

(12) 按照专利合作条约所公布的国际申请

更正本

(19) 世界知识产权组织
国际局



WIPO | PCT



(10) 国际公布号

WO 2025/130671 A9

(43) 国际公布日
2025 年 6 月 26 日 (26.06.2025)

(51) 国际专利分类号:
G06V 10/764 (2022.01) G06N 3/0464 (2023.01)
G06V 10/82 (2022.01)

(21) 国际申请号: PCT/CN2024/137703

(22) 国际申请日: 2024 年 12 月 9 日 (09.12.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202311747933.2 2023 年 12 月 19 日 (19.12.2023) CN

(71) 申请人: 北京京东远升科技有限公司 (BEIJING JINGDONG YUANSHEG TECHNOLOGY CO., LTD.) [CN/CN]; 中国北京市大兴区北京经济技术开发区科创十一街 18 号院 1 号楼 12 层 1204 (北京自贸试验区高端产业片区亦庄组团) 100176 (CN)。

(72) 发明人: 姚卓坤 (YAO, Zhuokun); 中国北京市大兴区北京经济技术开发区科创十一街 18 号院 1 号楼 12 层 1204 (北京自贸试验区高端产业片区亦庄组团) 100176 (CN)。

(74) 代理人: 中原信达知识产权代理有限责任公司 (CHINA SINDA INTELLECTUAL PROPERTY LTD.); 中国北京市西城区金融街 19 号富凯大厦 B 座 11 层 100033 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD,

(54) Title: OBJECT DETECTION METHOD AND APPARATUS

(54) 发明名称: 一种目标检测的方法和装置

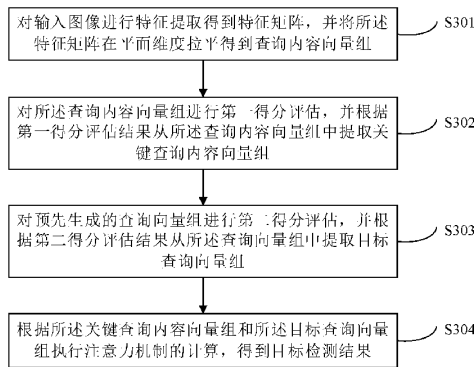


图 3

- S301 Perform feature extraction on an input image to obtain a feature matrix, and flatten the feature matrix in a planar dimension to obtain a query content vector set
- S302 Perform first score evaluation on the query content vector set, and extract a key query content vector set from the query content vector set on the basis of a first score evaluation result
- S303 Perform second score evaluation on a pre-generated query vector set, and extract a target query vector set from the query vector set on the basis of a second score evaluation result
- S304 Execute attention mechanism calculation on the basis of the key query content vector set and the target query vector set to obtain an object detection result

(57) Abstract: The present application relates to the technical field of artificial intelligence, and discloses an object detection method and apparatus. A specific embodiment of the method comprises: performing feature extraction on an input image to obtain a feature matrix, and flattening the feature matrix in a planar dimension to obtain a query content vector set; performing first score evaluation on the query content vector set, and extracting a key query content vector set from the query content vector set on the basis of a first score evaluation result; performing second score evaluation on a pre-generated query vector set, and extracting a target query vector set from the query vector set on the basis of a second score evaluation result; and executing attention mechanism calculation on the basis of the key query content vector set and the target query vector set to obtain an object detection result. According to the embodiment, the amount of attention mechanism calculation is remarkably reduced and the calculation speed is increased, without decreasing the feature resolution and reducing effective features, thereby achieving the effect of performing object detection in real time at a vehicle end.

[续页]

SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ,
UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚
(AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE,
BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR,
HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO,
PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN,
TD, TG)。

本国际公布:
— 包括国际检索报告(条约第21条(3))。

(48) 更正本的公布日:
2025 年 9 月 4 日 (04.09.2025)

(15) 更正内容:
见 2025 年 9 月 4 日 (04.09.2025) 公布的公告

(57) 摘要: 本申请公开了一种目标检测的方法和装置, 涉及人工智能技术领域。该方法的一具体实施方式包括: 通过对输入图像进行特征提取得到特征矩阵, 并将特征矩阵在平面维度拉平得到查询内容向量组; 对查询内容向量组进行第一得分评估, 并根据第一得分评估结果从查询内容向量组中提取关键查询内容向量组; 对预先生成的查询向量组进行第二得分评估, 并根据第二得分评估结果从查询向量组中提取目标查询向量组; 根据关键查询内容向量组和目标查询向量组执行注意力机制的计算, 得到目标检测结果。该实施方式在不降低特征分辨率、不减少有效特征的情况下显著减少了注意力机制的计算量, 提升了计算速度, 从而达到可以在车端实时的进行目标检测的效果。