

ICS 35.240
CCS L 77



**NATIONAL STANDARD OF THE
PEOPLE'S REPUBLIC OF CHINA**

GB/T 46700-2025

**Artificial intelligence - Audio, video and image analysis
algorithm interface**

人工智能 音视频及图像分析算法接口

Issued on: October 31, 2025

Implemented on: May 1, 2026

**Issued by: State Administration for Market Regulation;
Standardization Administration of the PRC**

Table of Contents

1.Scope	1
2 Normative References	1
3.Terms, Definitions, and Abbreviations	1
3.1 Terms and Definitions	1
3.2 Abbreviations	2
4.Application Environment	2
5.Algorithm Package Interface	3
5.1 General Rules	3
5.2 Requirement	3
5.3 Data Types	4
5.4 Parsing Algorithm Package Interface Functions	7
5.5 Comparison Algorithm Package Interface Functions	11
6.Algorithm Service Interface	17
6.1 General Rules	17
6.2 Requirement	17
6.3 Function	17
41 Appendix C (Normative) Description of Algorithm Service Data Types 44 C.1 Response Status Object 44 C.2 Response Status List Object 44 C.3 Authorized Object 45 C.4 Parsing the Task Object 45 C.5 Parse the task list object 48 C.6 Parse the task result object 51 C.7 Feature Object 53 C.8 Feature Comparison Objects 54 C.9 Feature Comparison Results Object 55 C.10 Feature Library List Object 55 C.11 Compare Task Object 56 C.12 Compare Task List Objects 57 C.13 Compare the task result object 57 C.14 Task Status Object.....	58

Foreword

This document is in accordance with GB/T 1.1-2020 "Standardization Work Guidelines Part

1.Structure and Drafting Rules of Standardization Documents". Drafting is scheduled.

Please note that some content in this document may involve patents. The issuing organization of this document assumes no responsibility for identifying patents. This document was proposed by the Ministry of Industry and Information Technology of the People's Republic of China. This document is under the jurisdiction of the National Technical Committee on Standardization of Audio, Video and Multimedia Systems and Equipment (SAC/TC242). This document was drafted by: Hangzhou Hikvision Digital Technology Co., Ltd., China Electronics Technology Standardization Institute, and Zhejiang

Dahua Technology Co., Ltd. Beijing Huawen Zhonghe Technology Co., Ltd., Dongguan Dingli Automation Technology Co., Ltd., Zhongyi Instech Technology Co., Ltd., Shandong Jishi Angle Technology Co., Ltd., Guangdong Weibang Technology Co., Ltd., Jinan Keming Digital Technology Co., Ltd., Zibo Vocational College Speed Technology Co., Ltd., Hangzhou Magic Point Technology Co., Ltd., Guangdong Chengtian Transportation Technology Development Co., Ltd., Beijing Chongyuan Technology Co., Ltd. The company, Wuhan Yixun Beidou Spacetime Technology Co., Ltd., Hangzhou Milluoxing Technology (Group) Co., Ltd., and Guangzhou Aoweiya Electronic Technology Co., Ltd. Limited Liability Company. The main drafters of this document are. Chen Zhe, Yang Yang, Ren Wenqi, Kong Weisheng, Wu Geng, Wen Hao, Wang Gongming, Zhou Zhiheng, Liu Zhiqiang, Wang Wei, and Zhang Ran. Zhang Wei, Shen Yunjie, Wei Shengyu, Guo Qing, Gao Yongchao, Wang Shuang, Tian Yonghui, Zhang Heng, Liu Zitao, Liu Zhenhua, Su Jie, Fu Tingjie, Yuan Gang, Xiao Tao Zheng Daoqin, Pang Zhibin, Lin Feihe, Cao Yijin, Chen Min, Wu Dongliang, Tang Lin, Huang Weibing, Chen Dahao, Yang Yang, Wang Haijun, Cheng Xiaolin, Tang Zhenyu Zheng Jun, Chen Lei, Xie Zechun, Luo Zhi, Zeng Honghan, Wang Xiaoqing, Zhang Nengfeng, Huang Hao, Zhou Kang, Li Dongya, Yang Peng, Luo Weijie, Yao Lichao, Pan Xiaoming Liu Zhenghua, Li Zhigang, Wang Pin, Zhang Zhiyu, Zhao Jin, Zhang Hui, Yang Ming, Wang Dongsheng, Zheng Song, Ji Haijiao, Chen Min, Sun Linjun, Zhao Dingjin, Chen Guangyan Li Zhongxi, Cui Xiaoxiu, Cao Shipeng, Xiang Dafeng, Huang Zanzhan, Guo Hongyan, Gang Yong, Chen Zhenjie, Pan Weixin, Chen Qingkui, Li Qiang, Jia Deshuang, Xu Zhongjian Xiao Chuanbao, Chen Nan, Zeng Tianle, Yao Daming, Fu Cheng, Liu Qiong, Zou Yingsi. AI audio, video and image analysis algorithm interface

1 Scope

GB/T 46700-2025 is the Chinese national standard covering the interface an analysis model exposes - how the media goes in and how the detections, classifications and tracks come back out, the metadata and confidence, the model description and versioning, so that algorithms from different vendors are interchangeable behind the same platform. At 33,000 words, first edition, under the Ministry of Industry and Information Technology. Issued on 31 October 2025, it has been in force since 1 May 2026.

This document specifies the technical parameters and requirements for the audio/video and image analysis algorithm interface, covering parsing algorithms and comparison algorithms within the algorithm package. This includes the application environment, data format definition, and interface interaction process under the algorithm service model. This document applies to the application of audio, video, and image analysis algorithms. It covers the system integration of intelligent application systems with algorithm packages and algorithm services. For development reference.

2 Normative references

The contents of the following documents, through normative references within the text, constitute essential provisions of this document. Dated citations are not included. For references to documents, only the version corresponding to that date applies to this document; for undated references, the latest version (including all amendments) applies. This document.

GB/T 28181-2022 Technical Requirements for Information Transmission, Exchange and Control of Public Security Video Surveillance Network Systems

3 Terms, Definitions and Abbreviations

3.1 Terms and Definitions The following terms and definitions apply to this document.

3.1.1 Artificial intelligence technology is integrated into specific software or hardware systems to solve problems in specific fields or perform specific tasks.

3.1.2 To process and analyze data, and to extract information, knowledge and patterns from it.

3.1.3 Techniques and methods for comparing and matching data with different features help identify and process similarities and differences in data.

3.1.4 algorithm package A collection of algorithm libraries that enable functions such as parsing and comparing audio, video, and image data.

3.1.5 algorithm service Based on audio, video, and image analysis algorithms, it provides network services for media data parsing and feature comparison.

3.1.6 feature A specific physical quantity extracted from audio, video, and image samples to represent that sample.