DOI: https://doi.org/10.33103/uot.ijccce.22.2.3

Convolutional Recurrent Neural Networks for Text Lecture Summarization

Muna Ghazi Abdulsahib[1], Matheel E. Abdulmunim[2]
[1,2]*Department of computer sciences, University of Technology, Baghdad, Iraq*
[1]*MUNA.G.Abdulsahib@ uotechnology.edu.iq*, [2]*Matheel.E.Abdulmunim@ uotechnology.edu.iq*

Abstract— *Text summarization can be utilized for variety type of purposes; one of them for summary lecture file. A long document expended long time and large capacity. Since it may contain duplicated information, more over, irrelevant details that take long period to access relevant information. Summarization is a technique which provides the primary points of the whole document, and in the same time it will indicates the majority of the information in a small amount of time. For this reason it can save user time, decrease storage, and increase transfer speed to transmit through the internet. The summarization process will eliminate duplicated data, unimportant information, and also replace complex expression with simpler expression. The proposed method is using convolutional recurrent neural network deep model as a method for abstractive text summarization of lecture file that will be great helpful to students to address lecture notes. This method proposes a novel encoder-decoder deep model including two deep model networks which are convolutional and recurrent. The encoder part which consists of two convolutional layers followed by three recurrent layers of type bidirectional long short term memory. The decoder part which consists of one recurrent layer of type long short term memory. And also using attention mechanism layer. The proposed method training using standard CNN/Daily Mail dataset that achieved 92.90% accuracy.*

Index Terms— *Text Summarization, Deep Learning, Convolutional Neural Network, Long Short Term Memory, Convolutional Recurrent Neural Network.*

I. INTRODUCTION

Summarization is the process which provides a short information over those whole media, that demonstrates fast thought of the greater part imperative or important majority of the data inside the original content in little amount of time thus it is save client time, lessen storage, furthermore increase transfer speed for transmission through that web. However, the summarization of video, audio, picture, and text is complex task because of its sequence-to-sequence nature [1]-[3]. In this way, this paper takes text from lecture file then performs text summarization. Text summarization will be a standout amongst the real ideas utilized within those fields of documentation. Long documents are troublesome will peruse and see all the likewise it expends quite a few periods. Outline solves this issue by giving an abbreviated outline of it for semantics it could be about extraordinary assistance for learners to record lecture notes from starting with classes, conferences and seminars [4].

Recently, deep learning methods have successfully effectively tended to issues in different fields, for example, picture classification, machine translation, summarization, discourse recognition, text-to-speech generation and Era furthermore other machine learning related ranges [5], [6]. Deep learning that is taken in analyses unpredictable issues

Received 31/August/2021; Accepted 31/October/2021

DOI: https://doi.org/10.33103/uot.ijccce.22.2.3

will encourage the choice to facilitate the decision-making process and mirror the thing that the human mind could attain by extracting characteristics during different levels for reflection. Typically, higher-level layers bring fewer points over lower-level layers. Hierarchical levels of deep learning could help support learning. The output of particular case layer will be the enter of the following layer. In addition, the amount about layers determines those deepness, which affects the level for learning [7], [8]. Content outline may be a key subject of Natural Language Processing (NLP) research region. Extractive furthermore abstractive are two major parts of quick text summarizations. This paper performs abstractive text summarization using deep learning. The contribution of this paper used hybrid deep learning model from CNN and RNN that is not used in any survey for text summarization.

This paper is organized as follows: In section 2, some related work are presented. Sections 3,4,5,6 present theoretical background of text summarization, RNN, CNN and word embedding respectively. The proposed method introduces in details in section 7. Section 8 introduces the experimental results and discussion of applying the proposed method and finally in section 9, the conclusions are given.

II. RELATED WORK

In 2018, Yao et al. suggested abstractive outline method using a dual encoding mechanisms. The mechanisms are copy and coverage. They reduce repetition of words by coverage mechanism. CNN/DailyMail and DUC 2004 datsets used for experimental results that show good performance. Also in the model used fix decoding length where they suggested in the future utilize dynamic length [9].

Over 2019, Egonmwan et al. suggested extractive and abstractive content summarization technique using self-attention mechanism and transformer. They utilized CNN/Daily Mail and Newsroom datsets and also for word embeddings used pre-trained Glove. Further more, for extractive summarization they proposed strategia for labelling sentences that improved the results [10].

In 2020, D. Patel et al. recommended abstractive quick outline looking into Google search results to reduce searching time in the web. Suggested a model that will set an interface utilizing Python library for web scratching of the content that brings about those. To get the data from that web in a robotized summarized way. Using Amazon Product Reviews and CNN /Daily news datasets. The information cleaning will be completed under preprocessing steps. They built dictionary for word embedding, abstractive outline created utilizing the sequence-to-sequence model which utilization Long Short-Term Memory (LSTM) encoding decoding for preparation those model to process the result. Three stacked LSTM may be assembled for that encoder [8].

Also in 2020, Peng et al. proposed text summarization using dual attention pointer network. The CNN/Daily Mail dataset used to test those rundowns handled by these models that show effective performance. They also proposed strategia to enhance coverage mechanism for minimize the repetition issue. Further more, they apply reinforcement learning and scheduled sampling to enhance the quality of the produced summary [11].

In 2021, F. Nan et al. recommended a set of new measurements that quantify the score of entity-level hallucination for created rundowns and indicate that the substance mind flight issue could a chance to be mitigated toward basically sifting the preparation information. More over they recommend a few systems including hallucination filtering, multi-task learning furthermore method to generate sequence that yield improvements the

Received 31/August/2021; Accepted 31/October/2021

ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.2.3

summarization performance. Utilized the pre-trained Bidirectional Encoder Representations from Transformers (BART-large) model with adjust on the 3 outline dataset: Newsroom, CNN/DailyMail and XSUM[12].

Table I is demonstrated the comparison methods of related works focuses on methodology, datasets, evaluation measure, strengths and weaknesses for abstractive text summarization models.

TABLE I. COMPARISON METHODS OF RELATED WORK

Authors	Methodology	Dataset	ROUGE-L Measure	Strengths	Weaknesses
Yao et al. in 2018[9]	Dual encoding mechanism copy and coverage	CNN/Daily Mail and DUC2004	37.13	Reduce repetition	Not balance between precision and recall because use fix decoding length
Egonmwan et al. in 2019 [10]	Self -attention mechanism and transformer	CNN/Daily Mail and Newsroom	38.92	Simple to train by use transformer	Repetition words
D. Patel et al. In 2020[8]	Long Short Term Memory	CNN/Daily Mail and Amazon	____	Reduce searching time in web	Lacks scalability
Peng et al. in 2020[11]	Dual attention pointer network	CNN/Daily Mail	37.35	Hybrid training model	_____
F. Nan et al. In 2021[12]	pre-trained BART	Newsroom, CNN/DailyMail and XSUM	____	Proposed technique for entity-level data filtering	_____

III. TEXT SUMMARIZATION

Text summarization is a technique of reducing the long text to a shorter text by removing the less useful data and at the same time retaining the primary focuses majority of the data, which aides the clients to discover the intrigued majority of the data rapidly. The inputs of text summarization have the chance to be a single document or multi documents alternately. While the outputs could be abstractive or extractive [7], [13].

There are two types of automatic text summaries: Extractive summaries and Abstractive summaries. Extractive rundowns need aid processed by extracting those subsets of existing sentences from the source content. Abstractive rundowns would generate by reformulating sentences of the origin text [14], [15].

Text summarization method can be utilized for different purposes such as in email summary, reviews from claiming movies, news abstraction, framework about learner notes, rundown data to specialist. Furthermore legislature officials, businessman, summarize the therapeutic information to doctors and summarize the authoritative archive [16].

A. Extractive Text Summarization

The first type of summarization is extractive text summarization. Its utilization a measurable based methodology with selects significant sentences, words, and passages from those unique record on the measure of imperativeness and concatenates these essential parts of the record to structure a summary document. It doesn't change the source text works as create a copy of some significant sentences. Statistical methodology might

Received 31/August/2021; Accepted 31/October/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

Iraqi Journal of Computers, Communications, Control & Systems Engineering (IJCCCE), Vol. 22, No. 2, June 2022

DOI: https://doi.org/10.33103/uot.ijccce.22.2.3

summarize those report utilizing features such as expression frequency, location, title, appointing weights of the keywords and then ascertaining that score of the sentence. Furthermore selecting the most noteworthy scored sentence under the outline judgment [15]- [17].

Detriment about extractive text outline is it can't constantly process the normal result due to its nature. It can't produce a summary similar to a human. Because of this disadvantage abstractive text summarization will be a whole lot powerful as opposed extractive text summarization [18]. *Fig. 1* shows extractive text summarization [16].

FIG.1. EXTRACTIVE TEXT SUMMARIZATION [16].

B. Abstractive Text Summarization

The second type of summarization is abstractive text summarization. Abstractive text summarization may be comparable of the approach people synopsis the archive. Abstractive script summarization needs a comprehension of the document and then produces the rundown. It rebuilds sentences by paraphrasing prompt presenting introducing new words, sentences, and repine. Abstractive script outline has a great compression rate besides lessens duplicated in this way produces an additional important and exact outline with decrease vagueness. Its employments machine learning strategies and far-reaching NLP [7],[16].

Accomplishing abstractive summarization is more mind boggling over text extraction. There need aid a few processes incorporate, such that script preprocessing, expression embedding, model build-up, train-test information evaluation, information acceptance etc. However, abstractive compress may be also superior to extractive compress since the summary will be an estimated representational of a human-generated summary, which makes it all the more serious [7], [18]. *Fig. 2* reveals to abstractive text summarization [16].

FIG.2. ABSTRACTIVE TEXT SUMMARIZATION [16].

IV. RECURRENT NEURAL NETWORK

Recurrent Neural Network (RNN) is a type of supervised deep learning model. Produced for artificial neurons for one alternately more feedback loops, those sentiment loops would repetitive circle. RNNs can efficiently procedure successions sequences of varying length inputs also outputs and were planned on have the capacity with handle

Received 31/August/2021; Accepted 31/October/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.2.3

consecutive data which goes over numerous structures for example, audio, video, text, and so on [19].

A repetitive neural system has the structure about various Feed-forward neural networks with associations around their hidden units. Each layer on the RNN has a distinct time step and the weights are shared over the long run. A basic RNN has three layers they are input, repetitive hidden and output layers, the hidden state goes about likewise the neural networks inside memory. Those organize is prepared should hold majority of the data looking into past data, which feed into future predictions. RNNs can work again sequences of vectors in both of input and the output. Currently, on text summarization RNN encoder-decoder architecture building design will be think about a sentence likewise a sequence of tokens (characters or words) and those enter to the RNN model will be those embedding words to treatment them with RNN. Tokens are typically transformed in successive order, from left to right, and the RNN is anticipated with "memorize" the entirety sequence in its inside states, and those yield will be the word embedding of the outline. In NLP, there are a few expression embedding models, such that Word2Vec, GloVe, FastText, also Bidirectional Encoder Representations from Transformers (BERT), which are those generally significantly utilized statement embedding models [20]. RNN model is illustrated as in *Fig. 3* [21].

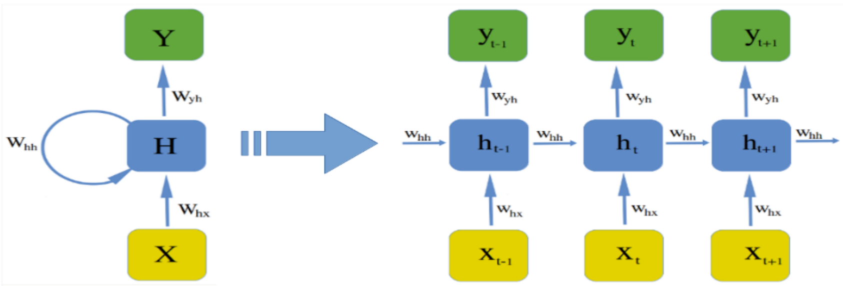


FIG.3. RNN MODEL [21].

RNNs endure from the thing that may be known as short-term memory. This wonder is created eventually by the vanishing gradient issue. The vanishing gradient issue is eventually driven by those nature of back-propagation. This might a chance to be a standout amongst the urgent reasons prompting challenges in the preparing for recurrent neural networks. Gating RNN is a technique utilized with unravels the issue for vanishing gradients, which happens when training a long sequence by RNN. Issue might make tackled eventually permitting the gradients should back-propagate along a straight way utilizing gates, whereas every gate need a weight also a bias. Gates aides the model control and change the measure of data that streams between hidden states. Throughout training, those weights furthermore biases of the gates would update. Two of the well-known gating are denominate Long Short-Term Memory (LSTM) Furthermore Gated Recurrent Unit (GRU) [7].

However, the LSTM unit confirms constant, non-vanishing also non-exploding error flow, in this way comprehending the issue from claiming RNN. In recent years, RNN have been effectively utilized once more to different applications for example: speech recognition, summarization, and translation. LSTM comprises of an input gate, output gate, a forget gate, also a memory cell. These three entryways monitor those memory substances and the cell states during the current timestamp. The construction modeling about LSTM cell is exhibited in *Fig. 4* [7].

Received 31/August/2021; Accepted 31/October/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.2.3

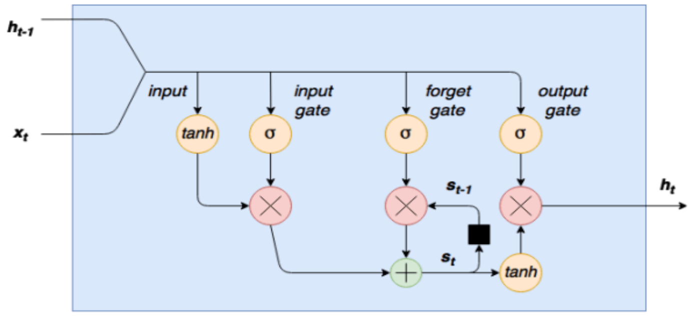


FIG.4. LSTM CELL [7].

V. CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Network (CNN) is a special type of multi-layer neural network. It found pervasiveness for handling issues connected with NLP assignments such as sentence classification, script classification, text summarization, conclusion analysis, machine translation and response relations. CNN is made of two real parts: Feature Extraction and Classification [21]-[23].

In feature extraction part, the system will perform a sequence of convolutions and pooling operations throughout which the characteristics are distinguished. The aggregate amount from claiming convolutional layers characterizes those downright amounts of features that have to extricate. The greater part utilized activation function clinched along side CNN will be ReLU abbreviation of (Rectified Linear Unit). This serves those neurons provide for correct pixel values likewise output to every positive value, while for all the negative values the output will be zero. The pooling operations will be used to lessen the size of the input and also those results from convolutions. Due to this, those number for parameters that necessity to be investigated is smaller. Hence, the computational time reduces [21],[22].

While another part, the output layer which will be a network of the fully connected layers that serve likewise a classifier on highest on these extracted features. The neurons in a fully connected layer bring full associations with every last bit activations in the past layer [21], [22]. CNN model is illustrated in *Fig. 5* [22].

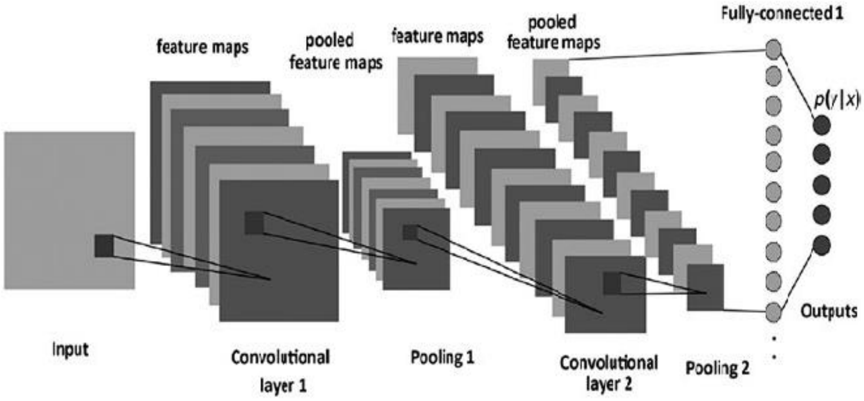


FIG.5. CNN MODEL[22].

Received 31/August/2021; Accepted 31/October/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.2.3

VI. WORD EMBEDDING

Word embedding is a kind of word representation it is also called a distributed semantic. It captures the semantic and syntax features of words. Word embedding is NLP task. This process allows words with identical meaning to be understood by machine learning algorithms. On other hand, it is involving mapping of words into vectors of numbers utilizing the neural network. There are several pretrained word embedding models available like Word2Vec (from Google), GloVe (from Stanford), FasTest (from Facebook), also Bidirectional Encoder Representations from Transformers (BERT). Word embedding can learned from scratch instead of utilizing a pretrained word embedding model but pretrained model have advantage than learned from scratch because it learned from huge of vocabulary. This paper will focus in Word2Vec pretrained model [7],[24].

Word2Vec is a two layers neural network that processes words. Its input is a word corpus and its output is a set of feature vectors that represent words in that corpus. So, if two words have identical features, then such words are related. Word2Vec model comprises from two architectures are Continuous Bag of Words (CBOW) and Skip Gram. Both types rely on the context window. A context window refers to the number of words appearing on the right and left of a current word. The words appearing in the context window are known as context (or neighboring words). While, GloVe represents the global vector instead of the context window. In CBOW, the current word is predicted using the neighboring words. whereas, in Skip Gram performs inverse of CBOW which infers that it predicts the neighboring words from the current word [7].

The challenge of Word2Vec and Glove is how to learn the representation for out-of-vocabulary words. That forced to use a random vector, which is not perfect.The vocabulary of Word2Vec contains of 3 million words while the vocabulary of GloVe is 0.4 million, so the possibility of out-of-vocabulary for Word2Vec considerably is much less than GloVe [7],[24].

VII. THE PROPOSED METHOD

This section provides the proposed method to perform text summarization of text obtained from lecture file, that will a chance to be great supportive to students to archive lecture file. The text summarization performed by using Convolutional Recurrent Neural Network (CRNN) deep learning model method. A brief description of the proposed method with a general block diagram will be discussed first, and then it will be followed by details about each step and algorithm.

The proposed system represents novel CRNN deep model of abstractive text summarization. CRNN including CNN and RNN network model. In order to handling issues of sequence-to-sequence with diverse lengths sequences where the input is long text and output is summary text is constructed Encoder-Decoder building design. In Encoder-Decoder design, the encoder produces vector representation, which is then used as the decoder initial hidden state to predict summary. This architecture is composed of two phases which are training phase and inference phase. *Fig. 6* illustrates the general block diagram of the proposed method.

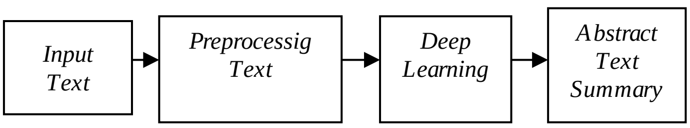


FIG.6. BLOCK DIAGRAM OF THE PROPOSED METHOD.

Received 31/August/2021; Accepted 31/October/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.2.3

The framework of the proposed CRNN deep learning model illustrates in *Fig. 7*.

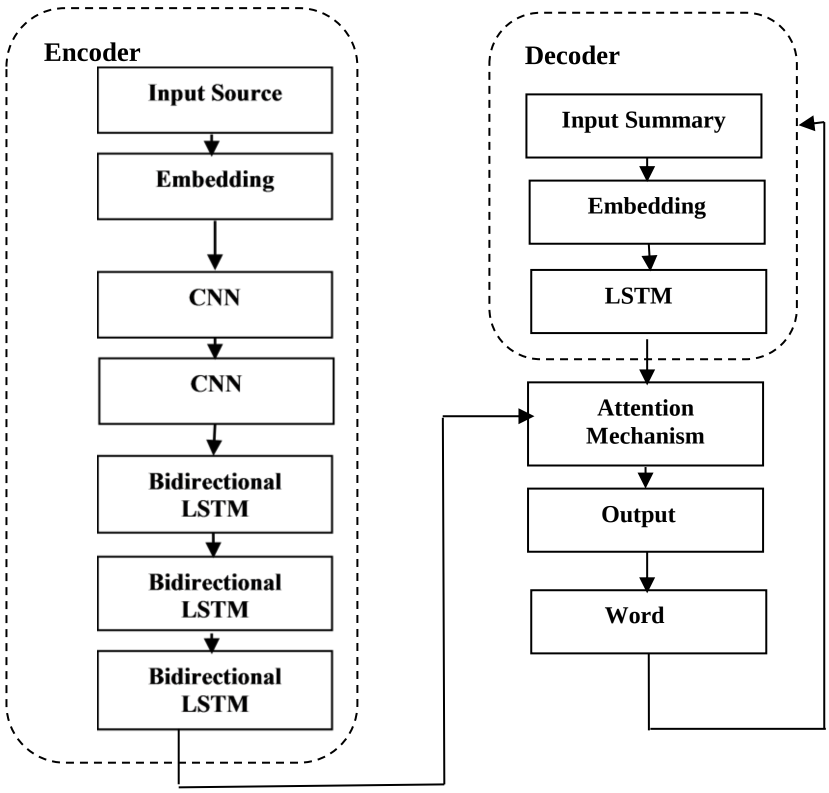


FIG.7. FRAMEWORK OF THE PROPOSED CRNN DEEP LEARNING MODEL.

The proposed method details are described in this section for each step.

- First in text preprocessing step, after loading dataset will perform preprocessing process which include split words, lowercasing, remove punctuations, remove stop words (stop words such as: “a”, “the” and so on), contractions, remove words that have a period <= 1 character. This preprocessing step makes the text summarization more accurate. Following the individuals operations split data under training and testing sets.
- The next step is embedding, on text summarization CRNN encoder-decoder building design will be contemplate about a sentence likewise a sequence of words, and those enter to the CRNN model will be those embedding words to produce embedding words of the summary. Which is a kind of word representational will get it toward neural network algorithms. Word2Vec will be used to make a word embedding of embedding layer. The embedding layer gives the output of settled length vector to the next layer in the models.
- After that, in the encoder will be encoded utilizing two one-dimensional convolutional neural network layers the CNN part followed by three bidirectional long short term memory layers (bidirectional LSTM comprises for forward LSTM and backward LSTM) the RNN part. Which produces vector representation, which is utilized as the decoder starting state.
- Then in the decoder it will be utilized one LSTM layer followed by attention mechanism. Across this mechanism, the decoder can use of intermediate hidden

Received 31/August/2021; Accepted 31/October/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.2.3

states in the encoder also utilize all that information to choose which word will be is next.

- Finally, the structural of the model make made for two stages which are training and inference stage. Afterward perform compiling and training of the encoder and decoder in training stage, the inference stage utilized those encoder model and decoder model on predict the target summary.

Algorithm (1) proposed text summarization using CRNN deep learning model.

Algorithm (1): Text Summarization Using CRNN
Input: Source Text.
Output: Summary Text.
Process:
Step 1: Load dataset.
Step 2: Preprocessing process which perform the following actions: split words, lowercasing, remove punctuations, remove stop words, contractions, remove words that have a period <= 1 character.
Step 3: Split data into training and testing sets.
Step 4: Training Phase- Encoder part:
Step 4.1: Input size of source text.
Step 4.2: Word embedding using pretrained Word2vec.
Step 4.3: Create the model architecture for two layers CNN of a one-dimensional Convolutional.
Step 4.4: Create the model architecture for three layers RNN of type Bidirectional LSTM.
Step 5: Training Phase- Decoder part:
Step 5.1: Input size of summary text.
Step 5.2: Word embedding using pretrained Word2vec.
Step 5.3: Create the architecture for one layer RNN of type Unidirectional LSTM.
Step 6: Perform attention mechanism for the output of encoder and decoder.
Step 7: Output dense layer for the output of decoder and attention.
Step 8: Compile the model.
Step 9: Training the model.
Step 10: Save encoder and decoder model.
Step 11: Inference phase: Predict encoder and predict decoder model.
Step 12: Generate summary.
Step 13: End.

VIII. DISCUSSIONS AND RESULTS

In this section, experimental discussions of the tests led were exhibited on show the capability of the suggested abstractive text summarization technique. It holds that vital information about the optimal values of the suggested technique parameters with a description for their impacts on the performance of the suggested technique furthermore additionally a comparison with some previous strategy.

The most effective suitable values of the method parameter used in the training of proposed deep model design are presented in Table II.

Received 31/August/2021; Accepted 31/October/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.2.3

TABLE II. THE VALUES OF DEEP MODEL HYPER PARAMETERS

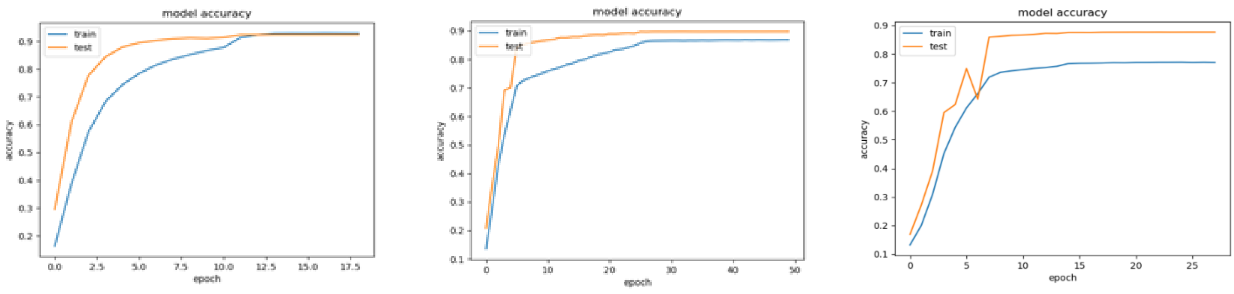
Hyper Parameter	**Value**
Embedding Dimensions	300 dimensions
Hidden Size	256 be in bidirectional 512
Batch Size	128
Epochs	100 with using early-stopping
Optimizer	Adam optimizer
Initial Learning Rate	0.001
Dropout	0.4

The performance of the deep model might have been evaluated based on the final accuracy of the model. Once executing the above hyper parameters, it can be concluded that the CRNN model achieved a higher accuracy over the just RNN model. The CRNN accuracy attained 92.90% whereas RNN accuracy attained 87.64 %. In addition, CRNN get result a training accuracy of 92.90 %, while test accuracy was 92.46%. This implies that suggested CRNN model not have overfitting issue.

Here are some concepts utilized within design deep model in order to prevent overfitting issue, additionally make an approach of improving the performance results:

1- Dropout some hidden units are used to prevent overfitting.
2-Reduce learning rate of portion factor 0.1 if the loss not reduce for some specific epochs.
3-Early-stop when stop learning is something like that saves best performance of the model.
4- Using attention mechanism.

Fig. 8 illustrated results accuracy differences for training and testing sets of the proposed CRNN model to notice the behavior of the model in different epochs and batch size.



(a) (b) (c)

FIG.8. THE ACCURACY OF TRAINING AND TESTING IN THE PROPOSED MODEL. (A) IN 30 EPOCHS WITH BATCH SIZE 128. (B) IN 50 EPOCHS WITH BATCH SIZE 200. (C) IN 100 EPOCHS WITH BATCH SIZE 200.

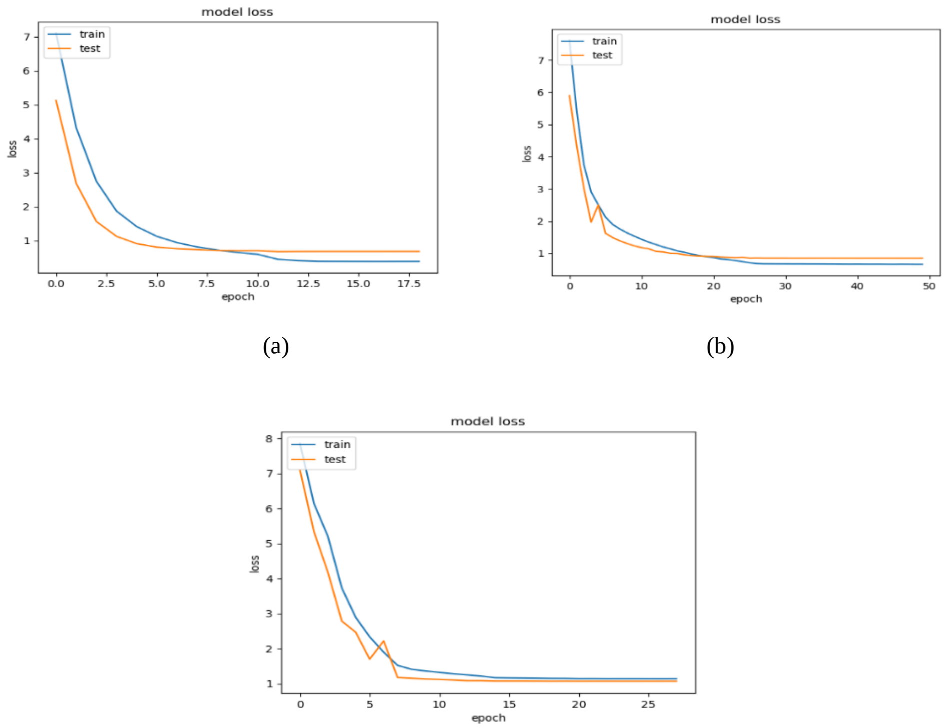
According to *Fig. 8*, have been induce that in (a) and (c) there is not enhance in the accuracy after epoch 19 and 28 respectively. So, stop training the model after those epochs that the result of using early-stop concept that using when stop training and so as should get those best model.

Fig. 9 illustrated results loss differences for training and testing sets of the proposed CRNN model to notice the behavior of the model in different epochs and batch size.

Received 31/August/2021; Accepted 31/October/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

Iraqi Journal of Computers, Communications, Control & Systems Engineering (IJCCCE), Vol. 22, No. 2, June 2022

DOI: https://doi.org/10.33103/uot.ijccce.22.2.3



(a) (b)

(c)

FIG.9. THE LOSS OF TRAINING AND TESTING IN THE PROPOSED MODEL. (A) IN 30 EPOCHS WITH BATCH SIZE 128. (B) IN 50 EPOCHS WITH BATCH SIZE 200. (C) IN 100 EPOCHS WITH BATCH SIZE 200.

As stated by *Fig. 9*, inducing that the loss keep reduce until there is a slight increment in the loss after some epoch. So, stop training the model as the result of using early-stop idea. Furthermore, pointed to the accuracy of the model depended on the loss.

The big challenges in design deep learning model including require long time in training and require high powerful hardware. The performance might decline according to the hardware limitation(constraint). When used powerful hardware can increasing the hyper parameters values that result improved model performance further more reduce training time.

The CNN - Daily Mail dataset have been used to train and test the proposed CRNN deep model for abstractive text summarization, becausee it can generate multisentence summaries for long document, which have been selected 98401 files. Whereas, the others datasets have single sentence summary for a document. The CNN - Daily Mail dataset can be downloading from this site (https://cs.nyu.edu/~kcho/DMQA/) which has become a standard dataset used for summarization tasks. This dataset consists of 92.000 documents files from the news articles of CNN and 219.000 documents files from the news articles of Daily Mail.Whereas, each file composed of paired of article text source and multi sentences summaries. Its source text has average of 780 words or around 40 sentences, while the summaries have average of 56 words or around 4 sentences.

Table III is demonstrated the comparison of suggested model with other models focuses on encoder and decoder components, optimizer, datasets and measure rouge which are used to train those summarization models.

Received 31/August/2021; Accepted 31/October/2021
 ISSN (Print) 1811-9212 *ISSN (Online) 2617-3352*

DOI: https://doi.org/10.33103/uot.ijccce.22.2.3

TABLE III. COMPARISON OF DIFFERENT ABSTRACTIVE TEXT SUMMARIZATION MODELS

	Encoder Components	**Decoder Components**	**Optimizer**	**Dataset**	**ROUGE-L Measure**
Yao et al. in 2018[9]	Bidirectional GRU	Unidirectional GRU	Adagrad	CNN/Daily Mail and DUC2004	37.13
Egonmwan et al. in 2019 [10]	Unidirectional GRU	Unidirectional GRU	Adam	CNN/Daily Mail and Newsroom	38.92
Peng et al. in 2020[11]	Bidirectional LSTM	Unidirectional LSTM	Adam	CNN/Daily Mail	37.35
The proposed method	Two-Convolutional and Three-Bidirectional LSTM	Unidirectional LSTM	Adam	CNN/Daily Mail	38.45

IX. CONCLUSIONS

This paper introduced another strategy for abstractive text summarization of lecture document. This method performed by utilizing CRNN deep learning that including two deep network model: CNN and RNN. In RNN, the proposed method utilized LSTM in order to handling issue of long sequence that suffered in RNN.

The paper exhibited general framework for the method and the most effective and suitable model hyper parameters utilized within diverse steps of the proposed model are presented. Also, introduced the concepts utilized to enhance the performance. Furthermore, introduced the comparison of proposed model with other models.

The big challenges in deep learning model it requires powerful hardware. When utilized powerful hardware can increasing those values of hyper parameters that accomplish a higher accuracy. Experimental results indicate that proposed model achieved 92.90% accuracy. In the future might like to take advantage of transfer learning technique.

REFERENCES

[1] N. Raphal, H. Duwarah, and P. Daniel,"Survey on abstractive text summarization." In *2018* International Conference on Communication and Signal Processing (ICCSP), pp. 0513-0517. IEEE, 2018.
[2] A. A. A. Karim and R. A. Sameer, “Static and Dynamic Video Summarization”, *eijs*, vol. 60, no. 7, pp. 1627-1638, Jul. 2019.
[3] R. F. Ghani, S. A. Mahmood, Y. N. Jurn and L. Al-Jobouri, "Key Frames Extraction Using Spline Curve Fitting for Online Video Summarization," 2019 11th Computer Science and Electronic Engineering (CEEC*)*, pp. 69-74, doi: 10.1109/CEEC47804.2019.8974340,2019.
[4] A. Vinnarasu, and D. V. Jose. "Speech to text conversion and summarization for effective understanding and documentation." International Journal of Electrical and Computer Engineering 9, no. 5, pp. 3642-3648,2019.
[5] M. M. Mahmoud, A. R. Nasser, "Dual Architecture Deep Learning Based Object Detection System for Autonomous Driving". IRAQI JOURNAL OF COMPUTERS, COMMUNICATIONS, CONTROL AND SYSTEMS ENGINEERING, vol.21, no.2, pp. 36-43,2021.
[6] W.M. S. Abedi, I. Nadher, and A. T. Sadiq. "Modification of Deep Learning Technique for Face Expressions and Body Postures Recognitions.", Vol. 29, No. 3s, pp. 313-320, 2020.
[7] D. Suleiman, and A.Awajan. "Deep learning based abstractive text summarization: Approaches, datasets, evaluation measures, and challenges." Mathematical Problems in Engineering 2020 , Hindawi Vol. 2020, doi: 10.1155/2020/9365340.2020.
[8] D. Patel, N. Shah, V. Shah and V. Hole, "Abstractive Text Summarization on Google Search Results," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 538-543, doi: 10.1109/ICICCS48265.2020.9120998,2020.

Received 31/August/2021; Accepted 31/October/2021

DOI: https://doi.org/10.33103/uot.ijccce.22.2.3

[9] K. Yao, L. Zhang, D. Du, T. Luo, L. Tao, and Y. Wu, “Dual encoding for abstractive text summarization,” IEEE Transactions on Cybernetics, pp. 1–12, 2018.

[10] E. Egonmwan and Y. Chali, “Transformer-based model for single documents neural summarization,” in Proceedings of the 3rd Workshop on Neural Generation and Translation, pp. 70–79, Hong Kong, 2019.

[11] Li, Z. Peng, S. Tang, C. Zhang, and H. Ma, “Text summarization method based on double attention pointer network,” IEEE Access, vol. 8, pp. 11279–11288, 2020.

[12] F. Nan, R. Nallapati, Z. Wang, C. N. d. Santos, H. Zhu, D. Zhang , K. McKeown and B. Xiang,” Entity-level Factual Consistency of Abstractive Text Summarization”, Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics, pages 2727–2733 April 19 - 23, 2021.

[13] R. M. Hadi, S. H. Hashem, and A. T. Maolood. "An effective preprocessing step algorithm in text mining application." Engineering and Technology Journal 35, no. 2 Part (B) Scientific , pp. 126-131,2017.

[14] P. Shah and N. P. Desai, "A survey of automatic text summarization techniques for Indian and foreign languages," 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT*)*, pp. 4598-4601, 2016, doi: 10.1109/ICEEOT.2016.7755587.

[15] K. Wanjale, P. Marathe, V. Patil, S. Lokhande, and H. Bhamare. "Comprehensive Survey on Abstractive Text Summarization.",International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181 ,Vol. 9 ,no. 09, September,2020. doi: 10.17577/IJERTV9IS090466.

[16] P. Batra, S. Chaudhary, K. Bhatt, S. Varshney and S. Verma, "A Review: Abstractive Text Summarization Techniques using NLP," 2020 International Conference on Advances in Computing, Communication & Materials (ICACCM), 2020, pp. 23-28, doi: 10.1109/ICACCM50413.2020.9213079

[17] Naser, E. F. Naser, “Word Retrieval based on FREAK Descriptor to Identify the Image of the English Letter that Corresponds to the First Letter of the Word ", Engineering and Technology Journal, Vol. 38, No. 3B , pp. 150-160,2020, doi: 10.30684/etj.v38i3B.1511.

[18] M. A. I. Talukder, S. Abujar, A. K. M. Masum, S. Akter and S. A. Hossain, "Comparative Study on Abstractive Text Summarization," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT*)*, 2020, pp. 1-4, doi: 10.1109/ICCCNT49239.2020.9225657.

[19] T. Shi, Y. Keneshloo, N. Ramakrishnan, and C. K. Reddy, "Neural abstractive text summarization with sequence-to-sequence models." ACM Transactions on Data Science 2, no. 1, pp. 1-37,2021.

[20] A. Conneau, H. Schwenk, L.Barrault, and Y. Lecun. "Very deep convolutional networks for text classification.", arXiv preprint arXiv:1606.01781, 2016.

[21] M. Andersson, A. Maja, and S. Hedar. "Sketch Classification with Neural Networks: A Comparative Study of CNN and RNN on the Quick, Draw! data set.",Uppsala universitet, 2018.

[22] H. Singh, and Y.A. Lone,” Deep Neuro-Fuzzy Systems with Python”, Berkeley: Apress, 2020, doi: 10.1007/978-1-4842-5361-8.

[23] H. N. Abdullah, and H. K. Abduljaleel. "Deep CNN based skin lesion image denoising and segmentation using active contour method." Engineering and Technology Journal,vol. 37, no. 11A ,pp. 464-469,2019, doi: 10.30684/etj.37.11A.3.

[24] T.-H. Le, “Neural Methods for Sentiment Analysis and Text Summarization. Document and Text Processing”, Université de Lorraine, English. ffNNT : 2020LORR0037ff. fftel-02929745,2020.

Received 31/August/2021; Accepted 31/October/2021

ISSN (Print) 1811-9212 *ISSN (Online) 2617-3352*