
# A Hybrid Temporal Lambda Layer Embedded in Autoencoder Neural Network for Fake News Detection

**T. V. Divya[1,*], Figlu Mohanty[2]**

[1]Research Scholar, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Aziz Nagar, Hyderabad, Telangana, India
[2]Associate Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Aziz Nagar, Hyderabad, Telangana, India
Email: md2829@gmail.com; figlu@klh.edu.in

**Abstract**

Many social media applications use different animated or morphed images to make fake news viral. Recognition of text from images for their classification as real or fake requires a neural network. BERT (Bidirectional Encoder Representation Transformer) or MLP-based (Multi-Layer Perceptron) algorithms are successful when working with textual data alone. However, the system needs to extract the sequential text from the images to identify the semantic meaning of the content before the classification process. The dataset utilized was acquired from The Indian Fake News Dataset (IFND) contains text and visual data from 2013 to 2021. The data includes both visual and textual information, as well as 126k data points obtained from millions of users. In the proposed model, a squeezed lambda is implemented to process the data in the three forms of verbal tenses, i.e., past to future and future to past. In the lambda layer, temporal classification is performed by applying two bidirectional LSTM (Long Short Term Memory) layers based on the retuning sequences of the character list available in the dataset. It also computes the batch cost of every iteration and reduces them based on the ratio of prediction and input class labels available. To ensure that the suggested technique is more accurate than the current approach, a validation was undertaken, resulting in a +0.5 increase in accuracy over the BERT (Bidirectional Encoder Representation Transformer) model. Hence, the proposed method has achieved higher accuracy than existing algorithms. Than existing algorithms.

**Keywords:** Lambda Layer; Temporal Classification; Sequential Data; Encoders; Decoders; Sigmoid

## 1. Introduction

Identification of fake news based on the images is the complicated task. CNN (Convolution Neural Network)s (Convolution Neural Network)are popular for working with images but the dataset in this research contains textual information in the form of images. So, the major task of this research is to extract the text from images, which is achieved using the LSTM (Long Short-Term Memory) (Long Short Term Memory) networks in the previous works but in textual extraction to identify the sequence in between the words and to preserve the context of sentence, it is better to implement bi-directional LSTM (Long Short-Term Memory) instead of traditional techniques. In recent trends, many researchers used Bi-LSTM approaches, which are inefficient because of the huge number of layers involved in it. With the increase in the sentences the layers in the network also increases. The proposed research decreases the number of layers by replacing the dense layers with lambda layers. Traditional approaches use CNN (Convolution Neural Network) to identify fake news based on the text extracted from the images, but more layers are



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

required for implementation. Since every image contains textual information, high computational resources are needed to predict fake news content from images [22]. In the proposed architecture, instead of alone CNN (Convolution Neural Network), the model has implemented other layers and reduced the number of layers by implementing the power layers known as "Bidirectional" and "Lambda". Figure 1 has presented all the layers implemented in the proposed methodology.

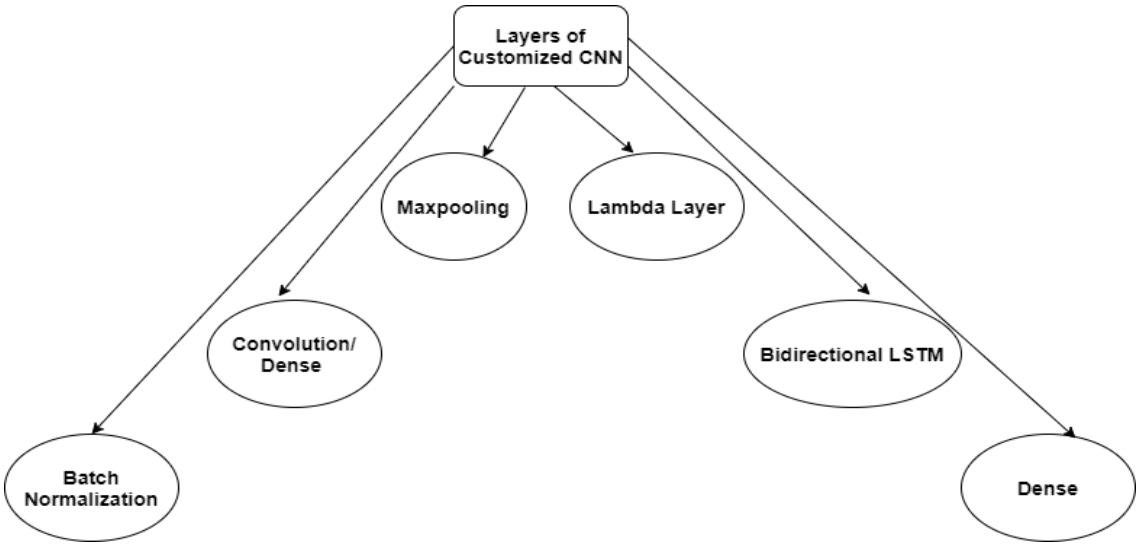


**Figure 1.** Layers Implemented in Proposed Neural Network

1. Bidirectional LSTM (Long Short-Term Memory): The bi-directional LSTM (Long Short-Term Memory) constructs a bag of words based on the TF-IDF (Term Frequency-Inverse Document Frequency) score. This layer can process the sequential information in both directions. In traditional RNN (Recurrent Neural Network), patterns are extracted, and data is processed in a loop by generating feedback from every iteration. The major drawback of this method is that iterating over the same data more times makes it complicated to remember the sequences exactly. It also needs to improve solving the interdependency between the new and old information. The system has resolved this problem by moving the sequences from past to future and future to past.

2. CTC (Connectionist Temporal Classification) Lambda Layer: After localizing the text in the image using the di-directional layers, the CTC (Connectionist Temporal Classification) decoder converts the image format string to raw textual information. It is famous for editing the input data as per the requirements. It transforms the data by defining the customized functions with the help of regular expressions. The lambda layer ignores the length and width of the characters that helps the model to reduce the operations related to the size computations. The proposed architecture applies to squeeze functionality before applying the normalization layers. The squeeze functionality extracts the context associated with the sentence by defining the key value pairs where the essential word is represented as key and the content followed by context is represented as value.

3. Convolution Layer: This layer balances the computational load of the network, and it performs dot product operation between the learnable parameters and weights [13]. The learnable parameters are passed as "Kernels", traversing the RGB image along the height and Width to find the Region of Interest (RoI). The computation is shown in figure (2).

The convolution layer assumes a critical function in the neural network architecture under consideration by effectively managing the computational burden and extracting features from the input images. An overview of the fundamental operations of the convolution layer. The convolution layer processes the input image using tiny matrices called kernels or filters. The kernels execute dot product operations on the regions of the image that overlap, thereby capturing critical attributes including edges, textures, and patterns. Activation map or feature map refers to the output of the convolution operation. The map effectively emphasises the regions where particular features are detected, thereby enhancing the predictive capabilities of the subsequent layers.



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

The kernels utilise the dot product operation at each stage to identify regions of interest as they traverse the height and breadth of the RGB image. This operation enables the network to concentrate on particular attributes present in localised areas of the image.

The network significantly diminishes its computational complexity in comparison to completely connected layers through the implementation of convolutional layers. Limiting the quantity of connections between neurons accomplishes this reduction, consequently diminishing the number of parameters that require training.

By employing padding, the convolution operation is able to maintain the image dimensions without reducing critical edge information or causing context loss at the edges.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} * \begin{bmatrix} 10 & 20 \\ 30 & 40 \end{bmatrix} =$$

$$\begin{bmatrix} 1*10+2*20+5*30+6*40 & 2*10+3*20+6*30+7*40 & 3*10+4*20+7*30+8*40 \\ 5*10+6*20+9*30+10*40 & 6*10+7*20+10*30+11*40 & 7*10+8*20+11*30+12*40 \\ 9*10+14*20+13*30+14*40 & 10*10+11*20+14*30+15*40 & 11*10+12*20+15*30+16*40 \end{bmatrix}$$

**Figure 2.** Working of Convolution Layer

The output produced by this operation is known as the "Activation Map". But during the traverse, the Size of the activation map changes, as shown in equation (1)

$$Output\_size = \frac{Width - Spatial + 2*padding}{Stride} + 1 \text{- (1)}$$

Where,

a. Width represents the Width of the input image whose dimension is represented Width*Height

b. Spatial represents the spatial Size of the kernel, and the Stride size of the kernel is represented by Stride

c. Padding represents the number of pixels to be added to improve the quality of the images. All the padded pixels are filled with zero values.

The proposed research wants to process the image with the same Size so preservation of borders is taken care of with the help of padding. In case of the same padding attribute to retain the Size of the input image, the padding value should be half of the spatial filter size. By half the size of the spatial filter, the dimensions of the original picture are kept, and the dense layer is able to handle the feature maps clearly.

4. Dense Layer: The major goal of this layer is to change the output image size by connecting every neuron in the previous layer with every other neuron in the next successive layer. The computation of the dot product is presented in figure (3)

$$\begin{bmatrix} 10 & 2 & 3 \\ 4 & 50 & 6 \\ 7 & 8 & 90 \end{bmatrix} * \begin{bmatrix} 1 \\ 2 \\ 30 \end{bmatrix} = \begin{bmatrix} 10*1+2*2+3*30 \\ 4*1+50*2+6*30 \\ 7*1+8*2+90*30 \end{bmatrix} = \begin{bmatrix} 104 \\ 284 \\ 2723 \end{bmatrix}$$

**Figure 3.** Working of Dense Layer



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

The dense layer connects with the convolution layer and classifies the images to predict the class label. The flattening process is applied before the prediction process to form a single connected layer based on the weights and bias value [12]. The Size of the dense layer depends on the classifier activation function attached to it. In the proposed research, the dimensionality of the input image is 32*128, which is flattened as a single vector of size 4096. The initial layers are assumed as 64 then the numbers of trainable parameters are 262,144.

5. Max Pooling Layer: Adding max pooling after the convolution layer reduces the image size for the next layer by decreasing the number of pixels. Every pooling layer needs to define the stride size of the kernel and the new matrix obtained after applying the pooling size is reduced to a factor of k, which represents the size of the stride. The major focus of the pooling layer is to extract information from the image irrespective of its state and transformed position. “Spatial Variance” [13] performs this operation. The output of this layer is known as the "Pooled Feature Map". The computation of max pooling is shown in figure (4)

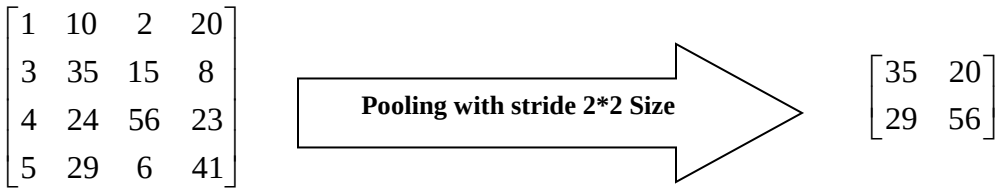


**Figure 4.** Feature Map Extraction

6. Batch Normalization: Many of the existing neural networks are performing better on the training data, but when it comes to the test data, these systems fail because of the variance in data distribution. This results in the failure of data balancing with different ranges of values. This layer helps the model to transform the data into uniformly distributed values [14]. The neural network converts records into batches and applies normalization and standardization techniques. These layers apply a sigmoid function on the input and pass the normalized values to the next layer. The normalization computation is shown in equation (2)

$$Batch\_norm = \frac{\sum_{i=1}^{n} input_i - \mu}{\sqrt{\frac{(\sum_{i=1}^{n} input_i - \mu)^2}{m}} + \varepsilon} \quad \text{- (2)}$$

where,

a. input represents the pixel values of the image
b. μ represents the mean value of the input column vector
c. m represents the median value of the input column vector
d. ε represents the deviation in the pixel values

To guarantee the robustness and accuracy of our suggested model completed a rigorous validation procedure. The data set was divided into validation and training subsets. An evaluation of the model's performance was conducted for each epoch of the multiple epochs that occurred during the training phase.

The suggested techniques present a unique way to detecting false news by retrieving textual information from images using a hybrid temporal lambda layer inserted in an autoencoder neural network. This technique solves the difficulty of distinguishing false news from images carrying textual information, which typical CNN (Convolution Neural Network)-based approaches struggle with owing to their emphasis on processing visual material. The suggested technology is unusual in its ability to effectively extract text from images while reducing the number of words required for reliable fake news identification. Existing methods mostly use RNN (Recurrent Neural Network) or LSTM (Long Short-Term Memory) NN to cover the meaning of words. The proposed model, on the other hand, changes the processing function based on the text’s current weights, which makes it better at finding fake news. In addition, the model uses bidirectional LSTM (Long Short-Term Memory) to recognise word sequences and conducts



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

dimensionality reduction of collected text using both decoder and encoder modules. This novel mix of methodologies distinguishes the proposed methodology from earlier efforts and puts it as a possible answer to the issues involved with detecting false news images includes textual information.

The remaining of this work is structured as follows:

- Section 2 presents analytical review of the state-of-the-art related studies.
- Section 3 provides a detailed methodology of the proposed system.
- Section 4 debates the performed experiments and the corresponding results.
- Section 5 concludes our contributions and findings.

## 2. Related Work

[7] Anurag Goswami et al. introduced an algorithm for identifying fake news using the BERT (Bidirectional Encoder Representation Transformer) approach. News can spread any data. It can be real or fake. Identifying fake content in the article takes time to overcome. The author has combined two approaches with high data filtering, removing the unwanted data in three different stages for detecting the news. The raw data is initially considered, and then it moves to the CNN (Convolution Neural Network) process, where many layers are built into it. In this process, it first crosses five layers to find unrelated data. Then it is mapped to the words that have been repeated rapidly. Now it reverts to connecting two-layer so it is easy for the algorithm to find and mark the content or modify it to original data. The model is to verify that the approach reaches efficiency compared to existing, so validations are chosen, and therefore the achieved result is 98.90% efficient. Many layers can cause a long time to execute, and memory space is consumed more; consequently, the purchase cost will be high.

[12] B. B. Gupta et al. have proposed deep learning and multi-feature approaches. The approach mainly detects social hub networks where fake news is spread. The working process starts with raw data collected, particularly from Facebook. Here, each profile is verified and analyzed to detect any news recognized in that profile. Then complete profile is verified, and check the forwarded content and the process will be repeated. Now that particular data is extracted to multi-feature process and retrieve the fake and unusual data. As known, the author was using ML (Machine Learning) and DP the whole data is sent for classification. Whether the data verified is mock-up or genuine. Machine Learning is the process of creating statistical models and algorithms that allow computers to carry out particular tasks without explicit guidance. By comparing approaches, only DP has high efficiency and the best features for evaluation. The efficiency rate is 99.4%, distinguished from previous methods.

[16] Sonal Garg et al. has worked on the dataset IFND that is a mock-up of news. This data contains pictures and text. An approach is used for identifying data that helps to read the scripted language and change it to machine-known language. Images are difficult to analyze because they are high in pixel size or quality, then the Size of the raw data is high. After the collection of a script, the extraction of a mock-up comes into existence. The LDA approach is mainly used for news. Therefore, the LDA approach has detected fake scripts, and the ML (Machine Learning) technique is the best approach. In this, the high efficiency is 94%. In picture recognition using the LSTM (Long Short-Term Memory) method if the Size is 226*226 then 74%, 32*32 then 66% have been obtained. The new approach, VGG, consumes more time, and detecting data from the social hub takes much work.

[8] Asif Ekbal et al. has introduced AMFB for scripts and captured recognition in fake news. The combination of content and image is in raw data where to initial separation of image and content needs to be performed. Then the process should be carried out with each path for each path preprocessing is taken. In preprocessing, the whole content is analyzed and evaluated. From that, ML (Machine Learning) approaches are considered, BiLSTM (Bidirectional Long Short-Term Memory) is chosen for the script, and ABM is for a picture. The combination of the script and pictures for that MFB is considered, then finally, the author's approach comes into existence, and the separation happens accordingly. The efficiency of the approach is 98% achieved. Here the separation of NLP and Image is very tuff, and it consumes more time.

[21] Yin Zhang et al. introduced an FND-SCTI approach to retrieve the best data from two datasets with 17k and 17k records. Firstly, text data and visualization are the data contained in datasets. Then the data should be separated into three cases, and only two modules should be selected to work with the proposed method. VGG can represent images, and then hierarchical will learn the script data. An encoder is used for multinomial extraction. All these are combined to train the news; for that, DNN is used to know which will give high efficiency. Now the trained and tested methods need to be applied to the 2 datasets and check the results Twitter has achieved 74% and Weibo has achieved 84%.



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

Finally, the proposed method has the best efficiency rate in multinomial methods. However, the main issue is the proposed method of image reorganization for poor-quality images, and if the pixel is high, it is not easy to store more data.

[19] Ratish Agarwal et al. has been proposed an approach for a Diversified dataset to derive mock-up news using ML (Machine Learning). Raw data is a combination of visual and script, here image content will be different from the text. So far, 126k data has been retrieved from n millions of people's accounts. The approaches to detecting the data are also different, so many exist. The complete data is divided into 4 types, and each module has different features to extract. Then training and test cases have been arranged to evaluate the news. Moreover, the author has used DP methods i.e., bi-lstm, to get the best result. Finally, comparing all approaches with one another, 95% is the best efficient rate obtained with the proposed approach. The issue with this approach is text evaluation.

[4] Anshika Choudhary et al. has proposed a Linguistic approach that can obtain language model features. This approach is working on two finite datasets regarding political media. In phase one, the data is preprocessed in that only the English script is recognized with different aspects. In phase two, the proposed approach reads the data and evaluates sentences according to the phrases. Then the partitions form a model, and the driven method is combined with ML (Machine Learning) and Neural approach to get the best result. In the last layer, many hidden layers are formed to extract the unreal news easily, and data will be retrieved mock-up or genuine. The final process is also known as the nn approach. Thence estimation of the approach is a combination of linguistic and driven models; therefore, 86% of efficiency has been achieved.

[6] Iraklis Varlamis et al. have achieved RNN (Recurrent Neural Network), and CNN (Convolution Neural Network) approaches to retrieve high efficiency. Considering raw data, the first preprocessing is to be taken place and remove all unmatched data with the help of NLP. Then an approach of conv1D took place to separate the prospective ways to analyze words and their related meaning. Pooling is conducted between 1d3 and 1d, so a separation ratio occurs, and the unreal news will be deleted or formed into layers. The hybrid technique is used for the mapping articles, and then all these layers will be applied finite approach to gain genuine and unreal news. Therefore, this approach is applied to two datasets, and data is retrieved with different efficiencies and time consumptions. Hence, the two datasets' results are 60% and 99%. Here neural method gives the best result compared to the hybrid approaches, retrieving data in less time.

Hager Saleh et al [14] has developed OPCNN methodology for the detection of fake news in different sectors. Four different area datasets are considered with different sections were all datasets hold similar qualities i.e., text. The data has to be pre-processed for removing unwanted data then splitting works in by training and test phase. As the data is trained, further process is to extract the features. Two kind of feature extraction was chosen embedding and N-gram TF-IDF (Term Frequency-Inverse Document Frequency). In word embedding the text data is divided in vector was similar words are considered and meaning was extracted. Later the optimization was done hyperopt method which was efficient while extraction of words. Now the proposed methodology phases with five layers were dropout layer which avoids overfitting followed with convolution layer decreases the complexity the next pooling utilizes the feature mapping and reduce the unwanted features the next flatten which convert text to single dimensional array and finally, output layer whether the news is original or not. To check the performances of the new model the validation techniques are applied and has proven that the performances is more efficient than previous approaches.

Dilip Kumar Sharma et al [17] has introduced a LSTM (Long Short-Term Memory) methodology for the huge period of detection fake news. The IFND dataset holds 2k03 to 2k21 news related to the all categories, which may be fake or original. The dataset holds text and images where the text can extract best ML (Machine Learning) and the images extraction DL (Deep Learning). For any data extraction, the augmentation is performed with intelligent techniques and the LDA is used for the modelling for categorization of several news. As the baseline dataset is now pre-processed for extracting meaning full text and images. Now the data was distributed in three categories image, text, and multimodal. For extraction of text, ML (Machine Learning) has high and efficient. Two techniques are derived TF-IDF (Term Frequency-Inverse Document Frequency) and Tokenization was applied on the text data and six ML (Machine Learning) methodologies are applied. For image extraction, the DL (Deep Learning) methodologies are applied VGG-16 and ResNet-50. Finally, for multimodal the combination of ML (Machine Learning) and DL (Deep Learning) are applied those are LSTM (Long Short-Term Memory) and VGG16. The VGG-16 holds five sets of conv and pooling layers. Followed with dense and finally outcome. In resnet-50 the input, conv, pooling, four conv, gap, dense and finally outcome was designed. For every approach the performances was evaluated and efficiency was derived.



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

Dr. M. Anusha et al [1] has focused on identification of fake news with DL (Deep Learning) and ML (Machine Learning) technologies. Twitter data was considered in the year 2k09 with four attributes and ids included. Two classes are present positive contains 8l and same with negative. The data contains 9 worlds contain most counts. The data was pre-processed to unwanted data, remove symbols, drop stopwords, erase usernames, exclude emoji/symbols, apply stemming are removed. Data was analyzed because twitter contains more text data compared to images. To extract the words Glove embedding vectors are used with the clusters. To categories are chosen word2vec was utilized. The model was designed with BiLSTM (Bidirectional Long Short-Term Memory) and CNN (Convolution Neural Network) was the model was designed with input, embedded; LSTM (Long Short-Term Memory), conv1d, gobal max pooling, dropout, dense and finally dense is used. The model was validated using two-word extractor methods there is no much differences in accuracy.

Muhammad Asad Arshed et al [2] has implemented a CNN (Convolution Neural Network) with VGG-16 for the detection of deepface detection fake and true images. The dataset was derived from Kaggle site. Online sources, diffusion stable, and styleGAN2, which are different sources. The data was in text phase where the process of text is transmitted to images. This process holds text data like little description that text was encoded to initial some noise with junction. Later the diffusion model is applied to it and decoding is done while the image is derived. The dataset holds four kinds of images original, GAN_fake, Stable & GAN, and diffusion fake images. The images are transmitted in patches form with some fame size then the linear projections are made with conjunction with patch and positional embeddings. In encoders phase the normalization is performed with MSA and to it, the conjection was derived then normalization was performed where it can directly transmitted to conjunction or MLP method is utilized and finally MLP head is provided. The data is CLS block is derived and finally the four phases are derived similar to datasets categories.

Pundlik Jadhav et al [5] has introduced a BERT (Bidirectional Encoder Representation Transformer) methodology for the categoring original news from the fake. The dataset contains two categories of news authentic and fraudulent, which has text data. As the data is pre-processed for removing unwanted data and updating the dataset is very essential for improving it. Transformer framework-based BERT (Bidirectional Encoder Representation Transformer) understands context in every direction. An encoder stack captures bidirectional word linkages using attention techniques. The model evaluates incoming sequence words using self-attention techniques. BERT (Bidirectional Encoder Representation Transformer) is pre-trained on large corpora using ML (Machine Learning) goals and prediction of subsequent sentences tasks. Fine-tuning for various activities involves adding appropriate layers to the framework and utilizing activity-specific data to fine-tune the overall strategy. Along with this two, more methods are examined and process across those proposed had high performances and efficiency. However, the GPT has closer results with proposed methodology and has similar predictions.

Nida Aslam et al [3] has developed a Bi-LSTM-GRU-dense methodology for the detection of text based fake news. The dataset was collected from LIAR, which hold two categories of data fake and true news. Along with those, three more datasets are utilized for combining the features. The data was divided into two paths were first part holds statements transmitted to NLP processing and the words are embedded to form the vectors of words and then Bi-LSTM GRU is utilized. On other hand, the remaining features are transmitted to pre-process the data and DL (Deep Learning) dense was initialized. The both processes are combined and transmitted to ensemble phase and finally the validation are performed. The process was implemented in the python domain and the implementation was easy. The performances were not accurate and level of efficiency has to be increased too.

C. Vinotheni et al [20] has designed a Fast-RNN methodology for the easy detection of fake news. The dataset contains Tamil text, which contains two kinds of data fake and real. Now the dataset is divided in two phases in one phase the pre-processing is applied by resizing the text data to one size. Now the train data was applied to the segmentation of data in two phase Bi-LSTM and other is FRNN. In Bi-LSTM the data contains sentences or text embedding form it contains vectors formed data. The other phase contains segmentation with holds the FRNN method, which can accurate mode. Where the other path data undergoes with pre-processing and the test data was examined with 20% of performances. Now the data is connected to segmentation and both techniques are applied. The segmented data is transmitted to sentences or segmented text and finally the data is analyzed with the parametric. The Bi-LSTM holds four stages embedded, forward, Backward, and finally flatten or concatenate phase.

Saif Saad Alnuaimi et al [11] has developed a hybrid model for detection of fake news from the videos. Two different datasets UCF 50, 101 and ISOT, which contains different categories of information related to social media and many more. Generally, the video data was split in five phases for images, ETE, factorized utilizes CNN (Convolution Neural



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

Network), RNN (Recurrent Neural Network), and hybrid that is the combined structure of CNN (Convolution Neural Network) & RNN (Recurrent Neural Network). In image the frame level CNN (Convolution Neural Network) are considered for the extraction of features and transmitted to previous approaches. In ETE, the 3d kernels are utilized for the extraction of info with relative features. Now to classify the videos the process was in three simple steps inputting, extraction of features, and predict based on the proposed classification. The validation was performed with limited aspects of accuracy and loss was derived. Which has achieved with best-optimized values.

**Table 1:** Existing Research Works Analysis

| **Author** | **Algorithm** | **Merits** | **Demerits** | **Accuracy** |
|---|---|---|---|---|
| Anurag Goswami et al. | BERT-based deep learning | It is the best combination of 1D CNN (Convolution Neural Network). | Filtering with n number of layers can take a long time for results | 98.90% |
| B. B. Gupta et al. | Deep Learning approaches use multiple associated features. | Using ML (Machine Learning) and DP, the data is retrieved easily, and efficiency is high | With too many profiles, searching is difficult when news is detected. | 99.4% |
| Sonal Garg et al. | Machine Learning Algorithms for IFND dataset | Using random forest image reorganization is easy and high accuracy is acquired | The raw data may not contain only images. So, too many algorithms are being used to get the best accuracy. | 94% |
| Asif Ekbal et al. | AMFB | The images contain less pixel rate, which acquires good accuracy. | If the dataset contains separate image and text parts, it is easy. | 98% - spot fake |
| Yin Zhang et al. | Deep learning | FND-SCTI data sets have low image quality, so n number of records can be retrieved. | It refuses some images for quality purposes, even if it's fake. | 84% - Weibo 74% - Twitter |
| Ratish Agarwal et al. | ML and DP on Datasets | Bi-LTSM has the best accuracy compared to the remaining algorithms, and it has a very easy-use visualization. | Text recognition is tuff | 95% |
| Anshika Choudhary et al. | Linguistic Feature-Based Learning | By combining two approaches, the best output has been achieved. | Only the English language is used, and the remaining languages are removed at the preprocessing stage | 86% |
| Iraklis Varlamis et al | Hybrid CNN (Convolution Neural Network)-RNN | The combined neural approach has high retrieving efficiency. | LSTM has given the best accuracy compared to the remaining algorithms | 60% of ISOT 100% in FA-KES |
| Hager Saleh et al | OPCNN | Even the cross-validation performances has achieved high efficiency. | Huge dataset has less accuracy compared to remaining datasets. | 99.9% |



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

| | | | | |
|---|---|---|---|---|
| Dilip Kumar Sharma et al | DL, LSTM (Long Short-Term Memory), Bi-LSTM, VGG-16 | The dataset was huge even the performances was efficient. | More time consumption while the training comparison. | 92.7% |
| Dr. M. Anusha et al | BiLSTM + 1DCNN | The model was examined with two techniques. | The performances has to be increased. | 83% |
| Muhammad Asad Arshed et al | CNN | Multiple classes are derived accurately. | Works with some datasets only. | 99% |
| Pundlik Jadhav et al | BERT | By combining methods, the complexity has reduced. | The reading score was less compared to remaining methods. | 95% |
| Nida Aslam et al | Bi-LSTM-GRU-Dense | Multiple techniques are utilized for accurate predictions. | The performances has to be increased. | 89.8% |
| C. Vinotheni et al | Fast-RNN, Bi-LSTM | The process for implementation is very accurate. | No need of applying cleaner for segmentation process | 92.88% |
| Saif Saad Alnuaimi et al | CNN + RNN (Recurrent Neural Network) | The model has overcome the overfitting issue. | The extraction of features has less in take. | |

### 2.1. Research Gaps Identified

1. Identifying fake news from the image content is difficult because existing neural networks can work on the images alone. We need to explore the content of the images.

2. Few researchers worked on vast amounts of content stored as csv files in the distributed environment using the standard RNN (Recurrent Neural Network) approaches. RNN (Recurrent Neural Network) just identifies the temporal sequences it cannot store the previous words and takes necessary actions.

3. Many standard neural networks implemented CNN (Convolution Neural Network) layers only to identify fake news images. They have also implemented more layers to improve accuracy, which in turn has improved the cost of resources.

### 2.2. Novelty

The proposed methodology extracts the text from the images using the lambda-embedded neural network. Later, the system implements bidirectional LSTM (Long Short-Term Memory) to recognize the sequence between the words. Based on the sequence, one must reduce the number of words extracted for efficient detection of fake news. The dimensionality reduction of the text extracted is taken care of by the encoder and decoder modules of the network.

## 3. Proposed Methodology

The proposed model initially processes the images to reshape them to the same Size. The model maintains the aspect ratio in terms of Width and height. The processing of the image is performed under the following categories. If the height and Width of the images are less than the minimum number of pixels, then the model pads the zero values and concatenates with the original image. If the height and Width of the images are greater than the required value, then it performs the subtraction operation and expands the dimensions of the edges. Finally, all the pixel values are normalized to the standard range of values. Few images are transformed in the dataset to improve the system's efficiency by handling the augmented images generated by GAN cycles. The system to extract the text and images of all possible characters in different transformations is loaded into the model. The system fastens the process. It encodes the characters available in the dataset and performs a clustering technique to identify the annotations associated with the images as class labels. Since the classification is multiple class labels, the neural network applies a dense layer



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

with a softmax function as the output layer. Few symbols in the character dataset are unnecessary, so the proposed system applies the padding sequences on both train and test datasets. The padding sequences help the model reduce the number of images annotated based on the symbols. The overall architecture of the proposed model is presented in figure 5.

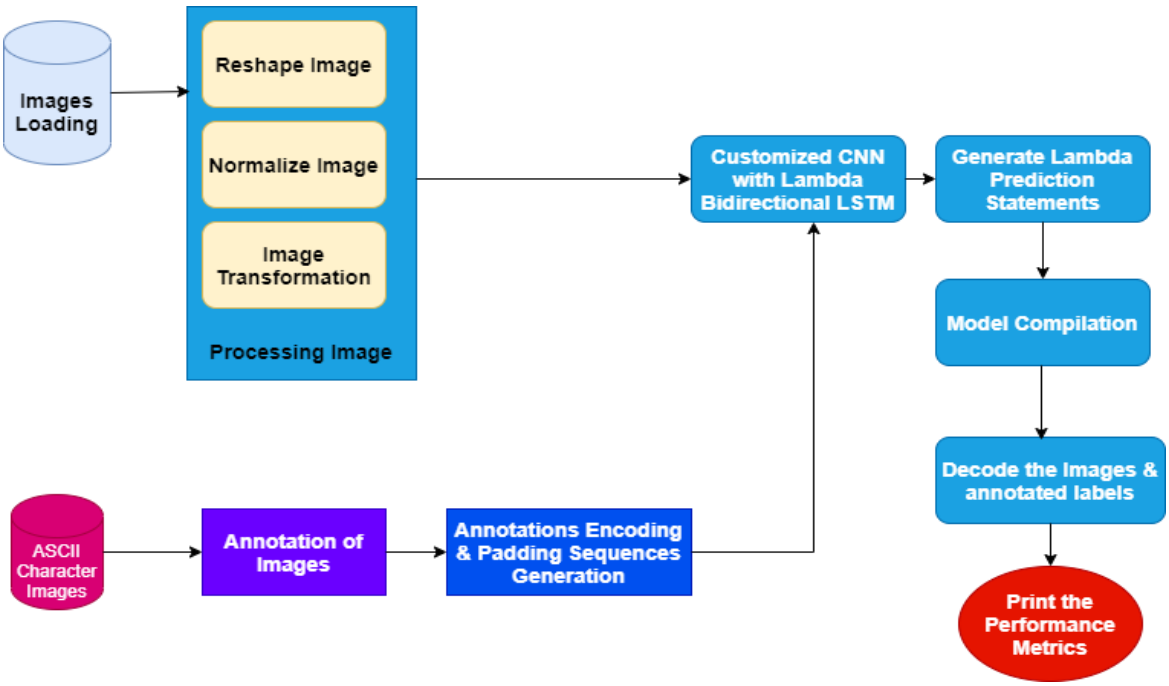


**Figure 5.** Proposed Architecture for Text Extraction from Images

The annotated labels and transformed images are passed as input to the customized neural network. The neural network contains 17 layers of different categories whose detailed count is presented in table 2.

**Table 2:** Description of Layers in the Customized CNN (Convolution Neural Network)

| f | Layers Name | Count | Activation Function |
|---|---|---|---|
| **1** | Convolution | 7 | - |
| **2** | Max pooling | 4 | - |
| **3** | Batch Normalization | 2 | - |
| **4** | Lambda | 1 | Sigmoid |
| **5** | Bidirectional | 2 | Relu |
| **6** | Dense | 1 | Softmax |

Pseudocode for Lambda Layered Bi-LSTM for Text Extraction from Images

Input: Semantic Images Generated from Dataset, SIGD

Output: Accuracy computations for text extraction

Begin:

1. Initialize width and height from image shape

2. new_w←32, new_h←h*(new_w/w)

3. for i in len(SIGD):



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

a. trans_image[i]← resize(SIGD[i],new_w,new_h)

b. gen_image[i]←subtract(trans_image[i],vector_value)

4. for c,i in enemrate(dir(gen_image)):

a. labels.append(c.index(i))

b. labels_splits←labels.split(‘ ‘)

c. train_pad_labels_split←pad_sequence(labels_split,value=len(labels))

5. for i in len(0,7):

a. proposed_layer← sequential.add(Dense[i],activation=’relu’)

b. squ_layer←Lambda(x:proposed_layer.squeeze(start_i:end_i),drop_out=0.7)

6. print accuracy & CTC (Connectionist Temporal Classification) loss

End

In the lambda layer, the model applies the sigmoid function, which helps the model to reduce the pixel intensity from -1 to 1 so that the mean of all the distributed data is surrounded by zero and the standard deviation is surrounded by one. This, in turn, helps the model compute the probability of every character in the text. Since the relation between the different sequencing elements might be linear or non-linear, the proposed system applies the ReLU (Rectified Linear Unit) activation function. The output obtained by the model is encoded using the CTC (Connectionist Temporal Classification) function to predict the class labels. After compiling and saving the best model, it is decoded to reconstruct the original image. The working of lambda layer is presented in figure 6.

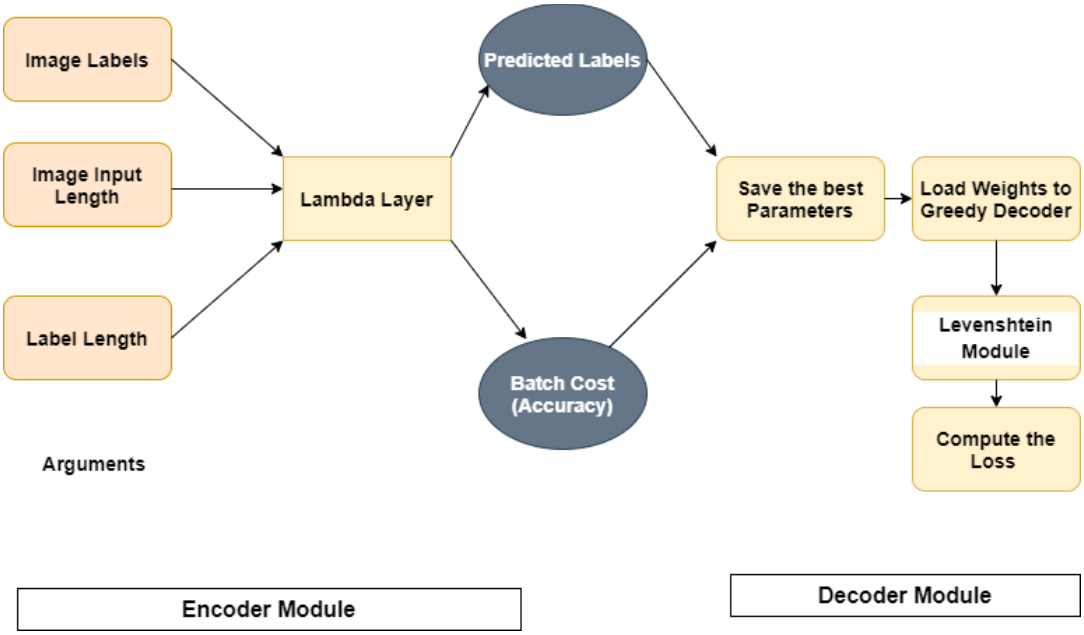


**Figure 6.** Flow Diagram for Encoder & Decoder Using Lambda Layer

## 4. Results & Discussion

The proposed model identifies the labels associated with each image by extracting the textual information from the images. Initially the model recognizes the dimensionality of labels and padding labels. It also describes the summary of the proposed neural network to identify the trainable parameters. Every epoch is presented to identify the loss and accuracy of every iteration.



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

```
train_labels[3001]
```

```
[76]
```

```
train_padded_label[310]
```

```
array([19])
```

```
train_padded_label.shape, valid_padded_label.shape
```

```
((3069, 1), (341, 1))
```

**Figure 7.** Dimensionality Analysis on the Dataset

In Figure 7, the proposed model presents the number of labels available after dividing them into train and valid. The data in the train and valid are further classified as padding labels because few image labels are similar in content. However, existing models treated them separately due to variations in their prefix or suffix values. In the proposed methodology, the annotations are represented as linked lists, so the size of the padding labels is significantly reduced. After padding, the labels in the training dataset are reduced to 3069, and the labels in the validation dataset are reduced to 341.

```
Model: "model_2"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_1 (InputLayer)         [(None, 32, 128, 1)]      0
_________________________________________________________________
conv2d_22 (Conv2D)           (None, 32, 128, 64)       640
_________________________________________________________________
max_pooling2d_8 (MaxPooling2 (None, 16, 64, 64)        0
_________________________________________________________________
conv2d_23 (Conv2D)           (None, 16, 64, 128)       73856
_________________________________________________________________
max_pooling2d_9 (MaxPooling2 (None, 8, 32, 128)        0
_________________________________________________________________
conv2d_24 (Conv2D)           (None, 8, 32, 256)        295168
_________________________________________________________________
conv2d_25 (Conv2D)           (None, 8, 32, 256)        590080
_________________________________________________________________
max_pooling2d_10 (MaxPooling (None, 4, 32, 256)        0
_________________________________________________________________
conv2d_26 (Conv2D)           (None, 4, 32, 512)        1180160
_________________________________________________________________
batch_normalization_4 (Batch (None, 4, 32, 512)        2048
_________________________________________________________________
conv2d_27 (Conv2D)           (None, 4, 32, 512)        2359808
_________________________________________________________________
batch_normalization_5 (Batch (None, 4, 32, 512)        2048
_________________________________________________________________
max_pooling2d_11 (MaxPooling (None, 2, 32, 512)        0
_________________________________________________________________
conv2d_28 (Conv2D)           (None, 1, 31, 512)        1049088
_________________________________________________________________
lambda_2 (Lambda)            (None, 31, 512)           0
_________________________________________________________________
bidirectional_10 (Bidirectio (None, 31, 512)           1574912
_________________________________________________________________
bidirectional_11 (Bidirectio (None, 31, 512)           1574912
_________________________________________________________________
dense_2 (Dense)              (None, 31, 79)            40527
=================================================================
Total params: 8,743,247
Trainable params: 8,741,199
Non-trainable params: 2,048
_________________________________________________________________
```

**Figure 8.** Summary of the Layered Architecture Implemented in the Proposed Methodology



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

Figure 8 represents the layers utilized to implement the proposed architecture and the shapes generated as output and parameters. The model has implemented lambda and bidirectional LSTM (Long Short-Term Memory) layers as additional layers, and internally it customized the CNN (Convolution Neural Network) layers with max pooling and batch normalization layers.

Every image in the CNN (Convolution Neural Network) has height, Width, and depth, along with batch size. The depth of the image represents the colour channels. In figure 8, the output shape has four parameters, in which the first represents the batch size of any n value because it is denoted as "None". The parameters are available for the convolution layers whose count is produced by multiplying the height and widths of the input and output channels.

```
Epoch 00016: val_accuracy did not improve from 0.63050
Epoch 17/20
614/614 [==============================] - 27s 45ms/step - loss: 0.4457 - accuracy: 0.7498 - val_loss: 0.8364 - val_accuracy:
0.6452

Epoch 00017: val_accuracy improved from 0.63050 to 0.64516, saving model to sgdo-25000r-20e-3069t-341v.hdf5
Epoch 18/20
614/614 [==============================] - 28s 46ms/step - loss: 0.3944 - accuracy: 0.7784 - val_loss: 0.7904 - val_accuracy:
0.6481

Epoch 00018: val_accuracy improved from 0.64516 to 0.64809, saving model to sgdo-25000r-20e-3069t-341v.hdf5
Epoch 19/20
614/614 [==============================] - 29s 47ms/step - loss: 0.3628 - accuracy: 0.7941 - val_loss: 0.8185 - val_accuracy:
0.6657
```

**Figure 9.** Initial Epochs Training For Text Extraction from Images

Figure 9 considers a sample epoch training of 20 iterations and presents the accuracy variation in every iteration. Suppose the current iteration has an improved accuracy than the previous iteration. In that case, the network saves it as the best model in the Hadoop version because it helps the big data environment utilize the architecture without starting from scratch. Saving the best model helps the classification process to get the proper weights of each important input vector. This method saves time and memory utilization. The classification process becomes simple with these re-trained weights. From the epochs training, it is clear that with the increase of epochs, the accuracy and loss are decreasing. The model can be claimed as the "Best Model".

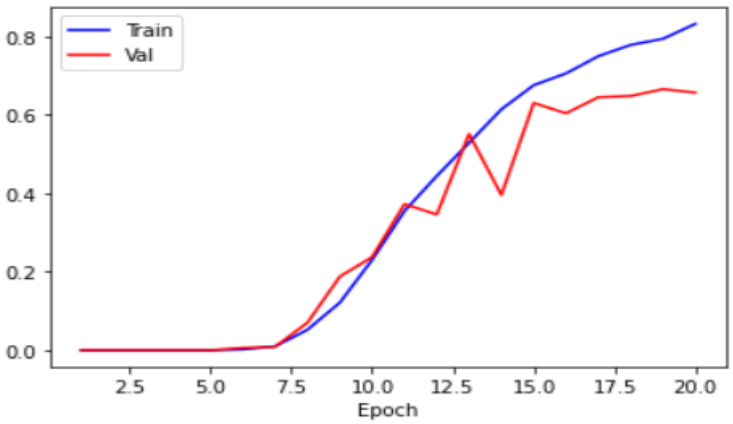


**Figure 10.** Accuracy Graph for Proposed Network

The dataset is divided into train and validation, and the accuracy graph is plotted for every iteration on both datasets. The accuracy nature of the datasets is gradually increasing with the number of iterations. In figure 10, epochs are plotted on X-axis, and the accuracy percentage is plotted on Y-axis. The training accuracy, presented as blue in colour, has a clear improvement slope from starting to ending epochs. The valid accuracy marked in red has up and downs at a few iterations but later stabilizes with the increasing iterations.



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

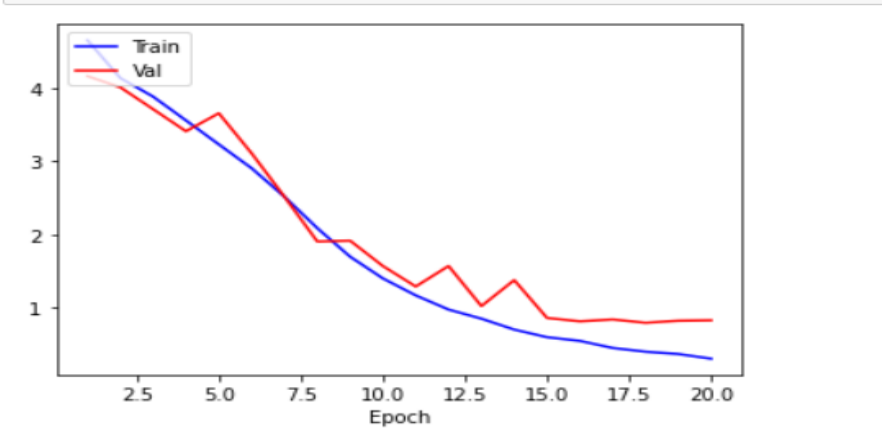


**Figure 11.** Loss Evaluation Graph over Train & Valid Data

It is essential to check accuracy and loss metrics to evaluate the CNN (Convolution Neural Network) model. Figure 11 represents the nature of loss on both datasets. It is also clear that the loss values are initially high, gradually decreasing, and almost near zero during the 20$^{th}$ epoch training. The clear downslope of the training loss metric proves that the proposed model performs better than the existing approaches.

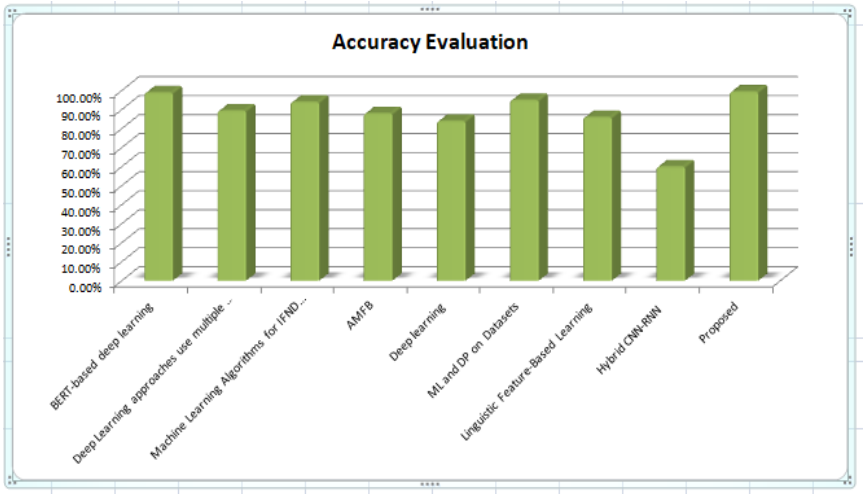


**Figure 12.** Accuracy Comparison between Existing & Proposed Methodologies

Figure 12 represents that the proposed methodology has achieved +0.5% more than BERT (Bidirectional Encoder Representation Transformer) based approach, which is stated as the first research work in the survey. Similarly, figure 12 compares and plots the accuracies obtained by different existing works stated in the literature survey, and it includes the proposed methodology to prove its efficiency. In the graph, the y-axis represents accuracy measurement in percentages, and the x-axis denotes the methodologies implemented in survey papers.

Efficiency Improvement

The proposed approach for finding fake news uses a two-way LSTM model along with techniques for reducing the number of dimensions to make finding fake news in image based material more accurate and useful. When comparison is done with effectively for proposed model to other methods that are already in use, a comparison was performed.

Comparison of Accuracy

Proposed Method: The proposed methodology got an accuracy rate of 86%, which is +0.5% better than the BERT-based approach.



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

Existing Method: In conventional method, including CNN & RNN, exhibit disparate levels of accuracy contingent upon the dataset & depends upon the configuration implemented. Based on evaluations of various datasets, method employing RNN & CNN structure, for instance recorded accuracy rates ranges.

### 4.1. Discussion

It is crucial to emphasise the potential implications of the suggested methodology on the domain of false news detection within the discussion section. The methods capacity to derive text from images in an efficient manner and reduce the word count to ensure precise detection of false news signifies a noteworthy progression in the domain, Moreover, the implementation of bidirectional LSTM (Long Short-Term Memory) as well as dimensionality reduction methods demonstrates the potential of the model to enhance the precision and effectiveness of identifying false news in image-based content.

Regarding the proposed methodology limitations, it is critical to recognise that the techniques may encounter difficulties when dealing with low-quality images or image with high pixels rates. Furthermore, text recognition might be problematic, especially when the information comprises both visual and textual material. These constraints should be noted while implementing the proposed method, which may need more study and refining to handle effectively.

## 5. Conclusion

Extracting text from the images is crucial for identifying the sequence between the old and new information. Many hackers and fake people in the yellow media try to change the essential and small words, changing the scenario's entire meaning. Particularly in regards to the extraction of textual information from images, the ramifications of this study extend to the wider domain of neural network applications. In the proposed system, the bidirectional LSTM (Long Short-Term Memory) helps the model transform the sentences into different tenses, tries to store the longest possible sequences using memory gates available, and updates the content timely. The system also takes care of the content with different styling notations. Social media people try to use a different style for the alphabet present in the crucial world to avoid recognizing the morphing. Utilizing an autoencoder neural network with an integrated hybrid temporal lambda layer to extract textual information from images, the proposed methodology has made a substantial contribution to the field of false news detection. The existing approaches focused on the semantic coverage of the sentences available in the images based on the RNN (Recurrent Neural Network) or LSTM (Long Short-Term Memory) neural networks. However, the proposed model customized the computation function based on the current weights available for the text. The outcomes of the comparison with established methodologies indicate that the accuracy rate of the proposed approach is +0.5% higher than that of the BERT (Bidirectional Encoder Representation Transformer)-based approach. This finding suggests that the proposed method has the capacity to surpass current techniques and establish a novel benchmark in the realm of detecting false news within image-based content. In addition, to extending, the envelope of what is currently possible in the detection of false news, the proposed methodology provides techniques and insights applicable to a vast array of image-based content analysis endeavours. The research findings have profound consequences across multiple domains, such as content moderation, media forensics and misinformation detection. In these fields, the effective extraction of textual information from images is of the utmost important.

## References

**[1]** Anusha, M., & Leelavathi, R. (n.d.). (2024). International Journal Of Intelligent Systems And Applications In Engineering Sentiment Analytics On Sarcasm Detection Using Bi-Lstm-1dcnn Model For Fake News Detection. In Original Research Paper International Journal of Intelligent Systems and Applications in Engineering IJISAE (Vol. 2024, Issue 5s). www.ijisae.org

**[2]** Arshed, M. A., Mumtaz, S., Ibrahim, M., Dewi, C., Tanveer, M., & Ahmed, S. (2024). Multiclass AI-Generated Deepfake Face Detection Using Patch-Wise Deep Learning Model. Computers, 13(1). https://doi.org/10.3390/computers13010031

**[3]** Aslam, N., Ullah Khan, I., Alotaibi, F. S., Aldaej, L. A., & Aldubaikil, A. K. (2021). Fake Detect: A Deep Learning Ensemble Model for Fake News Detection. Complexity, 2021. https://doi.org/10.1155/2021/5557784



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024

**[4]** Choudhary, Anshika; Arora, Anuja (2020). Linguistic Feature Based Learning Model for Fake News Detection and Classification. Expert Systems with Applications, 114171–. doi:10.1016/j.eswa.2020.114171

**[5]** Jadhav, P., Rajesh, D., & Shukla, K. (n.d.). (2024). International Journal Of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Deep Learning Analysis For Revealing Fake News Using Linguistic Complexity And Semantic Signatures. In Original Research Paper International Journal of Intelligent Systems and Applications in Engineering IJISAE (Vol. 2024, Issue 12s). www.ijisae.org

**[6]** Jamal Abdul Nasir;Osama Subhani Khan;Iraklis Varlamis; (2021). Fake news detection: A hybrid CNN (Convolution Neural Network)-RNN based deep learning approach. International Journal of Information Management Data Insights, (), –. doi:10.1016/j.jjimei.2020.100007

**[7]** Kaliyar, R. K., Goswami, A., & Narang, P. (2021). FakeBERT: Fake news detection in social media with a BERT (Bidirectional Encoder Representation Transformer)-based deep learning approach. Multimedia Tools and Applications, 80(8), 11765–11788. doi:10.1007/s11042-020-10183-2

**[8]** Kumari, R., & Ekbal, A. (2021). AMFB: Attention-based multimodal Factorized Bilinear Pooling for multimodal Fake News Detection. Expert Systems with Applications, 184, 115412. doi:10.1016/j.eswa.2021.115412

**[9]** Qian, S., Wang, J., Hu, J., Fang, Q., & Xu, C. (2021, July 11). Hierarchical Multimodal Contextual Attention Network for Fake News Detection. Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '21: The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. https://doi.org/10.1145/3404835.3462871.

**[10]** Raj, C., & Meel, P. (2021). ConvNet frameworks for multimodal fake news detection. Applied Intelligence. https://doi.org/10.1007/s10489-021-02345-y.

**[11]** Saad Alnuaimi, S., Hikmat Rasheed, B., Yuvaraj, D., Sundaravadivel, P., & Isaac, R. A. (n.d.). (2024). International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Hybrid Deep Learning Techniques for Large-Scale Video Classification. In Original Research Paper International Journal of Intelligent Systems and Applications in Engineering IJISAE (Vol. 2024, Issue 15s). www.ijisae.org

**[12]** Sahoo, S. R., & Gupta, B. B. (2020). Multiple features-based approaches for automatic fake news detection on social networks using deep learning. Applied Soft Computing, 106983. doi:10.1016/j.asoc.2020.106983

**[13]** Sahoo, S. R., & Gupta, B. B. (2021). Multiple features-based approaches for automatic fake news detection on social networks using deep learning. Applied Soft Computing, 100, 106983. https://doi.org/10.1016/j.asoc.2020.106983.

**[14]** Saleh, H., Alharbi, A., & Alsamhi, S. H. (2021). OPCNN-FAKE: Optimized Convolutional Neural Network for Fake News Detection. IEEE Access, 9, 129471–129489. https://doi.org/10.1109/ACCESS.2021.3112806

**[15]** Singh, B., & Sharma, D. K. (2021). Predicting image credibility in fake news over social media using a multimodal approach. Neural Computing and Applications. https://doi.org/10.1007/s00521-021-06086-4.

**[16]** Sharma, D.K., Garg, S. (2021), IFND: a benchmark dataset for fake news detection. Complex Intell. Syst. https://doi.org/10.1007/s40747-021-00552-1

**[17]** Sharma, D. K., & Garg, S. (2023). IFND: a benchmark dataset for fake news detection. Complex and Intelligent Systems, 9(3), 2843–2863. https://doi.org/10.1007/s40747-021-00552-1

**[18]** Sri Silpa Padmanabhuni and Pradeepini Gera, (2022) "Synthetic Data Augmentation of Tomato Plant Leaf using Meta Intelligent Generative Adversarial Network: Milgan" International Journal of Advanced Computer Science and Applications(IJACSA), 13(6), http://dx.doi.org/10.14569/IJACSA.2022.0130628

**[19]** Shubha Mishra, Piyush Shukla, Ratish Agarwal, (2022) "Analyzing Machine Learning Enabled Fake News Detection Techniques for Diversified Datasets", Wireless Communications and Mobile Computing, vol., Article ID 1575365, 18 pages, 2022. https://doi.org/10.1155/2022/1575365

**[20]** Vinotheni, C., & S., L. P. (2024). Fast Recurrent Neural Network with Bi-LSTM for Handwritten Tamil text segmentation in NLP. ACM Transactions on Asian and Low-Resource Language Information Processing. https://doi.org/10.1145/3643808

**[21]** Zeng, Jiangfeng; Zhang, Yin; Ma, Xiao (2020). Fake news detection for epidemic emergencies via deep correlations between text and images. Sustainable Cities and Society, (), 102652–. doi:10.1016/j.scs.2020.102652

**[22]** Zeng, J., Zhang, Y., & Ma, X. (2021). Fake news detection for epidemic emergencies via deep correlations between text and images. Sustainable Cities and Society, 66, 102652. https://doi.org/10.1016/j.scs.2020.102652.



DOI: https://doi.org/10.54216/JCIM.150124
Received: April 15, 2024 Revised: June 14, 2024 Accepted: August 15, 2024