

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2019年11月21日(21.11.2019)



(10) 国際公開番号

WO 2019/220975 A1

(51) 国際特許分類:
G06N 3/063 (2006.01)

(21) 国際出願番号: PCT/JP2019/018382

(22) 国際出願日: 2019年5月8日(08.05.2019)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2018-094184 2018年5月15日(15.05.2018) JP

(71) 出願人: 国立大学法人東京工業大学
(TOKYO INSTITUTE OF TECHNOLOGY) [JP/
JP]; 〒1528550 東京都目黒区大岡山 2 - 1 2 - 1 Tokyo (JP).

(72) 発明者: 中原 啓 貴 (NAKAHARA Hiroki);
〒1528550 東京都目黒区大岡山 2 - 1 2 - 1 国
立大学法人東京工業大学内 Tokyo (JP).

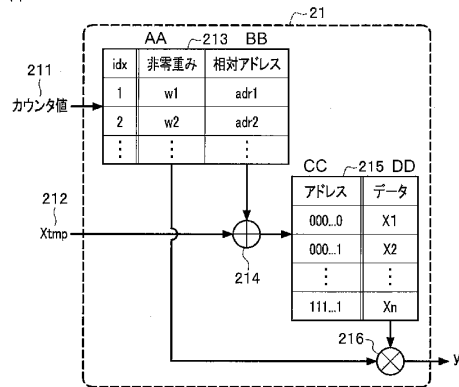
(74) 代理人: 特許業務法人磯野国際特許商
標事務所 (ISONO INTERNATIONAL PATENT
OFFICE, P.C.); 〒1050001 東京都港区虎ノ
門一丁目1番18号 ヒューリック
虎ノ門ビル Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保
護が可能): 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, JO, KE, KG, KH,

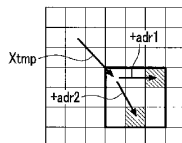
(54) Title: NEURAL NETWORK CIRCUIT DEVICE, NEURAL NETWORK PROCESSING METHOD, AND NEUR-
AL NETWORK EXECUTION PROGRAM

(54) 発明の名称: ニューラルネットワーク回路装置、ニューラルネットワーク処理方法
およびニューラルネットワークの実行プログラム

〔図4〕
(a)



(b)



211 Counter value
AA Non-zero weighting
BB Relative address
CC Address
DD Data

(57) Abstract: Provided are a neural network circuit device which can reduce memory volume and can be disposed on a chip, a neural network processing method, and an execution program of a neural network. A 3-state neural network circuit (200) comprises: a non-zero convolution circuit (21) that, in an intermediate layer, receives a weighting W_i and an input value X_i for performing convolution, and performs a convolution operation; a summing circuit (22) that obtains the sum

〔続葉有〕

WO 2019/220975 A1

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) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (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).

添付公開書類:

- 国際調査報告 (条約第21条(3))

of a bias W_0 and each computed value obtained by the convolution operation; and an activation function circuit (23) that uses an activation function $f(u)$ to convert a summed signal Y . The non-zero convolution operation circuit (21) skips a weighting of zero for the weighting W_i and performs a convolution operation on the basis of a non-zero weighting and an input value X_i corresponding to the non-zero weighting.

(57) 要約: メモリ量を削減してオンチップに配置可能なニューラルネットワーク回路装置、ニューラルネットワーク処理方法およびニューラルネットワークの実行プログラムを提供する。3状態ニューラルネットワーク回路(200)は、中間層の中で、畳み込みを行う入力値 X_i と、重み W_i を受け取り、畳み込み演算を行う非ゼロ畳み込み演算回路(21)と、畳み込み演算された各演算値とバイアス W_0 との総和を取る総和回路(22)と、総和を取った信号 Y を活性化関数 $f(u)$ で変換する活性化関数回路(23)と、を備える。非ゼロ畳み込み演算回路(21)は、重み W_i がゼロの重みをスキップし、非ゼロの重みと当該非ゼロの重みに対応する入力値 X_i をもとに畳み込み演算を行う。