

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

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
16 November 2023 (16.11.2023)



(10) International Publication Number
WO 2023/217370 A1

(51) International Patent Classification:
G06N 3/04 (2006.01) *G06N 3/08* (2006.01)

(21) International Application Number:
PCT/EP2022/062878

(22) International Filing Date:
12 May 2022 (12.05.2022)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **HUAWEI TECHNOLOGIES CO., LTD.**
[CN/CN]; Huawei Administration Building Bantian, Long-
gang District, Shenzhen, Guangdong 518129 (CN).

(72) Inventor; and

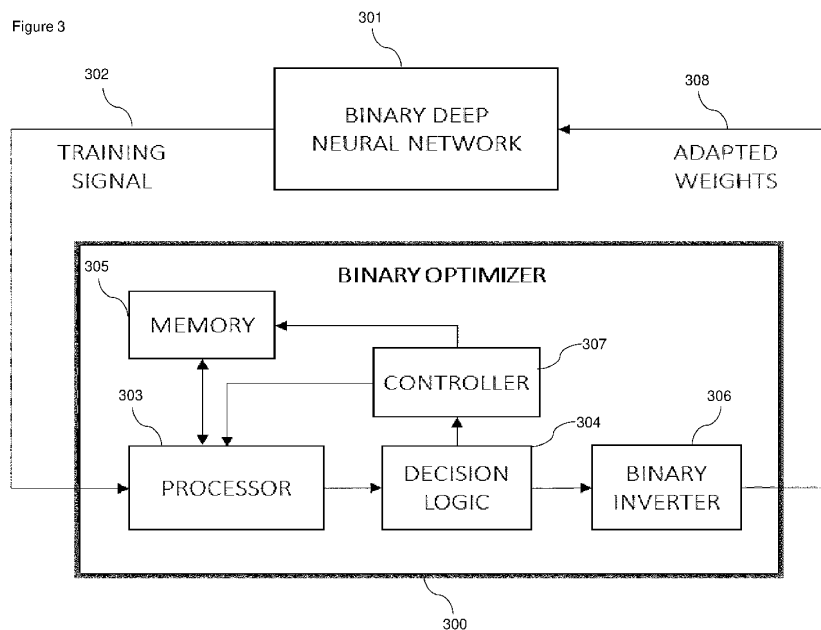
(71) Applicant (for MN only): **NGUYEN, Van Minh** [FR/FR];
Huawei Technologies Duesseldorf GmbH, Riesstr. 25,
80992 Munich (DE).

(72) Inventor: **LECONTE, Louis**; Huawei Technologies Dues-
seldorf GmbH, Riesstr. 25, 80992 Munich (DE).

(74) Agent: **KREUZ, Georg M.**; Huawei Technologies Dues-
seldorf GmbH, Riesstr. 25, 80992 Munich (DE).

(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, IQ, IR, IS, IT, JM, 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,

(54) Title: APPARATUS AND METHOD FOR TRAINING BINARY DEEP NEURAL NETWORKS



(57) Abstract: Disclosed is a device (600) for training a binary deep neural network, the device being configured to: generate (501) a training signal (302) in dependence on an error between an output of a prototype version of the binary deep neural network (301) and an expected output, the prototype version of the binary deep neural network having multiple binary weights each having a respective value; and in dependence on the training signal (302), output (502) for each binary weight of the prototype version of the binary deep neural network (301) a respective decision to invert or maintain the respective value of the respective binary weight. This may allow the device to train a deep neural network comprising binary parameters directly in the binary domain without the need for gradient processing methods. As a result, the training process may require reduced memory and computational power because it avoids memory-intensive gradient signals.

[Continued on next page]

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, SC, 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