

ЗАТВЕРДЖУЮ

Проректор з навчальної роботи

Національного технічного університету України

“Київський політехнічний інститут

імені Ігоря Сікорського”

к. філос. н., професор

Анатолій МЕЛЬНИЧЕНКО



ВИТЯГ

з протоколу № 26 від 14 березня 2024 року розширеного засідання кафедри Інженерії програмного забезпечення в енергетиці Навчально-наукового інституту атомної і теплової енергетики Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”

БУЛИ ПРИСУТНІ:

- з кафедри інженерії програмного забезпечення в енергетиці: в.о. завідувача кафедри інженерії програмного забезпечення в енергетиці д.т.н, проф. Коваль О.В., професор д.т.н., проф. Барабаш О.В., професор д.т.н., проф. Гаврилко Є.В., професор д.т.н, доц., Мусієнко А.П., професор д.т.н., доц. Недашківський О.Л., професор д.е.н., проф. Сігайов А.О., професор д.т.н., доц. Федорова Н.В., доцент к.т.н., доц. Гагарін О.О., доцент к.т.н., доц. Кузьмініх В.О., доцент к.т.н., доц. Стативка Ю.І., доцент к.т.н., доцент к.ф.-м.н., доц. Свинчук О.В., доцент к.т.н. Варава І.А., доцент к.е.н., Гусева І.І., доцент к.т.н. Шпурик В.В., ст. викл. Дацюк О.А., ст. викл. Колумбет В.П., ст. викл. Бандурка О.І., ас. Оленєва К.М., ас. Швайко В.Г., асп.Олексій А.О., асп. Гейко О.О., асп. Бочок В.О., зав. лаб. Гайдаржи В.І., інж. Вакулович О.В., зав. лаб. Беленцов О.С.

- з інших кафедр КПІ ім. Ігоря Сікорського:

завідувач кафедри обчислювальної техніки факультету інформатики та обчислювальної техніки, д.т.н., професор Стіренко Сергій Григорович

доцент кафедри цифрових технологій в енергетиці навчально наукового інституту атомної та теплової енергетики КПІ ім. Ігоря Сікорського, к.т.н.,
Залевська Ольга Валеріївна

Запрошені з інших організацій:

завідувач кафедри кібербезпеки та захисту даних приватної установи «Університету науки, підприємництва та технологій», д.т.н., професор Семенов Сергій Геннадійович.

завідувач кафедри Інженерії програмного забезпечення автоматизованих систем Державного університету інформаційно-телекомунікаційних технологій д.т.н., професор Сторчак Каміла Павлівна.

СЛУХАЛИ:

1. Повідомлення випускника аспірантури кафедри інженерії програмного забезпечення в енергетиці Мельниченка Артема Васильовича за матеріалами дисертаційної роботи “Методи та програмні засоби підвищення швидкодії моделей розпізнавання образів на основі машинного навчання”, поданої на здобуття ступеня доктора філософії з галузі знань 12 Інформаційні технології за спеціальністю 121 Інженерія програмного забезпечення.

Освітньо-наукова програма «Інженерія програмного забезпечення».

Тему дисертаційної роботи затверджено на засіданні Вченої ради теплоенергетичного факультету (протокол № 5 від 30 листопада 2020 року) та перезатвердженню на тему “ Методи та програмні засоби підвищення швидкодії моделей розпізнавання образів на основі машинного навчання ” на засіданні Вченої ради навчально-наукового інституту атомної та теплової енергетики Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського” (протокол № 6 від 25 грудня 2023 року).

Науковими керівниками затверджені: доктор технічних наук, доцент, НЕДАШКІВСЬКИЙ О. Л., кандидат технічних наук, доцент, ШАЛДЕНКО О.В.

2. Запитання до здобувача.

Запитання по темі дисертації ставили: д.т.н. професор Стіренко С.Г., к.т.н., доцент Залевська О.В., д.т.н. професор Сторчак К.П., д.т.н. професор Барабаш О.В.

3. Виступи за обговореною роботою.

В обговоренні дисертації взяли участь: д.т.н, проф. Коваль О.В., проф. д.т.н., проф. Барабаш О.В., доцент Залевська О.В., д.т.н. професор Сторчак К.П. к.т.н., доц. Шалденко О.В., д.т.н. проф. Недашківський О.Л.

УХВАЛИЛИ:

ПРИЙНЯТИ такий висновок про наукову новизну, теоретичне та практичне значення результатів дисертаційного дослідження:

1. Актуальність теми дослідження.

У сучасному високотехнологічному світі, нейронні мережі вийшли на передній план як ключова технологія. Ця варіація математичних моделей продемонструвала високу ефективність у багатьох задачах, що варіюються від комп'ютерного зору до розуміння природніх мов, тим самим ставши невід'ємною частиною щоденного життя. Втім, розгортання нейронних мереж у реальних сценаріях часто ускладнюється їхньою обчислювальною складністю та ресурсоемністю. Великий об'єм енергоспоживання, що потребується для навчання і використання великих моделей нейронних також має негативний вплив на навколишнє середовище.

Обчислювальна складність часто проявляється у вигляді великої кількості параметрів та глибоких архітектур, які вимагають значного об'єму обчислювальної потужності як для навчання, так і для подальшого використання на кінцевих пристроях. Ця складність є особливо проблематичною в застосуваннях нейронних мереж на пристроях Інтернету речей (IoT), де обчислювальні ресурси часто обмежені. Ресурсоемні характеристики включають в себе обчислювальну потужність і використання пам'яті. Це питання є особливо актуальним у мобільних та вбудованих пристроях, де пам'ять є обмеженим ресурсом. Більше того, затримка, спричинена нестачею ресурсів, часто є неприйнятною в ряді задач, що включає в себе системи автономного керування, де навіть невелика затримка в прийнятті рішень може мати серйозні наслідки.

Окреслене сформувало протиріччя, коли з одного боку, нейронні мережі є ефективним і перспективним інструментом для вирішення задач розпізнавання образів, і з насиченням даними і появою нових глибоких архітектур дозволяють отримати кращі результати при вирішенні багатьох задач, з іншого боку, нейронні мережі потребують великий обсяг обчислювальних ресурсів для тренуванні і запуску, і відповідно, методів оптимізації при використанні на кінцевих пристроях для досягнення бажаної швидкодії. Таке протиріччя може бути вирішено за допомогою розробки методів та програмних засобів підвищення швидкодії моделей розпізнавання образів на основі машинного навчання.

2. Зв'язок роботи з науковими програмами, планами, темами.

Дисертаційне дослідження відповідає вимогам статті 5 Закону України «Про пріоритетні напрями розвитку науки і техніки» від 11 липня 2001 року № 2623-III (зі змінами та доповненнями від 12.01.2023 р.), пункту першого розділу другого «Переліку пріоритетних тематичних напрямів наукових досліджень і науково-технічних розробок на період до 2023 року», затвердженого Постановою КМУ від 7 вересня 2011 р. № 942. Дисертаційна робота виконана відповідно з поточними та перспективними планами наукової та науково-технічної діяльності кафедри інженерії програмного забезпечення в енергетиці Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» для подальшого розвитку інженерії програмного забезпечення. Результати дисертаційної роботи є складовою науково-дослідної роботи: “Методи і алгоритми оптимізації розпізнавання образів на основі методів машинного навчання” (РК №0121U109207).

Особисто автором в НДР №0121U109207 запропоновано вдосконалений метод розрахунку критерію важливості ваг нейронної мережі при видаленні вагів перед навчанням для моделі розпізнавання обличчя, та моделі архітектури трансформер. Проведено експерименти для оцінки ефективності розробленого методу.

3. Наукова новизна отриманих результатів.

У дисертації вперше одержані такі нові наукові результати:

Удосконалено модель нейронної мережі для виявлення обличчя RetinaFace, яка на відміну від існуючих використовує метод прунінгу SNIP для оптимізації, що дозволяє використовувати розріджені матриці для зберігання і виконання мережі з метою подальшого удосконалення та збільшення швидкодії.

Удосконалено метод прунінгу SNIP для моделі виявлення обличчя RetinaFace, який на відміну від існуючих передбачає можливість виключення контекстних модулів з процесу прунінгу. Вдосконалений метод дозволяє досягти більшої точності при незмінній кількості виключених параметрів.

Вперше розроблено метод прунінгу перед навчанням для моделей архітектури трансформер, який на відміну від існуючих враховує важливість механізму «уваги». Використання розробленого методу дозволяє значно збільшити точність класифікації кінцевої моделі в порівнянні з методом SNIP.

Вперше розроблено архітектуру програмного забезпечення для моделювання та дослідження методів прунінгу перед навчанням нейронних мереж, яка на відміну від існуючих дозволяє приводити матриці вагових коефіцієнтів мережі до розрідженого формату, використовуючи запропонований механізм оцінки важливості вагів.

4. Теоретичне та практичне значення результатів роботи.

Оптимізована мережа RetinaFace містить на 68% параметрів менше ніж початкова мережа при втраті точності на лише 1.4%. Вдосконалений метод дозволив зменшити втрати точності з 1.4% до 0.7% порівняно з методом SNIP при порівнянні з необрізаною моделлю, при скороченні параметрів на 68%.

Реалізація методу прунінгу для архітектури трансформер дозволила натренувати мережу з покращенням точності до 37% порівняно з методом SNIP при порівнянні з необрізаною моделлю, при скороченні кількості параметрів на 90%.

Встановлено, що результати визначення критеріїв важливості вагів, отриманих розробленим алгоритмом, можуть бути використані для підвищення швидкодії нейронних мереж від 20% до 65% шляхом використання розріджених матриць формату 2:4, в залежності від графічного процесора.

Встановлено, що додаткові виходи для сіамських нейронних мереж, призначених для встановлення схожості двох зображень, не дають приросту в швидкості сходження і точності моделі.

Результати досліджень прийняті до впровадження в Товаристві з обмеженою відповідальністю «ВОТЧЕД» (акт від 20.10.2023р.); в навчальному процесі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» (акт впровадження від 19.01.2024р.) при викладанні дисципліни «Технології розпізнавання образів та машинний зір» для

студентів освітньо-кваліфікаційного рівня «Магістр» спеціальності 122 «Комп'ютерні науки».

5. Апробація/використання результатів дисертації. Результати досліджень, викладені у дисертаційній роботі, доповідались та обговорювались на міжнародних та всеукраїнських конференціях та семінарах: XVI Міжнародна науково-практична конференція «Modern scientific research: achievements, innovations and development prospects» (11-13 вересня 2022 року, Берлін, Німеччина), VII Міжнародна науково-практична конференція «Modern problems of science, education and society» (11-13 вересня 2023 Київ, Україна), I Міжнародна науково-практична конференція «Current challenges of science and education» (18-20 вересня 2023, Берлін, Німеччина), X Міжнародна науково-практична конференція «Innovations and prospects in modern science» (25-27 вересня 2023, Стокгольм, Швеція), I міжнародна науково-практична конференція «Сучасні аспекти інженерії програмного забезпечення» (14 грудня 2023, Київ, Україна), III Міжнародна науково-технічна конференція «Системи і технології зв'язку, інформатизації та кібербезпеки: актуальні питання і тенденції розвитку» (30 листопада 2023 року Київ, Україна).

6. Дотримання принципів академічної доброчесності

За результатами науково-технічної експертизи дисертація Мельниченка А.В. визнана оригінальною роботою, яка не містить елементів фальсифікації, компіляції, фабрикації, плагіату та запозичень.

7. Перелік публікацій за темою дисертації із зазначенням особистого внеску здобувача.

За результатами дисертаційної роботи опубліковано 10 наукових публікацій, у тому числі:

- 3 статей у наукових фахових виданнях України за спеціальністю;
- 7 тезах виступів на наукових конференціях.

1. Melnychenko, A., Shaldenko, O. Evaluation of a snip pruning method for a state-of-the-art face detection model. *Computational Problems of Electrical Engineering*, 2023, Vol. 12, №1, pp. 22-27

Особистий внесок – розробка вдосконаленої моделі RetinaFace та вдосконаленого методу прунінгу;

2. Melnychenko, A., Zdor K. Incorporating attention score to improve foresight pruning on transformer models. *Computer Science and Applied Mathematics*, 2023, №2, pp.22-28

Особистий внесок – розробка методу прунінгу з використанням оцінок уваги для розрахунку критеріїв важливості вагів;

3. Melnychenko, A., Zdor, K. Efficiency of supplementary outputs in siamese neural networks. *Advanced Information Systems*, 2023, Volume 7, №3, pp. 49–53..

Особистий внесок – оцінка методу оптимізації мережі за допомогою внесення в архітектуру моделі додаткових виходів.

Наукові праці, які засвідчують апробацію матеріалів дисертації:

4. Мельниченко, А., Шалденко, О. Особливості використання прунінгу перед тренуванням нейронної мережі для детекції обличчя, XX Міжнародна науково-практична конференція молодих вчених і студентів, 25–28 квітня 2023 року, Київ, Україна

Особистий внесок – розробка вдосконаленої моделі RetinaFace та вдосконаленого методу прунінгу;

5. Melnychenko A. Evaluating SNIP pruning method on the state-of-the-art face detection model. *Modern scientific research: achievements, innovations and development prospects*, XVI Міжнародна науково-практична конференція, 11-13 вересня 2022 року, Берлін, Німеччина. С. 68-72.

Особистий внесок – розробка вдосконаленої моделі RetinaFace та вдосконаленого методу прунінгу;

6. Melnychenko, A., Zdor, K. Applying classification and regression supplementary output in siamese neural network using fashion MNIST and plantvillage datasets, VII Міжнародна науково-практична конференція “Modern problems of science, education and society”, 11-13 вересня 2023 Київ, Україна, С. 126-129.

Особистий внесок – оцінка методу оптимізації мережі за допомогою внесення в архітектуру моделі додаткових виходів;

7. Melnychenko, A., & Zdor, K. Applying classification and regression supplementary outputs in siamese neural network using plantvillage dataset, I Міжнародна науково-практична конференція “Current challenges of science and education”, 18-20 вересня 2023, Берлін, Німеччина. С. 79-82.

Особистий внесок – оцінка методу оптимізації мережі за допомогою

внесення в архітектуру моделі додаткових виходів;

8. Melnychenko A., Zdor K. Applying classification and regression supplementary output in siamese neural network using fashion MNIST and plantvillage datasets, X Міжнародна науково-практична конференція “Innovations and prospects in modern science”, 25-27 вересня 2023, Стокгольм, Швеція. С. 87-92.

Особистий внесок – оцінка методу оптимізації мережі за допомогою внесення в архітектуру моделі додаткових виходів;

9. Мельниченко А., Здор К. Збільшення ефективності оптимізації моделей архітектури ViT перед навчанням шляхом включення активацій механізму самоуваги, I міжнародна науково-практична конференція “Сучасні аспекти інженерії програмного забезпечення”, 14 грудня 2023, Київ, Україна.

Особистий внесок – розробка методу прунінгу з використанням оцінок уваги для розрахунку критеріїв важливості вагів;

10. Мельниченко А.В., Здор К.А. Врахування механізмів самоуваги при прунінгу моделей нейронних мереж Vision Transformer. Збірник матеріалів III Міжнародної науково-технічної конференції “Системи і технології зв’язку, інформатизації та кібербезпеки: актуальні питання і тенденції розвитку”, 30 листопада 2023 року, Київ, Україна. С. 214 – 215.

Особистий внесок – розробка методу прунінгу з використанням оцінок уваги для розрахунку критеріїв важливості вагів;

Якість та кількість публікацій відповідають “Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженого Постановою Кабінету Міністрів України від 12 січня 2022 р. № 44”.

ВВАЖАТИ, що дисертаційна робота Мельниченка А.В. “Методи та програмні засоби підвищення швидкодії моделей розпізнавання образів на основі машинного навчання”, що подана на здобуття ступеня доктора філософії з галузі знань 12 Інформаційні технології за спеціальністю 121 Інженерія програмного забезпечення за своїм науковим рівнем, новизною отриманих результатів, теоретичною та практичною цінністю, змістом та оформленням повністю відповідає вимогам, що пред’являють до дисертацій на здобуття ступеня доктора філософії та відповідає напрямку наукового дослідження освітньо-наукової

програми КПІ ім. Ігоря Сікорського “Інженерія програмного забезпечення” зі спеціальності 121 Інженерія програмного забезпечення.

РЕКОМЕНДУВАТИ:

1. Дисертаційну роботу “Методи та програмні засоби підвищення швидкодії моделей розпізнавання образів на основі машинного навчання”, подану Мельниченком Артемом Васильовичем на здобуття наукового ступеня доктора філософії, до захисту у разовій спеціалізованій вченій раді.

2. Вченій раді КПІ ім. Ігоря Сікорського утворити разову спеціалізовану вчену раду у складі:

Голова:

доктор технічних наук, професор, завідувач кафедри обчислювальної техніки факультету інформатики та обчислювальної техніки, Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» СТИРЕНКО Сергій Григорович;

Члени ради:

Рецензенти:

доктор технічних наук, професор, в.о. завідувача кафедри інженерії програмного забезпечення в енергетиці навчально-наукового інституту атомної та теплової енергетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» КОВАЛЬ Олександр Васильович;

кандидат технічних наук, доцент кафедри цифрових технологій в енергетиці навчально-наукового інституту атомної та теплової енергетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» ЗАЛЕВСЬКА Ольга Валеріївна;

Опоненти:

доктор технічних наук, професор, завідувач кафедри кібербезпеки та захисту даних Приватної установи «Університету науки, підприємництва та технологій» СЕМЕНОВ Сергій Геннадійович;

доктор технічних наук, професор, завідувачка кафедри Інженерії програмного забезпечення автоматизованих систем Державного університету інформаційно-телекомунікаційних технологій СТОРЧАК Каміла Павлівна.

Головуючий на засіданні

В.О. завідувача кафедри інженерії програмного забезпечення в енергетиці навчально-наукового інституту атомної та теплової енергетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

доктор технічних наук, професор

«15» березня 2024 року



Олександр КОВАЛЬ

Гарант освітньо-наукової програми

завідувачка кафедри програмного забезпечення комп'ютерних систем факультету прикладної математики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

доктор технічних наук, доцент

«15» березня 2024 року



Євгенія СУЛЕМА

Учений секретар кафедри інженерії програмного забезпечення в енергетиці навчально-наукового інституту атомної та теплової енергетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

доктор технічних наук, доцент



Андрій МУСИЄНКО

«15» березня 2024 року