

(51) International Patent Classification:
G16B 40/10 (2019.01) *G16B 40/20* (2019.01)(21) International Application Number:
PCT/US2021/018422(22) International Filing Date:
17 February 2021 (17.02.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/979,385 20 February 2020 (20.02.2020) US
17/176,151 15 February 2021 (15.02.2021) US(71) Applicant: **ILLUMINA, INC.** [US/US]; 5200 Illumina Way, San Diego, California 92122 (US).(72) Inventors: **DUTTA, Anindita**; 5200 Illumina Way, San Diego, California 92122 (US). **VESSERE, Gery**; 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, 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, 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,

(54) Title: KNOWLEDGE DISTILLATION AND GRADIENT PRUNING-BASED COMPRESSION OF ARTIFICIAL INTELLIGENCE-BASED BASE CALLER

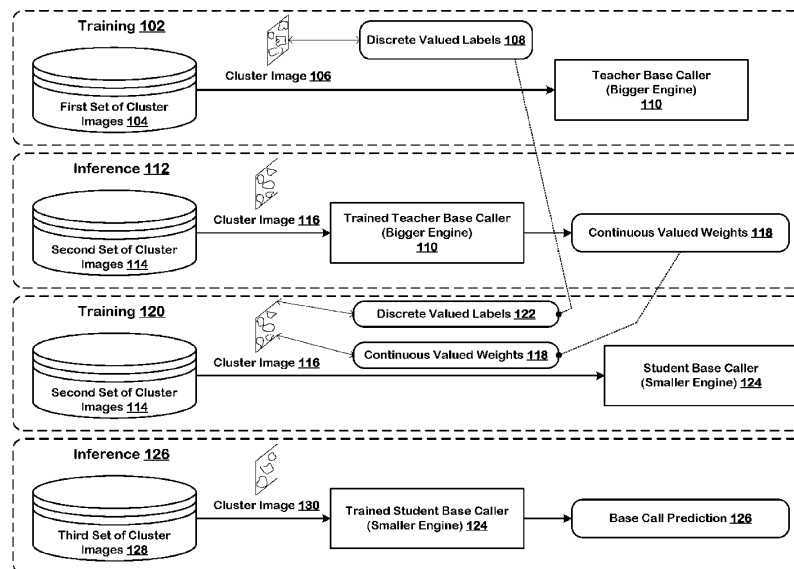


FIG. 1

(57) Abstract: The technology disclosed compresses a larger, teacher base caller into a smaller, student base caller. The student base caller has fewer processing modules and parameters than the teacher base caller. The teacher base caller is trained using hard labels (e.g., one-hot encodings). The trained teacher base caller is used to generate soft labels as output probabilities during the inference phase. The soft labels are used to train the student base caller.

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).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *with international search report (Art. 21(3))*