

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
02 December 2021 (02.12.2021)



(10) International Publication Number
WO 2021/240286 A1

(51) International Patent Classification:
G06N 3/063 (2006.01)

(21) International Application Number:
PCT/IB2021/054105

(22) International Filing Date:
13 May 2021 (13.05.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
16/884,128 27 May 2020 (27.05.2020) 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: **TSAL, Hsinyu**; IBM CORPORATION, 650 Harry Road, Almaden Research Center, San Jose, California 95120-6099 (US). **BURR, Geoffrey**; IBM CORPORATION, 650 Harry Road, Almaden Research Center, San Jose, California 95120-6099 (US). **NARAYANAN, Pritish**; IBM CORPORATION, 650 Harry Road, Almaden Research Center, San Jose, California 95120-6099 (US).

(74) Agent: **WILLIAMS, Julian**; IBM United Kingdom Limited, Intellectual Property Law, Hursley Park, Hursley, Winchester Hampshire SO21 2JN (GB).

(54) Title: EFFICIENT TILE MAPPING FOR ROW-BY-ROW CONVOLUTIONAL NEURAL NETWORK MAPPING FOR ANALOG ARTIFICIAL INTELLIGENCE NETWORK INFERENCE

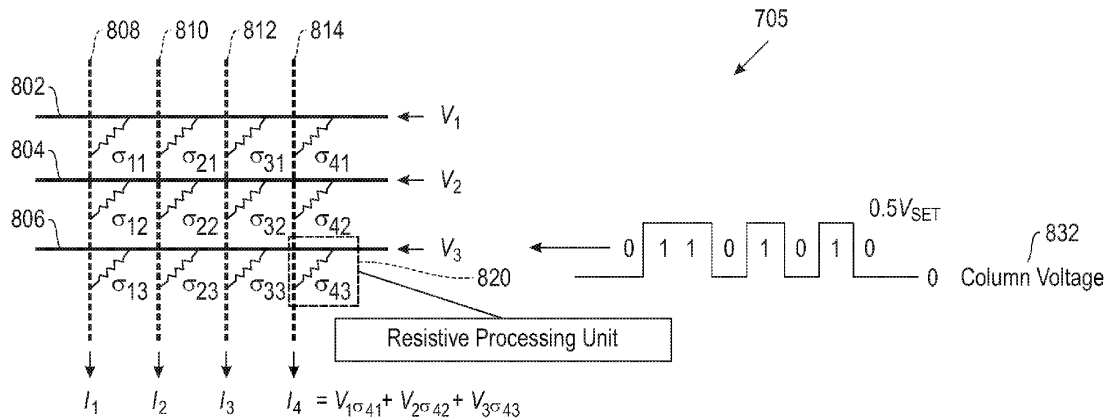


FIG. 8

(57) Abstract: Implementing a convolutional neural network (CNN) includes configuring a crosspoint array to implement a convolution layer in the CNN. Convolution kernels of the layer are stored in crosspoint devices of the array. Computations for the CNN are performed by iterating a set of operations for a predetermined number of times. The operations include transmitting voltage pulses corresponding to a subpart of a vector of input data to the crosspoint array. The voltage pulses generate electric currents that are representative of performing multiplication operations at the crosspoint device based on weight values stored at the crosspoint devices. A set of integrators accumulates an electric charge based on the output electric currents from the respective crosspoint devices. The crosspoint array outputs the accumulated charge after iterating for the predetermined number of times. The accumulated charge represents a multiply-add result of the vector of input data and the one or more convolution kernels.

[Continued on next page]

(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, 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))*
- *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*