



- (51) International Patent Classification:
G06N 3/045 (2023.01) *G06N 3/084* (2023.01)
- (21) International Application Number:
PCT/EP2024/058965
- (22) International Filing Date:
02 April 2024 (02.04.2024)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
63/606,590 05 December 2023 (05.12.2023) US
63/553,503 14 February 2024 (14.02.2024) US
- (71) Applicant: **DEEPMIND TECHNOLOGIES LIMITED**
[GB/GB]; 5 New Street Square, London EC4A 3TW (GB).
- (72) Inventors: **TEPLYASHIN, Denis**; 6 Pancras Square, London N1C 4AG (GB). **MANJUNATHA, Pranav**; 6 Pancras Square, London N1C 4AG (GB). **SAVINOV, Nikolay**; 6

Pancras Square, London N1C 4AG (GB). **ADLER, Jonas Anders**; 6 Pancras Square, London N1C 4AG (GB). **RAE, Jack William**; 6 Pancras Square, London N1C 4AG (GB).

(74) Agent: **FISH & RICHARDSON P.C.**; HighLight Business Towers Mies-van-der-Rohe-Str. 8, 80807 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, CV, 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, MG, MK, MN, MU, 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.

(54) Title: ATTENTION NEURAL NETWORKS WITH PARTIAL POSITION ENCODING

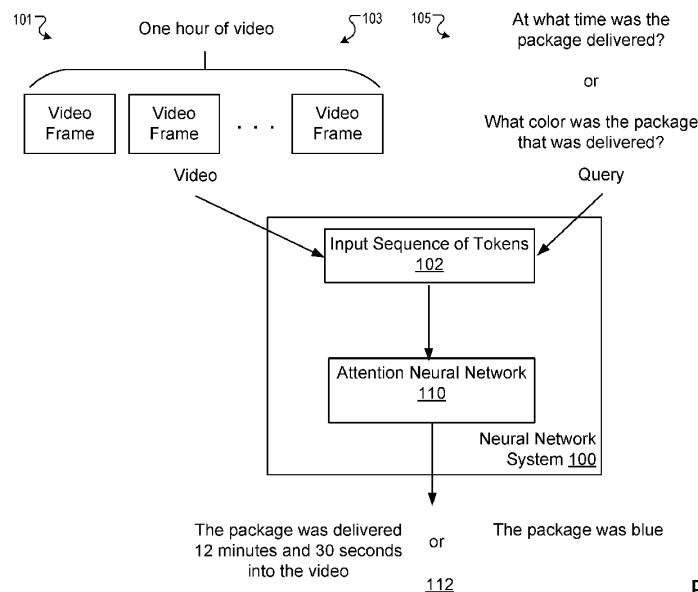


FIG. 1A

(57) Abstract: Methods, systems, and apparatus, including computer programs encoded on computer storage media, for processing input sequences using a neural network that uses a partial position encoding scheme. The neural network generally includes both global and local attention layers. In the partial position encoding scheme, while the local attention layers do use position encoding, (i) a subset of the global attention layers can apply an attention mechanism that does not use position encoding, or (ii) the subset of global attention layers can apply an attention mechanism that does not apply position encoding to one or more of the dimensions of the input to the attention mechanism.

[Continued on next page]

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, CV, 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, ME, 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*