

(19)



SUOMI - FINLAND
(FI)

PATENTTI- JA REKISTERIHALLITUS
PATENT- OCH REGISTERSTYRELSEN
FINNISH PATENT AND REGISTRATION OFFICE

(10) **FI/EP3764314 T3**
(12) **EUROOPPAPATENTIN KÄÄNNÖS**
ÖVERSÄTTNING AV EUROPEISKT PATENT
TRANSLATION OF EUROPEAN PATENT SPECIFICATION

(45) Käännöksen kuulutuspäivä - Kungörelsedag av översättning - **08.02.2024**
Translation available to the public

(97) Eurooppapatentin myöntämispäivä - Meddelandedatum för **29.11.2023**
det europeiska patentet - Date of grant of European patent

(51) Kansainvälinen patenttiluokitus - Internationell patentklassificering -
International patent classification
G06T 1/20 (2006 . 01)

(96) Eurooppapatenttihakemus - Europeisk patentansökan - **EP20192178.0**
European patent application

(22) Tekemispäivä - Ingivningsdag - Filing date **08.03.2018**

(97) Patenttihakemuksen julkiseksitilopäivä - Patentansökans **13.01.2021**
publiceringsdag - Patent application available to the public

(30) Etuoikeus - Prioritet - Priority
09.04.2017 US US201715482791

(73) Haltija - Innehavare - Holder
1• INTEL Corporation , 2200 Mission College Blvd. , Santa Clara, CA 95054 , (US)

(72) Keksijä - Uppfinnare - Inventor
1• Nurvitadhi, Eriko , 2111 NE 25th Ave , Hillsboro, Oregon 97124 , (US)
2• Vembu, Balaji , 1144 Bazio Court , Folsom, California 95630 , (US)
3• Lin, Tsung-Han , 2273 Central Park Dr. , Campbell, California 95008 , (US)
4• Sinha, Kamal , 713 Oreno Cir , Folsom, CA 95630 , (US)
5• Barik, Rajkishore , 1510 Vista Club Circle, 210 , Santa Clara, California 95054 , (US)
6• Galoppo von Borries, Nicolas C. , 1808 NE Knott St , Portland, California 97212 , (US)

(74) Asiamies - Ombud - Agent
Papula Oy , P.O.Box 981 , 00101 Helsinki , (FI)

(54) Keksinnön nimitys - Uppfinningens benämning - Title of the invention
KONEOPPIMISEN HARVALASKENTAMEKANISMI
MACHINE LEARNING SPARSE COMPUTATION MECHANISM

PATENTTIVAATIMUKSET

1. Menetelmä neuroverkon kerroskäsittelyn suorittamiseksi, joka menetelmä käsittää:

5 tallennetaan matriisielementtejä grafiikkaprosessorin (614) muistiin; ja

laukaistaan laskentaytimen suoritus grafiikka-prosessorissa (614), missä laskentaydin suorittaa matriisien elementeille konvoluutiokerrosoperaation, jota
10 seuraa bias-kerrosoperaatio konvoluutiokerrosoperaation lähdössä, jota seuraa korjattu lineaariyksikkökerrosoperaatio bias-kerrosoperaation lähdössä, missä lähdöt välitetään kerrosoperaatioiden välillä, missä operaatiot suoritetaan useilla grafiikkaprosessorin (614)
15 suoritusyksiköillä (720), missä suoritusyksiköt (720) on jaettu useisiin ryhmiin ((720(1)-720(10)); (720(11)-720(13)); (720(14)-720(16)); (720(17)- 720(19))), missä useiden ryhmien ryhmälle allokoitua suoritusyksiköt (720) suorittavat saman kerroksen operaatioita.

20

2. Ei-transitorinen koneellisesti luettava väline, joka sisältää käskyjä, jotka yhden tai useamman prosessorin suorittamana aiheuttavat mainitun yhden tai useamman prosessorin suorittaa patenttivaatimuksen 1
25 mukaisen menetelmän operaatiot.

3. Tietojenkäsittelyjärjestelmä, joka käsittää:

muistilaitteen; ja

30 yhden tai useamman prosessorin, joka on konfiguroitu suorittamaan muistilaitteeseen tallennettuja käskyjä, missä käskyt aiheuttavat mainitun yhden tai useamman prosessorin suorittaa neuroverkon kerroskäsittelyä, missä mainittu yksi tai useampi prosessori sisältää grafiikkaprosessorin (614) ja käskyt aiheuttavat
35 mainitun yhden tai useamman prosessorin:

tallentaa matriisielementit grafiikkaprosessorin (614) muistiin; ja

laukaista laskentaytimen suoritus grafiikkaprosessorissa (614), missä laskentaydin aiheuttaa
5 grafiikkaprosessorin (614) suorittaa matriisien elementeille konvoluutiokerrosoperaation, jota seuraa bias-kerrosoperaatio konvoluutiokerrosoperaation lähdössä, jota seuraa korjattu lineaariyksikkökerrosoperaatio bias-kerrosoperaation lähdössä, missä lähdöt välitetään
10 kerrosoperaatioiden välillä, missä kerrosoperaatiot suoritetaan useilla grafiikkaprosessorin (614) suoritusyksiköillä (720), missä suoritusyksiköt (720) on jaettu useisiin ryhmiin ((720(1)-720(10)); (720(11)-720(13)); (720(14)-720(16)); (720(17)- 720(19))), missä
15 useiden ryhmien ryhmälle allokoitua suoritusyksiköt (720) suorittavat saman kerroksen operaatioita.