AI-based Examination Proctoring System Using Pose Estimation Introduction

忠 信 笃 敬

章廷玮 | Zhang Tingwei

Contents

1. Research Background and Significance

Research background

The proctoring problem caused by the popularization of online examinations

According to the 55th Statistical Report on the Development of China's Internet, by December 2024, the number of online education users in my country will reach 355 million. The popularity of online education has made computer-based exams an important means of assessing learners' abilities, but the traditional proctoring model has problems such as high labor costs, high missed detection rates, and inability to respond in real time.

Existing automated proctoring systems have limitations

Most of the time, a single detection technology is used, which is difficult to cover complex cheating behaviors, easily leads to missed detections and false detections, and has high hardware requirements.

1. Research Background and Significance

Research significance

Multi-task fusion solves complex cheating behavior detection problems

The multi-task collaborative detection mechanism can cover a wider range of abnormal test behaviors, such as turning the head, raising hands, using illegal items, etc.

Lightweight design suitable for low computing power devices

Through lightweight model design and performance optimization, it can run in real time in ordinary hardware environments, reducing dependence on high-cost hardware.

2. Core Challenges and Difficulties

Technical Challenges

Multi-task collaborative detection

Through the multi-task collaborative detection mechanism, combined with posture estimation and target detection, the accuracy and comprehensiveness of abnormal behavior recognition are improved.

Lightweight deployment

The system design needs to take into account the limitations of ordinary hardware environments, and through lightweight models and performance optimization, ensure that it can run efficiently on low-computing power devices.

Engineering Challenges

Real-time guarantee

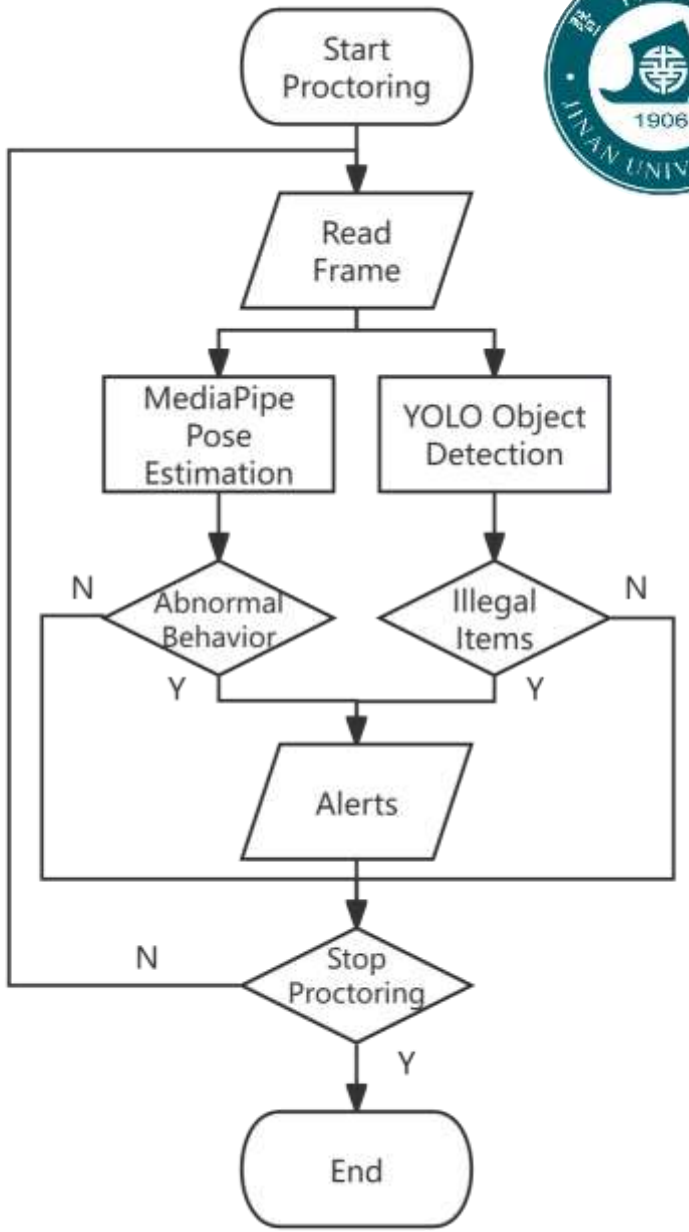
The system needs to detect abnormal behaviors and illegal items in real time, and output alarm information through the interface to ensure that the invigilator can respond in time.

3. System Framework and Methods

Overall architecture

The system uses the real-time video stream of the camera as input, reads the original BGR format frame data through OpenCV, and then performs dual-path preprocessing.

One path directly inputs the BGR image into YOLO11n for target detection, and the other path converts the original BGR image into RGB format for MediaPipe posture estimation, extracts 33 key points, combines the YOLO detection results with the MediaPipe key point coordinates, and triggers the alarm logic through spatial relationship and timing analysis.

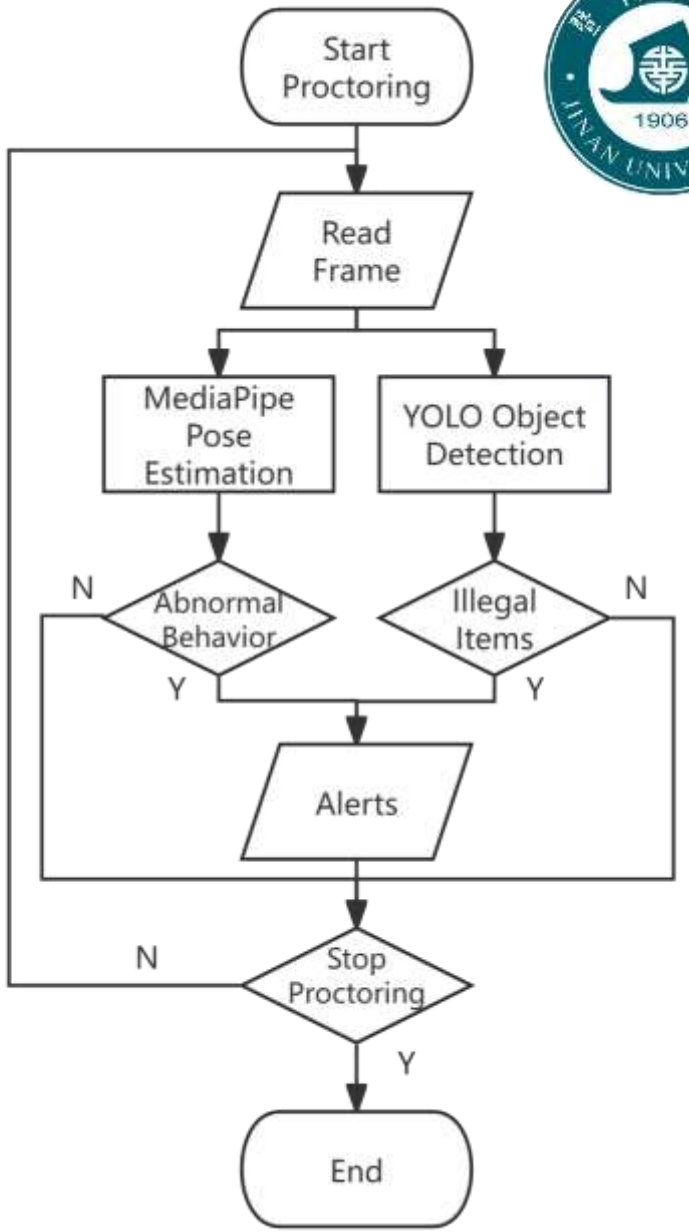


3. System Framework and Methods

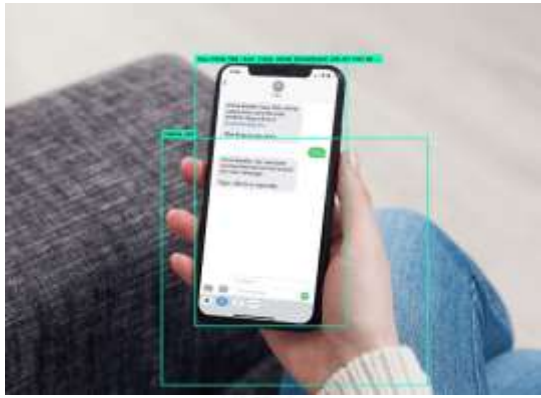
Overall architecture

The alarm logic is based on the detection results. If abnormal behavior or illegal items are found, the system will trigger an alarm in real time and overlay the detection box and key point connection lines on the image. The annotated image is output in real time through the PyQt5 interface, and text alarms are issued at the same time to ensure that the invigilator can quickly identify abnormal behavior.

The system automatically saves the current frame image to the screenshots directory while warning, and records the alarm information (including timestamps and specific events) through the logging module for subsequent tracing and analysis.



3. System Framework and Methods



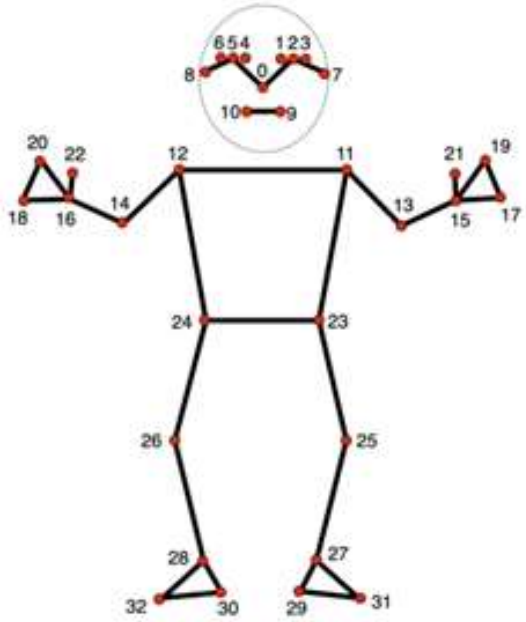
Key technology analysis

Object Detection Module

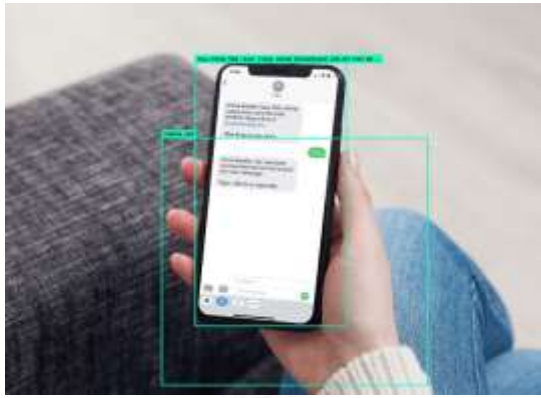
Violation detection uses the YOLO11n model to infer the frame image and output the target category in the current frame. If the violation target category exists and lasts for more than 2 seconds, a warning is triggered.

Attitude Estimation Module

Abnormal behavior recognition is based on the normalized coordinates of 33 key points output by MediaPipe, which are mapped to pixel space. Alarms are triggered by calculating the spatial relationship between key points and combining the duration of abnormal states.



3. System Framework and Methods



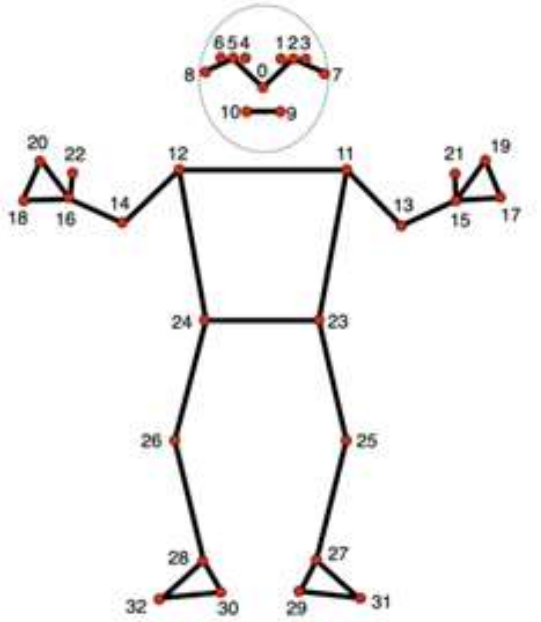
Key technology analysis

Performance Optimization

Lightweight models and multi-threaded optimization strategies enable the system to achieve real-time detection on ordinary devices.

The system is deployed through CUDA acceleration and TensorRT to achieve end-to-end latency ≤ 100ms.

Reduce invalid calculations through confidence threshold setting to balance accuracy and speed.



4. Innovation and Advantages

Algorithm innovation

Multi-task collaborative framework

The system combines YOLO11n target detection and MediaPipe posture estimation to achieve multi-task collaborative detection, which can simultaneously identify abnormal behaviors and illegal items, improving the accuracy and comprehensiveness of abnormal behavior identification.

Abnormal judgment logic

The system improves the accuracy of recognition by combining the spatial relationship and time threshold of abnormal objects and abnormal postures through judgment logic based on spatiotemporal relationships.

Engineering innovation

Performance optimization

The system achieves real-time detection with end-to-end latency ≤100ms through model lightweighting and confidence threshold optimization, CUDA acceleration and TensorRT deployment, adapting to low computing power environments.

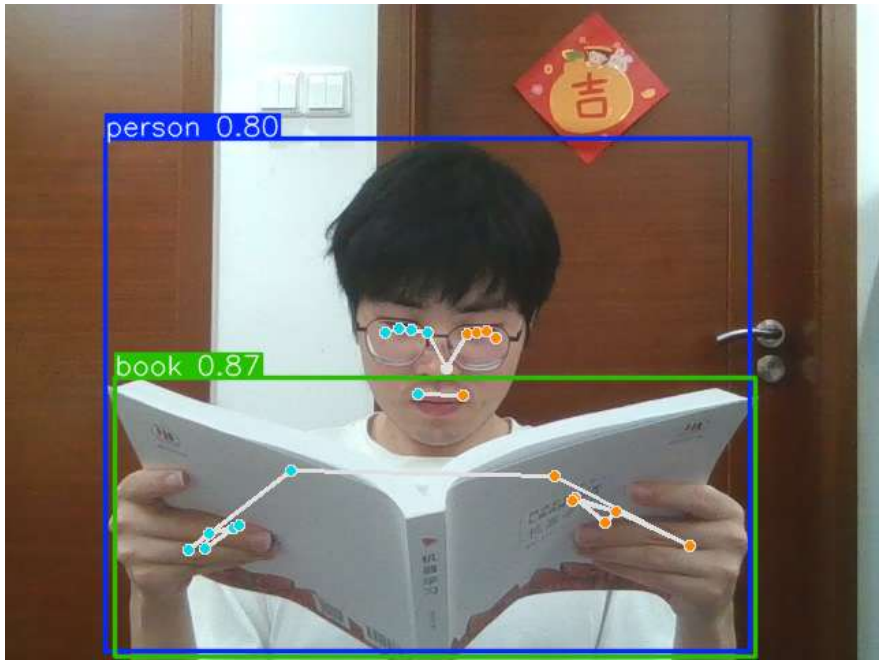
5. Experimental Results and Demonstration

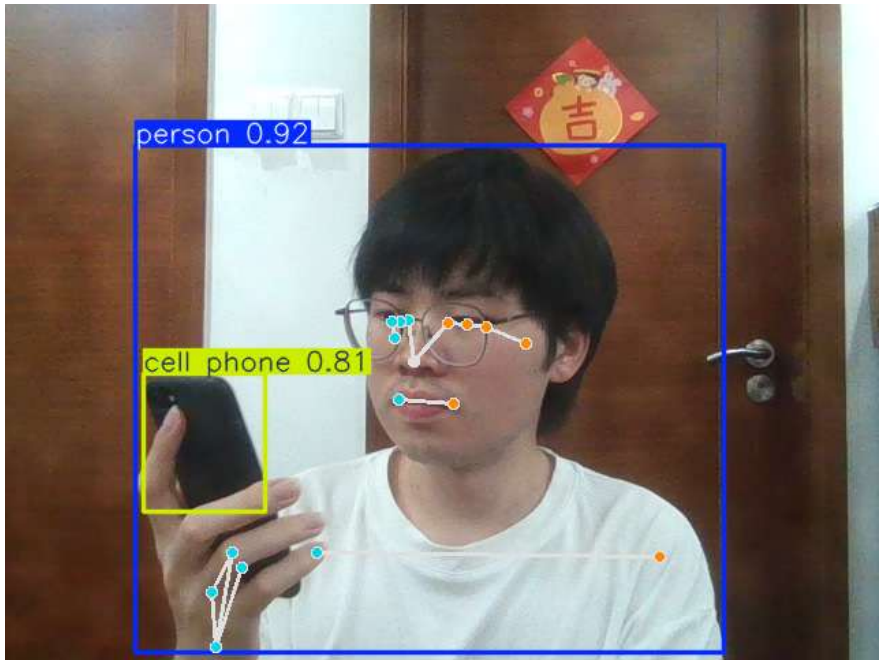
Single frame delay

Number	YOLO Latency (ms)	MediaPipe Latency (ms)
1	12. 817	29. 566
2	13. 375	30. 369
3	13. 445	29. 928
4	12. 978	30. 179
5	12. 343	29. 737
6	12. 252	29. 014
7	13. 304	29. 805
8	12. 473	29. 513
9	12. 501	29. 878
10	13. 311	30. 347
Average	12. 8799	29. 8336

System delay: end-to-end ≤100ms

5. Experimental Results and Demonstration





2025-05-07 05:25:54 - screenshots/20250507_052554_971.png | 异常物品：书本！ 异常姿态：低头！
2025-05-07 05:25:56 - screenshots/20250507_052556_002.png | 异常物品：书本！ 异常姿态：低头！
2025-05-07 05:25:57 - screenshots/20250507_052557_082.png | 异常物品：书本！ 异常姿态：低头！
2025-05-07 05:25:58 - screenshots/20250507_052558_110.png | 异常物品：书本！ 异常姿态：低头！
2025-05-07 05:25:59 - screenshots/20250507_052559_170.png | 异常物品：书本！ 异常姿态：低头！
2025-05-07 05:26:00 - screenshots/20250507_052600_217.png | 异常物品：书本！ 异常姿态：低头！
2025-05-07 05:26:01 - screenshots/20250507_052601_258.png | 异常物品：书本！ 异常姿态：低头！
2025-05-07 05:26:02 - screenshots/20250507_052602_300.png | 异常物品：书本！ 异常姿态：低头！
2025-05-07 05:26:03 - screenshots/20250507_052603_323.png | 异常物品：书本！ 异常姿态：低头！
2025-05-07 05:26:04 - screenshots/20250507_052604_357.png | 异常物品：书本！ 异常姿态：低头！

2025-05-07 05:31:52 - screenshots/20250507_053152_392.png | 异常物品：手机！ 异常姿态：转头！
2025-05-07 05:31:57 - screenshots/20250507_053157_941.png | 异常物品：手机！ 异常姿态：转头！
2025-05-07 05:32:01 - screenshots/20250507_053201_311.png | 异常物品：手机！ 异常姿态：转头！
2025-05-07 05:32:02 - screenshots/20250507_053202_395.png | 异常物品：手机！ 异常姿态：转头！
2025-05-07 05:32:05 - screenshots/20250507_053205_376.png | 异常物品：手机！ 异常姿态：转头！
2025-05-07 05:32:06 - screenshots/20250507_053206_445.png | 异常物品：手机！ 异常姿态：转头！
2025-05-07 05:32:07 - screenshots/20250507_053207_514.png | 异常物品：手机！ 异常姿态：转头！
2025-05-07 05:32:08 - screenshots/20250507_053208_580.png | 异常物品：手机！ 异常姿态：转头！
2025-05-07 05:32:09 - screenshots/20250507_053209_645.png | 异常物品：手机！ 异常姿态：转头！
2025-05-07 05:32:10 - screenshots/20250507_053210_709.png | 异常物品：手机！ 异常姿态：转头！

6. Summary and Outlook

Summary

This project designed and implemented an AI proctoring system based on multi-task fusion, which solved the core problems of multi-task collaborative detection and low computing power adaptation, and provided a cost-effective proctoring solution for online examinations.

Future work

Expand detection categories

Enable the system to identify more prohibited items.

Enhance interface interaction

Dark theme interface, illegal item selection.

Remote monitoring

Realize remote transmission of video streams and centralized management of multiple examination room images, and combine cloud storage to achieve real-time synchronization and backtracking of alarm records.

Thank you for watching!

忠信笃敬

章廷玮 | Zhang Tingwei