Design of an Intelligent Grading System Based on Artificial Intelligence

Yuhao Zhang[1], Hong Zhou[2], Jiaqi Wu[1], Weiping Deng[3], Haibo Yi[1,*]

[1]School of Artificial Intelligence, Shenzhen Polytechnic, Shenzhen, China

[2]Shenzhen Saixi Information Technology Co., Ltd, Shenzhen, China

[3]WeBank, Shenzhen, China

[*]Correspnding author

Keywords: AI, intelligent grading system, education, web design

Abstract: Artificial intelligence has been widely applied in image recognition tasks, and its applications span across various domains. This paper presents an intelligent grading system based on artificial intelligence (AI) that significantly improves the efficiency and accuracy of grading standardized tests. The proposed system utilizes deep learning techniques, including convolutional neural networks (CNN) and recurrent neural networks (RNN), to recognize Chinese characters and words in test papers. The system involves collecting labeled images containing Chinese characters and words, preprocessing the images to remove noise and standardize their size and orientation, extracting features from the preprocessed images, and training the deep learning model using CNN and RNN architectures for character and word recognition. Various metrics were used to evaluate the performance of the trained model, which achieved high accuracy and precision in identifying characters and words. The proposed system has significant practical implications for automating the grading process and reducing human bias.

1. Introduction

Manual grading is a widely used method for subjective assessments in various industries and educational institutions. However, in some fields such as standardized tests and multiple-choice exams, there is a growing preference for automated grading systems.

While manual grading has its advantages, including the ability to account for unique circumstances, it is also time-consuming, subjective, and prone to inconsistencies, which may result in errors or bias. On the other hand, automated grading systems can provide greater efficiency and consistency, although they may not capture all nuances and complexities involved in subjective assessments [1-3].

Including the use of artificial intelligence to recognize English letters and Chinese characters involves different implementations and practical applications. English letters are easy to be recognized by cnn due to their simple structure, while Chinese characters require a more complex model due to their complex structure and large character set.

Overall, the current situation of manual grading varies depending on the industry and context, with some sectors moving towards automation while others continue to rely on human judgement [4].

The main technologies of artificial intelligence (AI) include [5-10]:

Machine Learning: A branch of AI that involves training computer algorithms to learn patterns in data and make predictions or decisions without being explicitly programmed.

Natural Language Processing (NLP): The ability of computers to understand, interpret, and generate human language, including speech recognition, machine translation, and sentiment analysis.

Computer Vision: The ability of computers to interpret and understand visual information, including image and video recognition, object detection, and scene reconstruction.

Robotics: The use of machines with artificial intelligence and sensing capabilities to perform

DOI: 10.25236/icmmct.2023.003

tasks normally done by humans, such as industrial automation, autonomous vehicles, and drones.

Expert Systems: Computer programs that mimic the decision-making abilities of a human expert in a particular domain, using knowledge-based rules and inference engines.

Neural Networks: A type of machine learning algorithm inspired by the structure and function of the human brain, used for tasks involving pattern recognition, prediction, and decision-making.

Reinforcement Learning: A type of machine learning where an agent learns how to interact with an environment by receiving feedback in the form of rewards or punishments, allowing it to optimize its behavior over time.

Overall, these various technologies of artificial intelligence enable computers to learn, reason, and act like humans in various contexts, leading to increased efficiency, accuracy, and innovation in many industries.

Artificial intelligence (AI) has a wide range of applications across various industries, including [11-15]:

Natural language processing (NLP), image and speech recognition, predictive analytics, robotics, personalization, fraud detection, healthcare, and autonomous vehicles. AI-powered NLP is used to analyze written or spoken language to understand sentiment, intent, and context, and it is commonly used in chatbots, virtual assistants, and other conversational interfaces. AI can recognize images and speech, enabling applications such as facial recognition, voice assistants, and image search. AI algorithms can analyze large datasets to identify patterns and make predictions about future outcomes, which is useful in areas like financial forecasting, risk management, and supply chain optimization. AI-powered robots are increasingly being used in manufacturing, healthcare, and other industries to perform tasks that are difficult or dangerous for humans. AI can also be used to personalize experiences for individual users, detect fraudulent activity, improve patient care, and enable self-driving cars and other autonomous vehicles to navigate their environment and make decisions in real-time. These are just some of the many applications of AI, and the technology is constantly evolving to create new possibilities.

Machine learning is a rapidly evolving field, and there are many ongoing research projects related to various aspects of it. Here are some current areas of active research [16-20]:

These include deep learning, which explores new architectures for neural networks and optimization techniques. Reinforcement learning focuses on developing algorithms for agents to learn optimal behavior through trial-and-error. Other areas of research include transfer learning, explainable AI, fairness in machine learning, privacy-preserving machine learning, multimodal learning, few-shot learning, active learning, and meta-learning. These projects aim to improve the performance and capabilities of machine learning models, while addressing issues such as fairness, privacy, and reducing the need for labeled data.

2. System Design

There are several popular frameworks for image recognition using deep learning algorithms. Some of the most well-known ones include:

TensorFlow: Developed by Google, TensorFlow is an open-source framework for building machine learning models, including those used for image recognition. It includes a range of pre-built models and tools that make it easy to get started with image classification, object detection, and other tasks.

PyTorch: This is another popular open-source framework for deep learning, developed by Facebook. PyTorch has gained popularity due to its ease of use and flexibility, making it a popular choice for researchers and developers alike.

Keras: Keras is a high-level neural networks API written in Python. It can be used on top of TensorFlow or Theano. Its simplicity and ease of use make it a popular choice for beginners and researchers who want to quickly prototype and test models.

Caffe: Developed by Berkeley Vision and Learning Center (BVLC), Caffe is a deep learning framework that is particularly popular for computer vision tasks like image recognition. It has a number of pre-trained models available for use, and its architecture is optimized for speed and efficiency.

Microsoft Cognitive Toolkit (CNTK): Formerly known as Computational Network Toolkit, CNTK is an open-source toolkit developed by Microsoft for deep learning. It supports several programming languages and provides a range of tools for training and deploying deep learning models, including those used for image recognition.

These frameworks provide a range of functionality for image recognition, from basic classification to more advanced tasks like segmentation and object detection. Choosing the right framework depends on your specific needs and goals, as each one has its own strengths and weaknesses.

Table 1 outlines the components of an AI-based grading system:

Table 1 components of an AI-based grading system.

Component	Description
Input Data	The data used to train the AI model, which could include previous student work or a large dataset of graded assignments.
Pre-processing	Steps taken to prepare the input data for analysis, such as normalization, filtering, or feature extraction.
Feature selection	Choosing which features (e.g., grammar, syntax, word choice) to include in the analysis, and how to represent them numerically.
Algorithm	The machine learning algorithm used to analyze the input data and generate grades. This could be a supervised learning algorithm, such as linear regression or neural networks, or an unsupervised learning algorithm, such as clustering or anomaly detection.
Training Data	A subset of the input data that is used to train the AI model. This is typically labeled data, where each sample has a corresponding grade assigned by a human grader.
Validation Data	Another subset of the input data that is held out during training and used to evaluate the performance of the AI model.
Hyperparameters	Settings that control how the AI algorithm learns from the training data, such as the learning rate, regularization, or number of hidden layers. These are often tuned using a separate validation set to optimize performance.
Output	The final output of the AI grading system, which could be a numerical grade, letter grade, or other assessment metric.
Post-processing	Any additional processing steps taken after the output is generated, such as rounding, scaling, or applying rubrics to adjust grades.
Human Review	Depending on the context, it may be necessary to have a human review the AI-generated grades to ensure accuracy and fairness. This could be done through random sampling or targeted review of specific assignments.

English letter recognition can be performed using a variety of AI techniques, including deep learning and traditional machine learning. Here's a general overview of the process involved in using AI to recognize English letters:

Data collection: The first step is to collect a large dataset of images containing English letters. This could include handwritten letters or typed text.

Preprocessing: The images are then preprocessed to make them suitable for analysis. This may involve resizing, cropping, or converting the images to grayscale.

Feature extraction: Next, features are extracted from the images that can be used to distinguish between different letters. Common features include edge detection, texture analysis, and blob detection.

Model training: Using the preprocessed data and extracted features, a machine learning model is trained to recognize the letters. This could involve supervised learning, where the model is trained on labeled examples of letters, or unsupervised learning, where the model learns to classify the letters based on their intrinsic characteristics.

Evaluation: Once the model is trained, it is evaluated on a separate test set to gauge its accuracy and performance. Additional tuning may be necessary to optimize the model's performan

Deployment: Finally, the trained model can be deployed to recognize English letters in new images or documents.

Table 2 describes and compares the differences between AI based English letter and Chinese recognition.

Table 2 differences between AI based English letter and Chinese recognition.

Aspect	English Letter Recognition	Chinese Character Recognition
Data Collection	Collections of labeled images are often used to train the model, consisting of typed or handwritten letters.	Large collections of labeled images are required to train the model, consisting of a wide variety of characters written in different styles and fonts.
Feature Extraction	Features such as edge detection, blob detection, and texture analysis may be used to identify characteristics unique to individual letters.	Character strokes and components, spatial relationships, and semantic meanings must all be considered when extracting features from Chinese characters.
Algorithm	A range of algorithms can be used for letter recognition, including traditional machine learning (e.g., SVM or kNN) as well as deep learning approaches (e.g., convolutional neural networks or recurrent neural networks).	Deep learning algorithms are widely used for Chinese character recognition, including convolutional neural networks and recurrent neural networks.
Pre-Processing	Images may undergo pre-processing steps to normalize lighting, scale, or orientation.	Pre-processing steps include binarization, noise reduction, and normalization of character size and position.
Human Review	In some cases, human review may be necessary to ensure accuracy and consistency of the recognition results.	Human review is often required to ensure the accuracy of recognition results, particularly for complex characters or those written in unusual styles.
Performance Metrics	Accuracy, precision, and recall are commonly used metrics to evaluate performance.	Metrics such as character error rate, stroke accuracy, and overall recognition rate are used to evaluate performance.

3. AI Design

The recognition process of letters and Chinese characters based on deep learning involves several steps:

Data collection: Large datasets of labeled images containing letters or Chinese characters are collected. These images may be handwritten or typed, and can be obtained from a variety of sources.

Preprocessing: The images are preprocessed to remove noise, normalize lighting and contrast, and standardize the size and orientation of the characters. This could involve techniques such as binarization, smoothing, and filtering.

Feature extraction: Features are extracted from the preprocessed images that can be used to distinguish between different characters. In deep learning, this is often done automatically by the neural network, which learns to recognize patterns in the input data.

Model training: Using the preprocessed data and extracted features, a deep learning model is trained to recognize the characters. This typically involves using convolutional neural networks (CNNs) or recurrent neural networks (RNNs), which are optimized for image and sequence recognition, respectively. The model is trained using backpropagation and gradient descent to minimize a loss function.

Evaluation: Once the model is trained, it is evaluated on a separate test set to measure its accuracy and performance. Various metrics such as accuracy, precision, and recall can be used to evaluate the model's performance.

Deployment: Finally, the trained model can be deployed to recognize new letters or characters. This typically involves feeding new images into the model and using its output to make predictions about the identity of the characters.

Table 3 describes the model parameters for deep learning recognition of Chinese characters.

Table 3 the model parameters for deep learning recognition of Chinese characters.

Parameter	Description
Network Architecture	This refers to the structure of the neural network used for character recognition. Popular architectures include convolutional neural networks (CNNs), recurrent neural networks (RNNs), and their combinations (e.g., CRNN).
Number of Layers	The number of layers in the neural network architecture, which can vary depending on the complexity of the recognition task. For example, a simple CNN may have just a few convolutional and pooling layers, while a more complex architecture like ResNet or Inception may have dozens of layers.
Filter Size	The size of the convolutional filters used in the CNN layers. Larger filter sizes typically capture more complex features, but also increase the number of parameters and computation required.
Stride	The stride length used in the CNN layers, which determines how much the filter is shifted for each convolution operation. A larger stride reduces computation at the cost of potentially missing important features.
Activation Function	The non-linear activation function used in the neural network, such as ReLU, sigmoid, or tanh. This introduces non-linearity into the model and helps it learn more complex functions.
Dropout	Dropout is a regularization technique used to prevent overfitting by randomly dropping out units from the network during training. The dropout rate controls the probability of dropping out each unit.
Recurrent Connections	RNNs use recurrent connections to maintain a memory of previous inputs, allowing them to process sequences of data. The type of recurrent connection used, such as Gated Recurrent Units (GRUs) or Long Short-Term Memory (LSTM) units, can affect the performance of the model.
Batch Size	The number of samples processed in each batch during training. Larger batch sizes typically result in faster training, but require more memory and may lead to worse generalization performance.
Learning Rate	The learning rate controls the step size taken during gradient descent optimization. A smaller learning rate results in slower but more accurate convergence, while a larger learning rate can cause the model to overshoot the optimal solution.
Optimization Algorithm	The optimization algorithm used for gradient descent, such as stochastic gradient descent (SGD), Adam, or RMSprop. These algorithms control how the model updates its parameters during training.

4. Result

The steps for programming and implementing an AI-based grading system are:

Define the problem: Determine the task that the AI-based grading system is intended to perform, such as grading essays or programming assignments.

Collect data: Gather a large dataset of previously graded work as training data for the AI-based grading system.

Preprocess data: Clean and preprocess the collected data to ensure consistency and standardization.

Choose appropriate algorithms: Select AI algorithms, such as machine learning or deep learning, that are appropriate for the grading task.

Train the model: Train the AI-based grading model on the preprocessed data using the chosen algorithms.

Test the model: Evaluate the performance of the AI model on new, unseen data by comparing its predictions with human graders.

Refine the model: Adjust the model parameters, algorithms, or preprocessing steps to improve its accuracy and efficiency.

Implement the model: Integrate the AI-based grading system into the existing grading workflow, such as an online learning management system.

Monitor and update: Continuously monitor the performance of the AI-based grading system and update it as necessary to ensure its continued accuracy and effectiveness.

5. Conclusion

The innovation of AI-based grading systems lies in their ability to automate and streamline the grading process, providing several benefits over traditional manual grading methods. Here are some of the key innovations of AI-based grading systems:

Efficiency: AI-based grading systems can grade large volumes of work quickly and consistently, reducing the workload for human graders and allowing for faster feedback to students.

Objectivity: AI-based grading systems can grade work objectively, without bias or personal preferences that can affect human grading.

Consistency: AI-based grading systems can provide consistent grading across multiple graders, ensuring that all students receive fair and equitable grades.

Personalization: AI-based grading systems can provide personalized feedback to students based on their individual strengths and weaknesses, helping them improve their learning outcomes.

Scalability: AI-based grading systems can scale easily to accommodate growing class sizes or increasing workloads, without adding more human graders.

Cost-efficient: AI-based grading systems can reduce the cost of grading by minimizing the need for human graders, which can be expensive and time-consuming.

Feedback granularity: AI-based grading systems can provide more granular and detailed feedback to students, beyond just a final grade, which can help them better understand their strengths and weaknesses.

Overall, AI-based grading systems have the potential to transform the way we grade student work, making the process faster, more objective, and more efficient. While there are still some limitations and challenges associated with these systems, they represent an innovative and promising approach to grading in education.

Acknowledgment

We acknowledge the funding support from 2022 Guangdong Provincial Ordinary University Characteristic Innovation Projects (No. 2022KTSCX308).

References

[1] Trehern W, Ortiz-Ayala R, Atli K C et al. Data-driven shape memory alloy discovery using Artificial Intelligence Materials Selection (AIMS) framework. Acta Materialia, 2022, 228:117751-.

[2] Zhao S, Xu W, Chen L. The modeling and products prediction for biomass oxidative pyrolysis based on PSO-ANN method: An artificial intelligence algorithm approach. Fuel, 2022, 312:122966-.

[3] Julius Chapiro C D D. Translating artificial intelligence from code to bedside: The road towards AI-driven predictive biomarkers for immunotherapy of hepatocellular carcinoma. Journal of Hepatology, 2022.

[4] L Craxì. Letter to the Editor: Comment on "The Application of Artificial Intelligence for the Diagnosis and Treatment of Liver Diseases". Hepatology, 2021, 74(3).

[5] Gupta R, Mittal A, Agrawal V et al. OdoriFy: A conglomerate of Artificial Intelligence-driven prediction engines for olfactory decoding. Journal of Biological Chemistry, 2021:100956.

[6] Browning L, Colling R, Rakha E et al. Digital pathology and artificial intelligence will be key to supporting clinical and academic cellular pathology through COVID-19 and future crises: the PathLAKE consortium perspective. Journal of Clinical Pathology, 2021(7).

[7] House, Patrick M.Kopelyan, MikhailBraniewska, NataliaSilski, BartoszChudzinska, AgataHolst, BrigitteSauvigny, ThomasMartens, TobiasStodieck, StefanPelzl, Sirko. Automated detection and segmentation of focal cortical dysplasias (FCDs) with artificial intelligence: Presentation of a novel convolutional neural network and its prospective clinical validation. Epilepsy research, 2021, 172(1).

[8] J Martín-Pérez, Molner N, Malandrino F et al. Choose, not Hoard: Information-to-Model Matching for Artificial Intelligence in O-RAN. IEEE Communications Magazine, 2022.

[9] Kugler L. Artificial Intelligence, Machine Learning, and the Fight Against World Hunger. Communications of the ACM, 2022(2):65.

[10] Moya A, RJ López-Sastre. Stellar mass and radius estimation using artificial intelligence. Astronomy & Astrophysics, 2022, 663:A112.

[11] Kim J H, Kim Y J, Lee Y J et al. Automated histopathological evaluation of pterygium using artificial intelligence. The British journal of ophthalmology, 2022.

[12] Duran-Frigola M, Cigler M, Winter G E. Advancing Targeted Protein Degradation via Multiomics Profiling and Artificial Intelligence. Journal of the American Chemical Society, 2023.

[13] Rudolph J, Huemmer C, Ghesu F C et al. Artificial Intelligence in Chest Radiography Reporting Accuracy Added Clinical Value in the Emergency Unit Setting Without 24/7 Radiology Coverage. Investigative radiology. 2022(2):57.

[14] Hua W, Chen Y, Qadrdan M et al. Applications of blockchain and artificial intelligence technologies for enabling prosumers in smart grids: A review. Renewable and Sustainable Energy Reviews, 2022, 161:112308-.

[15] He Z, Guo W, Zhang P. Performance prediction, optimal design and operational control of thermal energy storage using artificial intelligence methods. Renewable and Sustainable Energy Reviews, 2022, 156:111977-.

[16] Envelope H, Envelope A, Envelope N et al. Expert systems and the prospects of artificial intelligence for the automatic supervisory control of salinity gradient solar ponds. Solar Energy, 2022, 246:281-293.

[17] L Cobo-Sánchez, M Domínguez-Rodrigo, M Pizarro-Monzo et al. Determining the diagenetic paths of archaeofaunal assemblages and their palaeoecology through artificial intelligence: an application to Oldowan sites from Olduvai Gorge (Tanzania). Journal of Quaternary Science, 2022, 37(3):543-557.

[18] Kim D, Ho C H, Park I et al. Untangling the contribution of input parameters to an artificial intelligence PM2.5 forecast model using the layer-wise relevance propagation method. Atmospheric Environment, 2022, 276:119034-.

[19] Shoaib M, Naz I, Khan M I, et al. Artificial intelligence knacks-based stochastic paradigm to study lie group analysis with the impact of electric field on MHD Prandtl–Eyring fluid flow system. International Journal of Modern Physics B, 2022, 36(30).

[20] Marsden M, Perkins Z, Marsh W, et al. 92Evaluation of an Artificial Intelligence (AI) System to Augment Clinical Risk Prediction of Trauma Induced Coagulopathy: A Prospective Observational Study. British Journal of Surgery, 2022(Supplement_1):Supplement_1.