



- (51) **International Patent Classification:**
G06N 3/06 (2006.01) *G06N 20/00* (2006.01)
- (21) **International Application Number:**
PCT/SG2019/050440
- (22) **International Filing Date:**
03 September 2019 (03.09.2019)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (71) **Applicant:** AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH [SG/SG]; 1 Fusionopolis Way, #20-10 Connexis North Tower, Singapore 138632 (SG).
- (72) **Inventors:** PARAMASIVAM, Vishnu; c/o Industry Development, Institute of Microelectronics, 2 Fusionopolis Way, #08-02 Innovis Tower, Singapore 138634 (SG). DO, Anh Tuan; c/o Industry Development, Institute of Microelectronics, 2 Fusionopolis Way, #08-02 Innovis Tower,

Singapore 138634 (SG). KOH, Eng Kiat; c/o Industry Development, Institute of Microelectronics, 2 Fusionopolis Way, #08-02 Innovis Tower, Singapore 138634 (SG). PU, Junran; c/o Industry Development, Institute of Microelectronics, 2 Fusionopolis Way, #08-02 Innovis Tower, Singapore 138634 (SG). LI, Fei; c/o Industry Development, Institute of Microelectronics, 2 Fusionopolis Way, #08-02 Innovis Tower, Singapore 138634 (SG). MANI, Aarth; c/o Industry Development, Institute of Microelectronics, 2 Fusionopolis Way, #08-02 Innovis Tower, Singapore 138634 (SG).

(74) **Agent:** VIERING, JENTSCHURA & PARTNER LLP; P.O. Box 1088, Rochor Post Office, Rochor Road, Singapore 911833 (SG).

(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,

(54) **Title:** NEURAL NETWORK PROCESSOR SYSTEM AND METHODS OF OPERATING AND FORMING THEREOF

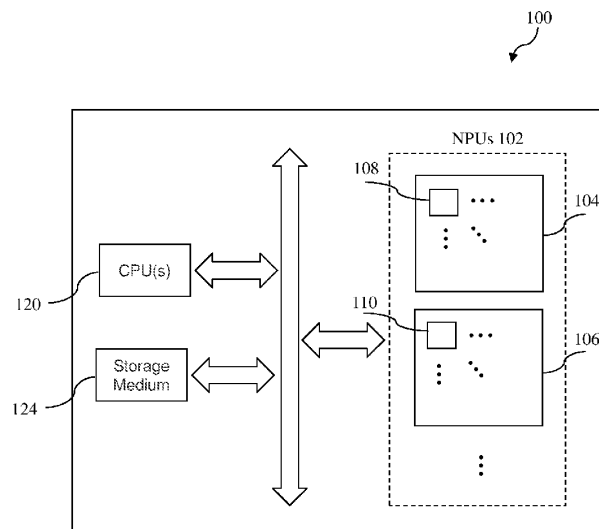


FIG. 1

(57) **Abstract:** There is provided a neural network processor system including: a plurality of neural processing units, including a first neural processing unit and a second neural processing unit, whereby each neural processing unit comprises an array of neural processing core blocks, each neural processing core block comprising a neural processing core; and at least one central processing unit communicatively coupled to the plurality of neural processing units and configured to coordinate the plurality of neural processing units for performing neural network computations. In particular, the first and second neural processing units have a different structural configuration to each other. There is also provided a corresponding method of operating and a corresponding method of forming the neural network processor system.

[Continued on next page]

DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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