

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号
WO 2025/139962 A1

(43) 国际公布日
2025 年 7 月 3 日 (03.07.2025)

- (51) 国际专利分类号:
G06V 30/19 (2022.01) **G06V 30/18** (2022.01)
G06T 11/60 (2006.01)
- (21) 国际申请号: PCT/CN2024/140309
- (22) 国际申请日: 2024 年 12 月 18 日 (18.12.2024)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202311813675.3 2023年12月27日 (27.12.2023) CN
- (71) 申请人: 苏州镁伽科技有限公司 (SUZHOU MEGAROBO TECHNOLOGIES CO., LTD.) [CN/CN];
中国江苏省苏州市中国 (江苏) 自由

- 贸易试验区苏州片区苏州工业园区玲珑街88号 215000 (CN)。
- (72) 发明人: 徐海俊 (XU, Haijun); 中国江苏省苏州市中国 (江苏) 自由贸易试验区苏州片区苏州工业园区玲珑街88号 215000 (CN)。韩晓 (HAN, Xiao); 中国江苏省苏州市中国 (江苏) 自由贸易试验区苏州片区苏州工业园区玲珑街88号 215000 (CN)。
- (74) 代理人: 北京集佳知识产权代理有限公司 (UNITALEN ATTORNEYS AT LAW CO., LTD.); 中国北京市朝阳区建国门外大街 22 号赛特广场七层 100004 (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,

(54) Title: IMAGE PROCESSING MODEL TRAINING METHOD, AND IMAGE PROCESSING METHOD AND APPARATUS

(54) 发明名称: 图像处理模型的训练方法、图像处理方法及装置

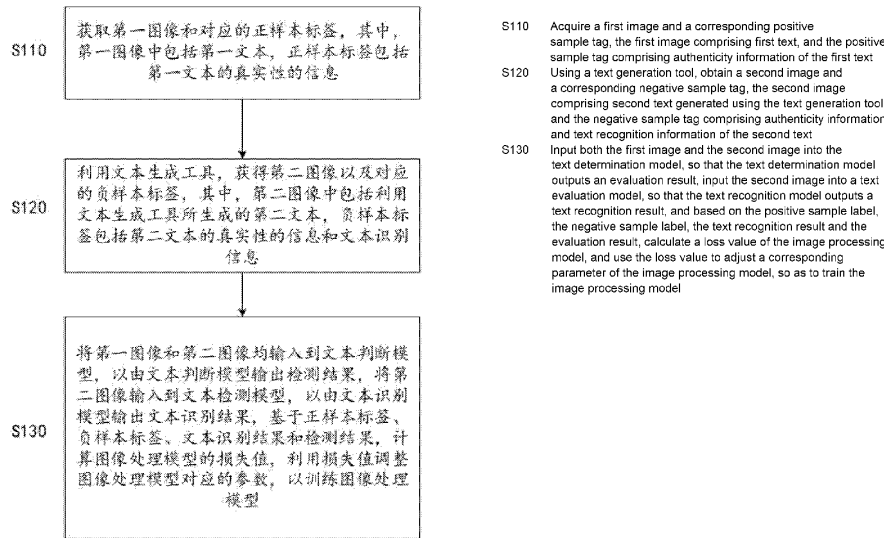


图 1

(57) Abstract: The present application provides an image processing model training method, and an image processing method and apparatus. The image processing model comprises a text recognition model and a text determination model, the text recognition model being used to perform text recognition on an image, and the text determination model being used to evaluate the authenticity of the text in the image. The method comprises: acquiring a first image and a corresponding positive sample tag; using a text generation tool, obtaining a second image and a corresponding negative sample tag; inputting both the first image and the second image into the text

[见续页]

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, 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))。

determination model, so that the text determination model outputs an evaluation result, inputting the second image into a text evaluation model, so that the text evaluation model outputs a text evaluation result, and, based on the positive sample tag, the negative sample tag, the text recognition result, and the evaluation result, calculating a loss value of the image processing model, and using the loss value to train the image processing model. The solution effectively saves the energy of workers, and increases model training speed.

(57) 摘要: 本申请提供了一种图像处理模型的训练方法、图像处理方法及装置。图像处理模型包括文本检测模型和文本判断模型, 文本检测模型用于对图像进行文本识别, 文本判断模型用于检测图像中的文本的真实性。该方法包括: 获取第一图像和对应的正样本标签; 利用文本生成工具, 获得第二图像以及对应的负样本标签; 将第一图像和第二图像均输入到文本判断模型, 以由文本判断模型输出检测结果, 将第二图像输入到文本检测模型, 以由文本检测模型输出文本识别结果, 基于正样本标签、负样本标签、文本识别结果和检测结果, 计算图像处理模型的损失值, 利用损失值训练图像处理模型。该方案有效地节省了工作人员的精力, 加速了模型训练的速度。