



- (51) **International Patent Classification:**
G06F 40/177 (2020.01)
- (21) **International Application Number:**
PCT/IB2021/052167
- (22) **International Filing Date:**
16 March 2021 (16.03.2021)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
16/847,792 14 April 2020 (14.04.2020) US
- (71) **Applicant:** **INTERNATIONAL BUSINESS MACHINES CORPORATION** [US/US]; New Orchard Road, Armonk, New York, 10504 (US).
- (71) **Applicants (for MG only):** **IBM UNITED KINGDOM LIMITED** [GB/GB]; PO Box 41, North Harbour, Portsmouth Hampshire PO6 3AU (GB). **IBM (CHINA) INVESTMENT COMPANY LIMITED** [CN/CN]; 25/F, Pangu Plaza, No.27, Central North 4th Ring Road, Chaoyang District, Beijing 100101 (CN).
- (72) **Inventors:** **WANG, Xin Ru**; IBM CORPORATION, 650 Harry Road, Almaden Research Center, San Jose, Califor-

nia 95120-6099 (US). **BURDICK, Douglas, Ronald**; IBM CORPORATION, 650 Harry Road, Almaden Research Center, San Jose, California 95120-6099 (US). **ZHENG, Xinyi**; IBM CORPORATION, Intellectual Property Law, Almaden Research Center, 650 Harry Road, San Jose, California 95120-6099 (US).

- (74) **Agent:** **SHAW, Anita**; IBM UNITED KINGDOM LIMITED, Intellectual Property Law, Hursley Park, Hursley, Winchester Hampshire So21 2jn (GB).

(81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ,

(54) **Title:** VISION-BASED CELL STRUCTURE RECOGNITION USING HIERARCHICAL NEURAL NETWORKS AND CELL BOUNDARIES TO STRUCTURE CLUSTERING

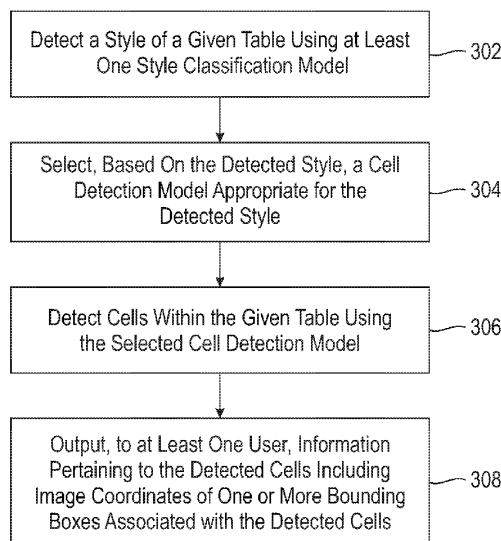


FIG. 3

(57) **Abstract:** Methods, systems, and computer program products for vision-based cell structure recognition using hierarchical neural networks and cell boundaries to structure clustering are provided herein. A computer-implemented method includes detecting a style of the given table using at least one style classification model; selecting, based at least in part on the detected style, a cell detection model appropriate for the detected style; detecting cells within the given table using the selected cell detection model; and outputting, to at least one user, information pertaining to the detected cells comprising image coordinates of one or more bounding boxes associated with the detected cells.

UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, 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).

Published:

— *with international search report (Art. 21(3))*