models in practical scenarios [198], [203].

X. CONCLUSION

The increasing volume of people using communication platforms has opened the door for the spread of fake news. Fake news can influence readers in many aspects, and it is crucial to understand this phenomenon and study mechanisms that allow its early detection. Deep learning has shown its potential in various tasks, including natural language processing, and our systematic literature review highlights its effectiveness in fake news detection.

From our findings, the main categories of algorithms used for FND are CNN, RNN, GNN, Attention-based mechanisms, and BERT. Among these, the most frequently used are RNN-based models, which include the Bi(LSTM). We also found that Liar, ISOT, PHEME, and FakeNewsNet are the publicly available datasets most frequently used in fake news detection. These datasets are a central aspect because selecting a proper dataset is crucial. In effect, the data selection will have an important impact on the detection effectiveness.

Finally, we found that transfer learning and the class imbalance problem are not widely explored in fake news detection studies, even though these techniques have shown promising results in increasing detection accuracy in many fields. Overall, our systematic review highlights the potential of deep learning in fake news detection and identifies important areas for future research. We also provide a comprehensive list of the main gaps and open issues in this domain to guide the next steps of research in this area.

REFERENCES

[1] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," *J. Econ. Perspect.*, vol. 31, no. 2, pp. 211–236, May 2017.

[2] M. R. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, "Depression detection from social network data using machine learning techniques," *Health Inf. Sci. Syst.*, vol. 6, no. 1, pp. 1–12, Dec. 2018.

[3] H. Gao and H. Liu, "Data analysis on location-based social networks," in *Mobile Social Networking: An Innovative Approach*, 2014, pp. 165–194.

[4] K. Sharma, F. Qian, H. Jiang, N. Ruchansky, M. Zhang, and Y. Liu, "Combating fake news: A survey on identification and mitigation techniques," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 3, pp. 1–42, May 2019.

[5] L. Wu, J. Li, X. Hu, and H. Liu, "Gleaning wisdom from the past: Early detection of emerging rumors in social media," in *Proc. SIAM Int. Conf. Data Mining*, 2017, pp. 99–107.

[6] L. Wu, F. Morstatter, K. M. Carley, and H. Liu, "Misinformation in social media: Definition, manipulation, and detection," *ACM SIGKDD Explorations Newslett.*, vol. 21, no. 2, pp. 80–90, Nov. 2019.

[7] J. Ma, W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K.-F. Wong, and M. Cha, "Detecting rumors from microblogs with recurrent neural networks," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, 2016, pp. 3818–3824.

[8] S. K. Bharti, R. Pradhan, K. S. Babu, and S. K. Jena, "Sarcasm analysis on twitter data using machine learning approaches," in *Trends in Social Network Analysis: Information Propagation, User Behavior Modeling, Forecasting, and Vulnerability Assessment*, 2017, pp. 51–76.

[9] S. Helmstetter and H. Paulheim, "Weakly supervised learning for fake news detection on Twitter," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Aug. 2018, pp. 274–277.

[10] S. Kumar and N. Shah, "False information on web and social media: A survey," 2018, *arXiv:1804.08559*.

[11] K. Shu, L. Cui, S. Wang, D. Lee, and H. Liu, "DEFEND: Explainable fake news detection," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2019, pp. 395–405.

[12] R. K. Kaliyar, A. Goswami, P. Narang, and S. Sinha, "FNDNet—A deep convolutional neural network for fake news detection," *Cognit. Syst. Res.*, vol. 61, pp. 32–44, Jun. 2020.

[13] S. Keele, "Guidelines for performing systematic literature reviews in software engineering," Tech. Rep., 2007.

[14] S. Jalali and C. Wohlin, "Systematic literature studies: Database searches vs. Backward snowballing," in *Proc. ACM-IEEE Int. Symp. Empirical Softw. Eng. Meas.*, Sep. 2012, pp. 29–38.

[15] J. Babineau, "Product review: Covidence (Systematic review Software)," *J. Can. Health Libraries Assoc. J. de l'Association des bibliothèques de la santé du Canada*, vol. 35, no. 2, p. 68, Aug. 2014.

[16] S. Girgis, E. Amer, and M. Gadallah, "Deep learning algorithms for detecting fake news in online text," in *Proc. 13th Int. Conf. Comput. Eng. Syst. (ICCES)*, Dec. 2018, pp. 93–97.

[17] Q. Abbas, M. U. Zeshan, and M. Asif, "A CNN-RNN based fake news detection model using deep learning," in *Proc. Int. Seminar Comput. Sci. Eng. Technol. (SCSET)*, Jan. 2022, pp. 40–45.

[18] A. Abdullah, M. Awan, M. Shehzad, and M. Ashraf, "Fake news classification bimodal using convolutional neural network and long short-term memory," *Int. J. Emerg. Technol. Learn*, vol. 11, pp. 209–212, Jul. 2020.

[19] A. Abedalla, A. Al-Sadi, and M. Abdullah, "A closer look at fake news detection: A deep learning perspective," in *Proc. 3rd Int. Conf. Adv. Artif. Intell.*, Oct. 2019, pp. 24–28.

[20] M. Al-Sarem, A. Alsaeedi, F. Saeed, W. Boulila, and O. AmeerBakhsh, "A novel hybrid deep learning model for detecting COVID-19-related rumors on social media based on LSTM and concatenated parallel CNNs," *Appl. Sci.*, vol. 11, no. 17, p. 7940, Aug. 2021.

[21] M. N. Alenezi and Z. M. Alqenaei, "Machine learning in detecting COVID-19 misinformation on Twitter," *Future Internet*, vol. 13, no. 10, p. 244, Sep. 2021.

[22] N. M. AlShariah and A. Khader, "Detecting fake images on social media using machine learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 12, pp. 170–176, 2019.

[23] M. Z. Asghar, A. Habib, A. Habib, A. Khan, R. Ali, and A. Khattak, "Exploring deep neural networks for rumor detection," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 4, pp. 4315–4333, Apr. 2021.

[24] M. C. Buzea, S. Trausan-Matu, and T. Rebedea, "Automatic fake news detection for Romanian online news," *Information*, vol. 13, no. 3, p. 151, Mar. 2022.

[25] M. K. Elhadad, K. F. Li, and F. Gebali, "An ensemble deep learning technique to detect COVID-19 misleading information," in *Proc. 23rd Int. Conf. Network-Based Inf. Syst. Adv. Networked-Based Inf. Syst. (NBiS-2020)*. Springer, 2020, pp. 163–175.

[26] K. M. Fouad, S. F. Sabbeh, and W. Medhat, "Arabic fake news detection using deep learning," *Comput., Mater. Continua*, vol. 71, no. 2, pp. 3647–3665, 2022.

[27] Md. Z. H. George, N. Hossain, Md. R. Bhuiyan, A. K. M. Masum, and S. Abujar, "Bangla fake news detection based on multichannel combined CNN-LSTM," in *Proc. 12th Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT)*, Jul. 2021, pp. 1–5.

[28] M. H. Goldani, R. Safabakhsh, and S. Momtazi, "Convolutional neural network with margin loss for fake news detection," *Inf. Process. Manage.*, vol. 58, no. 1, Jan. 2021, Art. no. 102418.

[29] S. Gonwirat, A. Choompol, and N. Wichapa, "A combined deep learning model based on the ideal distance weighting method for fake news detection," *Int. J. Data Netw. Sci.*, vol. 6, no. 2, pp. 347–354, 2022.

[30] Y.-F. Huang and P.-H. Chen, "Fake news detection using an ensemble learning model based on self-adaptive harmony search algorithms," *Expert Syst. Appl.*, vol. 159, Nov. 2020, Art. no. 113584.

[31] V.-I. Ilie, C.-O. Truica, E.-S. Apostol, and A. Paschke, "Context-aware misinformation detection: A benchmark of deep learning architectures using word embeddings," *IEEE Access*, vol. 9, pp. 162122–162146, 2021.

[32] K. Ivancová, M. Sarnovský, and V. Maslej-Krcsňáková, "Fake news detection in Slovak language using deep learning techniques," in *Proc. IEEE 19th World Symp. Appl. Mach. Intell. Informat. (SAMI)*, Jan. 2021, pp. 000255–000260.

[33] Y. Ji, "Fake news detection based on a bi-directional LSTM with CNN," in *Proc. 3rd Int. Conf. Comput. Data Sci. (CONF-CDS)*. Springer, 2022, pp. 36–44.

[34] R. Kaliyar, A. Goswami, and P. Narang, "A hybrid model for effective fake news detection with a novel COVID-19 dataset," in *Proc. 13th Int. Conf. Agents Artif. Intell.*, 2021, pp. 1066–1072.

[35] R. K. Kaliyar, A. Mohnot, R. Raghhul, V. Prathyushaa, A. Goswami, N. Singh, and P. Dash, "Multideepfake: Improving fake news detection with a deep convolutional neural network using a multimodal dataset," in *Proc. 10th Int. Conf. Adv. Comput. (IACC)*, Panaji, Goa, India. Springer, 2021, pp. 267–279.

[36] R. K. Kaliyar, R. Singh, S. N. Laya, M. S. Sudharshan, A. Goswami, and D. Garg, "Rumeval2020-an effective approach for rumour detection with a deep hybrid C-LSTM model," in *Proc. 10th Int. Conf. Adv. Comput. (IACC)*, Panaji, Goa, India. Springer, Dec. 2021, pp. 300–312.

[37] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter, "Detection and resolution of rumours in social media: A survey," *ACM Comput. Surveys*, vol. 51, no. 2, pp. 1–36, Mar. 2019.

[38] V. L. Rubin, Y. Chen, and N. K. Conroy, "Deception detection for news: Three types of fakes," in *Proc. Assoc. Sci. Technol.*, Jan. 2015, vol. 52, no. 1, pp. 1–4.

[39] J. Brummette, M. DiStaso, M. Vafeiadis, and M. Messner, "Read all about it: The politicization of 'fake news' on Twitter," *Journalism Mass Commun. Quart.*, vol. 95, no. 2, pp. 497–517, 2018.

[40] A. Marlatt. *Records Suggest 2020 Election Conspiracy Involved 80m People*. Accessed: Nov. 11, 2022. [Online]. Available: https://www.satirewire.com/claim-anti-trump-conspiracy-involved-80-million-people/

[41] A. J. Keya, S. Afridi, A. S. Maria, S. S. Pinki, J. Ghosh, and M. F. Mridha, "Fake news detection based on deep learning," in *Proc. Int. Conf. Sci. Contemp. Technol. (ICSCT)*, Aug. 2021, pp. 1–6.

[42] P. M. Konkobo, R. Zhang, S. Huang, T. T. Minoungou, J. A. Ouedraogo, and L. Li, "A deep learning model for early detection of fake news on social media," in *Proc. 7th Int. Conf. Behavioural Social Comput. (BESC)*, Nov. 2020, pp. 1–6.

[43] R. Kozik, S. Kula, M. Choraś, and M. Woźniak, "Technical solution to counter potential crime: Text analysis to detect fake news and disinformation," *J. Comput. Sci.*, vol. 60, Apr. 2022, Art. no. 101576.

[44] V. M. Kresnáková, M. Sarnovský, and P. Butka, "Deep learning methods for fake news detection," in *Proc. IEEE 19th Int. Symp. Comput. Intell. Informat. 7th IEEE Int. Conf. Recent Achievements Mechatronics, Autom., Comput. Sci. Robot. (CINTI-MACRo)*, Nov. 2019, pp. 000143–000148.

[45] E. Masciari, V. Moscato, A. Picariello, and G. Sperlí, "Detecting fake news by image analysis," in *Proc. 24th Symp. Int. Database Eng. Appl.*, Aug. 2020, pp. 1–5.

[46] J. A. Nasir, O. S. Khan, and I. Varlamis, "Fake news detection: A hybrid CNN-RNN based deep learning approach," *Int. J. Inf. Manage. Data Insights*, vol. 1, no. 1, Apr. 2021, Art. no. 100007.

[47] A. Priya and A. Kumar, "Deep ensemble approach for COVID-19 fake news detection from social media," in *Proc. 8th Int. Conf. Signal Process. Integr. Netw. (SPIN)*, Aug. 2021, pp. 396–401.

[48] C. Raj and P. Meel, "ConvNet frameworks for multi-modal fake news detection," *Appl. Intell.*, vol. 51, no. 11, pp. 1–17, 2021.

[49] S. P. Ramya and R. Eswari, "Attention-based deep learning models for detection of fake news in social networks," *Int. J. Cognit. Informat. Natural Intell.*, vol. 15, no. 4, pp. 1–25, Jan. 2022.

[50] H. Saleh, A. Alharbi, and S. H. Alsamhi, "OPCNN-FAKE: Optimized convolutional neural network for fake news detection," *IEEE Access*, vol. 9, pp. 129471–129489, 2021.

[51] M. Samadi, M. Mousavian, and S. Momtazi, "Persian fake news detection: Neural representation and classification at word and text levels," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 21, no. 1, pp. 1–11, Jan. 2022.

[52] M. Sarnovskỳ, V. Maslej-Krešňáková, and K. Ivancová, "Fake news detection related to the COVID-19 in Slovak language using deep learning methods," *Acta Polytechnica Hungarica*, vol. 19, no. 2, pp. 43–57, 2022.

[53] I. K. Sastrawan, I. P. A. Bayupati, and D. M. S. Arsa, "Detection of fake news using deep learning CNN–RNN based methods," *ICT Exp.*, vol. 8, no. 3, pp. 396–408, Sep. 2022.

[54] K. L. Tan, C. Poo Lee, and K. M. Lim, "FN-net: A deep convolutional neural network for fake news detection," in *Proc. 9th Int. Conf. Inf. Commun. Technol. (ICoICT)*, Aug. 2021, pp. 331–336.

[55] M. P. Thilakarathna, V. A. Wijayasekara, Y. Gamage, K. H. Peiris, C. Abeysinghe, I. Rafaideen, and P. Vekneswaran, "Hybrid approach and architecture to detect fake news on Twitter in real-time using neural networks," in *Proc. 5th Int. Conf. Inf. Technol. Res. (ICITR)*, Dec. 2020, pp. 1–6.

[56] F. Torgheh, M. R. Keyvanpour, B. Masoumi, and S. V. Shojaedini, "A novel method for detecting fake news: Deep learning based on propagation path concept," in *Proc. 26th Int. Comput. Conf., Comput. Soc. Iran (CSICC)*, Iran, Mar. 2021, pp. 1–5.

[57] A. Wani, I. Joshi, S. Khandve, V. Wagh, and R. Joshi, "Evaluating deep learning approaches for COVID19 fake news detection," in *Proc. Int. Workshop Combating Online Hostile Posts Regional Lang. During Emergency Situation*. Springer, 2021, pp. 153–163.

[58] Z. Wang, Z. Yin, and Y. A. Argyris, "Detecting medical misinformation on social media using multimodal deep learning," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 6, pp. 2193–2203, Jun. 2021.

[59] F. Xing and C. Guo, "Mining semantic information in rumor detection via a deep visual perception based recurrent neural networks," in *Proc. IEEE Int. Congr. Big Data (BigDataCongress)*, Jul. 2019, pp. 17–23.

[60] A. Zervopoulos, A. G. Alvanou, K. Bezas, A. Papamichail, M. Maragoudakis, and K. Kermanidis, "Deep learning for fake news detection on Twitter regarding the 2019 Hong Kong protests," *Neural Comput. Appl.*, vol. 34, no. 2, pp. 969–982, Jan. 2022.

[61] Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, and J. Gao, "EANN: Event adversarial neural networks for multi-modal fake news detection," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2018, pp. 849–857.

[62] R. K. Kaliyar, "Fake news detection using a deep neural network," in *Proc. 4th Int. Conf. Comput. Commun. Autom. (ICCCA)*, Dec. 2018, pp. 1–7.

[63] O. Ajao, D. Bhowmik, and S. Zargari, "Fake news identification on Twitter with hybrid CNN and RNN models," in *Proc. 9th Int. Conf. Social Media Soc.*, Jul. 2018, pp. 226–230.

[64] L. Wu, Y. Rao, H. Yu, Y. Wang, and A. Nazir, "False information detection on social media via a hybrid deep model," in *Proc. 10th Int. Conf. Social Inform. (SocInfo)*, St. Petersburg, Russia. Springer, Sep. 2018, pp. 323–333.

[65] K. Popat, S. Mukherjee, A. Yates, and G. Weikum, "DeClarE: Debunking fake news and false claims using evidence-aware deep learning," 2018, *arXiv:1809.06416*.

[66] S. Alyoubi, M. Kalkatawi, and F. Abukhodair, "The detection of fake news in Arabic tweets using deep learning," *Appl. Sci.*, vol. 13, no. 14, p. 8209, Jul. 2023.

[67] G. Güler and S. Gündüz, "Deep learning based fake news detection on social media," *Int. J. Inf. Secur. Sci.*, vol. 12, no. 2, pp. 1–21, Jun. 2023.

[68] Y. Doke, "Deep fake detection through deep learning," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 5, pp. 861–866, May 2023.

[69] Y. Lu and H. Ye, "Detection method of fake news spread in social network based on deep learning," in *Proc. Int. Conf. Adv. Hybrid Inf. Process.* Springer, 2022, pp. 473–488.

[70] F. Mira, "Deep learning technique for recognition of deep fake videos," in *Proc. IEEE IAS Global Conf. Emerg. Technol. (GlobConET)*, May 2023, pp. 1–4.

[71] C. Mallick, S. Mishra, and M. R. Senapati, "A cooperative deep learning model for fake news detection in online social networks," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 4, pp. 4451–4460, Apr. 2023.

[72] M. Madani, H. Motameni, and R. Roshani, "Fake news detection using feature extraction, natural language processing, curriculum learning, and deep learning," *Int. J. Inf. Technol. Decis. Making*, vol. 23, no. 3, pp. 1063–1098, May 2024.

[73] G. Mareeswari and E. V. Dinesh, "Deep neural networks based detection and analysis of fake tweets," in *Proc. 4th Int. Conf. Signal Process. Commun. (ICSPC)*, Mar. 2023, pp. 56–61.

[74] M. Samadi and S. Momtazi, "Fake news detection: Deep semantic representation with enhanced feature engineering," *Int. J. Data Sci. Analytics*, pp. 1–12, Mar. 2023.

[75] J. Alghamdi, Y. Lin, and S. Luo, "Does context matter? Effective deep learning approaches to curb fake news dissemination on social media," *Appl. Sci.*, vol. 13, no. 5, p. 3345, Mar. 2023.

[76] F. W. R. Tokpa, B. H. Kamagaté, V. Monsan, and S. Oumtanaga, "Fake news detection in social media: Hybrid deep learning approaches," *J. Adv. Inf. Technol.*, vol. 14, no. 3, pp. 606–615, 2023.

[77] O. Prakash and R. Kumar, "Fake news detection in social networks using attention mechanism," in *Proc. Int. Conf. Cogn. Intell. Comput. (ICCIC)*. Springer, 2023, pp. 453–462.

[78] S. Kumar, A. Kumar, A. Mallik, and R. R. Singh, "OptNet-fake: Fake news detection in socio-cyber platforms using grasshopper optimization and deep neural network," *IEEE Trans. Computat. Social Syst.*, early access, 2023, doi: 10.1109/TCSS.2023.3246479.

[79] Q. Zhang, Z. Guo, Y. Zhu, P. Vijayakumar, A. Castiglione, and B. B. Gupta, "A deep learning-based fast fake news detection model for cyber-physical social services," *Pattern Recognit. Lett.*, vol. 168, pp. 31–38, Apr. 2023.

[80] A. Kishwar and A. Zafar, "Fake news detection on Pakistani news using machine learning and deep learning," *Expert Syst. Appl.*, vol. 211, Jan. 2023, Art. no. 118558.

[81] P. K. Verma, P. Agrawal, V. Madaan, and R. Prodan, "MCred: Multi-modal message credibility for fake news detection using BERT and CNN," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 8, pp. 10617–10629, Aug. 2023.

[82] A. K. Yadav, S. Kumar, D. Kumar, L. Kumar, K. Kumar, S. K. Maurya, M. Kumar, and D. Yadav, "Fake news detection using hybrid deep learning method," *Social Netw. Comput. Sci.*, vol. 4, no. 6, p. 845, Nov. 2023.

[83] A. Saeed and E. A. Solami, "Fake news detection using machine learning and deep learning methods," *Comput., Mater. Continua*, vol. 77, no. 2, pp. 2079–2096, 2023.

[84] A. Y. Umar, I. S. Ahmad, and K. Muhammad, "Fake news detection using CNN/GRU deep learning model," *Sule Lamido Univ. J. Sci. Technol.*, vol. 7, no. 1, pp. 57–67, 2023.

[85] Y. Singh, "Fake news detection using LSTM in TensorFlow and deep learning," *J. Appl. Sci. Educ. (JASE)*, vol. 3, no. 2, pp. 1–14, 2023.

[86] P. Sharma and R. Sahu, "Fake news detection using deep learning based approach," in *Proc. Int. Conf. Circuit Power Comput. Technol. (ICCPCT)*, Aug. 2023, pp. 651–656.

[87] M. R. H. Shezan, M. N. Zawad, Y. A. Shahed, and S. Ripon, "Bangla fake news detection using hybrid deep learning models," in *Applied Informatics for Industry 4.0*. Boca Raton, FL, USA: CRC Press, 2023, pp. 46–60.

[88] C. Nandhakumar, C. Kowsika, R. Reshema, and L. Sandhiya, "Fake news detection using machine learning and deep learning classifiers," in *Proc. Int. Conf. Inf. Commun. Technol. Intell. Syst.* Springer, 2023, pp. 165–175.

[89] P. M. Subhash, D. Gupta, S. Palaniswamy, and M. Venugopalan, "Fake news detection using deep learning and transformer-based model," in *Proc. 14th Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT)*, Jul. 2023, pp. 1–6.

[90] A. Jaiswal, H. Verma, and N. Sachdeva, "Swarm optimized fake news detection on social-media textual content using deep learning," in *Proc. Int. Conf. Adv. Comput., Commun. Appl. Informat. (ACCAI)*, May 2023, pp. 1–8.

[91] I. Ennejjai, A. Ariss, N. Kharmoum, W. Rhalem, S. Ziti, and M. Ezziyyani, "Artificial intelligence for fake news," in *Proc. Int. Conf. Adv. Intell. Syst. Sustain. Develop.* Springer, 2022, pp. 77–91.

[92] O. Ngada and B. Haskins, "Investigating fake news detection by means of deep learning on a limited data set," in *Proc. IEEE Asia–Pacific Conf. Comput. Sci. Data Eng. (CSDE)*, Dec. 2022, pp. 1–6.

[93] Z. Wang, "Deep learning methods for fake news detection," in *Proc. IEEE 2nd Int. Conf. Data Sci. Comput. Appl. (ICDSCA)*, 2022, pp. 472–475.

[94] N. Jayakody, A. Mohammad, and M. N. Halgamuge, "Fake news detection using a decentralized deep learning model and federated learning," in *Proc. 48th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2022, pp. 1–6.

[95] D. R. Collen, L. K. Nyandoro, and K. Zvarevashe, "Fake news detection using 5L-CNN," in *Proc. 1st Zimbabwe Conf. Inf. Commun. Technol. (ZCICT)*, Zimbabwe, Nov. 2022, pp. 1–7.

[96] A. Qdroo and M. Baykara, "A new approach to detect fake news related to COVID-19 pandemic using deep neural network," *J. Appl. Sci. Technol. Trends*, vol. 3, no. 2, pp. 81–88, Dec. 2022.

[97] Z. A. Jawad and A. J. Obaid, "Combination of convolution neural networks and deep neural networks for fake news detection," 2022, *arXiv:2210.08331*.

[98] S. Suratkar and F. Kazi, "Deep fake video detection using transfer learning approach," *Arabian J. Sci. Eng.*, vol. 48, no. 8, pp. 9727–9737, Aug. 2023.

[99] J. Del Ser, M. N. Bilbao, I. Laña, K. Muhammad, and D. Camacho, "Efficient fake news detection using bagging ensembles of bidirectional echo state networks," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2022, pp. 1–7.

[100] A. Sedik, A. A. Abohany, K. M. Sallam, K. Munasinghe, and T. Medhat, "Deep fake news detection system based on concatenated and recurrent modalities," *Expert Syst. Appl.*, vol. 208, Dec. 2022, Art. no. 117953.

[101] K. Sangeeta, "Fake news detection using feature selection and deep learning," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 6, pp. 3878–3887, Jun. 2022.

[102] J. Rautela, V. Ramalingam, and H. Makhdoomi, "Fake news detection through deep learning techniques," *Int. J. Health Sci.*, vol. 6, no. 5, p. 2107—2111, 2022.

[103] V. Kandasamy, Š. Hubálovský, and P. Trojovský, "Deep fake detection using a sparse auto encoder with a graph capsule dual graph CNN," *PeerJ Comput. Sci.*, vol. 8, p. e953, May 2022.

[104] Y. Tashtoush, B. Alrababah, O. Darwish, M. Maabreh, and N. Alsaedi, "A deep learning framework for detection of COVID-19 fake news on social media platforms," *Data*, vol. 7, no. 5, p. 65, May 2022.

[105] R. Muppidi and D. V. Biksham, "Deep convolutional neural network for fake news detection over online social networks," *Interantional J. Sci. Res. Eng. Manage.*, vol. 6, no. 4, pp. 1–10, Apr. 2022.

[106] Q. Hu, Q. Li, Y. Lu, Y. Yang, and J. Cheng, "Multi-level word features based on CNN for fake news detection in cultural communication," *Pers. Ubiquitous Comput.*, vol. 24, no. 2, pp. 259–272, Apr. 2020.

[107] S. Deepak and B. Chitturi, "Deep neural approach to fake-news identification," *Proc. Comput. Sci.*, vol. 167, pp. 2236–2243, 2020.

[108] C. Kulkarni, P. Monika, S. Shruthi, M. D. Bharadwaj, and D. Uday, "COVID-19 fake news detection using glove and bi-LSTM," in *Proc. 2nd Int. Conf. Sustain. Expert Syst. (ICSES)*. Springer, 2022, pp. 43–56.

[109] R. Malhotra, A. Mahur, and Achint, "COVID-19 fake news detection system," in *Proc. 12th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence)*, Jan. 2022, pp. 428–433.

[110] Y. Tian, J. Gu, Y. Jia, and R. O. Sinnott, "An exploration of machine and deep learning models for fake news detection in social media," in *Proc. 8th Int. Conf. Behav. Social Comput. (BESC)*, 2021, pp. 1–6.

[111] H. Zhu and R. O. Sinnott, "A performance comparison of fake news detection approaches," in *Proc. IEEE Asia–Pacific Conf. Comput. Sci. Data Eng. (CSDE)*, Dec. 2021, pp. 1–7.

[112] G. S. Mahara and S. Gangele, "Fake news detection: A RNN-LSTM, bi-LSTM based deep learning approach," in *Proc. IEEE 1st Int. Conf. Data, Decis. Syst. (ICDDS)*, Dec. 2022, pp. 01–06.

[113] G. Anusha, G. Praveen, D. Mounika, U. S. Krishna, and R. Cristin, "Detection of fake news using recurrent neural network," in *Proc. IEEE Int. Conf. Distrib. Comput. Electr. Circuits Electron. (ICDCECE)*, Apr. 2022, pp. 1–5.

[114] P. K. Sree, G. R. Babu, P. B. V. R. Rao, P. V. Chintalapati, and M. Prasad, "Fake news detection using cellular automata based deep learning," in *Proc. 3rd Int. Conf. Comput. Inf. Technol. (ICCIT)*, Sep. 2023, pp. 167–171.

[115] H. Ali, M. Khan, A. AlGhadhban, M. Alazmi, A. Alzamil, K. Al-Utaibi, and J. Qadir, "Analyzing the robustness of fake-news detectors under black-box adversarial attacks," *IEEE Access*, vol. 9, pp. 81678–81692, 2021.

[116] E. Qawasmeh, M. Tawalbeh, and M. Abdullah, "Automatic identification of fake news using deep learning," in *Proc. 6th Int. Conf. Social Netw. Anal., Manage. Secur. (SNAMS)*, Oct. 2019, pp. 383–388.

[117] K. K. Kumar, S. H. Rao, G. Srikar, and M. B. Chandra, "A novel approach for detection of fake news using long short term memory (LSTM)," *Int. J.*, vol. 10, no. 5, 2021, Art. no. 1320800.

[118] T. Ahmad, M. S. Faisal, A. Rizwan, R. Alkanhel, P. W. Khan, and A. Muthanna, "Efficient fake news detection mechanism using enhanced deep learning model," *Appl. Sci.*, vol. 12, no. 3, p. 1743, Feb. 2022.

[119] N. Aslam, I. Ullah Khan, F. S. Alotaibi, L. A. Aldaej, and A. K. Aldubaikil, "Fake detect: A deep learning ensemble model for fake news detection," *Complexity*, vol. 2021, pp. 1–8, Apr. 2021.

[120] T. Chauhan and H. Palivela, "Optimization and improvement of fake news detection using deep learning approaches for societal benefit," *Int. J. Inf. Manage. Data Insights*, vol. 1, no. 2, Nov. 2021, Art. no. 100051.

[121] M.-Y. Chen, Y.-W. Lai, and J.-W. Lian, "Using deep learning models to detect fake news about COVID-19," *ACM Trans. Internet Technol.*, vol. 23, no. 2, pp. 1–23, May 2023.

[122] R. Garg, "Effective fake news classifier and its applications to COVID-19," in *Proc. IEEE Bombay Sect. Signature Conf. (IBSSC)*, Nov. 2021, pp. 1–6.

[123] V. Jain, R. K. Kaliyar, A. Goswami, P. Narang, and Y. Sharma, "AENeT: An attention-enabled neural architecture for fake news detection using contextual features," *Neural Comput. Appl.*, vol. 34, no. 1, pp. 771–782, Jan. 2022.

[124] T. Jiang, J. P. Li, A. U. Haq, and A. Saboor, "Fake news detection using deep recurrent neural networks," in *Proc. 17th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process. (ICCWAMTIP)*, 2020, pp. 205–208.

[125] N. Kanagavalli, S. B. Priya, and J. D, "Design of hyperparameter tuned deep learning based automated fake news detection in social networking data," in *Proc. 6th Int. Conf. Comput. Methodologies Commun. (ICCMC)*, Mar. 2022, pp. 958–963.

[126] J. Kumari, R. Choudhary, S. Kumari, and G. Krishna, "A deep learning based approach for classification of news as real or fake," in *Proc. Data Sci. Security (IDSCS)*. Singapore: Springer, 2021, pp. 239–246.

[127] D.-H. Lee, Y.-R. Kim, H.-J. Kim, S.-M. Park, and Y.-J. Yang, "Fake news detection using deep learning," *J. Inf. Process. Syst.*, vol. 15, no. 5, pp. 1119–1130, 2019.

[128] J. Liu, C. Wang, C. Li, N. Li, J. Deng, and J. Z. Pan, "DTN: Deep triple network for topic specific fake news detection," *J. Web Semantics*, vol. 70, Jul. 2021, Art. no. 100646.

[129] R. Mahesh, B. Poornika, N. Sharaschandrika, S. D. Goud, and P. U. Kumar, "Identification of fake news using deep learning architecture," in *Proc. 3rd Int. Conf. Inventive Res. Comput. Appl. (ICIRCA)*, Sep. 2021, pp. 1246–1253.

[130] B. Majumdar, Md. RafiuzzamanBhuiyan, Md. A. Hasan, Md. S. Islam, and S. R. H. Noori, "Multi class fake news detection using LSTM approach," in *Proc. 10th Int. Conf. Syst. Model. Advancement Res. Trends (SMART)*, Dec. 2021, pp. 75–79.

[131] S. Mengji, "Fake news detection using RNN-LSTM," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 9, no. 10, pp. 1731–1737, Oct. 2021.

[132] A. Mohapatra, N. Thota, and P. Prakasam, "Fake news detection and classification using hybrid BiLSTM and self-attention model," *Multimedia Tools Appl.*, vol. 81, no. 13, pp. 18503–18519, May 2022.

[133] U. Narayan, A. Kumar, and K. Kumar, "Fake news detection using hybrid of deep neural network and stacked LSTM," in *Proc. 3rd Int. Conf. Adv. Comput., Commun. Control Netw. (ICAC3N)*, Dec. 2021, pp. 385–390.

[134] N. Rai, D. Kumar, N. Kaushik, C. Raj, and A. Ali, "Fake news classification using transformer based enhanced LSTM and BERT," *Int. J. Cognit. Comput. Eng.*, vol. 3, pp. 98–105, Jun. 2022.

[135] R. R. Rajalaxmi, L. V. Narasimha Prasad, B. Janakiramaiah, C. S. Pavankumar, N. Neelima, and V. E. Sathishkumar, "Optimizing hyperparameters and performance analysis of LSTM model in detecting fake news on social media," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, Mar. 2022, doi: 10.1145/3511897.

[136] F. Sadeghi, A. J. Bidgoly, and H. Amirkhani, "Fake news detection on social media using a natural language inference approach," *Multimedia Tools Appl.*, vol. 81, no. 23, pp. 33801–33821, Sep. 2022.

[137] S. R. Sahoo and B. B. Gupta, "Multiple features based approach for automatic fake news detection on social networks using deep learning," *Appl. Soft Comput.*, vol. 100, Mar. 2021, Art. no. 106983.

[138] Y. Seo and C.-S. Jeong, "FaGoN: Fake news detection model using grammatic transformation on neural network," in *Proc. 13th Int. Conf. Knowl., Inf. Creativity Support Syst. (KICSS)*, Nov. 2018, pp. 1–5.

[139] P. Shrivastava and D. K. Sharma, "Fake content identification using pre-trained glove-embedding," in *Proc. 5th Int. Conf. Inf. Syst. Comput. Netw. (ISCON)*, Oct. 2021, pp. 1–6.

[140] T. E. Trueman, A. Kumar, P. Narayanasamy, and J. Vidya, "Attention-based C-BiLSTM for fake news detection," *Appl. Soft Comput.*, vol. 110, Oct. 2021, Art. no. 107600.

[141] M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, G. S. Choi, and B.-W. On, "Fake news stance detection using deep learning architecture (CNN-LSTM)," *IEEE Access*, vol. 8, pp. 156695–156706, 2020.

[142] P. Ushashree, A. Naik, S. Gurav, A. Kumar, S. Chethan, and B. Madhumala, "Fake news detection using neural network," in *Proc. IEEE Int. Conf. Integr. Circuits Commun. Syst. (ICICACS)*, Aug. 2023, pp. 1–5.

[143] A. Anand, R. Kulkarni, and P. Agrawal, "Fake news identification: An effective combined approach using ML and DL techniques," in *Proc. 2nd Int. Conf. Paradigm Shifts Commun. Embedded Syst., Mach. Learn. Signal Process. (PCEMS)*, Apr. 2023, pp. 1–6.

[144] K. Saini and R. Jain, "A hybrid LSTM-BERT and glove-based deep learning approach for the detection of fake news," in *Proc. 3rd Int. Conf. Smart Data Intell. (ICSMDI)*, Mar. 2023, pp. 400–406.

[145] A. Khoudi, N. Yahiaoui, and F. Rebahi, "Detect misinformation of COVID-19 using deep learning: A comparative study based on word embedding," in *Proc. 1st Int. Conf. Adv. Innov. Smart Cities (ICAISC)*, Jan. 2023, pp. 1–5.

[146] S. M. Bankar and S. K. Gupta, "Fake news detection using LSTM-based deep learning approach and word embedding feature extraction," in *Proc. Int. Conf. Commun., Electron. Digital Technol.* Singapore: Springer, 2023, pp. 129–141.

[147] A. Matheven and B. V. D. Kumar, "Fake news detection using deep learning and natural language processing," in *Proc. 9th Int. Conf. Soft Comput. Mach. Intell. (ISCMI)*, Nov. 2022, pp. 11–14.

[148] J. A. Reshi and R. Ali, "Online fake news detection using pre-trained embeddings," in *Proc. 5th Int. Conf. Multimedia, Signal Process. Commun. Technol. (IMPACT)*, 2022, pp. 1–5.

[149] M. Madani, H. Motameni, and H. Mohamadi, "Fake news detection using deep learning integrating feature extraction, natural language processing, and statistical descriptors," *Secur. Privacy*, vol. 5, no. 6, p. e264, 2022.

[150] P. Katariya, V. Gupta, R. Arora, A. Kumar, S. Dhingra, Q. Xin, and J. Hemanth, "A deep neural network-based approach for fake news detection in regional language," *Int. J. Web Inf. Syst.*, vol. 18, nos. 5–6, pp. 286–309, Dec. 2022.

[151] A. Divija, "Fake news classifier," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 6, pp. 1716–1722, Jun. 2022.

[152] J. Soni, "An efficient LSTM model for fake news detection," *Comput. Sci. Eng., Int. J.*, vol. 12, no. 2, pp. 1–10, Apr. 2022.

[153] E. Amer, K.-S. Kwak, and S. El-Sappagh, "Context-based fake news detection model relying on deep learning models," *Electronics*, vol. 11, no. 8, p. 1255, Apr. 2022.

[154] N. Xiang, "Deep learning-based fake information detection and influence evaluation," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–8, Feb. 2022.

[155] S. M. Jaybhaye, V. Badade, A. Dodke, A. Holkar, and P. Lokhande, "Fake news detection using LSTM based deep learning approach," in *Proc. ITM Web Conf.*, vol. 56, 2023, p. 03005.

[156] R. S. Aziz, A. T. Sadiq, M. Kherallah, and A. Douik, "Arabic fake news detection for COVID-19 using deep learning and machine learning," *Periodicals Eng. Natural Sci. (PEN)*, vol. 11, no. 6, p. 56, Dec. 2023.

[157] M. B. Narayanan, A. K. Ramesh, K. S. Gayathri, and A. Shahina, "Fake news detection using a deep learning transformer based encoder–decoder architecture," *J. Intell. Fuzzy Syst.*, vol. 45, no. 5, pp. 8001–8013, Nov. 2023.

[158] S. Malik, A. K. Chakraverti, and A. I. Abidi, "Enhancing fake news detection using classification algorithms and deep learning," in *Proc. 10th IEEE Uttar Pradesh Sect. Int. Conf. Electr., Electron. Comput. Eng. (UPCON)*, Dec. 2023, pp. 780–787.

[159] A. Chabukswar and P. D. Shenoy, "Fake news detection using optimized deep learning model through effective feature extraction," in *Proc. Int. Conf. Recent Adv. Inf. Technol. for Sustain. Develop. (ICRAIS)*, Nov. 2023, pp. 118–123.

[160] N. Ahuja and S. Kumar, "S-HAN: Hierarchical attention networks with stacked gated recurrent unit for fake news detection," in *Proc. 8th Int. Conf. Rel., INFOCOM Technol. Optim. (Trends Future Directions) (ICRITO)*, Jun. 2020, pp. 873–877.

[161] G. Sarin and P. Kumar, "ConvGRUText: A deep learning method for fake text detection on online social media," in *Proc. Pasific Asia Conf. Inf. Syst. (PACIS)*, 2020, pp. 60–74.

[162] A. Verma, V. Mittal, and S. Dawn, "FIND: Fake information and news detections using deep learning," in *Proc. 12th Int. Conf. Contemp. Comput. (IC)*, Aug. 2019, pp. 1–7.

[163] M. Dong, L. Yao, X. Wang, B. Benatallah, Q. Z. Sheng, and H. Huang, "DUAL: A deep unified attention model with latent relation representations for fake news detection," in *Proc. 19th Int. Conf. Web Inf. Syst. Eng. (WISE)*, Dubai, United Arab Emirates. Cham, Switzerland: Springer, Nov. 2018, pp. 199–209.

[164] S. Taheri, S. H. Hashemi, A. Y. Zomaya, and J. Yong, "Sequence graph transform: A general approach to substructure-aware sequence encoding," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, 2018, pp. 117–126.

[165] P. Velickovic, G. Cucurull, A. Casanova, A. R. P. Liò, and Y. Bengio, "Graph attention networks," in *Proc. ICLR*, 2018, pp. 1–12.

[166] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1025–1035.

[167] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.

[168] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans. Neural Netw.*, vol. 20, no. 1, pp. 61–80, Jan. 2008.

[169] N. Bai, F. Meng, X. Rui, and Z. Wang, "Rumour detection based on graph convolutional neural net," *IEEE Access*, vol. 9, pp. 21686–21693, 2021.

[170] F. B. Mahmud, M. Md. S. Rayhan, M. H. Shuvo, I. Sadia, and Md. K. Morol, "A comparative analysis of graph neural networks and commonly used machine learning algorithms on fake news detection," in *Proc. 7th Int. Conf. Data Sci. Mach. Learn. Appl. (CDMA)*, Mar. 2022, pp. 97–102.

[171] I. Pilkevych, D. Fedorchuk, O. Naumchak, and M. Romanchuk, "Fake news detection in the framework of decision-making system through graph neural network," in *Proc. IEEE 4th Int. Conf. Adv. Inf. Commun. Technol. (AICT)*, Sep. 2021, pp. 153–157.

[172] Y. Ren, B. Wang, J. Zhang, and Y. Chang, "Adversarial active learning based heterogeneous graph neural network for fake news detection," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, 2020, pp. 452–461.

[173] M. Sun, I. A. Hameed, H. Wang, and M. Pasquine, "Perceiving the narrative style for fake news detection using deep learning," in *Proc. IEEE 23rd Int Conf High Perform. Comput. Communications; 7th Int Conf Data Sci. Systems; 19th Int Conf Smart City; 7th Int Conf Dependability Sensor, Cloud Big Data Syst. Appl. (HPCC/DSS/SmartCity/DependSys)*, Dec. 2021, pp. 1195–1202.

[174] B. Upadhayay and V. Behzadan, "Hybrid deep learning model for fake news detection in social networks (student abstract)," in *Proc. AAAI Conf. Artif. Intell.*, 2022, vol. 36, no. 11, pp. 13067–13068.

[175] P. Hiremath, S. S. Kalagi, and Mohana, "Analysis of fake news detection using graph neural network (GNN) and deep learning," in *Proc. 2nd Int. Conf. Autom., Comput. Renew. Syst. (ICACRS)*, Dec. 2023, pp. 1805–1811.

[176] E. Y. Okano, Z. Liu, D. Ji, and E. E. S. Ruiz, "Fake news detection on fake. BR using hierarchical attention networks," in *Proc. 14th Int. Conf. Comput. Process. Portuguese Language (PROPOR)*, Evora, Portugal. Springer, Mar. 2020, pp. 143–152.

[177] A. Al Obaid, H. Khotanlou, M. Mansoorizadeh, and D. Zabihzadeh, "Multimodal fake-news recognition using ensemble of deep learners," *Entropy*, vol. 24, no. 9, p. 1242, Sep. 2022.

[178] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.

[179] S. M. Isa, G. Nico, and M. Permana, "Indobert for Indonesian fake news detection," *ICIC Exp. Lett.*, vol. 16, no. 3, pp. 289–297, 2022.

[180] B. Palani, S. Elango, and V. Viswanathan K, "CB-fake: A multimodal deep learning framework for automatic fake news detection using capsule neural network and BERT," *Multimedia Tools Appl.*, vol. 81, no. 4, pp. 5587–5620, Feb. 2022.

[181] S. Sharma, M. Saraswat, and A. K. Dubey, "Fake news detection using deep learning," in *Proc. 3rd Iberoamerican Conf. 2nd Indo-American Conf. Knowl. Graphs Semantic Web (KGSWC)*, Kingsville, Texas, USA, Nov. 2021. Springer, pp. 249–259.

[182] M. Kanchana, V. M. Kumar, T. P. Anish., and P. Gopirajan, "Deep fake BERT: Efficient online fake news detection system," in *Proc. Int. Conf. Netw. Commun. (ICNWC)*, Apr. 2023, pp. 1–6.

[183] R. H. Khan, A. Shihavuddin, M. M. M. Syeed, R. U. Haque, and M. F. Uddin, "Improved fake news detection method based on deep learning and comparative analysis with other machine learning approaches," in *Proc. Int. Conf. Eng. Emerg. Technol. (ICEET)*, Oct. 2022, pp. 1–6.
[184] A. Kumar, J. P. Singh, and A. K. Singh, "COVID-19 fake news detection using ensemble-based deep learning model," *IT Prof.*, vol. 24, no. 2, pp. 32–37, Mar. 2022.
[185] R. K. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach," *Multimedia Tools Appl.*, vol. 80, no. 8, pp. 11765–11788, Mar. 2021.
[186] M. Q. Alnabhan and P. Branco, "Evaluating deep learning for cross-domains fake news detection," in *Proc. Int. Symp. Found. Pract. Secur.* Cham, Switzerland: Springer, 2023, pp. 40–51.
[187] J. V. Tembhurne, M. M. Almin, and T. Diwan, "Mc-DNN: Fake news detection using multi-channel deep neural networks," *Int. J. Semantic Web Inf. Syst.*, vol. 18, no. 1, pp. 1–20, Feb. 2022.
[188] R. K. Kaliyar, P. Kumar, M. Kumar, M. Narkhede, S. Namboodiri, and S. Mishra, "DeepNet: An efficient neural network for fake news detection using news-user engagements," in *Proc. 5th Int. Conf. Comput., Commun. Secur. (ICCCS)*, Oct. 2020, pp. 1–6.
[189] F. Zhou, Y. Hu, and X. Shen, "MSANet: Multimodal self-augmentation and adversarial network for RGB-D object recognition," *Vis. Comput.*, vol. 35, no. 11, pp. 1583–1594, Nov. 2019.
[190] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," *Array*, vol. 16, Dec. 2022, Art. no. 100258.
[191] S. K. Hamed, M. J. A. Aziz, and M. R. Yaakub, "A review of fake news detection models: Highlighting the factors affecting model performance and the prominent techniques used," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 7, pp. 379–390, 2023.
[192] I. Ahmad, M. Yousaf, S. Yousaf, and M. O. Ahmad, "Fake news detection using machine learning ensemble methods," *Complexity*, vol. 2020, pp. 1–11, Oct. 2020.
[193] A. Zubiaga, M. Liakata, and R. Procter, "Learning reporting dynamics during breaking news for rumour detection in social media," 2016, *arXiv:1610.07363*.
[194] W. Y. Wang, "'Liar, liar pants on fire': A new benchmark dataset for fake news detection," 2017, *arXiv:1705.00648*.
[195] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "FakeNewsNet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media," *Big Data*, vol. 8, no. 3, pp. 171–188, Jun. 2020.
[196] M. Nirav Shah and A. Ganatra, "A systematic literature review and existing challenges toward fake news detection models," *Social Netw. Anal. Mining*, vol. 12, no. 1, p. 168, Dec. 2022.
[197] B. Cao, L. Hua, J. Cao, J. Gui, B. Liu, and J. Tin-Yau Kwok, "No place to hide: Dual deep interaction channel network for fake news detection based on data augmentation," 2023, *arXiv:2303.18049*.
[198] S. Warjri, P. Pakray, S. A. Lyngdoh, and A. K. Maji, "Fake news detection using social media data for khasi language," in *Proc. Int. Conf. Intell. Syst., Adv. Comput. Commun. (ISACC)*, Feb. 2023, pp. 1–6.
[199] S. Helmstetter and H. Paulheim, "Collecting a large scale dataset for classifying fake news tweets using weak supervision," *Future Internet*, vol. 13, no. 5, p. 114, Apr. 2021.
[200] K. D. K. Parimala and A. G. Mala, "An optimal detection of fake news from Twitter data using dual-stage deep capsule autoencoder," *J. Experim. Theor. Artif. Intell.*, vol. 36, no. 2, pp. 287–313, Feb. 2024.
[201] S. Kato, L. Yang, and D. Ikeda, "Domain bias in fake news datasets consisting of fake and real news pairs," in *Proc. 12th Int. Congr. Adv. Appl. Informat. (IIAI-AAI)*, Jul. 2022, pp. 101–106.
[202] H. F. Villela, F. Corrêa, J. S. D. A. N. Ribeiro, A. Rabelo, and D. B. F. Carvalho, "Fake news detection: A systematic literature review of machine learning algorithms and datasets," *J. Interact. Syst.*, vol. 14, no. 1, pp. 47–58, Mar. 2023.
[203] S. Suryavardan, S. Mishra, M. Chakraborty, P. Patwa, A. Rani, A. Chadha, A. Reganti, A. Das, A. Sheth, M. Chinnakotla, A. Ekbal, and S. Kumar, "Findings of factify 2: Multimodal fake news detection," 2023, *arXiv:2307.10475*.
[204] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *J. Big Data*, vol. 3, no. 1, pp. 1–40, 2016.
[205] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
[206] C. Wang and S. Mahadevan, "Heterogeneous domain adaptation using manifold alignment," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, 2011, vol. 22, no. 1, p. 1541.
[207] Y. Zhu, Y. Chen, Z. Lu, S. Pan, G.-R. Xue, Y. Yu, and Q. Yang, "Heterogeneous transfer learning for image classification," in *Proc. AAAI Conf. Artif. Intell.*, 2011, vol. 25, no. 1, pp. 1304–1309.
[208] B. Kulis, K. Saenko, and T. Darrell, "What you saw is not what you get: Domain adaptation using asymmetric kernel transforms," in *Proc. CVPR*, Jun. 2011, pp. 1785–1792.
[209] F. Arslan, N. Hassan, C. Li, and M. Tremayne, "A benchmark dataset of check-worthy factual claims," in *Proc. Int. AAAI Conf. Web Social Media*, vol. 14, 2020, pp. 821–829.
[210] M. Harel and S. Mannor, "Learning from multiple outlooks," 2010, *arXiv:1005.0027*.
[211] J. Nam and S. Kim, "Heterogeneous defect prediction," in *Proc. 10th Joint Meeting Found. Softw. Eng.*, 2015, pp. 508–519.
[212] P. Prettenhofer and B. Stein, "Cross-language text classification using structural correspondence learning," in *Proc. 48th Annu. Meeting Assoc. Comput. Linguistics*, Jul. 2010, pp. 1118–1127.
[213] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, 2014, pp. 1–9.
[214] Z. Alyafeai, M. S. AlShaibani, and I. Ahmad, "A survey on transfer learning in natural language processing," 2020, *arXiv:2007.04239*.
[215] J. Howard and S. Ruder, "Universal language model fine-tuning for text classification," 2018, *arXiv:1801.06146*.
[216] M. Masum, H. Shahriar, and H. M. Haddad, "A transfer learning with deep neural network approach for network intrusion detection," *Int. J. Intell. Comput. Res.*, vol. 12, no. 1, pp. 1087–1095, Jun. 2021.
[217] C. Käding, E. Rodner, A. Freytag, and J. Denzler, "Fine-tuning deep neural networks in continuous learning scenarios," in *Proc. Int. Workshops Comput. Vis. Workshops*, Taipei, Taiwan. Cham, Switzerland: Springer, Nov. 2017, pp. 588–605.
[218] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," 2014, *arXiv:1409.4842*.
[219] Microsoft Corporation. (2018). *MicrosoftML: A Package for Machine Learning With R*. [Online]. Available: https://microsoft.github.io/MicrosoftML/
[220] Microsoft. *Microsoft Machine Learning*. Accessed: May 5, 2024. [Online]. Available: https://docs.microsoft.com/en-us/machine-learning/
[221] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
[222] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[223] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. pattern Recognit. (CVPR)*, 2016, pp. 770–778.
[224] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," 2013, *arXiv:1301.3781*.
[225] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global vectors for word representation," in *Proc. Empirical Methods Natural Lang. Process. (EMNLP)*, 2014, pp. 1532–1543.
[226] W. Ferreira and A. Vlachos, "Emergent: A novel data-set for stance classification," in *Proc. NAACL-HLT*, 2016, pp. 1163–1168.
[227] B. Palani and S. Elango, "CTrL-FND: content-based transfer learning approach for fake news detection on social media," *Int. J. Syst. Assurance Eng. Manage.*, vol. 14, no. 3, pp. 903–918, Jun. 2023.
[228] B. M. Pavlyshenko, "Analysis of disinformation and fake news detection using fine-tuned large language model," 2023, *arXiv:2309.04704*.
[229] C. Seiffert, T. M. Khoshgoftaar, J. Van Hulse, and A. Napolitano, "A comparative study of data sampling and cost sensitive learning," in *Proc. IEEE Int. Conf. Data Mining Workshops*, 2008, pp. 46–52.
[230] R. Longadge and S. Dongre, "Class imbalance problem in data mining review," 2013, *arXiv:1305.1707*.
[231] P. Branco, L. Torgo, and R. P. Ribeiro, "A survey of predictive modeling on imbalanced domains," *ACM Comput. Surveys*, vol. 49, no. 2, pp. 1–50, Jun. 2017.
[232] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002.

[233] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," in *Proc. IEEE Int. Joint Conf. Neural Netw. (IEEE World Congr. Comput. Intell.)*, Jun. 2008, pp. 1322–1328.
[234] N. V. Chawla, N. Japkowicz, and A. Kotcz, "Editorial: Special issue on learning from imbalanced data sets," *ACM SIGKDD Explorations Newslett.*, vol. 6, no. 1, pp. 1–6, Jun. 2004.
[235] T. Bhatia, B. Manaskasemsak, and A. Rungsawang, "Detecting fake news sources on Twitter using deep neural network," in *Proc. 11th Int. Conf. Inf. Educ. Technol. (ICIET)*, Mar. 2023, pp. 508–512.
[236] K. Ghosh, C. Bellinger, R. Corizzo, P. Branco, B. Krawczyk, and N. Japkowicz, "The class imbalance problem in deep learning," *Mach. Learn.*, vol. 113, no. 7, pp. 4845–4901, Jul. 2024.
[237] J.-G. Gaudreault, P. Branco, and J. Gama, "An analysis of performance metrics for imbalanced classification," in *Proc. 24th Int. Conf. Discovery Sci. (DS)*, Halifax, NS, Canada. Cham, Switzerland: Springer, Oct. 2021, pp. 67–77.
[238] Z. Yan, Y. Zhang, X. Yuan, S. Lyu, and B. Wu, "DeepfakeBench: A comprehensive benchmark of deepfake detection," in *Proc. Adv. Neural Inf. Processing Syst.*, 2023, vol. 2, no. 6, pp. 1–32.
[239] P. Shrivastava and D. K. Sharma, "COVID-19 fake news detection using pre-tuned BERT-based transfer learning models," in *Proc. 11th Int. Conf. Syst. Model. Advancement Res. Trends (SMART)*, Dec. 2022, pp. 64–68.
[240] W. Tang, Z. Ma, H. Sun, and J. Wang, "Learning sparse alignments via optimal transport for cross-domain fake news detection," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Aug. 2023, pp. 1–5.
[241] I. Y. Agarwal and D. P. Rana, "Fake news and imbalanced data perspective," in *Data Preprocessing, Active Learning, and Cost Perceptive Approaches for Resolving Data Imbalance*. Hershey, PA, USA: IGI Global, 2021, pp. 195–210.
[242] A. J. Keya, M. A. H. Wadud, M. F. Mridha, M. Alatiyyah, and M. A. Hamid, "AugFake-BERT: Handling imbalance through augmentation of fake news using BERT to enhance the performance of fake news classification," *Appl. Sci.*, vol. 12, no. 17, p. 8398, Aug. 2022.
[243] J. W. W. Muigai, "Understanding fake news," *Int. J. Sci. Res. Publications*, vol. 9, no. 1, pp. 29–38, 2019.

MOHAMMAD Q. ALNABHAN received the bachelor's and master's degrees in computer science with a specialization in data mining. Currently, he is pursuing the Ph.D. degree with the University of Ottawa, Canada. His major field of study is focused on deep learning for security matters and fake news detection on social media.

He brings a wealth of expertise to his academic pursuits through a diverse range of work experiences. He has been a dedicated part-time Professor with the Computer Science Department, University of Ottawa, since 2022, where he imparts his knowledge and passion for the subject to students. In addition to his academic role, he has proven his skills as a full-stack Programmer and a Microsoft Certified Trainer Specialist, showcasing his proficiency in industry-standard technologies.

PAULA BRANCO received the Ph.D. degree in computer science from the Faculty of Sciences, University of Porto, Portugal, in 2018.

She joined the School of Electrical Engineering and Computer Science, University of Ottawa, as an Assistant Professor, in January 2020. Her research interests include artificial intelligence, machine learning, KDD, data mining, and in particular, imbalanced domain problems, rare extreme values forecast, health and cybersecurity applications, and privacy.