

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

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
26 August 2021 (26.08.2021)



(10) International Publication Number
WO 2021/168356 A1

- (51) International Patent Classification:
G16B 40/20 (2019.01)
- (21) International Application Number:
PCT/US2021/018913
- (22) International Filing Date:
19 February 2021 (19.02.2021)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
62/979,399 20 February 2020 (20.02.2020) US
62/979,411 20 February 2020 (20.02.2020) US
17/179,395 18 February 2021 (18.02.2021) US
17/180,480 19 February 2021 (19.02.2021) US
17/180,513 19 February 2021 (19.02.2021) US
- (71) Applicant: **ILLUMINA, INC.** [US/US]; 5200 Illumina Way, San Diego, CA 92122 (US).
- (72) Inventors: **VESSERE, Gery**; 5200 Illumina Way, San Diego, California 92122 (US). **PARNABY, Gavin Derek**; 5200 Illumina Way, San Diego, California 92122 (US). **DUTTA, Anindita**; 5200 Illumina Way, San Diego, California 92122 (US). **KASHEFHAGHIGHI, Dorna**; 5200 Illumina Way, San Diego, California 92122 (US). **JAGANATHAN, Kishore**; 5200 Illumina Way, San Diego, California 92122 (US). **KIA, Amirali**; 811 N El Camino Real, Apt. 302, San Mateo, California 94401 (US).
- (74) Agent: **BEFFEL, JR., Ernest J.** et al.; Haynes Beffel & Wolfeld LLP, P.O. Box 366, 637 Main Street, Half Moon Bay, California 94019 (US).
- (81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM,

(54) Title: DATA COMPRESSION FOR ARTIFICIAL INTELLIGENCE-BASED BASE CALLING

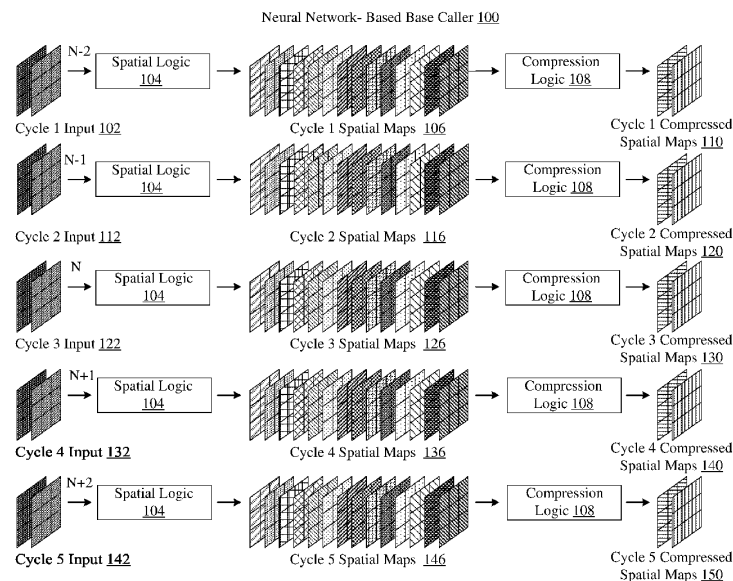


Figure 1A

(57) Abstract: The technology disclosed relates to an artificial intelligence-based method of base calling. In particular, it relates to processing, through a spatial network of a neural network-based base caller, a first window of per-cycle analyte channel sets in for a first window of sequencing cycles of a sequencing run, and generating respective sequences of spatial output sets for respective sequencing cycles in the first window of sequencing cycles, processing, through a compression network of the neural network-based base caller, respective final spatial output sets in the respective sequences of spatial output sets, and generating respective compressed spatial output sets for the respective sequencing cycles in the first window of sequencing cycles, and generating, based on the respective compressed spatial output sets, base call predictions for one or more sequencing cycles in the first window of sequencing cycles.

[Continued on next page]

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))