WILEY | Hindawi

Research Article

Intelligent Learning Algorithm for English Flipped Classroom Based on Recurrent Neural Network

Qi Shan [ID]

Jiangsu Union Technical Institute, Xuzhou Jiangsu 221000, China

Correspondence should be addressed to Qi Shan; shanqi60032@163.com

Reading and writing are the foundations of English learning as well as an important method of instruction. With the advancement of network technology and the onset of the information age, an increasing number of students have lost interest in traditional English reading and writing instruction in the classroom. Flipped classrooms have emerged as a result of this situation and have become the focus of research in one fell swoop. As a result, flipped classroom research at home and abroad has primarily focused on the theory and practical application of flipped classrooms, and flipped classroom application practice is primarily based on the overall classroom, with few separate discussions on the effects of flipped classroom students' self-learning. As a result, we developed a recurrent neural network-based intelligent assisted learning algorithm for English flipped classrooms. There are two main characteristics of the model. First, it is a gated recurrent unit based on a variant structure of the recurrent neural network. The double-gating mechanism fully considers the context and selects memory through weight assignment, and on this basis, it integrates the novel LeakyReLU function to improve the model's training convergence efficiency. Second, by overcoming time-consuming problems in the medium, the adoption of the connection sequence classification algorithm eliminates the need for prior alignment of speech and text data, resulting in a direct boost in model training speed. The experimental results show that in the English flipped classroom's intelligent learning mode, students explore and discover knowledge independently, their enthusiasm and interest in learning are greatly increased, and the flipped classroom's teaching effect is greatly improved.

1. Introduction

Education experts advocate a new teaching model in the context of the new curriculum reform [1–3], in which students are the main body and teachers are the leaders. This concept, however, cannot be fully reflected in classroom teaching in specific practical situations. The flipped classroom [4–6] came into being under this situation, which is in line with the requirements of contemporary educational concepts. Flipped classroom advocates autonomous learning. In flipped classroom teaching, teachers should not only teach students professional knowledge but also provide personalized guidance in the classroom to teach students the ability to learn. This is not only the requirement of the times but also the direction of the future development of national education.

Secondly, the current situation of literacy class teaching is unsatisfactory, which leads to worrying students' literacy skills [7, 8]. First, in the process of English reading and writing teaching [9], teachers and students lack sufficient attention to reading ability, writing ability, and communication ability. Second, the atmosphere in the teaching of English reading and writing is tense and fast. The English books that students can access every day are extremely limited. In addition to English textbooks, there are volumes of exercises and sets of exercises. As a result, the current stage of English reading and writing course teaching content has significant limitations and has failed to effectively and successfully broaden the scope of knowledge [10]. Many English reading and writing teachers do not know how to organize classroom activities, and some teachers do know how to organize classroom activities, but the effect is clearly difficult to

achieve expectations. To carry out classroom activities in English reading and writing classes, it is necessary to combine the flipped classroom teaching mode [11–13].

The ability of reading and writing is one of the main abilities of English learning, which occupies a vital position. For many years, the traditional teaching of English reading and writing has been based on students' passive acceptance, which has led to many students' boredom of reading and writing, resulting in the decline of students' learning confidence [5, 14], which makes it difficult for students to improve their English scores. Many scholars are studying how to change this situation. In English reading and writing teaching, using modern information technology as an auxiliary teaching means, improve students' learning confidence and interest, via class activities, read presses to write, to write gong of reading, speaking, reading, and writing as a whole, guide students to learn actively.

However, the speech signal of English reading aloud [15] has its complexity. It needs to take into account the students' different accents, volume, speed, pauses, and other factors. The input data has specificity due to different users, making the input of speech data more complicated and more complicated. Diversification adds a lot of difficulty to the recognition and training process of the model. According to the different types of speech data processed, it can be divided into isolated word recognition [16, 17], small vocabulary continuous speech recognition [18, 19], and large vocabulary continuous speech recognition [20, 21]. This article mainly focuses on the research of medium and long sentences with large vocabulary. In the structure of long sentences with large vocabulary, the context contains greater relevance. This relevance has a great effect on the learning process of the model. Many training models often ignore such relevance and only learn the current state. A single mapping relationship leads to low efficiency. Combining the characteristics of recurrent neural networks in machine learning [22], this article believes that assigning a certain "memory" weight to the relevant information of the context will help the judgment of the current output and more accurately convert the time sequence information of speech into the corresponding correct characters. This article will focus on the research and analysis of the variant structure GRU under the recurrent neural network branch [23, 24], make some improvements to its internal calculation structure, and combine it with a popular time series classification method [25, 26] to reduce the error rate in recognition results.

The main innovations of this paper are as follows:

(1) I built an intelligent assisted learning algorithm for English flipped classroom based on recurrent neural network, and the experimental results proved that in the intelligent learning mode of English flipped classroom, students autonomously explore and discover knowledge, and their learning enthusiasm and interest are greatly enhanced, flipping the classroom. The teaching effect has also been greatly improved

(2) To solve the existing problems in the field of English speech recognition, a hybrid cyclic neural network model incorporating LeakyReLU functions is proposed in this paper. The characteristics of the model are mainly based on a variant structure gated loop unit of the neural network. The relevance of the context is fully considered through the double-gated mechanism, selective memory is carried out through weight assignment, and the joining time sequence classification algorithm is introduced to improve the speed of model training

The remainder of this paper is organized as follows: Section 2 analyzes some related work. Some details of the principles of the proposed algorithm and related submodules are introduced in Section 3. Section 4 provides details of experimental results. In Section 5, the conclusion based on this study is given.

2. Related Work

Robert modelized the flipped classroom teaching process and clearly divided flipped classrooms. Specifically, it can be divided into two stages: before class and during class. The preclass mainly includes students watching instructional videos for target teaching and completing well-designed preclass exercises [27–29]; the inclass stage includes a small amount of teacher evaluation, student problem-solving, problem feedback, and evaluation of four steps. Through the analysis of the current popular flipped classroom teaching model, it can be found that most of the flipped classroom teaching process models built at home and abroad are based on Robert's structural model diagram as the main prototype and then combine different learning theories to target specific teaching situations. Make changes and start research. Aaron Sams believes that teachers can not only use the excellent teaching videos of other teachers but also use the Camtasia Studio screen recording software and other screen recording tools to make their own teaching videos, but they should ensure that the videos are short, vivid, and clear. With the deepening of flipped classroom research and practice, educators believe that to a large extent, the learning effect of students' preclass teaching videos determines the effect of classroom teaching. As a result, more focus is placed on assessment, as well as teachers and students, in order to communicate with preclass video learning impacts. Chung et al. used the "first guiding principle" of meta-design theory to design a flipped classroom method and conducted a two-stage study in two middle schools. Participants had a total of 382 students and 5 teachers, who came from four subject areas, namely, mathematics, physics, Chinese, and information and communication technology (ICT). Based on the experience of the pilot study, the author improved the flipped classroom model and tested its effect through the quasiexperimental design in the main study. The results found that the learning outcomes of the students in the flipped ICT course were similar to those of the nonflip ICT courses, but the student performance levels of the other three courses (namely,

mathematics, physics, and Chinese) improved after the flip, with small- and medium-sized effects. Finally, some practical suggestions are put forward in the article, and ideas for further research are put forward.

In addition, Sarah et al. look at effective classroom activities in English as a foreign language (EFL) and propose a student response system (SRS) to support teachers in organizing classroom activities in flipped classes. In order to investigate the effectiveness of this method, a quasiexperiment was conducted in an EFL classroom of an engineering school. The experimental group used SRS for classroom activities, while the control group followed the usual method. The results show that the use of SRS can improve students' motivation and self-efficacy in learning English grammar and improve their participation in classroom activities during flipped learning. In addition, the results of the questionnaire survey show that students accept SRS as a teaching method in EFL flipped classroom. Hernan et al. [30] described a teaching intervention that included a range of creative activities aimed at improving students' oral and writing English, particularly those who showed interest or attention deficits. At first, the participants did not appear to be particularly interested in learning the language. In the end, they were more willing and motivated to participate in chain games, creative writing, and screenwriting exercises after hearing about the proposed approach. These activities have aided students in improving their spoken and written English fluency as well as their comprehension of English grammar and structure.

3. Methodology

3.1. Definition of Flipped Classroom. The so-called flipped classroom is relative to the current teaching method of teacher explanations, students listening to lectures, and student homework after class; it refers to the use of information technology to facilitate teachers to record the explanations of knowledge points into short and succinct teaching methods. The video, accompanied by other learning materials and advanced homework, is sent to students through the learning management platform. Under the guidance of the teacher, the students will conduct self-study and complete the advanced homework; based on the information on the learning platform, the teacher is grasping the learning situation in detail. Under the circumstance, the classroom will focus on the targeted explanation, solve the problems with the students, and complete the homework. In the flipped classroom teaching model, the acquisition of knowledge mainly relies on students' self-study before class and no longer takes up students' valuable classroom time for teachers to teach. Students can use a variety of channels: online discussions with other students, watching related video materials, reading required image materials, etc. In the time in the classroom, students can learn more actively based on the information obtained before class to gain a deeper understanding. Teachers can also make full use of the remaining spare time to communicate with everyone. After class teaching, students plan their own learning content, learning pace, and learning style and adopt appropriate

knowledge presentation methods; teachers use cooperation and teaching methods to meet the needs of students, promote students' personalized inquiry learning, and let students get a real learning experience in self-inquiry.

The essence of the flipped classroom is to return the dominance of learning to the students, lead the students' subjective teaching methods, and have the energy to create education for the future. And flipped classroom as a kind of teaching concept and teaching mode is influencing and changing the traditional classroom teaching. It makes use of Internet technology and information technology to break down traditional classroom barriers, expand classroom teaching time and space, optimize students' learning processes, improve students' learning abilities, and realize indepth integration of information technology and curriculum teaching, all of which promote students' deep learning. At the same time, the flipped classroom model, as a part of the education reform movement, will completely subvert the traditional printing-based classroom teaching structure and teaching process and trigger a series of changes in the role of teachers, curriculum models, and management models.

Autonomous learning, in contrast to traditional acceptance learning, places a greater emphasis on students' ability to learn independently. As the subject of learning, students achieve their own learning goals through independent analysis, exploration, practice, and creation. We should change the current situation of curriculum implementation that emphasizes too much on learning, rote memorization, and mechanical learning, advocate students' active participation, willing to explore and diligent in doing, and cultivate students' ability to collect and process information, acquire new knowledge, analyze and solve problems, and communicate and cooperate. The concept of autonomous learning is based on inquiry-based learning, in which students are presented with situations in which they must conduct their own research, solve problems, and gain knowledge in the field. As a form of discovery learning, students in the classroom are provided with more opportunities to "experience and interact" with knowledge. It is promising to use the current popular neural network technology [31–33] to assist English learning.

3.2. English Speech Recognition

3.2.1. Language Model. There are often multiple combinations of output sequences that can meet the original signal input conditions in the speech recognition process. An external tool is required to assist the machine in constraining the discriminative category of words when judging the output, resulting in a more logical collocation situation. The role of the language model is to summarize the most probable text sequence on the output of the acoustic model, just like a dictionary. In the method of processing continuous speech recognition with large vocabulary, the commonly used language model is the N-gram language model, which is characterized by using the relationship between N adjacent words in the context to calculate the current maximum probability output. The N-gram model

is based on the assumption that the current output is only related to the previous $N-1$ words and has no relationship with other words. Assuming a word sequence is $W = (W_1, W_2, \cdots, W_N)$, then its probability calculation process is as follows:

$$P(W) = P(W_1, W_2, \cdots, W_N) = \prod_{i=1}^{N} P(W_i \mid W_{i-N+1}, \cdots, W_{i-1}),$$
$$(1)$$

where N is the number of words in the sequence and W_i is the i-th word in the sequence W.

There are certain restrictions on the selection of the N value in N-gram. Generally, 1~5 is more appropriate. If the value is too large, it will cause data sparseness. The more commonly used are binary language model (bi-gram) and the ternary language model (tri-gram). Among them, bigram only considers the impact of its previous word on the current judgment, and its probability calculation equation is as follows:

$$P(W) = P(W_1, W_2, \cdots, W_N) = \prod_{i=1}^{N} P(W_i \mid W_{i-1}), \quad (2)$$

where $P(W_i \mid W_{i-1})$ can be calculated by maximum likelihood estimation:

$$P(W_i \mid W_{i-1}) = \frac{\text{count}(W_{i-1} W_i)}{\text{count}(W_{i-1})}, \quad (3)$$

where $\text{count}(W_{i-1} W_i)$ is the frequency of occurrence of the phrase $W_{i-1} W_i$ in the transcription text set and $\text{count}(W_{i-1})$ is the frequency of occurrence of the word W_{i-1}. Similarly, tri-gram considers the influence of the first two words on the current judgment.

3.2.2. Pronunciation Dictionary. The pronunciation dictionary keeps track of the various characters to which each pronunciation can respond. It is generally identified by various linguistics official records and can be used directly in recognition methods to aid learning. In Chinese, the corresponding relationship is pinyin and Chinese characters. For example, when learning a Chinese character, we know how to pronounce it, but we do not know which character it corresponds to. We need to use tools like "Xinhua Dictionary" to assist in learning. Similarly, in the process of voice recognition, it also needs the participation of such a tool. Following the previous example, in the pinyin part, "ta" in the pronunciation dictionary can correspond to the Chinese characters "it," "he," "she," etc., which compresses the recognition results within certain constraints. The model is then used to match the most likely characters using the process of learning context information. The function of the pronunciation dictionary, from this perspective, is to match the phoneme information from the acoustic model with the word information from the language model. It is a highly effective tool for acoustic recognition.

3.3. Our Model

3.3.1. Gated Recurrent Unit. The intuitive improvement of gated recurrent unit is to merge the three gate units in the LSTM network into two gate units, which is more concise and refined in structure and is a good deformed structure. As a modified structure of LSTM, LSTM can solve the long dependency problem in RNN, as can GRU. On this basis, GRU also reduces the number of parameters, avoiding the problem of slower training speed due to more and more parameters in the longer-term training and learning of LSTM.

Figure 1 is the unit state of the previous unit, h_{t-1} is the input of the current unit, and they are collectively used as the calculation input of the current unit, and h_{t-1} carries the information passed by the previous unit selection. GRU turns the input gate, forget gate, and output gate in LSTM into two gates: update gate z_t and reset gate r_t. Among them, the update gate is mainly responsible for judging the degree to which the transfer information of the previous unit at the previous moment is calculated into the current state. The larger the value of the update gate, the deeper the state information previously transferred is calculated into the current state. The reset gate is mainly responsible for controlling how much information of the previous unit state is written to the candidate set of the current unit. The smaller the reset gate, the less information of the previous state is written. Finally, the unit state and output are combined into a state h_t, which makes the model have fewer parameters and higher processing efficiency.

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z), \quad (4)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r), \quad (5)$$

$$h'_t = \tanh(W_h x_t + U_h(h_{t-1} r_t) + b_h), \quad (6)$$

$$h_t = z_t h_{t-1} + (1 - z_t) h'_t. \quad (7)$$

The place where GRU is more efficient than LSTM is that it uses one step to complete the two actions of selective memory and selective forgetting.

3.3.2. Connectionist Temporal Classification. The connectionist temporal classification (CTC) was first proposed by Graves et al., as a temporal classification algorithm. Mainly used for the alignment of speech signals and text labels, CTC can help save the time cost of human alignment and improve training efficiency. Traditionally used cross entropy loss requires that the tags of the training data be aligned at the frame level before processing the speech sequence data. The alignment operation requires a certain amount of work, and the model needs to know the corresponding tags of each frame before training, which reduces work efficiency. In the process of training, CTC can automatically optimize the mapping relationship between the input sequence and the output sequence for direct training, which greatly improves the speed of decoding.

We use CTC to process the GRU output information. The output length of GRU is the same as the input length.
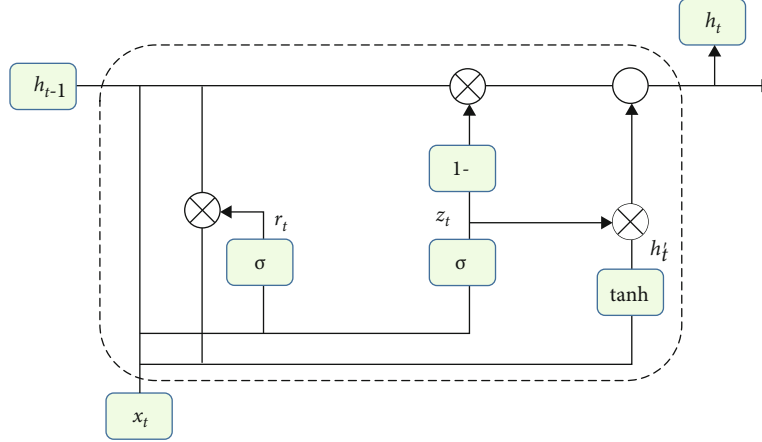
FIGURE 1: Schematic diagram of GRU unit structure.

With N voice signal inputs, there are N probability vector outputs, including the probability of each character corresponding to the dictionary. When the voice signal X is input,

$$\bar{Y} = \underset{Y}{\operatorname{argmax}} P(Y \mid X), \tag{8}$$

where $\bar{Y}$ is the output with the highest conditional probability and $P(Y \mid X)$ is the posterior probability of Y. During training, I hope to obtain its maximum value, that is, the most likely output. In recognition processing, the nearest similar output can be found more quickly.

As shown in Figures 2 and 3, the CTC algorithm does not require the input and output to be strictly aligned. Multiple output paths may correspond to one output result. Understanding the corresponding relationship between input and output can help us better understand the calculation method of loss function and the calculation method used in the test. The most prominent feature of CTC is the introduction of a placeholder blank node, mainly to model the parts without effective information, such as silence and pause, to represent the output state of the network when predicting uncertain information. After the alignment of the blank, it needs to be deleted. This process is called F transformation. If an output sequence of the network can be mapped to the correct annotated sequence through F transformation, then the output sequence is a CTC path. The process of F transformation is as follows: first, delete the repeated annotations between adjacent blank nodes in the sequence and then delete the blank nodes.

3.3.3. LeakyReLU Improved GRU. The calculation equation of LeakyReLU function is as follows:

$$\operatorname{LeakyReLU}(x) = \begin{cases} x, & x \geq 0, \\ ax, & x < 0, \end{cases} \tag{9}$$

where a is the fixed value of the slope selected before training, with the purpose of making the nonpositive part have a small slope. Compared with the training characteristics of sigmoid and tanh function, it is more consistent with
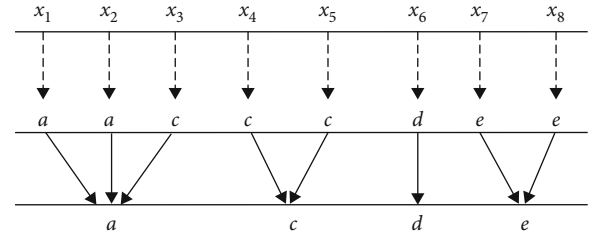

FIGURE 2: Schematic diagram of traditional alignment.
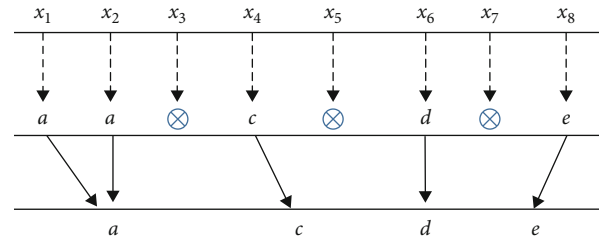

FIGURE 3: Schematic diagram of the alignment process of CTC placeholders.

the convergence effect that the training wants to achieve, so that the training process of the model can be stable and efficient.

4. Experiments

4.1. Parameter Settings. The experiments with all the algorithms were performed on a computer equipped with a single NVIDIA GTX1080TI GPU (11 GB). We have implemented the model construction through the PyTorch deep learning library, the programming language we use is Python, and we batch processed 100 samples each time.

4.2. Datasets. The THCHS30 database can be used to build a comprehensive Chinese speech recognition benchmark system, including performance under high noise conditions. The voice data from the dataset was recorded by 40 people in a quiet office environment through a carbon particle microphone. The total time is 35 hours. Most of those

involved in the recording were college students who spoke Mandarin. The recorded content text is from large-capacity news, the sampling frequency is 16 kHz, the sampling size is 16 bits, and the format is WAV. The composition of the dataset is shown in Table 1.

4.3. Evaluation Methods. The evaluation methods I use are character error rate, sentence error rate, and word error rate. The calculation equations are as follows:

$$CER = \frac{S_C + D_C + I_C}{N_C} \times 100\%, \quad (10)$$

$$SER = \frac{S_W + D_W + I_W}{N_W} \times 100\%, \quad (11)$$

where S_C is the number of replacement characters, D_C is the number of deleted characters, I_C is the number of inserted characters, and N_C is the number of characters. S_W is the number of replacement words, D_W is the number of deleted words, I_W is the number of inserted words, and N_W is the number of words.

4.4. Experimental Results. To test the sexual superiority of GRU in relation to phonological recognition patterns in other languages, LSTM-HMM, LSTM-CTC, and GRU-CTC were selected as the models for the comparative experiment. Through the comparison of CER and SER of the processing results of the same data set of each model in the experimental environment, the smaller the value, the better the performance in the experiment.

From the data in Table 2, the recognition effect of the hybrid model with CTC is lower than the CER and SER of the traditional HMM hybrid model. This set of comparison results shows that the introduction of the CTC method can help the performance of the model to a certain extent. The recognition effect has been improved to a certain extent, mainly due to the fact that the CTC method can directly model the entire sentence when the input and output are not aligned, which improves the completeness and relevance of the entire sentence recognition. The CER and SER of the GRU-CTC model, like the previous control group LSTM-CTC, combined with the CTC approach, are the best among the three deep learning-based models selected in the experiment and LSTM. CER and SER are both reduced as compared to CTC. Combining the above results and the characteristics of CTC and GRU, it can be concluded that CTC, as an efficient method that does not require manual label alignment, and the combination of LSTM, GRU, and other cyclic neural network variants have achieved considerable results. The improved double-door variant structure is more reasonable in the deletion and optimization of the structure, and the control of the relevance of the sequence context information is more efficient, and the recognition performance of the model has been improved. In summary, the hybrid cyclic neural model of GRU and CTC is a relatively efficient method of speech recognition.

4.5. Ablation Experiment of Activation Function. In order to further verify the role played by the activation function in

TABLE 1: Components of THCHS30 dataset.

Features	Train	Test
Audio duration	27.23	6.25
Number of speakers	30	10
Male	8	1
Female	22	9
Age	20-55	19-50
Total words	215995	49085

TABLE 2: Comparative experiment results.

Methods	Train (CER)	Train (SER)	Test (CER)	Test (SER)
LSTM-HMM	0.3525	0.4352	0.2874	0.3325
LSTM-CTC	0.3415	0.3952	0.2956	0.3425
GRU+CTC (ours)	0.3325	0.3699	0.2811	0.3155

TABLE 3: Ablation experiment results.

Methods	Train (CER)	Train (SER)	Test (CER)	Test (SER)
ELU	0.3625	0.4158	0.2858	0.3214
ReLU	0.3398	0.4012	0.2915	0.3396
LeakyReLU (ours)	0.3325	0.3699	0.2811	0.3155

the proposed algorithm, we added an ablation experiment. We selected LeakyReLU, ReLU, and ELU, which are the three activation functions for experiments. The experimental results are shown in Table 3.

It can be clearly seen from Table 3 that after using the three activation functions, respectively, the performance of LeakyReLU is the best, and secondly, the performance of ReLU is better than that of ELU. Therefore, it is reasonable and effective to choose the LeakyReLU activation function in this paper.

5. Conclusion

In this paper, we propose a recurrent neural network-based intelligent assisted learning algorithm for creating an English flipped classroom. There are two main characteristics of this model. For starters, it is a gated recurrent unit based on a recurrent neural network's variant structure. On this basis, the dual-gating mechanism fully considers the context's relevance, selects memory through weight allocation, and integrates the novel LeakyReLU function, which can improve the model's training convergence efficiency. Second, the connection sequence classification algorithm can be used to solve the general model. The time-consuming problem of processing input and output alignment is moderate, and model training speed has increased. The experimental results show that in the English flipped classroom's intelligent learning mode, students explore and discover

knowledge independently, and their enthusiasm and interest in learning are greatly enhanced, as is the flipped classroom's teaching effect.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author does not have any possible conflicts of interest.

References

[1] P. Zheng, X. Wang, and J. Li, "Exploration and practice of curriculum ideological and political construction reform——take "information security" course as an example," *ASP Transactions on Computers*, vol. 1, no. 1, pp. 1–5, 2021.

[2] F. Moller and T. Crick, "A university-based model for supporting computer science curriculum reform," *Journal of Computers in Education*, vol. 5, no. 4, pp. 415–434, 2018.

[3] K. Pyhältö, J. Pietarinen, and T. Soini, "Dynamic and shared sense-making in large-scale curriculum reform in school districts," *The Curriculum Journal*, vol. 29, no. 2, pp. 181–200, 2018.

[4] G. Akçayır and M. Akçayır, "The flipped classroom: a review of its advantages and challenges," *Computers & Education*, vol. 126, pp. 334–345, 2018.

[5] I. T. Awidi and M. Paynter, "The impact of a flipped classroom approach on student learning experience," *Computers & Education*, vol. 128, pp. 269–283, 2019.

[6] P. Strelan, A. Osborn, and E. Palmer, "The flipped classroom: a meta-analysis of effects on student performance across disciplines and education levels," *Educational Research Review*, vol. 30, p. 100314, 2020.

[7] N. Amalia, *Narrative Writing Intervention Plan: Analysis of Students' Literacy Learning Needs*, Muhammadiyah University Press, Indonesia, 2017.

[8] J. Castellví, M. C. Díez-Bedmar, and A. Santisteban, "Pre-service teachers' critical digital literacy skills and attitudes to address social problems," *Social Sciences*, vol. 9, no. 8, p. 134, 2020.

[9] L. Liang, Q. Yin, and C. Shi, "Exploring proper names online and its application in English teaching in university," *ASP Transactions on Computers*, vol. 1, no. 1, pp. 24–29, 2021.

[10] A. R. Kothari, J. J. Bickford, N. Edwards, M. J. Dobbins, and M. Meyer, "Uncovering tacit knowledge: a pilot study to broaden the concept of knowledge in knowledge translation," *BMC Health Services Research*, vol. 11, no. 1, pp. 1–10, 2011.

[11] Z. Turan and B. Akdag-Cimen, "Flipped classroom in English language teaching: a systematic review," *Computer Assisted Language Learning*, vol. 33, no. 5-6, pp. 590–606, 2020.

[12] M. Y. Abdullah, S. Hussin, and K. Ismail, "Implementation of flipped classroom model and its effectiveness on English speaking performance," *International Journal of Emerging Technologies in Learning*, vol. 14, no. 9, 2019.

[13] B. Ayçiçek and T. Y. Yelken, "The effect of flipped classroom model on students' classroom engagement in teaching English," *International Journal of Instruction*, vol. 11, no. 2, pp. 385–398, 2018.

[14] S. Karabatak and H. Polat, "The effects of the flipped classroom model designed according to the ARCS motivation strategies on the students' motivation and academic achievement levels," in *Education and Information Technologies*, pp. 1–21, Springer, 2019.

[15] E. P. Johnson, B. F. Pennington, J. H. Lowenstein, and S. Nittrouer, "Sensitivity to structure in the speech signal by children with speech sound disorder and reading disability," *Journal of Communication Disorders*, vol. 44, no. 3, pp. 294–314, 2011.

[16] T. Dijkstra, A. Wahl, F. Buytenhuijs et al., "Multilink: a computational model for bilingual word recognition and word translation," *Bilingualism: Language and Cognition*, vol. 22, no. 4, pp. 657–679, 2019.

[17] P. Karageorgos, T. Richter, M. B. Haffmans, J. Schindler, and J. Naumann, "The role of word-recognition accuracy in the development of word-recognition speed and reading comprehension in primary school: a longitudinal examination," *Cognitive Development*, vol. 56, p. 100949, 2020.

[18] V. Mitra, G. Sivaraman, C. Bartels et al., "Joint modeling of articulatory and acoustic spaces for continuous speech recognition tasks," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5205–5209, New Orleans, LA, USA, 2017, March.

[19] N. H. Gebreegziabher and A. Nürnberger, "An Amharic syllable-based speech corpus for continuous speech recognition," in *In International Conference on Statistical Language and Speech Processing*, pp. 177–187, Springer, Cham, 2019.

[20] K. Audhkhasi, B. Kingsbury, B. Ramabhadran, G. Saon, and M. Picheny, "Building competitive direct acoustics-to-word models for English conversational speech recognition," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4759–4763, Calgary, AB, Canada, 2018, April.

[21] Z. Zeng, Y. Khassanov, V. T. Pham, H. Xu, E. S. Chng, and H. Li, "On the end-to-end solution to Mandarin-English code-switching speech recognition," 2018, http://arxiv.org/abs/1811.00241.

[22] W. Chu, P. S. Ho, and W. Li, "An adaptive machine learning method based on finite element analysis for ultra low-k chip package design," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, pp. 1–1, 2021.

[23] R. Parthiban, R. Ezhilarasi, and D. Saravanan, "Optical character recognition for English handwritten text using recurrent neural network," in *In 2020 International Conference on System, Computation, Automation and Networking (ICSCAN)*, pp. 1–5, Pondicherry, India, 2020, July.

[24] J. Wang, "Speech recognition in English cultural promotion via recurrent neural network," *Personal and Ubiquitous Computing*, vol. 24, no. 2, pp. 237–246, 2020.

[25] J. Zhang, W. Wang, C. Lu, J. Wang, and A. K. Sangaiah, "Lightweight deep network for traffic sign classification," *Annals of Telecommunications*, vol. 75, no. 7-8, pp. 369–379, 2020.

[26] J. Zhang, J. Sun, J. Wang, and X. G. Yue, "Visual object tracking based on residual network and cascaded correlation filters," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 8, pp. 8427–8440, 2020.

[27] A. T. Taylor, E. L. Olofson, and W. R. Novak, "Enhancing student retention of prerequisite knowledge through pre-class activities and in-class reinforcement," *Biochemistry and Molecular Biology Education*, vol. 45, no. 2, pp. 97–104, 2017.

[28] J. Jovanovic, N. Mirriahi, D. Gašević, S. Dawson, and A. Pardo, "Predictive power of regularity of pre-class activities in a flipped classroom," *Computers & Education*, vol. 134, pp. 156–168, 2019.

[29] E. Han and K. C. Klein, "Pre-class learning methods for flipped classrooms," *American Journal of Pharmaceutical Education*, vol. 83, no. 1, p. 6922, 2019.

[30] C. J. Hernan, T. A. Collins, J. Q. Morrison, and S. D. Kroeger, "Decreasing inappropriate use of mobile devices in urban high school classrooms: comparing an antecedent intervention with and without the good behavior game," *Behavior Modification*, vol. 43, no. 3, pp. 439–463, 2019.

[31] Y. Ding, X. Zhao, Z. Zhang, W. Cai, and N. Yang, "Multiscale graph sample and aggregate network with context-aware learning for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4561–4572, 2021.

[32] Y. Tong, L. Yu, S. Li, J. Liu, H. Qin, and W. Li, "Polynomial fitting algorithm based on neural network," *ASP Transactions on Pattern Recognition and Intelligent Systems*, vol. 1, no. 1, pp. 32–39, 2021.

[33] Z. Chu, M. Hu, and X. Chen, "Robotic grasp detection using a novel two-stage approach," *ASP Transactions on Internet of Things*, vol. 1, no. 1, pp. 19–29, 2021.