

Article

EFND: A Semantic, Visual, and Socially Augmented Deep Framework for Extreme Fake News Detection

Muhammad Imran Nadeem ¹, Kanwal Ahmed ¹, Dun Li ^{1,*}, Zhiyun Zheng ¹, Hend Khalid Alkahtani ², Samih M. Mostafa ³, Orken Mamyrbayev ^{4,*} and Hala Abdel Hameed ^{5,6}

¹ School of Computer and Artificial Intelligence, Zhengzhou University, Zhengzhou 450001, China

² Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia

³ Computer Science Department, Faculty of Computers and Information, South Valley University, Qena 83523, Egypt

⁴ Institute of Information and Computational Technologies, Almaty 050010, Kazakhstan

⁵ Faculty of Computer and Information Systems, Fayoum University, Fayoum 63514, Egypt

⁶ Khaybar Applied College, Taibah University, Medina 42353, Saudi Arabia

* Correspondence: ielidun@zzu.edu.cn (D.L.); morkenj@mail.ru (O.M.)

Abstract: Due to the exponential increase in internet and social media users, fake news travels rapidly, and no one is immune to its adverse effects. Various machine learning approaches have evaluated text and images to categorize false news over time, but they lack a comprehensive representation of relevant features. This paper presents an automated method for detecting fake news to counteract the spread of disinformation. The proposed multimodal EFND integrates contextual, social context, and visual data from news articles and social media to build a multimodal feature vector with a high level of information density. Using a multimodal factorized bilinear pooling, the gathered features are fused to improve their correlation and offer a more accurate shared representation. Finally, a Multilayer Perceptron is implemented over the shared representation for the classification of fake news. EFND is evaluated using a group of standard fake news datasets known as “FakeNewsNet”. EFND has outperformed the baseline and state-of-the-art machine learning and deep learning models. Furthermore, the results of ablation studies have demonstrated the efficacy of the proposed framework. For the PolitiFact and GossipCop datasets, the EFND has achieved an accuracy of 0.988% and 0.990%, respectively.

Keywords: multimodal fake news detection; Multimodal Factorized Bilinear pooling; natural language processing; social sensing; misinformation/disinformation



Citation: Nadeem, M.I.; Ahmed, K.; Li, D.; Zheng, Z.; Alkahtani, H.K.; Mostafa, S.M.; Mamyrbayev, O.; Abdel Hameed, H. EFND: A Semantic, Visual, and Socially Augmented Deep Framework for Extreme Fake News Detection.

Sustainability **2023**, *15*, 133. <https://doi.org/10.3390/su15010133>

Academic Editors: Alfonso Chaves-Montero, Javier Augusto Nicoletti, Francisco José García-Moro and Walter Federico Gadea-Aiello

Received: 22 November 2022

Revised: 4 December 2022

Accepted: 5 December 2022

Published: 22 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Context-aware methods, i.e., content-based models, and social context-aware methods, i.e., social context-based models, are two of the most used techniques for detecting fake news [1–4]. Content-based models focus on the content of news, i.e., title, body, image, and video. While, socially aware methods take user creation time, engagements, connections, comments, and reposts into consideration. The socially aware methods further extend their expertise in measuring the propagation patterns and comparing them with fake news propagation patterns to detect anomalies, known as propagation structure-based methods. Furthermore, comments, likes, and retweets of a post are also examined to detect irregularities. These methods are known as post-based methods [5–13].

The content-based techniques offer a simpler and more realistic method for detecting fake news, especially in the initial stages, but unimodal content-based fake news detection techniques are inefficient at identifying false news since they employ distinct textual [14–27] and visual characteristics [28–34]. However, users are purposefully led astray on social media by fake news that is packaged in a variety of genuine facts. Therefore, additional

measures such as social context are considered for the accurate detection of fake news. For news articles, we have also introduced a similarity measure between the title and body of news. As the majority of fake news titles are just clickbait and the body of the news does not match the title [35]. This will provide very crucial information about news articles' authenticity and support the process of fake news detection. Furthermore, stance detection is incorporated for social media news, this is an important measure to determine a public standpoint and judgment towards a user's social media post for fake news detection [9].

The socially aware methods are targeted and effective, but data collection, noisy data, irrelevant data, and missing data pose a lot of challenges. Therefore, a multi-model approach for fake news detection is proposed using socially aware methods including user profile associations, user engagements stance, and context-aware methods including textual, visual features, and similarity measures. Compared to the previous works [36–60], we have added a wider range of news-related and social context features. We attempt to uncover fake news within a few minutes of inception. Our research is the first of its kind to use both credibility and stance in a multimodal automated fake news detection system. The primary objective of this study is to integrate content-based approaches with the social context to significantly boost the model's effectiveness.

The following is the primary contribution that this study has made:

- i. A similarity measure for the news article title and body for the credibility of the article.
- ii. User credibility based on a multi-feature of a user profile.
- iii. Fusing of textual and visual features via multi-modal factorized bilinear pooling.
- iv. A multimodal approach for identifying fake news based on news content and social context.
- v. Evaluation, findings, and a critical examination of the proposed framework.

The various sections of the paper are structured as follows: A review of the literature is provided in Section 2. In Section 3, the proposed model and its components are described. Section 4 describes the experimental setup. Section 5 incorporates the results and discussions. Finally, Section 6 concludes and explores future directions.

2. Literature Review

In this section, we provide a high-level summary of works that are pertinent to the proposed model. The researchers used content-based, social context-based, and hybrid features in multimodal fake news detection methods to verify the authenticity of the news. The following subsections provide descriptions of these techniques.

2.1. Content-Based Fake News Detection

The bulk of textual and visual data is utilized to create content-based characteristics. Textual qualities display the author's thoughts and ideas, additional it also exhibits their favored writing style [61,62]. Modeling and primarily expressing textual representations with deep neural systems [63–65] and tensor factorization [66–68] has been shown to be effective in detecting fake news. Various parts of fake news broadcasts can be discovered by extracting visual characteristics from visual components such as pictures and videos [35,69,70].

The framework provided by [71] merged textual and visual aspects into a unified totality. The authors utilized a hierarchical attention network with four layers to achieve their objective. They discovered hidden patterns in the title and body material of the news section. A unique component of the recommended method was the creation of a visual summary. The authors analyzed the semantic similarity between the produced visual summary and title with the content of the news article. They proved that their proposed strategy produced superior results compared to the current best practices. Moreover, a content-based study [72] offered a semi-supervised text fake news classification system utilizing a convolutional neural network that replicates temporal patterns. The authors trained the proposed technique by applying convolutional filters of different sizes on the titles and body of news items and then concatenating the generated feature vectors.

The analysis of testing findings revealed the promising performance of the suggested approach for evaluating whether or not news items were manufactured relative to their legitimate sources.

Vishwakarma et al. [70] suggested an image-to-text converter, entity extractor, web-scraping tool, and processing node to authenticate bogus news. First, it alters a news article's picture to extract its content. The second element of the system detects and removes text. The third component searches Google for suitable connections using entity strings. Finally, the fourth collects the links' text and estimates the proportion of entities shared by the image and the summed material. The proportion also reflects connection credibility. Finally, the ratio of trustworthy to untrustworthy connections determines news credibility.

Other studies examined content-based models using reinforcement learning, attention-residual networks, and fact-checking URL recommendation [73,74] "Hoax News Inspector" involves data collection and categorization [73]. The first module's query is the news article assertions. The second core module includes URL filtering, processing, and classification. URL filtering removes unwanted URLs. After collecting the most valuable URLs, the processing unit retrieves the characteristics needed to recognize fake news. A classification model predicts using all feature sets.

The Elementary Discoursed Unit, developed by Wang et al. [44], has a level of detail between the word and the sentence, making it ideal for the early detection of fake news. Mishra et al. [36] identified fake news by employing a probabilistic latent semantic analysis. Knowledge graph-based document representations can achieve state-of-the-art performance when combined with existing contextual representations, as demonstrated by Koloski et al. [37]. Dynamic fake news detection using a knowledge graph was proposed by Abdelnabi, Hasan, and Fritz [38] and Sun et al. [39]. Significant gains have been seen with some unimodal approaches to detecting fake news. Yet the majority of content published on social media and in the news is of the multimodal variety. So, it's clear that a detection method based solely on unimodal features is inadequate.

2.2. User Credibility Based Fake News Detection

Veracity in social media statistics is an urgent and modern problem. Given the sheer volume of information shared in the social media sphere, the authenticity of such information is especially important when individuals' personal details are concerned [75,76]. There are a number of proposed methods for assessing social media credibility [77–85]. There is a strong correlation between social network topology and user trustworthiness [86]. Using the strength of the ties between a user's Facebook friends, Podobnik et al. [83] offers a model to ascertain the level of trust between those friends. In addition, Agarwal and Zhou [82] provide an approach for gauging a social media user's reliability that makes use of a heterogeneous network in which each actor in the Twitter domain is represented by a distinct vertex type. An evaluation of reliability was conducted utilizing a regressive spread approach. However, the value of a weighting method and the passage of time are ignored in that work. The believability of each edge category should be evaluated independently, hence a weighting mechanism is required. Incorporating a temporal/time dimension is important since the value placed on trustworthiness changes throughout time. Aghdam et al. [87] and Al-Qurishi et al. [88] both go into further detail on the subject of credibility and the inclusion of network structure. Kožuh & Čakš [89] explored the topic of news credibility. They claimed that individuals' characteristics and level of interest in the news are the decisive factors in establishing credibility in social media news. The research also established a link between NFC and both confidence in the media and active participation in that trust.

Few studies [2,7–13,40–43] have tried to employ user profile characteristics for fake news detection. Wu et al. [12] identified bogus news by employing an LSTM network along propagation pathways and obtaining user personal information included from social media. To learn a representation for each tweet, Ma et al. [10] developed a recursive neural model that takes advantage of tree topologies in neural networks. To uncover the spread

of false information, Liu et al. [7] developed a time series classifier model using RNN and CNN. To better detect false news, Guo et al. [2] looked into the HSA-BLSTM model, which gathers information from both the text and the social context. One effective strategy for rumor detection was developed by Ma et al. [9] and Li et al. [8], which takes into account the user's perspective during multi-task learning. News circulation trends were graphically recorded by Wu et al. [13]. Unsupervised learning is utilized in the UbCadet model developed by Savyan and Bhanu [11] to identify compromised Twitter accounts.

The approach of rumor identification presented by Chen, Zhou, Trajcevski, and Bonsangue [40] makes use of multi-view learning and attention from several users. This method has the ability to learn and combine the representations of multiple users' perspectives throughout the tweet's propagation channel. Quantitative argumentation is the basis for Chi and Liao's [41] proposed QA-AXDS, a rumor-detection and user-interaction system that use a dialogue tree as its explanation model. Two parts make up the transformer framework-based model proposed by Raza and Ding [42]: an encoder element to extract representations from the fake news data and a decoder component to detect behavior based on previous data. To identify fake information on social media, Jarrahi and Safari [43] used CNN with three-dimensional input. They have concentrated their attention on the usefulness of the features offered by publishers.

In this study, we analyze credibility as a complicated attribute used by publishers to identify fake news on social media and to present a multi-modal framework with a high level of performance.

2.3. Multimodal Fake News Detection

Deep neural networks have seen widespread application in multimodal data-dependent tasks in recent years, including the answering of visual questions [28], the captioning of images [53], and the identification of fake news [54,56,57,60]. Chen et al. [17] developed an attention-based RNN model that extracts and uses an attention mechanism to blend aspects of a text, image, and social context. For use in a variety of internet-of-things (IoT) applications, Singh et al. [55] developed a model known as an extreme learning machine (ELM). Yang et al. [59] analyzed both the text and the images and then used the adaptive tag (AT) algorithm to derive user-interested tags. The Text Image-CNN model proposed by Yang et al. [60] gathers information that is both overt and covert from both the text and the images to identify instances of fake news. Wang et al. [58] introduced the Event Adversarial Neural Network, a comprehensive framework for the identification of misleading information and event discriminators (EANN). Textual and visual characteristics were retrieved in the multimodal feature extractor section using the Text-CNN and VGG-19 models, respectively. Unfortunately, there is no clear method for using this methodology to uncover intermodal relationships. Khattar et al. [54] suggested a comparable framework, named Multivariational Autoencoder, for the identification of fake news (MVAE). An encoder module is responsible for teaching the MVAE model, the multimodal information's common representation or latent vector, which includes both textual and visual components. This latent vector is used by the decoder to recreate the original samples. SpotFake is a multimodal system for detecting false news that was developed by Shivangi et al. [57]. This model avoids the extra tasks of EANN and MVAE and achieves a greater detection accuracy increase. The BERT model for representing textual features and a pre-trained CNN model using the Imagenet database (VGG-19) for representing visual features, SpotFake delivers a reasonable accuracy improvement over EANN and MVAE compared to past efforts [54,58]. Shivangi et al. [56] created SpotFake+, an enhanced version of SpotFake [57]. This suggested architecture has the advantage of being able to manage a dataset including full-length articles. This model outperformed previous efforts [54,57,58] because it makes use of transfer learning to recognize a news item's written and visual characteristics.

As a means of exploiting both the visual and textual content of news articles, Zhou et al. [45] presented the FND-CLIP framework. A ResNet-based encoder and a BERT-based encoder were used to combine the deep-learning features of text and images,

respectively. Article classification has been improved by applying scaled dot-product attention to a fine-grained fusion of image and text data, as performed by Wang et al. [46]. Their technique focused on associations between visual characteristics and collected multimodal feature interdependence. Shivangi et al. [47] developed a method to selectively extract useful data from the dominant modality while discarding irrelevant data from the weaker modalities. Using a contrastive learning strategy, Chen et al. [48] have trained variational autoencoders (VAE) to compress pictures and texts and minimize the Kullback-Leibler (KL) divergence for news containing valid image-text pairs. The multimodal characteristics are then reweighted based on the matching cross-modal ambiguity score. An implementation of a two-stage network is provided by Wei et al. [49], which initially trains two unimodal networks to learn cross-modal correlation via contrastive learning before fine-tuning the network for false news detection. The model developed by Das et al. [50] incorporates a wide variety of characteristics seen in social settings and in news articles. The dynamic analysis uses a recurrent neural network (RNN) to model the temporal evolution pattern of the propagation tree and the stance network.

Davoudi et al. [51] identified news articles by source, username, and URL domain. These attributes were employed as statistical characteristics in an ensemble model comprising pre-trained models, a statistical feature fusion network, a unique heuristic approach, and news article variables. Segura-Bedmar and Alonso-Bartolome [52] categorized fake news using unimodal and multimodal approaches. Their multimodal technique integrates text and image data based on CNN architecture. Images were beneficial for manipulative content, sarcasm, and misleading associations.

The following are some of the issues that current multimodal fake news detection systems are facing. Although the majority of them attain plenty of context information, they still:

- i. Lack of similarity for title and body text.
- ii. Effective integration of text and visual features.
- iii. lack of user context information.
- iv. Lack of stance analysis.

The objective of this research is to extract characteristics that are helpful and relevant from the substance of the news. Because we take into account a variety of modalities, our attention is focused on the extraction of features from the text and visual contents of a news item. Many sequence models exist for processing text, but they can't develop persistent associations between words or access the input phrase in order. As a result, BRNN equipped with an attention mechanism is employed to analyze text features in both directions.

In addition, earlier research has employed CNN to extract visual characteristics. Nevertheless, as a result of its pooling operation and longitudinal sensitivity, CNN cannot retrieve more informative information. CapsNet has been used to address the issue of information extraction at CNN. The Routing-by-agreement approach and the Margin loss function are utilized to single out the visual components inside the photos of news items that are considered to be the most essential.

Furthermore, to improve the overall effectiveness of the identification of fake news, the suggested model combines semantically significant characteristics with cosine similarity perspective, and social context information to produce an improved feature vector representation for the supplied news. The goal is to combine the retrieved characteristics of the image and the text to get the highest possible correlation between the two and provide a more accurate shared representation. The Multimodal Factorized Bilinear-pooling (MFB) method allows us to accomplish this goal. The increased feature vector is further sent to the multi-perceptron layer. The output indicates whether or not the news item or tweet contains false news.

3. Methodology

The architecture of the proposed multimodal for fake news detection is presented in this section. Figure 1 represents the workflow and related modules.

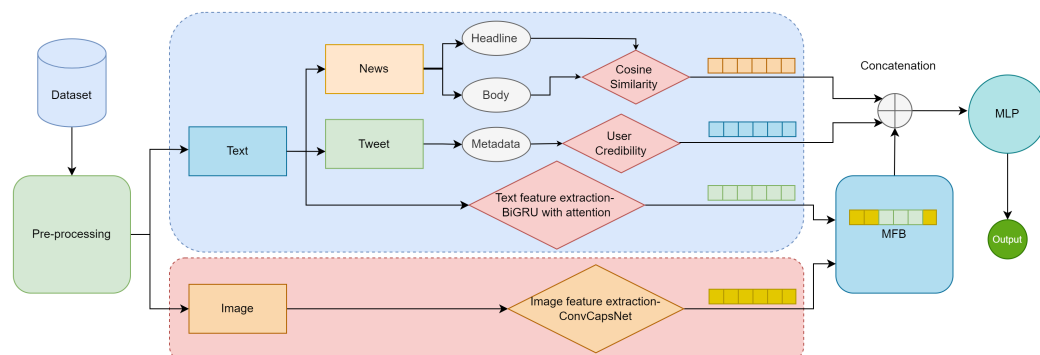


Figure 1. Architecture of the proposed multimodal for fake news detection.

Several text preparation techniques, like tokenization, phrasing, denoising, lemmatization, and stop-word removal, are used to turn documents into a representation appropriate for the classification model in the initial stage. The datasets feature photographs collected from various locations. Images have a high resolution, hence a robust system is needed to evaluate them in their native dimensions employing capsule neural networks. The processing of such high-quality images is time-consuming and expensive in all standard deep-learning models. We scaled all the pictures to 256×256 to overcome this problem. Image and text feature vectors are separately trained using neural networks. The news article credibility module calculates the similarity index for the item's title and body. Textual feature data from news articles is calculated with Semantic Encoding. Using metadata, the user credibility module ranks profiles. The fusion of textual characteristics and visual features is performed using multi-modal factorized bilinear pooling. Later the fused features are concatenated with text similarity and user credibility features. In the final stage, the concatenated features produced from the previous step are utilized as input vectors and fed into MLP for fake news classification.

3.1. Visual Encoding

The components of the image learning module are presented in Figure 2, followed by the description related to its processing.

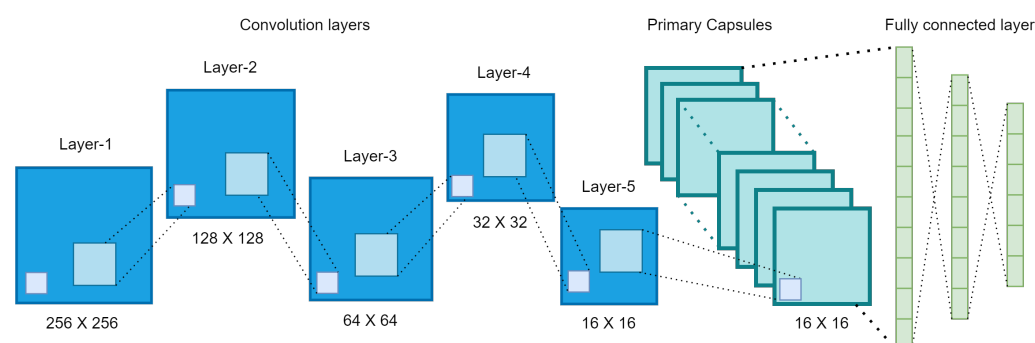


Figure 2. Image feature extraction using CapsNet with convolution.

To preserve item locations and attributes in a picture while modeling their hierarchical relationships, capsule networks have been constructed [90]. With the pooling layer, convolutional neural networks can extract the most insightful insights from their input. Since the data is pooled before being sent on to the next layer, likely, the network won't pick up on finer distinctions [91]. And the neural output that CNN generates is a scalar value. By

packing multiple neurons into each capsule, capsule networks provide a vectorial output of the same size but with distinct routings. [92] A vector's paths stand in for the pictures' settings. Scalar input activation functions like ReLU, Sigmoid, and Tangent are used by CNN. Instead, capsule networks employ a vectorial activation function described by the Equation (1) as squashing. If there is an item in the image, X_j squeezes short vectors toward 0 and long vectors towards 1 [93,94]. In capsule networks, the input value of capsule Y_j is calculated by adding the weighted sum of the prediction vectors $Z_{j|i}$ in the capsules of the lower layers, with the exception of the first layer. Multiplying the output Out_i of a lower-layer capsule by a weight matrix yields the prediction vector ($Z_{j|i}$) (W_{ij}).

$$X_j = \frac{|Y_j|^2}{1 + |Y_j|^2} \times \frac{Y_j}{|Y_j|}$$

$$Y_j = \sum_i N_{ij} Z_{j|i} \quad (1)$$

$$Z_{j|i=W_{ij}Out_i}; N_{ij} = \frac{f(M_{ij})}{\sum_k f(M_{ik})}$$

where, X_j represents capsule j 's output and Y_j its entire input. The dynamic routing procedure selects the coefficient N_{ij} . Logarithmic probability is a gift from M_{ij} . Log prior probability is calculated using SoftMax [95] and is equal to the total of the correlation coefficients between capsule i and capsules in the top layer. Objects of a certain class can be detected by calculating the margin loss in capsule networks using the Equation (2).

$$V_k = T_k \max(0, q^+ |X_k|)^2 + \alpha(1 - T_k) \max(0, |X_k - q^-|)^2 \quad (2)$$

If class k exists, then and only then does T_k equal 1. The loss is down-weighted when the hyperparameters, $q^+ = 0.9$, $q^- = 0.1$, are used [95]. Parameter information such as texture, color, location, size, etc. is contained in the direction of the vectors generated by the capsule networks, while the length of the vector reflects the likelihood of appearing in that region of the picture [94,96].

In this research, we present a capsule network with six convolution layers for classifying images of size 256×256 . The number of convolution layers is raised to improve the performance of the primary layer's feature map. The first layer has 16 filters of size 5×5 with a stride of 1. After each layer, a Max-pooling of size 2×2 is applied. The second, third, fourth, and fifth layers contain 32, 64, 128, and 256 filters with dimensions of 5×5 , 5×5 , 5×5 , and 9×9 , respectively. The sixth layer is the primary layer, and it has 512 filters with 32 capsules containing filters of size 9×9 .

3.2. Semantic Encoding

The text learning modules are represented in Figure 3, and the feature extraction process is discussed in the following paragraphs.

The proposed model seeks to learn information at the word, phrase, and document levels from various news articles and tweets. The word encoder is based on a bidirectional recurrent neural network (BRNN) [97], which allows the usage of variable-length contexts before and after the current word placement. Since we didn't want to use separate memory cells to keep tabs on the status of the input sequences, we turned to the Gated Recurrent Unit (GRU) [98]. It works well for determining correlation over broad temporal ranges. Both reset gate $reset_t$ and update gate $update_t$ are included in GRU. Both attempts to control the state's access to the most recent data. The GRU computes its new state at time t using the Equation (3). Using the new sequence information, this is a linear interpolation between the old state C_{t-1} and the candidate state C_t . The update gate $update_t$ is responsible for deciding what percentage of the previously stored data will be kept and what percentage of new data will be added. Here, $update_t$ is calculated using Equation (4).

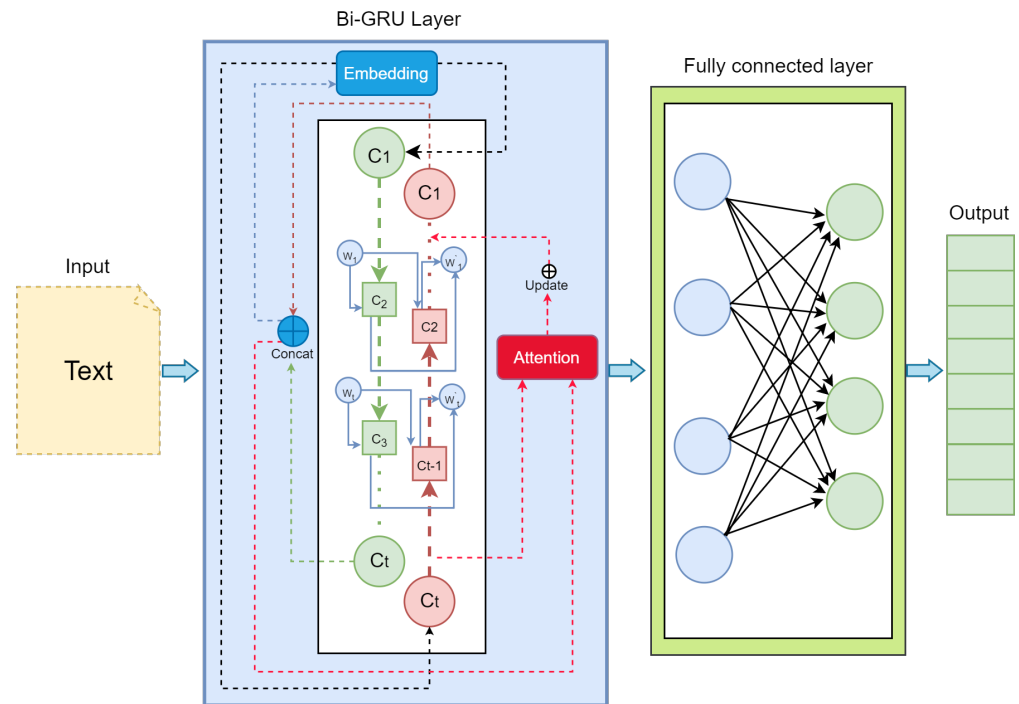


Figure 3. Text processing module using BiGRU and attention mechanism.

$$C_t = (1 - \text{update}_t) \tau C_{t-1} + \text{update}_t \tau \hat{C}_t \quad (3)$$

$$\text{update}_t = \sigma(W1_{\text{update}} \text{Emb}_t + W2_{\text{update}} C_{t-1} + W3_{\text{update}}) \quad (4)$$

where, Emb_t represents the embedding vector at time t . $W1$ represents the weight matrix, $W2$ and $W3$ represents bias matrices with the proper dimensions. The symbol σ indicates a sigmoid activation function, whereas the operation τ denotes elementwise multiplication. The current state is calculated as $\hat{C}_t$ represented in Equation (5).

$$\hat{C}_t = \tanh(W1_C \text{Emb}_t + \text{reset}_t \tau((W2_C C_{t-1}) + W3_n)) \quad (5)$$

where the reset gate reset_t is responsible for determining the amount of information from the previous state that is added to the current state. The bidirectional GRU employs hidden layers in both the forward and backward directions to perform an analysis of the input data, much like the unidirectional GRU does. The output is the result of adding together the computed values in both directions. Let $\vec{C}_t$ and $\overleftarrow{C}_t$ represent the forward and reverse outputs of the bidirectional GRU, respectively. The output is calculated by adding the forward and reverse outputs in order, such as $C_t = [\vec{C}_t, \overleftarrow{C}_t]$.

The sentence encoder takes the word representation as input and utilizes the embedding and bidirectional GRU layers to generate sentence-level vectors. After that, the sentence-level vectors are transformed into document-level vectors by the utilization of bidirectional GRU layers. There is a disparity in the amount of contribution made by individual words and sentences to the generative model. Consequently, the attention mechanism [90] is included in our effort to extract the crucial features of the model. Assume, the input text comprises M sentences, with T_i words per sentence. Let word_{it} and $t \in [1, T]$ represent the words in the i . sentence. The embedding layer and bidirectional GRU layer are responsible for the transformation of a word_{it} into the hidden state C_{it} . The transformation is described as Equation (6):

$$\begin{aligned}
\vec{C}_{it} &= \overrightarrow{GRU}(W1_e \text{ word}_{it}), t \in [1, T] \\
\overleftarrow{C}_{it} &= \overleftarrow{GRU}(W1_e \text{ word}_{it}), t \in [T, 1] \\
C_{it} &= [\vec{C}_{it}, \overleftarrow{C}_{it}]
\end{aligned} \tag{6}$$

where $W1_e$ represents the matrix of the embedding layer and $\overrightarrow{GRU}$ and $\overleftarrow{GRU}$ reflect the procedures described in the preceding section. Consequently, the attention weights of words α_{it} and sentences vectors s_i can be calculated as Equation (7):

$$\begin{aligned}
W2_{it} &= \tanh(W2_{\text{word}} C_{it} + W3_{\text{word}}) \\
\alpha_{it} &= \frac{\exp(W2_{it}^T W2_{\text{word}})}{\sum_t \exp(W2_{it}^T W2_{\text{word}})} \\
s_i &= \sum_t \alpha_{it} C_{it}
\end{aligned} \tag{7}$$

During the training phase, the context vector $W1_{\text{word}}$ receives a random starting point and is simultaneously updated with new information. This vector may be thought of as a high-level representation of a fixed input across words [99,100]. The sentence vectors s_i are then transformed into the hidden state C_i using a second bidirectional GRU layer, as shown in Equation (8).

$$\begin{aligned}
\vec{C}_t &= \overrightarrow{GRU}(s_i), i \in [1, Q] \\
\overleftarrow{C}_t &= \overleftarrow{GRU}(s_i), i \in [Q, 1] \\
C_i &= [\vec{C}_i, \overleftarrow{C}_i]
\end{aligned} \tag{8}$$

Afterward, the attention weights of words α_i and item vectors v are determined using the formulas in Equation (9).

$$\begin{aligned}
W2_i &= \tanh(W1_s h_i + W3_s) \\
\alpha_i &= \frac{\exp(W2_i^T W2_s)}{\sum_t \exp(W2_i^T W2_s)} \\
F &= \sum_i \alpha_i C_i
\end{aligned} \tag{9}$$

To represent the sentence-level context vector, $W2_s$ is given a random starting point and is then updated in the same way as $W2_{w1}$. Through the foregoing training procedure, the item vector F that is generated from a text contains multilevel contextual information derived from both the word-level and the sentence-level structures. Therefore, we refer to it as F_T in the next parts.

3.3. News Article Credibility Module

Based on research by Dong et al. [101] on detecting sensationalism in headlines and bodies of articles, we hypothesize that the degree to which these two elements are the same is a good indicator of an article's reliability. To determine the degree of resemblance, we first embed the article body and title onto the same space and then calculate the cosine distance between them. Since cosine similarity captures the angle of the documents rather than the magnitude, it is an excellent similarity metric for determining the relationship of the documents regardless of their size. It is a mathematical measure of the cosine of the angle formed by the projection of two vectors into space, represented in Equation (10).

$$\text{Similarity}(x, y) = \frac{x \cdot y}{\|x\| \|y\|} = \frac{\sum_1^n x_i y_i}{\sqrt{\sum_1^n x_i^2} \sqrt{\sum_1^n y_i^2}} \tag{10}$$

The dot product of the two vectors is represented as $x \cdot y = \sum_1^n x_i y_i = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$.

3.4. User Credibility Module

To locate socially trustworthy content, it is crucial to have a proper understanding of user interaction-based qualities. Examining the level of interest that users' followers have in their posts is a crucial part of this process. A feature-based ranking model is constructed using a measure that considers a number of critical characteristics shown in Figure 4.

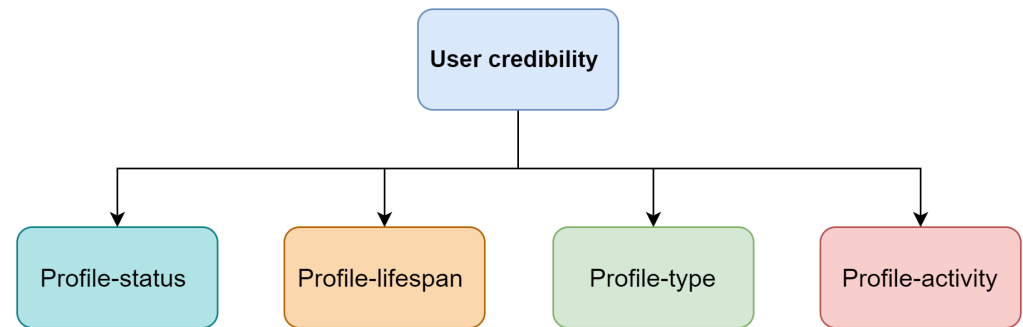


Figure 4. User credibility based on characteristics of user profile.

3.4.1. Profile Lifespan

It provides most information related to user credibility. Most of the time the misinformation is spread from user accounts that are not verified therefore the overall score for this feature is kept at the highest priority. The verified accounts get a 1 and the unverified accounts get a 0 P_{status} score.

3.4.2. Profile Status

It calculates the time information about the existence of user accounts on social media platforms. The variables under consideration are D_{signUp} is the signup date for the user, D_{start} is the date of creation of the social media platform, and D_{now} is today's date. The calculations are presented in Equation (11).

$$\begin{aligned}
 M_{User} &= D_{now} - D_{signUp} \\
 M_{Network} &= D_{now} - D_{start} \\
 P_{lifespan} &= M_{User} * M_{Network}
 \end{aligned}
 \tag{11}$$

where M_{User} is the number of months of the user profile, and $M_{Network}$ is the number of months after the creation of a social network.

3.4.3. Profile Type

Every who uses social media agrees with the fact that the number of followers and friends can be a huge factor in determining the credibility of a user. The users on a social network can be roughly classified into three types. The first ones are looking for information, they mostly scroll through the platform, follow people, and barely post their own updates. The second ones are content creators with few to the huge number of followers, they update quite often and keep their followers interested in their content. The third ones do not get their head into the social network but rather keep it balanced. They don't follow everyone and mostly interact with their circle of friends only. The profile type is calculated as Equation (12).

$$P_{type} = P_{followers} / P_{friends} \tag{12}$$

Here, the resultant P_{type} determines the type of the user. A score less than 0.7 indicates the user is a scroller, a score greater than 1.2 indicates the user is a content creator and a score between 0.7 and 1.2 indicates the user is a balanced user.

3.4.4. Profile Activity

The content of a post is very critical to gaining and losing followers. The number of times a user posts new content or retweets is an essential dimension of its credibility. In this study, we give less score to the retweet and more score to the original content posted by a user, represented in Equation (13).

$$P_{activity} = \left(\frac{P_{post}}{M_{User}} + \frac{P_{rt}}{4 * M_{User}} \right) \quad (13)$$

where $P_{activity}$ is the content score associated with the profile. P_{post} is the number of posts, P_{rt} is the number of retweets, and M_{User} is the age in the number of months for the user profile.

3.4.5. Total Credibility

The total user credibility C_{Total} score is computed by combining the profile status (50%), lifespan (20%), type (10%), and activity (20%) scores, represented in Equation (14). After that, a feature vector representing the user's trustworthiness is constructed by computing the average of the vector values associated with each location vector.

$$C_{Total} = (P_{status}) * 5 + (P_{lifeSpan}) * 2 + P_{type} + (P_{activity}) * 2 \quad (14)$$

3.5. Multi-Modal Factorized Bilinear Pooling (MFB)

To generate a common representation, MFB provides a phenomenon for the fusion of extracted features from semantic and visual encoders. Figure 5 represents the structure of MFB.

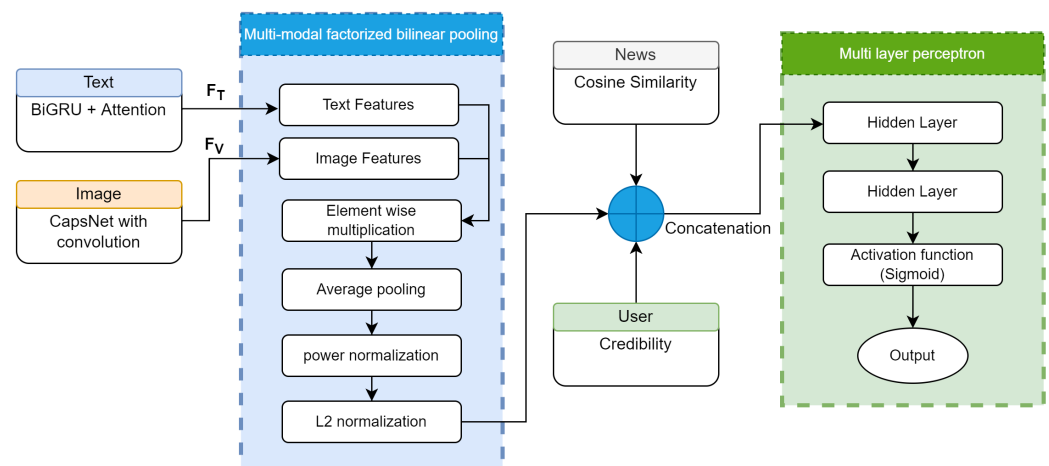


Figure 5. Representation of MFB, Concatenation, and MLP of the proposed model.

Using the MFB module, we combine the news article text feature (F_T) with visual (F_V) feature representations after acquiring them. MFB is preferred over regular concatenation for the reasons outlined below.

- Using a typical concatenation of data from many sources, it might be difficult to identify the endpoint of the derived features.
- Because features are piled one after the other after concatenation, it is possible that the association between picture and text feature representations will not be recognized.

Using the MFB module, these two issues may indeed be effectively addressed. Furthermore, using this fusion technique, the association between textual and visual components is strengthened. Let us suppose that the textual feature vector is represented by ($F_T \in F_m$) and the visual feature vector is represented by ($F_V \in F_n$). The fundamental multimodal bilinear model is thus specified by the following Equation (15).

$$F_{TV} = F_T^T W_i F_V \quad (15)$$

where $W_i \in F_{m \times n}$ is a projection matrix. The bilinear model's output is F_{TV} . Though it is effective at capturing pairwise interactions across feature dimensions, bilinear pooling introduces a large number of parameters, leading to a high processing cost and the possibility of over-fitting.

$$F_{TV} = \sum_{i=1} F_T^T U_i V_i^T F_V = 1^T (U_i^T F_T^T V_i^T F_V) \quad (16)$$

where k is the hidden dimensionality of the factored matrices $U_i = [U_1, \dots, U_k] \in F_{m \times k}$ and $V_i = [V_1, \dots, V_k] \in F_{n \times k}$ is the replication of two vectors, element by element, and $1 \in F_k$ is a vector of ones to obtaining the output feature F_T using Equation (16). First, we need to get familiar with two three-order tensors, $U = [U_1, \dots, U_x] \in F_{m \times k \times x}$ and $V = [V_1, \dots, V_x] \in F_{n \times k \times x}$, which will serve as weights for the output dimension. A further transformation into two-dimensional matrices is possible, $U' \in F_{m \times k \times x}$ and $V' \in F_{n \times k \times x}$ after which it may be rewritten as Equation (17):

$$\begin{aligned} F_{TV} &= \text{Pooling} \left(U'^T F_{TV} V'^T F_V \right) \\ F_{TV} &= \text{Sign}(F_{TV}) |F_{TV}|^{0.5} \\ F_{TV} &= \frac{F^T T V}{|F_{TV}|} \end{aligned} \quad (17)$$

3.6. Multi-Layer Perceptron (MLP)

In this step, we develop a multi-layer perceptron consisting of hidden layers and a sigmoid-activated sub-network. The input to this multi-layer perceptron network is a fusion of features from MFB concatenated with similarity features, and user credibility features. The final prediction probability of whether or not a news item or post is fake is calculated by mapping the input onto an objective space comprising two classes, shown in Figure 5. A binary cross-entropy loss between the ground truth and the predictions is designed as the optimal solution. The letters L and P in the Equation (18) stand for the original class and the predicted class, respectively.

$$MLP = \sum_i L \log P + (i - L) \log (i - P) \quad (18)$$

4. Experiment and Parameter Setup

The models are built, trained, tested, and evaluated all inside the confines of the Google Colab environment. Python is utilized to execute all coding strategies. The proposed multimodal is evaluated using the k-fold strategy for cross-validation. The TensorFlow and scikit-learn libraries are used to create machine learning models. CountVectorizer and the NLTK library are utilized for text preparation.

The news article's accompanying image is used in conjunction with convolutional CapsNet to generate a visual feature vector. The recommended batch size for training a convolutional CapsNet is 32, and the recommended number of epochs is 100. We used eight child capsules in the Primary capsule layer and two-parent capsules in the Child Capsule Layer. The number of capsules and the complexity of intermediary capsule layers determine the significance of the routing-by-agreement approach. The overall number of hyperparameters will vary depending on them, but it will be less than CNN. The capsule

connections in the CapsNet model are established between groups of neurons as opposed to individual neurons; hence, it has fewer parameters than CNN. In comparison to CNN, the Convolutional CapsNet model needs the least amount of time to learn entire sequence data. To produce the 32-dimensional visual feature vector, F_{visual} , we evaluated and modified a higher-capsule layer.

The proposed model combines 32-dimensional textual and visual feature vectors using Factorized bilinear pooling to produce a 32-dimensional multimodal feature vector, F_{TV} , with high-level informative features. These multimodal features along with other important features are given into MLP, which is utilized to distinguish bogus and true news based on anticipated probability values.

4.1. Dataset

For our research, we used the publicly available standard fake news dataset called FakeNewsNet. It includes two datasets Gossipcop and Politifact, which comprise news stories about politics and entertainment, respectively. The performance of the proposed model is measured by its effectiveness on these two datasets. The collection consists of news stories, both text, and visuals. The details of various important aspects of datasets are provided in Table 1.

Table 1. Detailed information of the FakeNewsNet dataset.

Features of Datasets	GossipCop Fake	GossipCop Real	PolitiFact Fake	PolitiFact Real
Total number of news articles	6048	16,817	432	624
Related to text contents	785	16,765	353	400
Related to social interactions	4298	2902	342	314
Related to news content having social interactions	675	2895	286	202
Total number of tweets	71,009	154,383	116,005	261,262
Related to tweets having interaction	3040	2546	6686	20,720
Related to tweets having likes	10,685	2264	18,453	52,082
Related to tweets having retweets	7614	5025	13,226	42,059

4.2. Evaluation Metrics

We employed the standard set of performance measures, including accuracy, recall, precision, and f-measure. Furthermore, the challenge of establishing the veracity of a news item is modeled after a classification issue. Here's a quick rundown of what each metric measures from Equations (19)–(22):

$$Accuracy = \frac{T^+ + T^-}{T^+ + F^+ + T^- + F^-} \quad (19)$$

$$Precision = \frac{T^+}{T^+ + F^+} \quad (20)$$

$$Recall = \frac{T^+}{T^+ + F^-} \quad (21)$$

$$f - measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (22)$$

where, False positive (F^+) means that fake news was correctly identified as such, whereas false negative (F^-) means that real news was correctly identified as fake. The *Accuracy* value in the challenge of identifying false news indicates the proportion of news pieces that were properly labeled. *Accuracy* is measured by the percentage of anticipated false

news stories that were accurately labeled. By counting how many false news articles were accurately identified as such, we may determine the recall or true positive rate (TPR). The *f* – measure is the harmonic mean of the *Accuracy* and *Recall*, and it is used to indicate the overall performance of the proposed model.

5. Results and Discussions

In the preceding paragraphs, we detailed our findings from an in-depth analysis of the experimental outcomes of the proposed model utilizing various indicators for measuring performance.

To access the performance of the suggested model, it is put up against FakeNewsNet, a publicly available benchmark dataset. Table 2 displays the collected data. The experimental results show that the proposed model has better accuracy, precision, recall, and f-measure than the baseline and state-of-the-art methods.

Table 2. Comparison of performance of the proposed model with baseline models on the FakeNews-Net dataset.

Models	Accuracy	GossipCop Precision	Recall	f-Measure	Accuracy	PolitiFact Precision	Recall	f-Measure
NB [22]	0.627	0.794	0.913	0.852	0.616	0.762	0.874	0.814
SVM [14,29]	0.494	0.467	0.914	0.613	0.582	0.467	0.911	0.613
RF [18]	0.858	0.984	0.85	0.916	0.847	0.896	0.845	0.872
VGG19 [54,56–58,60]	0.803	0.795	0.793	0.802	0.654	0.647	0.649	0.653
EANN [58]	0.915	0.904	0.899	0.918	0.747	0.728	0.734	0.741
MVAE [54]	0.775	0.759	0.767	0.769	0.673	0.657	0.659	0.652
SpotFake [57]	0.807	0.798	0.802	0.805	0.721	0.718	0.719	0.728
SpotFake+ [56]	0.856	0.832	0.828	0.851	0.846	0.835	0.829	0.842
Proposed	0.988	0.985	0.966	0.975	0.990	0.979	1.000	0.989

When compared to the textual model, it is abundantly clear that the visual model is responsible for producing superior results. This may be due to the fact that texts might occasionally include noisy and unstructured information, but images display evidence more clearly. It is possible to conclude from the findings that combining pictures and text is advantageous since it achieves superior performance when compared to either using images or text alone.

Furthermore, the proposed multimodal provides a complete solution for the fake news detection in news articles, since information like reposts, likes, shares, etc. are not available immediately after a news article is published, the actual content of the article is of utmost significance. Then content can be the only factor examined for detecting fake news. The proposed model uses both the textual and visual aspects of news articles as its input. Cosine similarity of the title and body of the news provides a concrete measurement for comparison of relatedness. Additionally, the information included inside the user profile as well as the behavioral features of the user was added to improve the efficacy of the proposed model.

To highlight the value of cosine similarity as a feature, we have attempted encoding the news headline and body together with their degree of similarity and found that this method beats encoding only the title and content. Evidently, the findings show a significant improvement in the reliability of the tests. To test the quality of our model and ensure that its findings are equivalent to those of other models using the same dataset, we have implemented a 10-fold Cross-validation resampling technique. The average loss and accuracy based on epochs are shown in Figure 6. Even though we used k-fold stratified cross-validation, there were still some misclassified test samples. The main reason for this is that it is difficult to tell the two groups apart due to the features that they share.

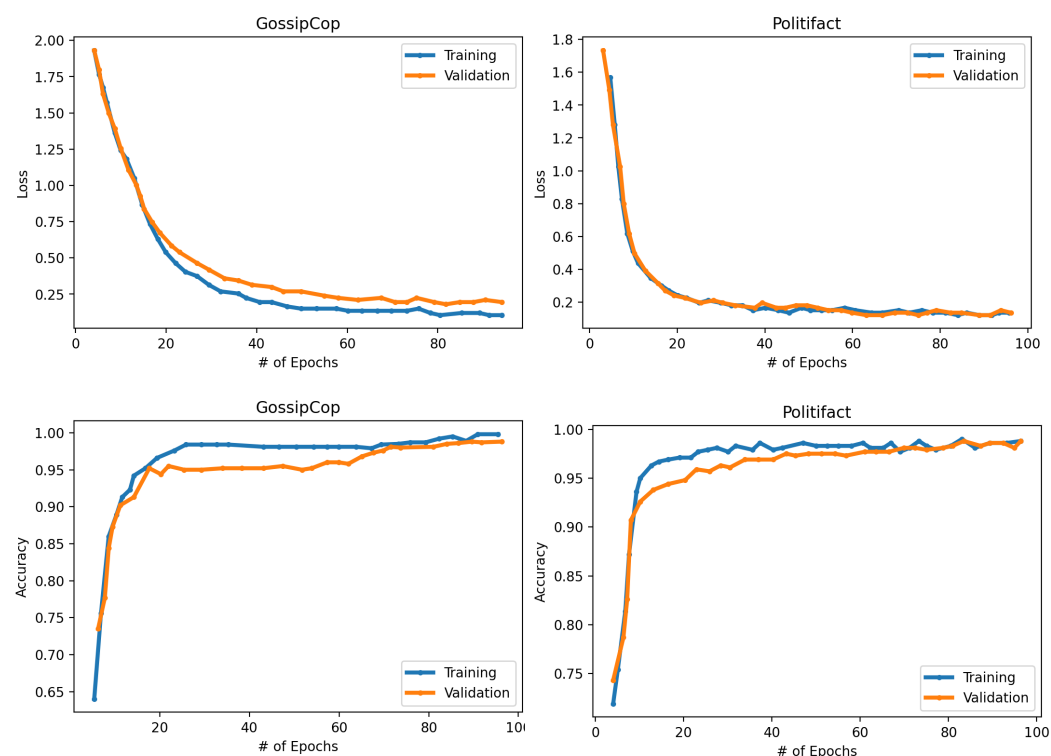


Figure 6. Average loss and accuracy for the proposed multimodal using FakeNewsNet.

As shown in Table 2, our proposed multimodal outperforms the state-of-the-art multimodal. The image features, cosine similarity, and routing-by-agreement method of the CapsNet architecture are crucial to the success of our suggested model. The accuracy improvement is also a reflection of the user credibility module's effectiveness. Despite the fact that textual characteristics are superior to visual features in unimodality mode, there are still some worries regarding textual features. Our suggested model achieves 7.3%, 21.5%, and 13.3% better performance than the current baseline models EANN [58], MVAE [54] and, SpotFake+ [56], respectively for the GossipCop dataset. Furthermore, for the PolitiFact dataset, our proposed model outperformed EANN [58], MVAE [54] and, SpotFake+ [56] with 24.5%, 32% and, 14.5% improved accuracy, respectively.

In the end, we also examined how well our suggested model performed in comparison to the most recent and cutting-edge techniques for identifying fake news. Table 3 represents the algorithms that were utilized for comparison.

The models compared here have the ability of early detection since they do not rely entirely on social interactions. OPCNN-FAKE combined the data from both sources into a single report. The outcomes in Table 4 show that the suggested model has the highest performance across all measures for both datasets. The comparison between the proposed multimodal and the state-of-the-art multimodal for the GossipCop and politifact datasets are shown in Figures 7 and 8, respectively.

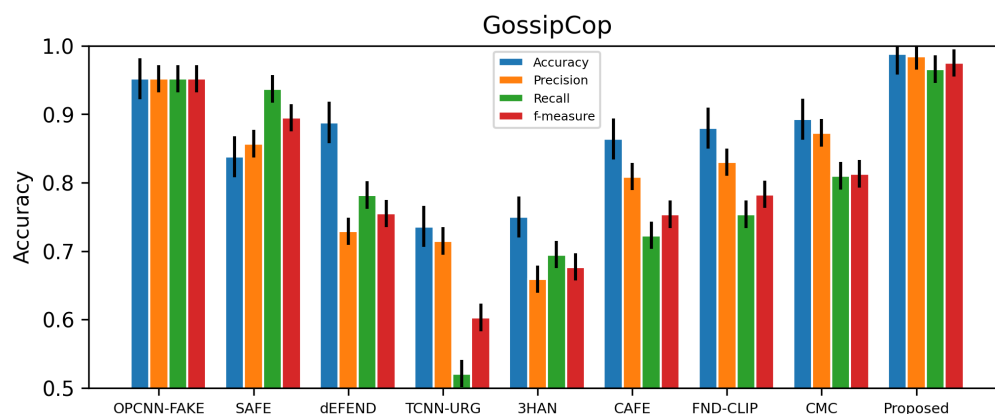
The proposed multimodal provides improved performance because we took the necessary steps to address the issues discovered by previous techniques. The combined feature representation can only be obtained using the approaches that are now considered state-of-the-art by concatenating textual and visual characteristics, which does not result in a strong connection between the picture and the text. In this work, features are extracted from a variety of models, and then those characteristics are combined to generate a common representation. In a later stage, the extra feature representations are enhanced by concatenating the additional features. We have conducted empirical research to explore and confirm the significance that pictures and social behavior play in the identification of false news. Figure 9 provides a selection of tweets and news stories that illustrate how well the suggested algorithm was able to classify their content.

Table 3. State of the art multimodals and their description.

Model	Description
OPCNN-FAKE [102]	It is a Convolutional Neural Network model that has been improved to detect fake news. The network parameters were optimized using grid search and hyperopic optimization methods.
SAFE [103]	It uses both textual and visual data from news articles. First, neural networks are utilized to separate visual and textual components of news coverage. The connection between the extracted attributes is then investigated across methods. Finally, a method for predicting fake news is developed by learning the correlation between textual and visual representations of news.
dEFEND [104]	Exploiting both news content and user comments, it collects the linked sentences and user remarks for fake news detection using a sentence-comment co-attention sub-network.
TCNNCURG [105]	When applied to text, the Two-Level Convolutional Neural Network with User Reaction Generator (TCNN-URG) builds a generative model of user response to news items based on past user replies while simultaneously capturing semantic information from the text at the sentence and word levels.
3HAN [106]	3HAN employs a hierarchical attention neural network architecture to analyze the textual contents of news articles to detect false news. It does this by encoding the textual contents using a hierarchical attention network that is composed of three levels: words, phrases, and headlines.
CAFE [48]	For news articles with appropriate image-text pairs, while learning to contrastively reduce the Kullback-Leibler (KL) divergence, variational autoencoders are trained to compress both the pictures and the texts. Multimodal characteristics are rebalanced based on the relevant cross-modal ambiguity score.
FND-CLIP [45]	The modal combines images and text from the news using the deep learning properties of text and images through the use of a ResNet-based encoder and a BERT-based encoder, respectively.
CMC [49]	This method employs a two-stage network, initially training two unimodal networks to learn cross-modal correlation via contrastive learning, and then fine-tuning the network to detect bogus news.

Table 4. Results of the proposed model compared to the most advanced multimodal.

Models	Accuracy	GossipCop Precision	Recall	f-Measure	Accuracy	PolitiFact Precision	Recall	f-Measure
OPCNN- FAKE [102]	0.952	0.952	0.952	0.952	0.952	0.952	0.952	0.952
SAFE [103]	0.838	0.857	0.937	0.895	0.874	0.889	0.903	0.896
dEFEND [104]	0.888	0.729	0.782	0.755	0.904	0.902	0.956	0.928
TCNN-URG [105]	0.736	0.715	0.521	0.603	0.712	0.711	0.941	0.810
3HAN [106]	0.750	0.659	0.695	0.677	0.844	0.825	0.899	0.860
CAFE [48]	0.864	0.809	0.723	0.754	0.867	0.809	0.848	0.828
FND-CLIP [45]	0.880	0.83	0.754	0.783	0.942	0.9285	0.9285	0.9285
CMC [49]	0.893	0.873	0.81	0.813	-	-	-	-
Proposed	0.988	0.985	0.966	0.975	0.990	0.979	1.000	0.989

**Figure 7.** Comparison between the proposed multimodal and the state-of-the-art multimodal for the GossipCop dataset.

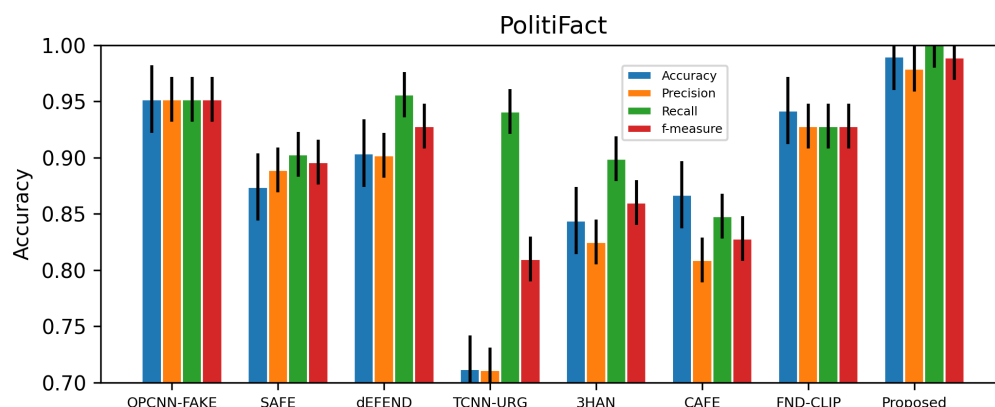


Figure 8. Comparison between the proposed multimodal and the state-of-the-art multimodal for the politifact dataset.



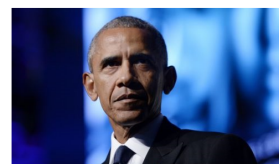
(a) HILLARY'S Campaign Likely FAKED Her Audience at NC Rally



(b) Anti-Trump Super PAC Launches Pre-Debate Video Featuring Hofstra Students



(c) DOJ: Chinese company tried to evade U.S. sanctions against North Korea



(d) Obama Claims It Will Be a Personal Insult If You Donate and Vote For Hillary

Figure 9. Examples of fake news correctly classified by the proposed model.

We have employed word, phrase, and document-level encoding for multilayer contextual information retrieval, which permits adjustable text length and simplifies the semantic encoder. In the instance of the visual encoder, we supply a six-layer convolutional network that is responsible for obtaining the most insightful insights and domain-specific characteristics. Some user-related qualities, such as cold start and unreliability, are particularly relevant in practical contexts. Because of the user's inexperience, very little information may be provided. In this research, we find that the cold start problem affects all of the attributes except for Credibility, Influence, and Sociality. It is not a major issue in the field of identifying fake news since content created by newcomers cannot be extensively disseminated on social media because of the absence of a considerable number of followers.

Furthermore, skepticism is crucial to uncovering fake news. This feature's unpredictability suggests it might be affected by the user's actions. It's possible that publishers will utilize this tactic to fool the system. Only the Sociality trait, out of all the ones we've studied, is suspect in this research. On the contrary, if a social influencer spreads misinformation or disinformation, it spreads quickly and widely. For this reason, we cannot recommend Sociality as a tool for identifying fake news. The median number of outlets sharing a given story shifted significantly among beats. There are more outlets that publish

political news than other types of news. Furthermore, while more outlets spread false celebrity news than fake political news, political fakery is produced by a smaller number of outlets. Accordingly, it's safe to say that publishers' online activities vary greatly depending on the type of news they're producing.

Ablation Study

The act of carefully assessing a framework in both the presence and absence of a certain component is referred to as an ablation study. This analysis is performed by individually removing and then grouping the framework's components. Identifying both the bottleneck and the unnecessary components lends a hand in the process of optimizing the design of the system. The ablation research is carried out to demonstrate the significance of the contributions made by each of the different modules as well as their level of efficacy. Text features, cosine similarity features, user trustworthiness features, and image features are included in the multi-modal that is being suggested. Experiments are being conducted with individual approaches, ensembles of two modules using the FakeNewsnet dataset with the same parameter settings as the overall proposed framework, shown in Figures 10 and 11.

Assessing the relative merits of alternative component arrangements within the framework, we made use of the top-1, top-5, and top-10 accuracies, as well as the reciprocal average rank (RAR) metric. The top-K accuracy measures how accurate the top-k projected scores are by calculating the percentage of correct labels within those scores, presented in Table 5. The RAR offers information on how far down the list the correct label is located.

Table 5. Results for the ablation study.

Models	Top-1	GossipCop			Top-1	PolitiFact		
		Top-5	Top-10	RAR		Top-5	Top-10	RAR
Image only	0.501	0.584	0.648	0.428	0.594	0.652	0.674	0.487
Text only	0.428	0.511	0.568	0.357	0.467	0.548	0.597	0.413
Credibility + Text	0.53	0.664	0.718	0.461	0.598	0.678	0.734	0.509
Similarity + Text	0.52	0.692	0.746	0.464	0.663	0.742	0.778	0.452
Credibility + Image	0.641	0.762	0.841	0.621	0.749	0.824	0.864	0.642
Proposed	0.727	0.951	0.988	0.791	0.751	0.967	0.990	0.826

Since we've taken the appropriate steps to address the issues uncovered by earlier methodologies, the ablation investigation also demonstrates the effectiveness of the proposed framework. picture and text features that have been obtained to maximize similarities and give a more reliable common representation. We can achieve this using the Multimodal Factorized Bilinear-pooling (MFB) technique. Furthermore, the proposed model integrates semantically significant characteristics, the cosine similarity perspective, and social context information to generate a better feature vector representation for the provided news, which in turn improves the overall effectiveness of the identification of fake news. Furthermore, the success of the user credibility module is seen in the increased precision.

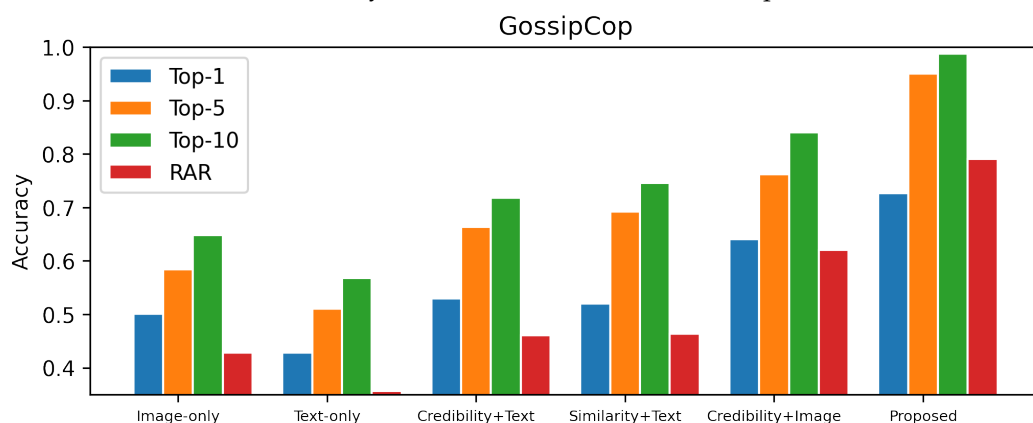


Figure 10. Comparison based on the ablation study for the GossipCop dataset.

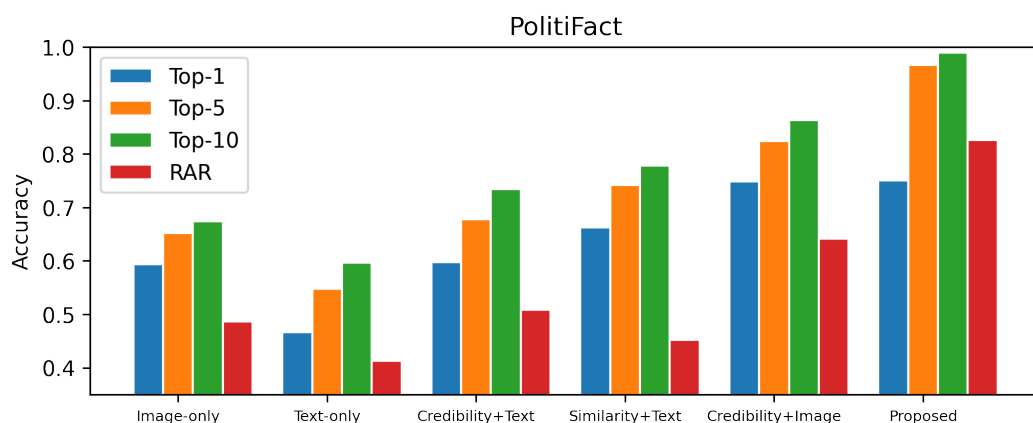


Figure 11. Comparison based on the ablation study for PolitiFact dataset.

This research provides a theoretical account of the steps involved in news classification using processed data, including the extraction of essential features from several modalities, the influence of social context and similarity characteristics, and the fusing of features to generate a common representation. We have conducted empirical research to verify the importance of cosine similarity in identifying fabricated articles. Second, our results provide light on hitherto unrecognized aspects of false news concerning social profiling and online behavior. Every one of these discoveries adds to the body of theoretical information on the subject.

The multimodal approach has several advantages, including the fact that it does not rely on a single data source, which is especially helpful in the case of the early identification of fake news on social media to halt the spread of disinformation. In its earliest stages, it just requires text and images as input, and based on these basic inputs, it derives the semantic and visual essential elements necessary to form a robust correlation. To further forecast whether or not a piece of news is true, it incorporates the cosine similarity properties. Based on our research results, multimodality is an effective technique for detecting bogus news. In this research, we show how to put a deep learning-based multimodal false news detection framework into practice.

In this study, we have developed a multimodal approach that considers the most essential information sources and extraction procedures for detecting fake news. In addition, we have resolved the issues of the current state of the art. However, our model also has some limitations: it does not support languages other than English because it has not been tested and calibrated for other languages. Due to the high association between visuals and accompanying language, complex and altered images with matching text descriptions can occasionally trick the framework. We could address these restrictions to generate a significant, long-lasting influence on the propagation and early identification of fake news. We can also add forgery detection, non-English languages, and meta information that may have a substantial impact on fake news detection. However, it requires the development of a suitable dataset, models, and experimental framework.

6. Conclusions

In this study, we created a methodology for identifying false news stories by combing through the piece's textual, contextual, social, and visual elements. Existing models for the fake news detection problem suffer from serious shortcomings due to the inability to acquire meaningful details from the text and its related images of news articles and social media posts. The proposed approach addresses this problem by combining textual, contextual, social, and visual data to learn a more accurate multimodal feature representation. In the proposed model, CapsNet is used to extract the most informative visual features from the image. It uses a BRNN with attention to extract linguistic aspects from texts as well. The cosine similarity between the headline and body of news articles is also calculated. In

addition, the user credibility module determines the user's relative social status. Visual and textual characteristics have been integrated using multi-modal factorized bilinear pooling, a common data representation has been generated, and further concatenated similarity and credibility features. Finally, the output is submitted to a multi-layer perceptron for classification as real or fake news. The effectiveness of the proposed model is measured with the help of the FakeNewsNet dataset, which is open to the public. The datasets come from the same sources as the news and social media sites, and they are called GossipCop and PolitiFact. Compared to other multimodal false news detection models, the suggested approach performs better in experiments. The results of this study suggest other research avenues worth exploring. The features solution domain may be successfully expanded by the extraction of various picture characteristics, which can then be used in the study of social media and news articles in the pursuit of false news identification.

Author Contributions: Conceptualization, M.I.N. and K.A.; methodology, M.I.N. and K.A.; software, H.K.A.; validation, D.L., S.M.M. and Z.Z.; formal analysis, K.A. and H.A.H.; investigation, D.L. and S.M.M.; resources, Z.Z.; data curation, H.A.H.; writing—original draft preparation, M.I.N.; writing—review and editing, K.A. and Z.Z.; visualization, M.I.N. and H.K.A.; supervision, D.L.; project administration, O.M.; funding acquisition, O.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This research has been funded by the Science Committee of the Ministry of Education and Science of the Republic Kazakhstan (Grant No. AP09259309).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MDPI	Multidisciplinary Digital Publishing Institute
NLP	Natural Language Processing
ML	Machine Learning
DL	Deep Learning
Conv	Convolutional Layer
GRU	Gated Recurrent Units
BiGRU	Bidirectional GRU
CapsNet	Capsule Neural Network
LSTM	Long Short-Term Memory
MFB	Multimodal Factorized Bilinear-pooling
MLP	Multilayer Perceptron

References

1. Bondielli, A.; Marcelloni, F. A survey on fake news and rumour detection techniques. *Inf. Sci.* **2019**, *497*, 38–55. [\[CrossRef\]](#)
2. Guo, B.; Ding, Y.; Yao, L.; Liang, Y.; Yu, Z. The Future of False Information Detection on Social Media: New Perspectives and Trends. *ACM Comput. Surv. (CSUR)* **2020**, *53*, 1–36. [\[CrossRef\]](#)
3. Shu, K.; Sliva, A.; Wang, S.; Tang, J.; Liu, H. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explor. Newsl.* **2017**, *19*, 22–36. [\[CrossRef\]](#)
4. Zhou, X.; Zafarani, R.; Shu, K.; Liu, H. Fake news: Fundamental theories, detection strategies and challenges. In Proceedings of the twelfth ACM International Conference on Web Search and Data Mining, Melbourne, VIC, Australia, 11–15. February 2019; pp. 836–837.
5. Zhou, X.; Jain, A.; Phoha, V.V.; Zafarani, R. Fake news early detection: A theory-driven model. *Digit. Threat. Res. Pract.* **2020**, *1*, 1–25. [\[CrossRef\]](#)

6. Guo, C.; Cao, J.; Zhang, X.; Shu, K.; Liu, H. Dean: Learning dual emotion for fake news detection on social media. *arXiv* **2019**, arXiv:1903.01728.
7. Liu, Y.; Wu, Y.F. Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks. In Proceedings of the AAAI conference on artificial intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32.
8. Li, Q.; Zhang, Q.; Si, L. Rumor detection by exploiting user credibility information, attention and multi-task learning. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July–2 August 2019; pp. 1173–1179.
9. Ma, J.; Gao, W.; Wong, K.F. Detect rumor and stance jointly by neural multi-task learning. In Proceedings of the Companion Proceedings of the the Web Conference, Lyon, France, 23–27 April 2018; pp. 585–593.
10. Ma, J.; Gao, W.; Wong, K.F. *Rumor Detection on Twitter with Tree-Structured Recursive Neural Networks*; Association for Computational Linguistics: Melbourne, Australia, 2018.
11. Savyan, P.V.; Bhanu, S.M.S. UbCadet: Detection of compromised accounts in twitter based on user behavioural profiling. *Multimed. Tools Appl.* **2020**, *79*, 1–37.
12. Wu, L.; Liu, H. Tracing fake-news footprints: Characterizing social media messages by how they propagate. In Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, Marina Del Rey, CA, USA, 5–9 February 2018; pp. 637–645.
13. Wu, K.; Yang, S.; Zhu, K.Q. False rumors detection on sina weibo by propagation structures. In Proceedings of the 2015 IEEE 31st International Conference on Data Engineering, Seoul, Republic of Korea, 13–17 April 2015; pp. 651–662.
14. Ahmed, H.; Traore, I.; Saad, S. Detection of online fake news using n-gram analysis and machine learning techniques. In Proceedings of the International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, Vancouver, BC, Canada, 25 October 2017; Springer: Cham, Switzerland, 2017; pp. 127–138.
15. Akyol, K.; Sen B. Modeling and predicting of news popularity in social media sources. *Cmc-computers Mater. Contin.* **2019**, *61*, 69–80. [\[CrossRef\]](#)
16. Asghar, M.Z.; Habib, A.; Habib, A.; Khan, A.; Ali, R.; Khattak, A. Exploring deep neural networks for rumor detection. *J. Ambient. Intell. Hum. Comput.* **2019**, *12*, 4315–4333. [\[CrossRef\]](#)
17. Chen, T.; Li, X.; Yin, H.; Zhang, J. Call attention to rumors: Deep attention based recurrent neural networks for early rumor detection. In Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, Melbourne, VIC, Australia, 3–6 June 2018; Springer: Cham, Switzerland, 2018; pp. 40–52.
18. Faustini, P.H.A.; Covões, T.F. Fake news detection in multiple platforms and languages. *Expert Syst. Appl.* **2020**, *158*, 113503. [\[CrossRef\]](#)
19. Kaliyar, R.K.; Goswami, A.; Narang P. FakeBERT: Fake news detection in social media with a BERT-based deep learning approach. *Multimed. Tools Appl.* **2021**, *80*, 11765–11788. [\[CrossRef\]](#)
20. Ma, J.; Gao, W.; Mitra, P.; Kwon, S.; Jansen, B.J.; Wong, K.F.; Cha, M. Detecting rumors from microblogs with recurrent neural networks. In Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI 2016), New York, NY, USA, 9–15 July 2016; pp. 3818–3824.
21. Ozbay, F.A.; Alatas, B. A novel approach for detection of fake news on social media using metaheuristic optimization algorithms. *Elektron. Elektrotehnika* **2019**, *25*, 62–67. [\[CrossRef\]](#)
22. Ozbay, F.A.; Alatas, B. Fake news detection within online social media using supervised artificial intelligence algorithms. *Phys. A Stat. Mech. Its Appl.* **2020**, *540*, 123174. [\[CrossRef\]](#)
23. Pérez-Rosas, V.; Kleinberg, B.; Lefevre, A.; Mihalcea, R. Automatic detection of fake news. *arXiv* **2017**, arXiv:1708.07104.
24. Shu, K.; Cui, L.; Wang, S.; Lee, D.; Liu, H. Defend: Explainable fake news detection. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 95–405.
25. Wang, W.Y. Liar, liar pants on fire: A new benchmark dataset for fake news detection. *arXiv* **2017**, arXiv:1705.00648.
26. Yin, L.; Meng, X.; Li, J.; Sun J. Relation extraction for massive news texts. *Comput. Mater Contin.* **2019**, *58*, 275–285. [\[CrossRef\]](#)
27. Yu, F.; Liu, Q.; Wu, S.; Wang, L.; Tan, T. A convolutional approach for misinformation identification. In Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI 2017), Melbourne, Australia, 19–25 August *IJCAI* **2017**, 3901–3907.
28. Antol, S.; Agrawal, A.; Lu, J.; Mitchell, M.; Batra, D.; Zitnick, C.L.; Parikh, D. VQA: Visual question answering. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 2425–2433.
29. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, V. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 2672–2680.
30. Gupta, M.; Zhao, P.; Han, J. Evaluating event credibility on twitter. In Proceedings of the 2012 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics, Anaheim, CA, USA, 26–28 April 2012; pp. 153–164.
31. Marra, F.; Gragnaniello, D.; Cozzolino, D.; Verdoliva, L. Detection of gan-generated fake images over social networks. In Proceedings of the 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Miami, FL, USA, 10–12 April 2018; pp. 384–389.
32. Qi, P.; Cao, J.; Yang, T.; Guo, J.; Li, J. Exploiting multi-domain visual information for fake news detection. In Proceedings of the 2019 IEEE International Conference on Data Mining (ICDM), Beijing, China, 8–11 November 2019; pp. 518–527.

33. Zeng, J.; Ma, X.; Zhou, K. Photo-realistic face age progression/regression using a single generative adversarial network. *Neurocomputing* **2019**, *366*, 295–304. [\[CrossRef\]](#)
34. Zhou, P.; Han, X.; Morariu, V.I.; Davis, L.S. Learning rich features for image manipulation detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp 1053–1061.
35. Varshney, D.; Vishwakarma, D.K. A unified approach for detection of Clickbait videos on YouTube using cognitive evidences. *Appl. Intell.* **2021**, *51*, 4214–4235. [\[CrossRef\]](#)
36. Mishra, S.; Shukla, P.; Agarwal, R. Analyzing machine learning enabled fake news detection techniques for diversified datasets. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 1–18. [\[CrossRef\]](#)
37. Koloski, B.; Perdihi, T.S.; Robnik-Šikonja, M.; Pollak, S.; Škrlj, B. Knowledge graph informed fake news classification via heterogeneous representation ensembles. *Neurocomputing* **2022**, *496*, 208–226. [\[CrossRef\]](#)
38. Abdelnabi, S.; Hasan, R.; Fritz, M. OpenDomain, Content-based, Multi-modal Fact-checking of Outof-Context Images via Online Resources. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–20 June 2022; pp. 14940–14949.
39. Sun, M.; Zhang, X.; Zheng, J.; Ma, G. DDGCN: Dual Dynamic Graph Convolutional Networks for Rumor Detection on Social Media. In Proceedings of the AAAI Conference on Artificial Intelligence, 22 February–1 March 2022; pp 4611–4619.
40. Chen, X.; Zhou, F.; Trajcevski, G.; Bonsangue, M. Multi-view learning with distinguishable feature fusion for rumor detection. *Knowl. Based Syst.* **2022**, *240*, 108085. [\[CrossRef\]](#)
41. Chi, H.; Liao, B. A quantitative argumentation-based Automated eXplainable Decision System for fake news detection on social media. *Knowl. Based Syst.* **2022**, *242*, 108378. [\[CrossRef\]](#)
42. Raza, S.; Ding, C. Fake news detection based on news content and social contexts: A transformer-based approach. *Int. J. Data Sci. Anal.* **2022**, *13*, 335–362. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Jarrahi, A.; Safari, L. Evaluating the effectiveness of publishers’ features in fake news detection on social media. *Multimed. Tools Appl.* **2022**, 1–27. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Wang, Y.; Wang, L.; Yang, Y.; Zhang, Y. Detecting fake news by enhanced text representation with multi-EDUstructure awareness. *arXiv* **2022**, arXiv:2205.15139.
45. Zhou, Y.; Ying, Q.; Qian, Z.; Li, S.; Zhang, X. Multimodal Fake News Detection via CLIP-Guided Learning. *arXiv* **2022**, arXiv:2205.14304.
46. Wang, J.; Mao, H.; Li, H. FMFN: Fine-Grained Multimodal Fusion Networks for Fake News Detection. *Appl. Sci.* **2022**, *12*, 1093. [\[CrossRef\]](#)
47. Singhal, S.; Pandey, T.; Mrig, S.; Shah, R.R.; Kumaraguru, P. Leveraging Intra and Inter Modality Relationship for Multimodal Fake News Detection. In Proceedings of the Web Conference, Virtual Event, Lyon France, 25–29 April 2022; pp. 726–734.
48. Chen, Y.; Li, D.; Zhang, P.; Sui, J.; Lv, Q.; Tun, L.; Shang, L. Cross-modal Ambiguity Learning for Multimodal Fake News Detection. In Proceedings of the ACM Web Conference, Lyon, France, 25–29 April 2022; pp. 2897–2905.
49. Wei, Z.; Pan, H.; Qiao, L.; Niu, X.; Dong, P.; Li, D. Cross-Modal Knowledge Distillation in Multi-Modal Fake News Detection. In Proceedings of the ICASSP 2022—2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 23–27 May 2022; pp. 4733–4737.
50. Das, S.D.; Basak, A.; Dutta, S. A heuristic-driven uncertainty based ensemble framework for fake news detection in tweets and news articles. *Neurocomputing* **2022**, *491*, 607–620. [\[CrossRef\]](#)
51. Davoudi, M.; Moosavi, M.R.; Sadreddini, M.H. DSS: A hybrid deep model for fake news detection using propagation tree and stance network. *Exp. Syst. Appl.* **2022**, *198*, 116635. [\[CrossRef\]](#)
52. Segura-Bedmar, I.; Alonso-Bartolome, S. Multimodal fake news detection. *Information* **2022**, *13*, 284. [\[CrossRef\]](#)
53. Karpathy, A.; Li, F.F. Deep visual-semantic alignments for generating image descriptions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3128–3137.
54. Khattar, D.; Goud, J.S.; Gupta, M.; Varma, V. MVAE: Multimodal variational autoencoder for fake news detection. In Proceedings of the The World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; pp. 2915–2921.
55. Singh, S.; Cha, J.; Kim, T.W.; Park, J. Machine learning based distributed big data analysis framework for next generation web in IoT. *Comput. Sci. Inf. Syst.* **2021**, *18*, 597–618. [\[CrossRef\]](#)
56. Singhal, S.; Kabra, A.; Sharma, M.; Shah, R.R.; Chakraborty, T.; Kumaraguru, P. Spofake+: A multimodal framework for fake news detection via transfer learning (student abstract). *Proc. AAAI Conf. Artificial Intell.* **2020**, *34*, 13915–13916. [\[CrossRef\]](#)
57. Singhal, S.; Shah, R.R.; Chakraborty, T.; Kumaraguru, P.; Satoh, S.I. Spofake: A multi-modal framework for fake news detection. In Proceedings of the 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM), Singapore, 11–13 September 2019; pp. 39–47.
58. Wang, Y.; Ma, F.; Jin, Z.; Yuan, Y.; Xun, G.; Jha, K.; Su, L.; Gao, J. EANN: Event adversarial neural networks for multi-modal fake news detection. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 August 2018; p. 849.
59. Yang, K.; Long, S.; Zhang, W.; Yao, J.; Liu, J. Personalized News Recommendation Based on the Text and Image Integration. *CMC-Comput. Mater. Contin.* **2020**, *64*, 557–570. [\[CrossRef\]](#)
60. Yang, Y.; Zheng, L.; Zhang, J.; Cui, Q.; Li, Z.; Yu, P.S. TI-CNN: Convolutional neural networks for fake news detection. *arXiv* **2018**, arXiv:1806.00749.

61. Choudhary, A.; Arora, A. Linguistic feature based learning model for fake news detection and classification. *Expert Syst. Appl.* **2021**, *169*, 114171. [\[CrossRef\]](#)
62. Potthast, M.; Kiesel, J.; Reinartz, K.; Bevendorff, J.; Stein, B. A stylometric inquiry into hyperpartisan and fake news. In Proceedings of the ACL 2018—56th Annual Meeting of the Association for Computational Linguistics, Melbourne, Australia, 15–20 July 2018; Volume 1, pp. 231–240. [\[CrossRef\]](#)
63. Long, Y.; Lu, Q.; Xiang, R.; Li, M.; Huang, C.-R. Fake news detection through multi-perspective speaker profiles. *Proc. Eighth Int. Jt. Conf. Nat. Lang. Process.* **2017**, *2*, 252–256.
64. Song, C.; Ning, N.; Zhang, Y.; Wu, B. A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks. *Inf. Process. Manag.* **2021**, *58*, 102437. [\[CrossRef\]](#)
65. Girgis, S.; Gadallah, M. Deep Learning Algorithms for Detecting Fake News in Online Text. In Proceedings of the 13th International Conference on Computer Engineering and Systems (ICCES), Cairo, Egypt, 18–19 December 2018; pp. 93–97. [\[CrossRef\]](#)
66. Hosseinimotlagh, S.; Papalexakis, E.E. Unsupervised Content-Based Identification of Fake News Articles with Tensor Decomposition Ensembles. In Proceedings of the Workshop on Misinformation and Misbehavior Mining on the Web (MIS2), Los Angeles, CA, USA, 9 February 2018.
67. Kaliyar, R.K.; Goswami, A.; Narang, P. DeepFake: Improving fake news detection using tensor decomposition-based deep neural network. *J. Supercomput.* **2021**, *77*, 1015–1037. [\[CrossRef\]](#)
68. Gupta, S.; Thirukovalluru, R.; Sinha, M.; Mannarswamy, S. CIMT Detect: A community infused matrix-tensor coupled factorization based method for fake news detection. In Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Barcelona, Spain, 28–31 August 2018. [\[CrossRef\]](#)
69. Liang, G.; He, W.; Xu, C.; Chen, L.; Zeng, J. Rumor identification in microblogging systems based on users' behavior. *IEEE Trans. Comput. Soc. Syst.* **2015**, *2*, 99–108. [\[CrossRef\]](#)
70. Vishwakarma, D.K.; Varshney, D.; Yadav, A. Detection and veracity analysis of fake news via scrapping and authenticating the web search. *Cogn. Syst. Res.* **2019**, *58*, 217–229. [\[CrossRef\]](#)
71. Meel, P.; Vishwakarma, D.K. HAN, image captioning, and forensics ensemble multimodal fake news detection. *Inf. Sci.* **2021**, *567*, 23–41. [\[CrossRef\]](#)
72. Meel, P.; Vishwakarma, D.K. A temporal ensembling based semi-supervised ConvNet for the detection of fake news articles. *Expert Syst. Appl.* **2021**, *177*, 115002. [\[CrossRef\]](#)
73. Varshney, D.; Vishwakarma, D.K. Hoax news-inspector: A real-time prediction of fake news using content resemblance over web search results for authenticating the credibility of news articles. *J. Ambient. Intell. Hum. Comput.* **2021**, *12*, 8961–8974. [\[CrossRef\]](#)
74. Vo, N.; Lee, K. The rise of guardians: Fact-checking URL recommendation to combat fake news. In Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval, Ann Arbor, MI, USA, 8–12 July 2018. [\[CrossRef\]](#)
75. Passant, A.; Kärgner, P.; Hausenblas, M.; Olmedilla, D.; Polleres, A.; Decker, S. Enabling trust and privacy on the social web. In Proceedings of the W3C workshop on the future of social networking, Barcelona, Spain, 15–16 January 2009; World Wide Web Consortium (W3C): Cambridge, MA, USA, 2009.
76. Majumder, N.; Poria, S.; Gelbukh, A.; Cambria, E. Deep learning-based document modeling for personality detection from text. *IEEE Intell. Syst. Appl.* **2017**, *32*, 74–79. [\[CrossRef\]](#)
77. Weng, J.; Lim, E.P.; Jiang, J.; He, Q. TwitterRank: Finding topic-sensitive influential twitterers. In Proceedings of the 3rd International Conference on Web Search and Data Mining, New York, NY, USA, 4–6 February 2010.
78. Silva, A.; Guimarães, S.; Meira, W. Jr.; Zaki, M. ProfileRank: Finding relevant content and influential users based on information diffusion. In Proceedings of the 7th workshop on social network mining and analysis, Chicago, IL, USA, 11–14 August 2013.
79. Yeniterzi, R.; Callan, J. Constructing effective and efficient topic-specific authority networks for expert finding in social media. In Proceedings of the 1st international workshop on social media retrieval and analysis, Gold Coast, QLD, Australia, 11 July 2014.
80. Kwak, H.; Lee, C.; Park, H.; Moon, S. What is Twitter, a social network or a news media? In Proceedings of the 19th international conference on World Wide Web, Raleigh, NC, USA, 26–30 April 2010.
81. Tsolmon, B.; Lee, K.-S. A graph-based reliable user classification. In Proceedings of the 1st International Conference on Advanced Data and Information Engineering (DaEng-2013): Lecture Notes in Electrical Engineering, Kuala Lumpur, Malaysia, 16–18 December 2013; Herawan, T., Deris, M., Abawajy, J., Eds.; Springer: Berlin, Germany, 2014; Volume 285, pp. 61–68.
82. Agarwal, M.; Zhou, B. Detecting malicious activities using backward propagation of trustworthiness over heterogeneous social graph. In Proceedings of the IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), Atlanta, GA, USA, 17–20 November 2013.
83. Podobnik, V.; Striga, D.; Jandras, A.; Lovrek, I. How to calculate trust between social network users? In Proceedings of the 20th international conference on software, telecommunications and computer networks (SoftCOM), Split, Croatia, 11–13 September 2012.
84. Sikdar, S.; Kang, B.; O'Donovan, J.; Höllerer, T.; Adah, S. Understanding information credibility on Twitter. In Proceedings of the International Conference on Social Computing, Alexandria, VA, USA, 8–14 September 2013.
85. Wu, H.; Arenas, A.; Gómez, S. Influence of trust in the spreading of information. *Phys. Rev. E* **2017**, *95*, 012301. [\[CrossRef\]](#)
86. Sherchan, W.; Nepal, S.; Paris, C. A survey of trust in social networks. *ACM Comput. Surv.* **2013**, *45*, 47. [\[CrossRef\]](#)
87. Aghdam, M.H.; Analoui, M.; Kabiri, P. Modelling trust networks using resistive circuits for trust-aware recommender systems. *J. Inf. Sci.* **2017**, *43*, 135–144. [\[CrossRef\]](#)

88. Al-Qurishi, M.; Rahman, S.M.M.; Alamri, A.; Mostafa, M.A.; Al-Rubaian, M.; Hossain, M.S.; Gupta, B.B. SybilTrap: A graph-based semi-supervised Sybil defense scheme for online social networks. *Concurr. Comp. Pract. E* **2018**, *30*, e4276. [\[CrossRef\]](#)
89. Kožuh, I.; Čakš, P. Explaining News Trust in Social Media News during the COVID-19 Pandemic-The Role of a Need for Cognition and News Engagement. *Int. J. Environ. Res. Public Heal.* **2021**, *18*, 12986. [\[CrossRef\]](#)
90. Sabour, S.; Frosst, N.; Hinton, G.E. Dynamic routing between capsules. *arXiv* **2017**, arXiv:1710.09829.
91. Dombetzki, L.A. An overview over capsule networks. In Proceedings of the Seminars Future Internet (FI) and Innovative Internet Technologies and Mobile Communication (IITM), Munich, Germany, 26 February–19 August 2018; pp. 89–95. [\[CrossRef\]](#)
92. Sezer, A.; Sezer, H.B. Capsule network-based classification of rotator cuff pathologies from MRI. *Comput. Electr. Eng.* **2019**, *80*, 106480. [\[CrossRef\]](#)
93. Lukic, V.; Brüggem, M.; Mingo, B.; Croston, J.H.; Kasieczka, G.; Best P.N. Morphological classification of radio galaxies- capsule networks versus CNN. *Mon. Not. R. Astron. Soc.* **2019**, *487*, 1729–1744. [\[CrossRef\]](#)
94. Beser, F.; Kizrak, M.A.; Bolat, B.; Yildirim, T. Recognition of sign language using capsule networks. In Proceedings of the 26th Signal Processing and Communications Applications Conference, Izmir, Turkey, 2–5 May 2018. [\[CrossRef\]](#)
95. Xu, Z.; Lu, W.; Zhang, Q.; Yeung, Y.; Chen, X. Gait recognition based on capsule network. *J. Vis. Commun. Image Represent.* **2019**, *59*, 159–167. [\[CrossRef\]](#)
96. Zhang, X.Q.; Zhao, S.G. Cervical image classification based on image segmentation preprocessing and a CapsNet network model. *Int. J. Imaging Syst. Technol.* **2018**, *29*, 19–28. [\[CrossRef\]](#)
97. Schuster, M.; Paliwal, K.K. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* **1997**, *45*, 2673–2681. [\[CrossRef\]](#)
98. Bahdanau, D.; Cho, K.; Bengio, Y. Neural Machine Translation by Jointly Learning to Align and Translate. *arXiv* **2014**, arXiv:1409.0473.
99. Sukhbaatar, S.; Weston, J.; Fergus, R. End-to-end memory networks. *Adv. Neural Inf. Process. Syst. (NeurIPS)* **2015**, *28*, 2440–2448.
100. Kumar, A.; Irsoy, O.; Ondruska, P.; Iyyer, M.; Bradbury, J.; Gulrajani, I.; Zhong, V.; Paulus, R.; Socher, R. Ask me anything: Dynamic memory networks for natural language processing. In Proceedings of the 33rd International Conference on Machine Learning (ICML), New York, NY, USA, 20–22 June 2016; pp. 1378–1387.
101. Dong, M.; Yao, L.; Wang, X.; Benatallah, B.; Huang, C. Similarity-aware deep attentive model for clickbait detection. In Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, Macau, China, 14–17 April 2019; Springer: Cham, Switzerland, 2019; pp. 56–69.
102. Saleh, H.; Alharbi, A.; Alsamhi, S.H. OPCNN-FAKE: Optimized convolutional neural network for fake news detection. *IEEE Access* **2021**, *9*, 129471–129489. [\[CrossRef\]](#)
103. Zhou, X.; Wu, J.; Zafarani, R. SAFE: Similarity-aware multi-modal fake news detection. *arXiv* **2020**, arXiv:2003.04981.
104. Cui, L.; Shu, K.; Wang, S.; Lee, D.; Liu, H. dEFEND: A system for explainable fake news detection. In Proceedings of the 28th ACM international conference on information and knowledge management, Beijing, China, 3–7 November 2019.
105. Qian, F.; Gong, C.; Sharma, K.; Liu, Y. Neural User Response Generator: Fake News Detection with Collective User Intelligence. *IJCAI* **2018**, *18*, 3834–3840.
106. Singhanian, S.; Fernandez, N.; Rao, S. 3HAN: A deep neural network for fake news detection. in nternational Conference on Neural Information Processing. In Proceedings of the International Conference on Neural Information Processing (ICONIP 2017), Long Beach, CA, USA, 4–9 December 2017.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.