[1,*]Xuemei Shi

[2]Xiaoyan Li

[3]Xiedong Song

[4]Dalei Zhang

Character Data Mining in Educational Scene

Journal of Electrical Systems

Abstract: Teaching and learning have been completely transformed by the quick growth of information technologies like big data, artificial intelligence, and the Internet of Things. Traditional educational methods can no longer satisfy the demands of modern fast-paced and lifelong learning, making the mining of educational data more urgent. Character mining in images is increasingly applied in educational settings. Artificial intelligence and machine learning algorithms, learning behaviors, such as CNN(Convolutional Neural Networks and RNN()Recurrent Neural Networks), have been used, to predict student performance, and provide personalized recommendations for educational resources. Therefore, research on the application of character mining in educational scenes should be conducted and implemented using OCR technology and the CRNN algorithm. The prediction and mining results are promising and hold commercial value. With the diversification of educational scenarios, the adaptability and flexibility of algorithms will become important research directions. Ultimately, the advancement of these technologies will further drive the digital transformation of education, providing learners with richer and more efficient learning resources.

Keywords: Character, data mining, educational, OCR, recognition

I. Introduction

Aulakh K [1] and Colleagues explored how Educational Data Mining (EDM) could be used in online learning environments. They discussed several EDM techniques, including regression, association rule mining, support vector machines, decision trees, K-means clustering, Naive Bayes, K-Nearest Neighbors, and Random Forest classification. These methods are frequently used to evaluate online learning environments. and understand learning behaviors, academic performance, and students' learning styles. They provide an assessment of students' learning performance, personalize student learning through course adjustments and learning recommendations, and evaluate online learning content. However, there is a dependency on specific learning environments, and a large amount of data is required to train and test models. It is used for identifying student complaints, analyzing dropout rates, conducting thorough performance evaluations of both students and teachers, predicting of student performance in higher education, assessment of learning, and assistance in teaching and research work. Alghamdi, Amnah Saeed, and Atta Rahman [2] introduced a network model and method called the algorithm inspired by the chaos-enhanced sine-wave (CESCA). The sample data from Wenzhou University, which comprised 702 cases and 12 features like gender, GPA, math courses, and English courses, was used to apply this model. The CESCA model demonstrated its advantages through comparison with other models, such as Random Forest and kernel methods. Reliance on data augmentation and feature extraction techniques is used to enhance performance. The model can be applied to analyze the leading factors of student success, such as semester grades, family income, parental occupation, and the prediction of high school students' success, among others. Lei Zhang and Han Yu[3] employed methods such as theoretical guidance and technical routes, evaluation data collection, evaluation data mining, and evaluation data analysis to study the application of the entire industry chain in university e-commerce applied undergraduate talent training based on collaborative education. This includes industry status analysis, market positioning, corresponding curriculum systems, and practical segments. The model offers research value and promise by integrating the fundamental components of conventional evaluation techniques. The application of big data in the quality assessment of talent training is still largely unexplored in research, largely because of the diversity of academic disciplines and different evaluation criteria. However, traditional methods of evaluation data collection are relatively simple and lack big data evaluation approaches. Roslan [4] and Colleagues used Decision Trees (DT) and Naive Bayes (NB) as the main data mining techniques to forecast how well children will succeed in math and English classes. These techniques are considered effective classifiers that can help understand and predict students' academic achievements. The method is easy to understand and

[1] Huainan Normal University, School of Computer Science, Huainan, Anhui, China
[2] Huainan Normal University, School of Computer Science, Huainan, Anhui, China
[3] School of Computer Science An And Engineering, Jining University, Jining, China
[4] Huainan Normal University, School of Computer Science, Huainan, Anhui, China
*Corresponding author: Xuemei Shi

very effective in certain applications, especially in real-time prediction and multi-category prediction. The study's limited sample size restricts how far the findings may be applied. The student samples lacked religious and ethnic diversity, and the selected schools were all urban public schools, which may affect the diversity of the research results. The application field involves analyzing students' performance in English and Mathematics subjects, as well as the correlation between these performances. By understanding these predictive factors, academic intervention measures can be optimized. Alkashami M [5] and others discussed the use of the ANFIS (Adaptive Neuro-Fuzzy Inference System) model to predict the early employment readiness of graduates in Middle Eastern countries, especially in Jordan. The study detailed the methodology of data preprocessing, ANFIS implementation, and testing phases. Using a confusion matrix, the accuracy of ANFIS was compared with other classifiers such as Decision Trees, SVM, Naive Bayes, and MLP. The results show that ANFIS is superior to other methods, highlighting the benefits of integrating neural networks with fuzzy logic in predictive models. ANFIS performs excellently in predicting the 'employment' category and also does quite well in predicting the 'unemployment' category. The method relies on data enhancement and feature extraction techniques to improve model performance. It is applied to analyze the main factors affecting student success, such as semester grades, family income, parental occupation, etc.

Rangineni, Sandeep, and Divya Marupaka [6] used classifier development operations to evaluate the answers of American students on the 2022 PISA problem-solving questions. The research utilized an array of data mining methodologies, such as association rule mining, classification, clustering, sentiment analysis, text mining, and image recognition methods. Feature selection was based on theoretical foundations, eliminating the need for additional feature selection algorithms, and providing recommendations for choosing classifiers when considering research topics, classifier interpretability, and simplicity. The sample size used in the study was relatively small, which may limit the general applicability of the results. The educational process was evaluated and improved by analyzing students' performance in an e-learning environment. The practical impact of data mining techniques was demonstrated across multiple industries, including healthcare, finance, and marketing.

Self-regulated learning (SRL) profiles in an online learning environment can be identified by applying Educational Data Mining (EDM) approaches, as explored by Araka E. [7] and colleagues. The usage of clustering algorithms like K-means was emphasized, as Expectation Maximization, and Agglomerative Hierarchical Clustering to categorize learners based on their interactive behaviors. By analyzing learners, targeted educational interventions are provided to improve SRL strategies. The study shows that students with poor self-regulation skills often fail or drop out of courses, while those with strong self-regulation skills tend to achieve higher grades. Learner profiles can be identified and targeted interventions designed to support SRL in an online learning environment. Tin Tin Ting and others[8] explored the moderating role of educational big data mining in analyzing the academic performance of adolescent criminal behavior. Through rigorous data preprocessing, the accuracy of the analysis was ensured. Only Pearson correlation and mediation analysis were used in the study, which may not fully reveal the complex relationships between variables. The methods of this study can be applied to the formulation of educational policies, the development of strategies for preventing juvenile delinquency, and social science research on adolescent behavior. Zhen Liu [9] and colleagues focused their research on the safety of vehicle driving at underground traffic interchanges, utilizing driving simulators and data mining analysis for the study. When conducting control experiments using this platform, the performance of the mechanical system was significantly improved. Although theoretical derivation methods have strong advantages, practical problems are more abstract, making theoretical development more challenging; therefore, this research method requires further study and understanding. This method is mainly applied in the design of underground roads and the development of driving simulators. Križanić S [10] explored the use of clustering analysis and decision tree techniques in Utilizing educational data mining to examine how students behave in an online learning environment. Clustering analysis was used to identify patterns in student behavior, and decision tree techniques were employed to predict exam performance. The ability to reveal the correlation between the frequency of students accessing instructional materials and their exam results emphasized the impact of e-course content management on learning outcomes. The study may rely on specific educational settings and student behavior patterns, which could limit its universal applicability in different contexts. The method is applied to analyze student behavior in an e-learning environment, predict students' exam performance, and facilitate timely educational interventions.

Based on the above research, we utilized CRNN, OCR technology, and CTPN to mine characters in the field of education and conducted research on data mining in educational scenarios. CRNN is used for recognizing and decoding text information in images, while CTPN is used for detecting text lines in images. With the OCR and

CTPN network, printed or handwritten text in paper documents, images, or videos can be converted into machine-encoded text. These technologies have significantly improved the efficiency and accuracy of character mining in the field of education. The research can be widely applied in areas such as educational intelligent scoring systems, document management, and blackboard writing recognition.

II. Summary

In educational settings, the application of character Mining typically involves two key components: the training model and the character detector. Here, ‘train_crnn’ refers to the training process of the CRNN, which is a deep learning model capable of recognizing and decoding textual information from images. The term ‘train_ctpn’ refers to the training of the Connectionist Text Proposal Network, an algorithm used for detecting lines of text in images. The combined use of these two technologies can significantly enhance the efficiency and accuracy of character Mining in educational settings. For example, in intelligent grading systems, train_crnn can be used to recognize and interpret the text in student answers, while train_ctpn can locate the specific answer areas on the answer sheets. Such systems can not only speed up the process of grading papers but also enhance the consistency and accuracy of scoring. Moreover, these technologies can be applied to the digitization of textbooks and notes, by identifying and transcribing printed or handwritten text, making the content easy to store, search, and share. With the continuous advancement of artificial intelligence technology, we can foresee that Shortly, character-mining technology will become more significant in the sphere of education. Using OCR[11] technology and the CTPN network, OCR is an important computer vision technology that can convert printed or handwritten text from paper documents, images, or videos into machine-encoded text. This technology is widely used in various scenarios, such as automatic data entry, document digitization, license plate Mining, and personal document management, among others. The core of OCR technology lies in its ability to recognize and process text of different fonts, sizes, and formats, even against complex backgrounds. It determines the shapes of characters by analyzing the patterns of dark and light in the image, then translates these shapes into computer-readable text. OCR speed and accuracy have increased dramatically with the advent of deep learning technologies. The performance of OCR systems is typically measured by indicators such as rejection rate, error rate, and recognition speed. To improve the accuracy of recognition, OCR systems may combine auxiliary information, such as language models and contextual information, to optimize the results. Additionally, OCR technology is continually evolving, and the emergence of new technologies like Intelligent Character Recognition (ICR) enables systems to better understand and process various complex text information. The CTPN network, short for Connectionist Text Proposal Network, is a deep learning model specifically designed for text detection. It effectively detects text information in natural images by combining the features of CNN [12] and LSTM [13]. The core of the CTPN model lies in using vertical anchors to predict the position of text lines; these anchors have a fixed width and predict their vertical positions and classification scores through the network. The general workflow of the CTPN is as follows:

Feature Extraction: Utilizes the VGG16[14] network as a base to extract image features.

Sliding Window: Applies a 3x3 sliding window to get spatial features, use the feature map.

Bidirectional LSTM: Inputs the feature map into a bidirectional LSTM to extract sequence features, which helps in understanding the context of the text.

Text Proposal Boxes: Generates text proposal boxes using a structure similar to the Region Proposal Network (RPN)[15].

Anchor Prediction: Predicts the position and height of each anchor, as well as whether they contain text.

Text Line Construction: Combines the detected text proposal boxes into complete text lines.

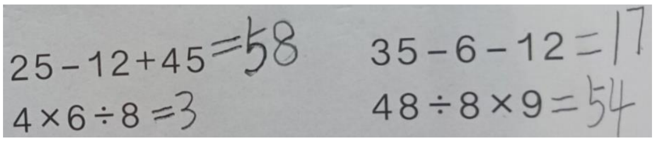
the crnn[16] algorithm is used to realize character Mining and output in educational images. First, input an education scene image, then extract the image's text, transform it into a text form save it to the file under the specified path, and output the results on the console.

An example of relevant education scene data is the image of the problem to be searched photographed by the student, as shown in Figure 1 below：

Crnn algorithm

Figure 1: Photo search

Photo image detection answer, as shown in Figure 2 below:



25-12+45=58 35-6-12=17

4×6÷8=3 48÷8×9=54

Figure 2: Photo Inspection

To solve the problem, we need to read all the contents of the image file, detect each small text area, first extract the features, then extract the sequence features, and finally get the prediction results.

III. Algorithm Analysis and Implementation

OCR technology uses the method of character Mining to translate it into words that can be recognized by the computer, scans the data, then processes and analyzes the image, and finally obtains the text and layout information. The performance indicators mainly include rejection rate, error rate, recognition speed, and feasibility.

ctpn ctpn[17] (connectionist text proposal network) algorithm: text detection essentially belongs to object detection, but text is quite different from conventional objects. Text is typically written from left to right, which is different from common objects, with the width between words being roughly the same. This writing style is standard in most European languages, such as English. This alignment is known as "left-aligned" in typesetting and page layout, where a paragraph's text is aligned with its left edge while its right side is asymmetrical. On websites throughout the world, this is the standard text alignment style for left-to-right text. Furthermore, text alignment does not affect the direction in which the text reads; yet, the direction of the text may dictate which alignment technique is most frequently employed for that particular script. Fixed width is enough to detect the height of the text, but it is still the RPN[18] method to deal with variable length sequences. The detected boxes can be combined and Form a sequence in the width direction.

ctpn network architecture: VGG[19] extracts features, blstm integrates context information, and completes detection based on RPN. In deep learning, pooling is a common operation used to reduce the dimensions of the feature map. It works by sliding a fixed-size window over the feature map and extracting certain statistical information (such as the maximum or average value) from each window. Reducing the spatial size of the data will result in a decrease in the number of parameters in the model, improving computational efficiency, and helping to prevent overfitting. After four pooling operations, one pixel in the feature map corresponds to 16 pixels in the original input, which means that each pooling operation halves the feature map's dimensions in width and height. If we assume that each pooling operation uses a 2x2 pooling kernel with a stride of 2 (meaning the size of the feature map is halved after each pooling), then after four pooling operations, each 4x4 area of the original input will be compressed into a single pixel in the feature map. This operation is very useful in convolutional neural networks because it allows the network to minimize the data's size while keeping crucial information. For example, if we have a 32x32 input image and we perform pooling operations on it four times, the final feature map will be of size 2x2. This means that each 16x16 region of the original image is compressed into a single pixel in the feature map. Moreover, pooling operations can also help the model achieve spatial invariance, meaning that even if objects in the image are translated, the model can still recognize them. This is because the pooling operation extracts statistical information within a local region, rather than specific pixel location information. Border adjustment can make the text detection box better.

The network structure has three parts, which are:

CNN convolution layer, can take input photos and extract feature sequences from them;

RNN cyclic layer, can forecast the feature sequence's label distribution after being acquired from the convolution layer;

The label re-integration and other operations derived from the circulation layer are transformed into the final prediction results via the CTC transcription layer.

crnn algorithm crnn absorbs the ctc+ LSTM modeling method in speech Mining but inputs the features of LSTM to transform the voice field's acoustic characteristics into the picture feature vector that the CNN network has extracted. The Crnn algorithm mainly integrates the advantages of CNN in image feature processing and LSTM in serialization recognition.

Through sequence recognition, it not only extracts robust features but also avoids the challenging single-character mining and segmentation found in standard methods. Of course, temporal dependency is also added by serialization recognition. The training image is equally scaled by CRNN during training. During testing, crnn preserves the input image's size proportion and requires the image height to be unified to 32 pixels to address issues such as character stretching's decreased recognition rate. The duration of the LSTM time is dependent on the size of the convolution feature image.

IV. Implementation Process and Analysis

In computer vision tasks, data reading and preprocessing are crucial steps. These procedures make sure that the input data's quality and format satisfy the specifications needed for model training. Below is a detailed explanation of each step you mentioned: Reading data: This is the first step in the preprocessing process, which involves loading images or other types of data from storage media such as hard drives, databases, or cloud storage. Obtaining labels: In supervised learning[20], each input data usually has a corresponding label, which represents the classification, attributes, or other relevant information of the data. The process of obtaining labels is to associate these pieces of information with the input data. Resizing: Since the original dimensions of the input data may vary or not meet the model’s input requirements, it is necessary to adjust the images to a uniform size. Generating candidate boxes: In object detection tasks, candidate boxes (also known as anchor boxes) are rectangular frames around potential target areas. These boxes define the areas that the model should focus on. Obtaining candidate boxes: This step involves calculating and selecting the candidate boxes that are most likely to contain the target. Filtering positive samples: Among the candidate boxes, some may contain the target (positive samples), while others may not (negative samples). The process of filtering positive samples is identifying and selecting candidate boxes that contain the target for further analysis. Data reading and preprocessing operations include reading data, obtaining labels, resizing, generating candidate boxes, obtaining candidate boxes, and filtering positive samples:

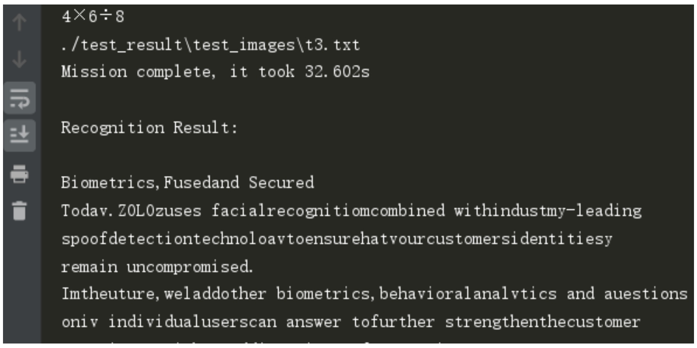
```
img_myname = self.img_mynames[idx]
my_img_path = os.path.join(self.datadir, img_myname)
img = cv2.imread(my_img_path)
#use default image for read error
if img is None:
print(img_path)
with open('error_imgs.txt','a') as f:
f.write('{}\n'.format(img_path))
img_myname = 'img_2647.jpg'
img_path = os.path.join(self.datadir, img_myname)
img = cv2.imread(img_path)
h, w, c = img.shape
rescale_fac = max(h, w) / 1600
if rescale_fac>1.0:
h = int(h/rescale_fac)
w = int(w/rescale_fac)
img = cv2.resize(img,(w,h))
```

Load the vgg16 model network to get the characteristic map, and add a 3*3 convolution layer to connect the full connection layer.

Build crnn image data, and create a text file, which includes the path where the image is located, the label corresponding to the image, etc.

Bidirectional sequence feature extraction is a powerful technique that allows the model to consider both forward (past context) and backward (future context) information of the input sequence simultaneously. This method is commonly used to enhance the performance of sequence labeling tasks, such as Named Entity Recognition (NER), Part-of-Speech tagging (POS), and others. crnn network architecture constructs convolution uses activation function to pool the maximum value, keeps the w direction unchanged, performs regularization, and creates sequential for bidirectional sequence feature extraction.

The operation results of the study are shown in Figure 3 and Figure 4 below:

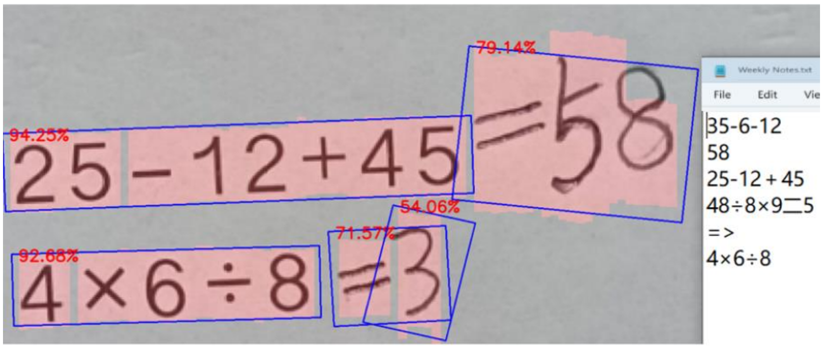


```
4×6÷8
./test_result\test_images\t3.txt
Mission complete, it took 32.602s

Recognition Result:

Biometrics,Fusedand Secured
Todav.Z0L0zuses facialrecognitiomcombined withindustmy-leading
spoofdetectiontechnoloavtoensurehatvourcustomersidentitiesy
remain uncompromised.
Imtheuture,weladdother biometrics,behavioralanalvtics and auestions
oniv individualuserscan answer tofurther strengthenthecustomer
```

Figure 3: operation results



```
35-6-12
58
25-12+45
48÷8×9二5
=>
4×6÷8
```

Figure 4: operation results

V. Conclusion

Using CRNN and OCR technology, combined with CTPN, character data can be effectively mined in educational scenarios. It can be seen from the running results that the characters in the image have been recognized, and the recognized characters have been output in the console and the specified files. These technologies significantly enhance the efficiency and accuracy of character mining, and can be extensively applied in fields such as educational intelligent scoring systems, document management, and blackboard writing recognition.

However, the accuracy of the predicted handwritten characters, formulas, and calculation results is not ideal and needs to be further studied.

With the continuous advancement of artificial intelligence technology, character mining technology is becoming increasingly important in the field of education. The development of OCR technology and the CTPN network will continuously improve the performance of character recognition systems.

In the future, further exploration of diverse mining and applications should be pursued. Optimizing deep learning models and algorithm parameters to adapt to different application scenarios is essential. Improving the recognition accuracy of handwritten characters, formulas, and calculation results will enhance predictive performance.

References

[1] Aulakh K, Roul RK, Kaushal M. E-learning enhancement through educational data mining with Covid-19 outbreak period in backdrop: A review. International journal of educational development. 2023 May 19:102814.

[2] Alghamdi, Amnah Saeed, and Atta Rahman. "Data mining approach to predict success of secondary school students: A Saudi Arabian case study." Education Sciences 13.3 (2023): 293.

[3] Zhang L, Yu H. Digital Marketing Evaluation of Applied Undergraduate Talent Training with E-commerce using Big Data Mining and Communication Technology Support. Computer-Aided Design and Applications. 2024 ;21:103-18.

[4] Roslan, Muhammad Haziq Bin, and Chwen Jen Chen. "Predicting students' performance in English and Mathematics using data mining techniques." Education and Information Technologies 28.2 (2023): 1427-1453.

[5] Alkashami M, Taamneh A, Khadragy S, Shwedeh F, Aburayya A, Salloum S. AI different approaches and ANFIS data mining: A novel approach to predicting early employment readiness in middle eastern nations. International Journal of Data and Network Science. 2023; 7(3):1267-82.

[6] Rangineni, Sandeep, and Divya Marupaka. "Data Mining Techniques Appropriate for the Evaluation of Procedure Information." International Journal of Management, IT & Engineering 13.9 (2023): 12-25.

[7] Araka E, Oboko R, Maina E, Gitonga R. Using educational data mining techniques to identify profiles in self-regulated learning: an empirical evaluation. The International Review of Research in Open and Distributed Learning. 2022 Feb 1; 23(1):131-62.

[8] Ting TT, Lim ET, Lee J, Wong JS, Tan JH, Tam RC, Chaw JK, Aitizaz A, Teoh CK. Educational big data mining: Mediation of academic performance in crime among digital age young adults. Online Journal of Communication and Media Technologies. 2024 Jan 1; 14(1):e202403.

[9] Liu Z, Yang Q, Wang A, Gu X. Vehicle Driving Safety of Underground Interchanges Using a Driving Simulator and Data Mining Analysis. Infrastructures. 2024 Feb 2; 9(2):28.

[10] Križanić S. Educational data mining using cluster analysis and decision tree technique: A case study. International Journal of Engineering Business Management. 2020 Feb 24; 12: 1847979020908675.

[11] Memon J, Sami M, Khan RA, Uddin M. Handwritten optical character recognition (OCR): A comprehensive systematic literature review (SLR). IEEE access. 2020 Jul 28; 8: 142642-68.

[12] Deb, Mainak, et al. "A CNN-based model to count the leaves of rosette plants (LC-Net)." Scientific Reports 14.1 (2024): 1496.

[13] Gozuoglu, Abdulkadir, Okan Ozgonenel, and Cenk Gezegin. "CNN-LSTM Based Deep Learning Application on Jetson Nano: Estimating Electrical Energy Consumption for Future Smart Homes." Internet of Things (2024): 101148.

[14] Ahmed, Fahad, et al. "Identification of kidney stones in KUB X-ray images using VGG16 empowered with explainable artificial intelligence." Scientific Reports 14.1 (2024): 6173.

[15] Zou W, Zhang Z, Peng Y, Xiang C, Tian S, Zhang L. SC-RPN: A strong correlation learning framework for region proposal. IEEE Transactions on Image Processing. 2021 Apr 5; 30:4084-98.

[16] Xu, Fan, et al. "A CRNN‐based method for Chinese ship license plate recognition." IET Image Processing 18.2 (2024): 298-311.

[17] Jiao, Li, and Hui Li. "Research on automatic identification algorithm of invoice information." International Conference on Algorithm, Imaging Processing, and Machine Vision (AIPMV 2023). Vol. 12969. SPIE, 2024.

[18] Fan Q, Zhuo W, Tang CK, Tai YW. Few-shot object detection with attention-RPN and multi-relation detector. InProceedings of the IEEE/CVF conference on computer vision and pattern recognition 2020 (pp. 4013-4022).

[19] Majib MS, Rahman MM, Sazzad TS, Khan NI, Dey SK. Vgg-scnet: A vgg net-based deep learning framework for brain tumor detection on mri images. IEEE Access. 2021 Aug 18; 9: 116942-52.

[20] Shwartz Ziv R, LeCun Y. To Compress or Not to Compress—Self-Supervised Learning and Information Theory: A Review. Entropy. 2024 Mar 12; 26(3): 252.