

Kahden optimointimenetelmän vertailu videoperusmallin koulutuksessa ja siirto-oppimisessä (aihe-esittely)

Johan Roos

16.09.2026

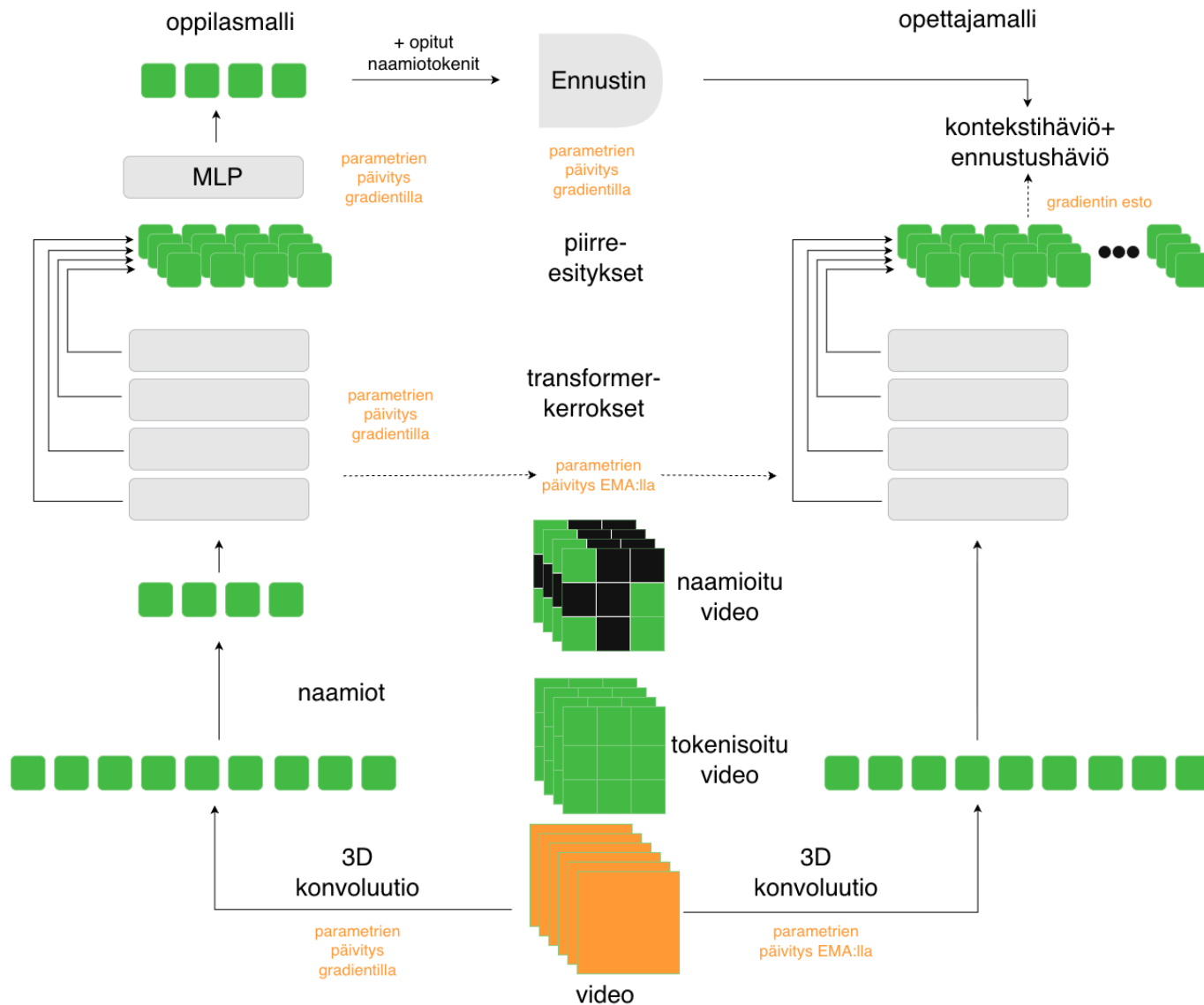
Ohjaaja: Ahti Salo

Valvoja: Ahti Salo

Työn saa tallentaa ja julkistaa Aalto-yliopiston avoimilla verkkosivuilla. Muilta osin kaikki oikeudet pidätetään.

Tausta (1/2)

- Itseohjattuun oppimiseen perustuvat JEPA-videomallit koulutetaan laajoilla video- ja kuva-aineistoilla, jotta niitä voidaan sovittaa myöhemmin eri tehtäviin
 - Siirto-oppimisessa mallin syötteen pohjalta tuottama piirre-esitys on tärkeä
- Koulutus
 - Opettajakamalli saa alkuperäisen videon syötteenä ja tuottaa siitä oman piirre-esityksen
 - Oppilasmalli muodostaa piirre-esityksen videon näkyvistä osista
 - Ennustin ennustaa sen perusteella opettajakamallin tuottamat peitettyjen alueiden piirre-esitykset



Tausta (2/2)

- Työssä verrataan kahta mallin koulutuksessa käytettävää optimointimenetelmää
 1. AdamW
 - Laskee gradienttien sekä gradienttien neliöiden liukuvien keskiarvojen pohjalta parametrikohittaiset päivitykset
 2. Muon, jolla voidaan optimoida matriisimuotoisia parametrejä
 - Ortogonalisoi likimääräisesti parametrien päivitysmatriisin Newton-Schulz-iteraatiolla
 - Laskettu päivitys keskittyy tasaisemmin useaan pääakselisuuntaan

Tavoitteet

- Tutkimuskysymykset
 1. Kuinka paljon laskentaa liukulukuoperaatioina (FLOP) AdamW ja Muon tarvitsevat saman suorituskyvyn saavuttamiseen?
 2. Kuinka samankaltaisia AdamW:llä ja Muonilla koulutettujen mallien piirre-esitykset ovat?
 3. Eroaako AdamW:llä ja Muonilla opittujen piirre-esitysten siirrettävyys uusiin videotehtäviin, kun mallien suorituskyky vakioidaan?

Rajaukset

- Tarkin vertailu saataisiin hienosäätämällä kaikki optimointiin vaikuttavat hyperparametrit kummallekin optimointimenetelmälle
 - Hyperparametrejä ovat esimerkiksi oppimisnopeus, oppimisnopeusaikataulu ja momentum
 - Optimiarvot riippuvat lisäksi mm. malliarkkitehtuurista, eräkoosta ja koulutuksen kestosta
 - Laskentabudjetti ei riitä parametrien säätämiseen
 - Siksi AdamW:n hyperparametreiksi valitaan kirjallisuudessa käytettyjä arvoja, ja Muonille käytetään niitä vastaavaa peruskonfiguraatiota
- Tavoitteena ei ole kilpailla uusimpia tutkimuksia vastaan
- Koska vertailu tehdään yhdellä pienellä mallilla, yleistä skaalautuvuutta esimerkiksi suurempien mallien tai aineistojen suhteen ei voida todentaa

Menetelmät

- Valitaan neuroverkko-arkkitehtuuri, jossa on parametrejä n. 25 miljoonaa
- Tokenisoitu koulutusaineisto on n. $1,5 \times 10^9$ tokenia
 - Something-Something-v2 (SSv2) videot
 - ImageNet-1k kuvat
- Kummankin optimointimenetelmän kokeet tehdään kolmella satunnaissiemenellä ja koulutetaan 20 jakson verran
 1. Suorituskyky mitataan SSv2-validointiaineistolla. Enkooderin painot jäädytetään ja sen piirre-esitysten perusteella koulutetaan lineaarinen luokitin
 2. AdamW- ja Muon-mallien piirre-esityksiä verrataan Centered Kernel Alignment-menetelmällä
 3. Siirto-oppimiskykyä mitataan kouluttamalla jäädytettyjen enkooderien piirre-esityksille lineaariset luokittimet Diving48-uimahyppyaaineistolla

Työkalut

- RTX5090 näytönohjain
- Keskeiset ohjelmistopaketit
 - PyTorch
 - Hydra
 - Wandb

Tietolähteet

- Lorenzo Mur-Labadia, Matthew Muckley, Amir Bar, Mido Assran, Koustuv Sinha, Mike Rabbat, Yann LeCun, Nicolas Ballas, and Adrien Bardes. “V-JEPA 2.1: Unlocking Dense Features in Video Self-Supervised Learning.” *arXiv Preprint arXiv:2603.14482*, 2026.
- Jeremy Bernstein. “Deriving Muon.” 2025.
<https://jeremybernste.in/writing/deriving-muon>.
- Keller Jordan, Yuchen Jin, Vlado Boza, Jiacheng You, Franz Cesista, Laker Newhouse, and Jeremy Bernstein. “Muon: An Optimizer for Hidden Layers in Neural Networks.” 2024.
<https://kellerjordan.github.io/posts/muon/>.
- Simon Kornblith, Mohammad Norouzi, Honglak Lee, and Geoffrey Hinton. “Similarity of Neural Network Representations Revisited.” *Proceedings of Machine Learning Research*, 2019.

Aikataulu

- Lähdekoodin viimeistely ja laskennallisten vertailujen tekeminen 9-10/26
- Työn kirjoittaminen 10-12/26
- Valmiin työn esittely 11/26