

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ
ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ
ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»

Кваліфікаційна наукова
праця на правах рукопису

ТРОЧУН ЄВГЕНІЙ ВОЛОДИМИРОВИЧ

УДК 004.8

ДИСЕРТАЦІЯ

МЕТОД ГІБРИДИЗАЦІЇ КЛАСИЧНИХ ТА НЕКЛАСИЧНИХ ОБЧИСЛЕНЬ
ДЛЯ ЗАВДАНЬ ШТУЧНОГО ІНТЕЛЕКТУ

121 Інженерія програмного забезпечення

12 Інформаційні технології

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей,
результатів і текстів інших авторів мають посилання на відповідне джерело

_____ Трочун Є.В.

Науковий керівник Гордієнко Юрій Григорович, доктор фізико-математичних
наук, професор, старший науковий співробітник

Київ – 2025

АНОТАЦІЯ

Трочун Є.В. Метод гібридизації класичних та некласичних обчислень для завдань штучного інтелекту. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 121 – Інженерія програмного забезпечення з галузі знань 12 – Інформаційні технології. – Національний Технічний Університет України «Київський Політехнічний Інститут імені Ігоря Сікорського», Київ, 2025.

Дисертаційна робота присвячена розробці комплексного методу гібридизації класичних та квантових обчислень для підвищення продуктивності та ефективності роботи систем штучного інтелекту, які будуються на основі глибоких нейронних мереж.

Вперше було розроблено комплексний метод підвищення ефективності роботи глибоких нейронних мереж за допомогою використання квантових обчислень.

Досліджено метод використання квантового пристрою у вигляді одного з прихованих рівнів гібридної глибокої нейронної мережі для підвищення швидкодії моделі штучного інтелекту. Підвищення швидкодії досягається за рахунок інкапсуляції частини обчислень у квантовий пристрій, що виконується на квантовому апаратному забезпеченні, яке надає значне прискорення виконання обчислень порівняно з класичним апаратним забезпеченням. Це робить класичну частину гібридної моделі менш глибокою та відповідно зменшує кількість класичних операцій, що необхідна для роботи нейронної мережі, що відповідно збільшує швидкодію моделі.

Розроблено метод використання квантового пристрою для попередньої обробки вхідних даних для збільшення варіативності тренувального набору даних та підвищення точності роботи моделі за рахунок залучення більш різноманітних даних для тренування моделі. Відповідно до методу, квантовий пристрій використовується для квантового «зашумлення» тренувальних даних

та виконує роль кроку штучного розширення тренувального набору даних (data augmentation, або ж скорочено DA) при підготовці моделі. Це дозволяє розширити тренувальний набір даних та відповідно досягнути вищих показників точності порівняно з референтними моделями, особливо у контексті задач, де тренувальний набір даних є обмеженим.

Розроблено метод використання квантової схеми у якості першого згорткового шару глибоких згорткових нейронних мереж для підвищення точності роботи моделі штучного інтелекту. Відповідно до розробленого методу, квантовий згортковий шар створює багатоканальне представлення вхідного зображення та модель використовує глибоку згорткову мережу у якості класичної частини гібридної моделі для опрацювання багатоканального варіанту зображення, що сприяє кращому виділенню характерних рис класів та сприяє підвищенню точності роботи моделі.

Розроблено набір програмних компонентів для виконання як класичних задач класифікації зображень, що використовуються для аналізу ефективності моделей штучного інтелекту, так і показано ефективність підходів для вирішення практико-орієнтованої задачі аналізу знімків місцевості, отриманих з супутника, для виявлення ділянок уражених природною катастрофою. Для експериментів було використано гібридні моделі, побудовані на основі глибоких згорткових нейронних мереж, як простої пірамідальної архітектури, так і складніші та більш сучасні рішення, такі як ResNet50 та EfficientNet.

Проведено аналіз результатів застосування запропонованих методів гібридизації класичних та некласичних обчислень. За результатами експериментів з застосуванням запропонованих методів, було продемонстровано їх ефективність у контексті підвищення точності роботи моделей з простішими архітектурами (згорткові нейронні мережі пірамідальної архітектури) на класичних та достатньо повних наборах даних, таких як CIFAR100 (покращення точності класифікації на 2.9% порівняно з референтною моделлю VGG-16). На практико-орієнтованій задачі класифікації зображень, представленої набором даних супутникових знімків, який є менш повним, збільшення ефективності було продемонстровано гібридними моделями на

основі як простіших, так і складних архітектур (таких як EfficientNet та ResNet50). Та було досягнуто нової найсучаснішої та найвищої «state-of-the-art» (SOTA) точності 98.81% (покращення на 0.31% порівняно з референтною моделлю) за допомогою гібридної квантово-класичної моделі, побудованої на основі EfficientNet. Також було доведено можливість та практичність використання методу гібридизації, що полягає у виокремленні частини прихованих шарів глибокої нейронної мережі у квантовий пристрій. Даний підхід гібридизації показав нижчу точність роботи моделі, досягнувши точності у 72.75% (зниження точності на 3.25% відносно референтної моделі) на наборі даних супутникових знімків. Проте це зниження точності може бути несуттєвим у ряді задач відносно пришвидшення операцій за рахунок перенесення частини обчислень на квантове апаратне забезпечення.

Розроблені методи є складовими комплексного методу гібридизації класичних та некласичних обчислень та дозволяють підготувати гібридні моделі на основі класичних нейронних мереж, залежно від типу задачі та вимог з підвищення швидкодії чи точності роботи моделей при вирішенні ряду різноманітних практичних задач.

Ключові слова: гібридні нейронні мережі, квантові обчислення, глибокі нейронні мережі, машинне навчання, класифікація зображень.

ABSTRACT

Trochun Y.V. Method of classical and non-classical computations hybridization for artificial intelligence. – Qualified scientific work on the rights of the manuscript.

Dissertation for the degree of Doctor of Philosophy in the specialty 121 – Software Engineering and 12 – Information Technologies. – National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, 2025.

The dissertation focuses on the development of a comprehensive method for hybridizing classical and quantum computations to enhance the performance and efficiency of artificial intelligence systems based on deep neural networks.

A comprehensive method to improve the efficiency of deep neural networks through quantum computing utilization was developed.

A method that involves a quantum device as one of the hidden layers in a hybrid deep neural network to enhance the speed of artificial intelligence models was investigated and evaluated. The acceleration is achieved by encapsulating part of the computations within the quantum device, which is executed on quantum hardware that provides significant computational speed compared to classical hardware. This reduces the depth of the classical portion of the hybrid model, thereby decreasing the number of classical operations required, which, in turn, increases the model's speed.

A method was also developed that uses a quantum device for preprocessing input data, increasing the variability of the training dataset and improving the model's accuracy by incorporating more diverse data. This involves utilizing the quantum device for quantum noise injection in training data, effectively serving as a "data augmentation" step during model preparation. This expands the training dataset and results in higher accuracy compared to reference models, particularly for tasks with limited training data.

Additionally, a method was designed to use quantum circuits as the first convolutional layer in deep convolutional neural networks to improve artificial intelligence model accuracy. This quantum convolutional layer creates multi-channel representations of input images, enabling the classical convolutional network in the

hybrid model to process these enriched representations, which improves feature extraction and boosts model accuracy.

A set of software components was developed to handle classical image classification tasks for benchmarking artificial intelligence models and to demonstrate the practical effectiveness of these approaches. These methods were applied to satellite imagery analysis for identifying areas affected by natural disasters. Hybrid models based on deep convolutional neural networks — ranging from simple pyramidal architectures to more advanced solutions like ResNet50 and EfficientNet — were used in experiments.

An analysis of the proposed methods of hybridizing classical and quantum computations demonstrated their effectiveness. Experiments showed an accuracy improvement of 2.9% on the CIFAR100 dataset using simpler architectures like pyramidal convolutional neural networks, compared to the reference VGG-16 model. For satellite imagery classification with a less complete dataset, hybrid models with both simpler and more complex architectures (such as EfficientNet and ResNet50) showed improvements, achieving a new state-of-the-art accuracy of 98.81% — an increase of 0.31% over the reference model — with a quantum-classical hybrid model based on EfficientNet.

The practicality of hybridizing neural network layers with quantum devices was also demonstrated. Although this method resulted in slightly lower accuracy (72.75%, a 3.25% decrease compared to the reference model) for satellite imagery classification, the trade-off for faster operations through quantum hardware could be valuable for specific tasks that value computation speed higher than classification accuracy.

These methods are components of a comprehensive approach to hybridizing classical and quantum computations, enabling the preparation of hybrid models based on classical neural networks tailored to task requirements for speed or accuracy improvements across various practical applications.

Keywords: hybrid neural networks, quantum computing, deep neural networks, machine learning, image classification.

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА

Наукові праці, в яких опубліковано основні наукові результати дисертації:

1. **Y. Trochun**, E. Pavlov, S. Stirenko and Y. Gordienko, “Impact of Hybrid Neural Network Structure on Performance of Multiclass Classification,” IEEE EUROCON 2021 - 19th International Conference on Smart Technologies, Lviv, Ukraine, 2021, pp. 152-156. IEEE. ISBN 978-1-6654-3299-3. DOI 10.1109/EUROCON52738.2021.9535586 (Web of Science).
2. **Y. Trochun** et al., “Hurricane Damage Detection by Classic and Hybrid Classic-Quantum Neural Networks,” 2021 International Conference on Space-Air-Ground Computing (SAGC), Huizhou, China, 2021, pp. 152-156. IEEE. ISBN: 978-1-6654-0012-1. DOI 10.1109/SAGC52752.2021.00033 (Scopus).
3. **Y. Trochun**, S. Stirenko, O. Rokovyi, O. Alienin, E. Pavlov and Y. Gordienko, “Hybrid Classic-Quantum Neural Networks for Image Classification,” 2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Cracow, Poland, 2021, pp. 968-972. IEEE. ISSN: 2770-4254. DOI 10.1109/IDAACS53288.2021.9661011 (Scopus).
4. **Y. Trochun**, S. Stirenko and Y. Gordienko, “Multichannel Quantum Data Preprocessing for Image Classification Problems”, Springer Book Series, Algorithms for Intelligent Systems, ISSN: 2524-7565 (Web of Science).
5. Yuri Gordienko, **Yevhenii Trochun** and Sergii Stirenko, “Multimodal Quantvolutional and Convolutional Neural Networks for Multi-class Image Classification”. Big Data and Cognitive Computing. 2024, 8(7), 75, ISSN: 2504-2289. DOI 10.3390/bdcc8070075 (Scopus Q2).
6. **Y. Trochun** and Y. Gordienko, “Effectiveness of Hybrid Quantum-Classical and Quantvolutional Neural Networks for image classification”, Inf. Comput. and Intell. syst. j., no. 5, pp. 68–79, Dec. 2024. ISSN 2708-

4930. DOI: 10.20535/2786-8729.5.2024.318743 (Фахове видання категорії «Б»).

7. Y. Gordienko, **Y. Trochun**, V. Taran, A. Khmelnytskyi, S. Stirenko, “HNN-QCn: Hybrid Neural Network with Multiple Backbones and Quantum Transformation as Data Augmentation Technique”. AI 2025, 6, 36, ISSN: 2673-2688. DOI 10.3390/ai6020036 (Scopus Q2).

ЗМІСТ

АНОТАЦІЯ.....	2
ABSTRACT.....	5
СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА	7
ЗМІСТ	9
Перелік умовних позначень.....	11
Вступ	12
Розділ 1 Технології штучного інтелекту та квантові обчислення	20
1.1. Квантові обчислення	20
1.1.1. Сучасний стан квантового апаратного забезпечення.....	20
1.1.2. Практичне застосування квантових обчислень	22
1.2. Гібридні квантово-класичні нейронні мережі	24
1.3. Сучасний стан використання квантових комп'ютерів.....	27
Висновки до розділу 1	32
Розділ 2 Особливості гібридизації класичних та некласичних обчислень	34
2.1. Вибір задачі	34
2.2. Вибір наборів даних для експериментів.....	35
2.2.1. Набір даних MNIST	36
2.2.2. Набір даних MNIST-Fashion	37
2.2.3. Набори даних CIFAR10 та CIFAR100	39
2.2.4. Набір даних SIHD	42
2.3. Визначення показників ефективності та метрик для експериментів	44
2.4. Огляд доступних архітектур квантового пристрою	46
2.5. Вибір інструментів розробки та оточення	52
Висновки до розділу 2	56
Розділ 3 Підходи побудови гібридних квантово-класичних нейронних мереж	58
3.1. Вибір алгоритмів побудови архітектур гібридних квантово-класичних нейронних мереж	58
3.2. Квантовий пристрій як один з прихованих шарів нейронної мережі.....	59

	10
3.3. Попередня квантова обробка даних.....	63
3.4. Квантовий пристрій як перший згортковий шар нейронної мережі	68
3.5. Багатоструктурна класична частина гібридної квантово-згорткової нейронної мережі	72
3.6. Комплексний метод гібридизації класичних та квантових обчислень для завдань штучного інтелекту.....	75
Висновки до розділу 3	78
Розділ 4 Аналіз результатів методів гібридизації класичних та неklasичних обчислень для побудови гібридних нейронних мереж	80
4.1. Аналіз впливу архітектури квантового пристрою на ефективність роботи гібридних квантово-класичних нейронних мереж	80
4.2. Аналіз методу гібридизації нейронних мереж, використовуючи квантовий пристрій як один з прихованих шарів нейронної мережі.....	86
4.3. Аналіз результатів попередньої квантової обробки даних.....	91
4.4. Аналіз методу гібридизації нейронних мереж, використовуючи квантовий пристрій як перший рівень нейронної мережі	101
4.5. Аналіз ефективності використання багатоструктурної класичної частини гібридної квантово-згорткової моделі	116
4.6. Порівняльний аналіз ефективності розглянутих методів гібридизації нейронних мереж	124
Висновки до розділу 4	131
Висновки.....	135
Список використаних джерел.....	137
Додаток А	150

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

ОКТ – Оптична Когерентна Томографія

ШІ – Штучний Інтелект

AUC – Area Under the Curve

AWS – Amazon Web Services

CIFAR – Canadian Institute For Advanced Research

CNN – Convolutional Neural Network

DA – Data Augmentation

GPT – Generative Pre-trained Transformer

GPU – Graphic Processing Unit

HNN – Hybrid Neural Network

HNN-QCn – HNN з багатоструктурною класичною частиною

LLM – Large Language Model

MNIST – Modified National Institute of Standards and Technology database

MV – Majority Voting

NN – Neural Network

QAOA – Quantum Approximation Optimization Algorithm

QC – Quantum Circuit

RGB – Red Green Blue

ROC – Receiver Operating Characteristic

SIHT – Satellite Imagery of Hurricane Damage

SOTA – State of the Art

TPU – Tensor Processing Unit

VGG – Visual Geometry Group

ВСТУП

Актуальність теми. Квантові обчислення є однією з найбільш актуальних тем досліджень останніх років, і ця галузь є надзвичайно цікавою з точки зору можливостей, які нові досягнення в ній можуть принести в різні сфери науки, такі як штучний інтелект, фізика, хімія тощо [1,2]. Досягнення в сфері квантових комп'ютерів обіцяють принести багато інновацій і революційних змін у багато різних сфер життя та досліджень. Однією з таких областей, яка може значно виграти від цього, є штучний інтелект (ШІ) і, зокрема, квантове машинне навчання. У контексті ШІ квантові комп'ютери можуть бути використані для підвищення ефективності операцій машинного навчання. Це реалізується у формі алгоритмів квантового машинного навчання, які дозволяють побудувати нейронну мережу, яка повністю працює на квантовому обладнанні [3,4], та гібридних квантово-класичних нейронних мереж, де частина моделі працює на класичному обладнанні, а інша частина виконується на квантовому обладнанні. Оскільки квантові комп'ютери здатні виконувати певні обчислення значно швидше і, крім того, є більш енергоефективними порівняно з класичними, їх використання для задач ШІ обіцяє бути дуже перспективним у майбутньому.

В сьогоденні, системи на базі ШІ широко використовуються у всіх сферах життя та часту виконують задачі критичної важливості, які є високочутливими до точності та ефективності їх роботи. Галузь ШІ є дуже динамічною та конкурентною сферою як з точки зору наукових досліджень, так і комерційних розробок.

Проте, хоч галузь ШІ є високорозвиненою та насиченою різноманітними ідеями та дослідженнями сферою, існує цілий ряд істотних проблем та задач в цій сфері. Над їх рішенням активно кипить робота та проводяться різноманітні дослідження. Є кілька основних груп проблем, що дотичні до тематики дисертаційного дослідження.

Першою з таких проблем є швидкодія систем ШІ, що базуються на нейронних мережах, адже нейронні мережі вимагають просто колосальної

кількості обчислень для їх роботи та з часом розмір архітектур нейронних мереж та їх складність тільки росте, що призводить до швидкого збільшення вимог з погляду апаратного забезпечення. Ця проблема стоїть надзвичайно гостро, особливо для задач, що вимагають опрацювання даних у наближеному до реального часу, зважаючи на те, що затримки в роботі критично важливих систем можуть призвести до серйозних наслідків.

Іншою важливою та дуже актуальною проблемою є високе споживання електроенергії нейронними мережами. Для прикладу, відома велика мовна модель GPT-3 має близько 175 мільярдів параметрів [5], в той час як її «наступник» GPT-4 – вже близько 1.7 трильйонів параметрів [6] та, відповідно, споживає надзвичайно багато енергії, щоб продукувати результат. Ця проблема також пов'язана з розвитком сфери ШІ та постійним збільшенням складності нових архітектур нейронних мереж. Наразі дана проблема має свій негативний вплив не лише на критично важливі системи, але і на будь-які інші системи на основі нейронних мереж, адже збільшення споживання енергії також означає як і значне зростання вартості кожного запиту до нейронної мережі, так і збільшення об'єму парникових газів, що продукуються в атмосферу в наслідок роботи центрів обробки даних, так званий карбоновий слід (carbon footprint) [7].

Ці обидві проблеми виникають через зростаючу складність найсучасніших архітектур моделей, які вимагають значних обчислювальних ресурсів. Крім того, проблема енергоефективності, ймовірно, стане більш гострою в майбутньому, враховуючи поширення застосувань систем на основі ШІ у різних сферах сучасного життя. Розширення використання ШІ для чутливих до часу застосунків підкреслює потенціал квантових обчислень для прискорення обчислень, що підвищить продуктивність критичних систем, які залежать від рішень на основі ШІ. Наприклад, аналіз супутникових зображень, таких як виявлення територій, постраждалих від природних чи антропогенних катастроф, значною мірою залежить від завдань сегментації та класифікації зображень, що виконуються за допомогою алгоритмів, що базуються на машинному навчанні. Такі застосування можуть суттєво виграти від прискореного оброблення та підвищеної ефективності моделей. Тому квантове машинне навчання і гібридні

квантово-класичні нейронні мережі вважаються дуже перспективними завдяки їхній вищій енергоефективності та значному прискоренню обчислень, що сприятиме розробці нових і більш екологічних рішень.

Гібридні квантово-класичні нейронні мережі дозволяють перенести частину обчислень нейронної мережі на квантовий пристрій, який працює на квантовому обладнанні. Це означає, що найскладніша частина нейронної мережі може високоефективно працювати за допомогою квантового комп'ютера, тоді як усі інші частини мережі обчислюються на класичному обладнанні. Квантово-класична гібридизація навантажень є багатообіцяючою в контексті останніх досягнень у гібридних системах, які використовують як CPU (Central Processing Unit, тобто центральний процесор), так і різні спеціалізовані прискорювачі, такі як GPU (Graphical Processing Unit, тобто графічні прискорювачі) та TPU (Tensor Processing Unit, тобто прискорювачі тензорних обчислень). Здійсненість та ефективність підходу гібридизації CPU-GPU-TPU була доведена в різних областях, що може вказувати на те, що квантово-класична гібридизація також може бути плідною в майбутньому [8, 9].

Зв'язок теми з науковими програмами, планами, темами. Запропоновані у дисертації методи були використані у науково-дослідній роботі за грантом Національного Фонду Досліджень України «Наука для відбудови України у воєнний та повоєнний періоди» (<https://nrfu.org.ua>), реєстраційний номер гранту: 2022.01/0199. Отримано запрошення до участі у міжнародному консорціумі COST Action OC-2024-1-27129 “Quantum Artificial Intelligence” (<https://www.cost.eu>), подано спільну заявку.

Мета і завдання дослідження. Метою дисертаційного дослідження є підвищення ефективності роботи систем ІІІ шляхом створення комплексного методу гібридизації нейронних мереж для використання квантового апаратного забезпечення для підвищення швидкодії або точності роботи моделей ІІІ.

Для досягнення мети було виділено наступні завдання:

- Провести огляд сучасних підходів гібридизації класичних та квантових обчислень для побудови гібридних моделей ІІІ.

- Виконати пошук та аналіз наборів даних та задач для аналізу ефективності роботи гібридних моделей ШІ.
- Розробити підхід гібридизації, що дозволить підвищити точність роботи моделей ШІ в контексті практичних задач, що мають недостатньо широкий тренувальний набір даних, на основі використання квантових обчислень.
- Розробити комплексний підхід гібридизації класичних та квантових обчислень для ефективного виконання обраних задач.
- Продемонструвати практичність використання гібридних систем ШІ для вирішення практичних задач.
- Виділити переваги та недоліки розроблених та розглянутих методів гібридизації класичних та квантових обчислень.

Об'єкт дослідження – способи гібридизації класичних та квантових обчислень для побудови глибоких гібридних квантово-класичних нейронних мереж.

Предмет дослідження – комплексні методи гібридизації класичних та квантових обчислень для побудови глибоких гібридних квантово-класичних нейронних мереж.

Методи дослідження. Методичною основою дослідження є аналіз методів та підходів застосування некласичних квантових обчислень для побудови гібридних квантово-класичних моделей на основі глибоких нейронних мереж та методів побудови гібридних квантово-класичних нейронних мереж для підвищення ефективності роботи моделей ШІ. В дослідженні було використано методи машинного навчання, трансферного навчання, техніки попередньої обробки даних, перехресної перевірки (cross-validation) метрик роботи моделей, метрики точності роботи моделей, техніка голосування більшістю (Majority Voting) та симулятори квантових обчислень.

Наукова новизна отриманих результатів:

- Вперше запропоновано комплексний метод гібридизації класичних та квантових обчислень, який на відміну від існуючих може бути

направлений на збільшення швидкодії моделей ШІ або на збільшення точності роботи моделей залежно від вимог задачі за рахунок залучення квантових обчислень у різних частинах гібридних нейронних мереж (у якості квантово-згорткового (quanvolutional) шару або ж у якості заміни частини прихованих повнозв'язних шарів глибокої нейронної мережі).

- Вперше запропоновано метод побудови гібридних систем ШІ, який на відміну від існуючих дозволяє залучати незмінні класичні SOTA моделі та дозволяє досягнути вищих значень точності роботи моделей порівняно з альтернативними підходами за рахунок залучення квантових обчислень у якості першого квантово-згорткового (quanvolutional) шару гібридних моделей.
- Набув подальшого розвитку метод використання трансферного навчання у контексті гібридних квантово-класичних нейронних мереж, що дозволяє зменшити обсяг ресурсів (час обчислень, енергія), потрібних для підготовки моделі та підвищити точність роботи гібридних квантово-класичних моделей ШІ.
- Набув подальшого розвитку метод використання багатоканальності (multi-channel) даних для вирішення задачі класифікації зображень, що дозволяє підвищити точність роботи гібридної моделі ШІ за рахунок використання квантового пристрою для створення додаткових каналів на основі вхідних даних.
- Набув подальшого розвитку метод штучного розширення різноманітності даних (data augmentation) для вирішення задачі класифікації зображень, що дозволяє підвищити стійкість по відношенню до невідомих даних (generalization) гібридної моделі ШІ за рахунок використання квантово-згорткового штучного розширення різноманітності даних (quanvolutional data augmentation) і використання результатів роботи квантового пристрою в якості додаткових каналів вхідних даних.
- Набув подальшого розвитку метод використання багатоструктурної (multi-backbone) моделі у якості класичної частини гібридної квантово-

згорткової (quanvolutional) моделі для вирішення задачі класифікації зображень, що дозволяє підвищити точність роботи гібридної моделі ШІ за рахунок незалежного паралельного опрацювання додаткових каналів, що були створені за допомогою квантового пристрою, спеціалізованими базовими моделями, що входять до складу багатоструктурної класичної частини моделі.

- Результати було продемонстровано на прикладі практичної задачі класифікації пошкоджень в зонах стихійних лих та в контексті використання систем ШІ загального застосування, особливо для задач, що вимагають надвисокої швидкодії.

Практичне значення отриманих результатів. Результати, отримані в ході дослідження, дозволяють застосовувати гібридні квантово-класичні моделі ШІ для вирішення різноманітних практичних задач, що мають підвищені вимоги до швидкодії алгоритмів або ж підвищені вимоги до точності роботи систем ШІ, особливо в контексті задач, що мають обмежений набір даних, які можна використати для тренування моделі. Розроблений метод пропонує два альтернативних підходи гібридизації нейронних мереж, що можуть бути використані залежно від вимог задачі. Перший з підходів полягає в зменшенні глибини класичної частини нейронної мережі за рахунок її заміни квантовим пристроєм, що надає приріст швидкодії системи. Другим підходом є створення додаткових каналів вхідних даних за допомогою квантового пристрою та використання SOTA класичних моделей для опрацювання додаткових каналів вхідних даних та залучає підхід «majority voting» (MV) для визначення фінального вердикту моделі на основі опрацювання додаткових каналів, створених квантовими перетвореннями.

Особистий внесок здобувача. Дисертація є результатом самостійних наукових досліджень, в результаті яких було розроблено авторський підхід гібридизації класичних та квантових обчислень для завдань ШІ. Основні результати та наукові положення, отримані в результаті проведення дисертаційного дослідження, були отримані здобувачем самостійно у процесі

науково-дослідної роботи. В роботах, опублікованих у співавторстві, дисертанту належить: [1] – аналіз впливу архітектури квантового пристрою на точність роботи гібридних нейронних мереж, побудованих на його основі; [2] – аналіз ефективності роботи підходу гібридизації класичних та квантових обчислень, що полягає у заміні частини прихованих шарів класичної нейронної мережі на квантовий пристрій, для вирішення практичної задачі на прикладі задачі аналізу супутникових зображень; [3] – аналіз ефективності роботи підходу гібридизації класичних та квантових обчислень, що полягає у заміні частини прихованих шарів класичної нейронної мережі на квантовий пристрій, для вирішення класичних проблем класифікації зображень; [4] – аналіз ефективності використання квантових обчислень для виконання кроку попередньої обробки вхідних даних, аналіз ефективності використання додаткових каналів вхідних даних, що створені на основі квантових обчислень; [5] – розроблений метод гібридизації класичних та квантових обчислень на основі використання квантового пристрою у якості першого згорткового шару гібридної згорткової глибокої нейронної мережі; [6] – аналіз ефективності розглянутих у дисертаційному дослідженні методів гібридизації класичних та квантових обчислень; [7] – аналіз ефективності використання підходу залучення багатоструктурної архітектури класичної частини гібридної квантово-згорткової глибокої нейронної мережі на прикладі набору даних CIFAR10.

Апробація результатів дисертації. Основні результати дисертаційного дослідження було представлено, обговорено та опубліковано на міжнародних наукових конференціях, а саме на: 19th International Conference on Smart Technologies, EUROCON 2021 (м. Львів, 2021 р.); International Conference on Space-Air-Ground Computing, SAGC 2021 (м. Хуейчжоу, Китай, 2021 р.); 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2021 (м. Краків, Польща, 2021 р.); The International Conference on Security, Fault Tolerance, Intelligence, ICSFTI 2021 (м. Київ, 2021 р.).

Публікації. За результатами дослідження опубліковано 7 наукових статей, що індексовані до наукометричних баз даних з міжнародним індексом цитування: Scopus – 5, Web of Science – 4.

Структура та обсяг роботи. Дисертаційна робота складається зі вступу, чотирьох розділів, загальних висновків, списку використаних джерел із 117 найменувань та додатків. Загальний обсяг дисертації становить 151 сторінок, з яких 149 сторінок основного тексту, 1 додаток на 2 сторінках. Дисертація містить 55 рисунків, 12 формул та 7 таблиць.

РОЗДІЛ 1

ТЕХНОЛОГІЇ ШТУЧНОГО ІНТЕЛЕКТУ ТА КВАНТОВІ ОБЧИСЛЕННЯ

1.1. Квантові обчислення

Квантові обчислення дозволяють використати феномен квантової механіки для виконання певних обчислень експоненційно швидше та ефективніше ніж того дозволяють класичні комп'ютери, та навіть більше – квантові обчислення дозволяють вирішувати задачі, які є невирішуваними за допомогою класичних комп'ютерів [1, 2]. Квантові комп'ютери будуються на основі кубітів – квантових бітів. В той час як звичайний «класичний» біт може перебувати лише в двох станах – 0 та 1 в кожен окремий момент часу, кубіт завжди перебуває в суперпозиції [3] обох станів. При замірі стану кубіту результатом заміру є ймовірність перебування кубіту в тому чи іншому стані, тобто значення отримане у результаті заміру стану кубіта є будь-яким раціональним числом в проміжку $[0; 1]$. Квантовий комп'ютер може впливати на стан кубіту за допомогою інтерференції хвиль та зміщувати результат заміру стану кубіту до заданого бажаного значення. Існують алгоритми для квантових комп'ютерів, які дозволяють використовувати дані особливості кубітів для надзвичайно швидкого та ефективного виконання певних задач.

1.1.1. Сучасний стан квантового апаратного забезпечення

Наразі виготовлення стабільних та якісних кубітів є надзвичайно складною задачею. Найскладнішим та найважливішим етапом виготовлення кубіту є його ізоляція від зовнішнього світу, адже якщо кубіт не є достатньо ізольованим, результати обчислень, що їх проводить кубіт, стають зашумлені.

Наразі активно проводяться дослідження, які спрямовані на пошук методів створення більш стабільних кубітів з нижчим рівнем помилок, які будуть

перебувати в когерентному стані довший час, що дозволить виготовляти багатоцільові квантові комп'ютери [10].

Однією з найперспективніших технологій виготовлення кубітів є іонні пастки, що дозволяють виготовляти кубіти на основі іонів, які утримуються електромагнітним полем [11]. Результати експериментів показали, що з сучасним рівнем розвитку обладнання та теорії можливо створити квантову схему, що містить 32 кубіти на основі іонів [12].

Іншою надзвичайно цікавою та перспективною є технологія виготовлення кубітів на основі надпровідників [13]. Надпровідники мають близько нульовий опір при неймовірно низьких температурах. Для виготовлення кубітів найчастіше використовують танталум та ніобілум [14]. Дану технологію наразі широко використовують техно-гіганти, такі як Google, IBM, Intel, Rigetti, що працюють над дослідженнями та розробкою квантових комп'ютерів. Більшість з сучасних версій QPU – квантових процесорів – працюють саме за допомогою кубітів, виготовлених за допомогою надпровідників. Найпотужніша квантова схема, побудована за допомогою даної технології, налічує 53 кубіти [15].

Найбільшою перешкодою до створення багатоцільових квантових комп'ютерів наразі є відсутність механізму автоматичної корекції помилок [16,17]. Даний механізм надзвичайно важливий і просто необхідний для проведення більш складних обчислень, що вимагають часу більшого ніж час втрати когерентності [18] кубітом [17]. Алгоритм квантової корекції помилок також повинен бути достатньо швидким, щоб мати змогу виправляти помилки та прибирати шум швидше, ніж його вносить процес втрати когерентності кубітів. І оскільки він теж вимагатиме квантового апаратного забезпечення для своєї роботи, з огляду на надзвичайно високі вимоги зі швидкодії, він також значно збільшує вимоги до самого квантового комп'ютера, збільшуючи необхідну для виконання операцій кількість кубітів. Отже, розробка ефективного та практичного механізму квантової корекції помилок є найбільшою перешкодою для створення багатоцільових та масштабованих квантових комп'ютерів [19].

1.1.2. Практичне застосування квантових обчислень

В наш час квантові комп'ютери є більше теоретичною сферою знань, ніж практичним інструментом. Квантове апаратне забезпечення має велику кількість невирішених на момент написання даної роботи проблем та ця сфера перебуває лише на перших початкових етапах розвитку технології. Проте навіть зараз за таких обставин дослідники вже використовують квантові обчислення для виконання певних вузькоспеціалізованих та важливих задач. Наразі квантові комп'ютери використовуються швидше як прискорювачі для специфічних задач, які квантова механіка допомагає вирішити, ніж звичайні класичні комп'ютери широкого профілю.

Основною сферою, де наразі на практиці використовуються квантові обчислення, є різноманітні симуляції квантових процесів [20]. Симуляції квантових процесів є не лише дослідницьким інтересом, але і мають важливі та практичні застосування. Квантові симуляції використовуються в таких сферах як хімія, матеріалознавство та фармакологія.

В 2022 році Джошуа Гоінгс продемонстрував потенційне практичне застосування квантових обчислень для аналізу та дослідження ензиму «cytochrome P450» [21]. Даний ензим надзвичайно важливий, адже відповідає за засвоєння близько 70% фармакологічних препаратів. Процес, у якому бере участь даний ензим, за своєю природою, є квантовим і це робить його надзвичайно складним для симуляції за допомогою класичного апаратного забезпечення.

Іншим важливим застосуванням квантових обчислень є моделювання різноманітних хімічних реакцій, як наприклад, реакції, що лежать в основі іншої новітньої та надзвичайно перспективної технології – секвестрації вуглецю [22]. Наразі проводяться дослідження з пошуку ефективних каталізаторів для перетворення вуглецю з повітря в твердий осад, який можна в стабільному стані зберігати протягом тривалого часу та таким чином суттєво зменшити об'єм парникових газів в атмосфері та позитивно вплинути на проблеми з кліматом, що

виникли внаслідок діяльності людини. Процес пошуку потрібних для даного процесу каталізаторів може бути значно пришвидшено за допомогою квантового апаратного забезпечення, яке зможе більш точно та швидко симулювати реакції вуглецю з різноманітними елементами.

Окрім секвестрації вуглецю, іншою важливою задачею в сфері хімії, в якій квантові комп'ютери будуть корисні, є пошук ефективного способу фіксації нітрогену для виробництва добрив, який не вимагатиме високої температури та тиску. Ця технологія обіцяє революціонізувати процес виробництва сільськогосподарських добрив, значно спростивши як сам процес їх виробництва, так і відповідно значно знизити їх ціну та підвищити доступність. Дані реакції неможливо точно та ефективно симулювати на класичних комп'ютерах. Симуляція, а отже і подальша розробка самої технології можлива за допомогою гібридної квантово-класичної системи [23].

Також квантові симуляції можуть бути надзвичайно важливими у сфері енергетики та екології. Наразі найбільш поширений тип батарей для усієї електроніки – літій-іонні батареї. Вони використовують кобальт для виготовлення катодів. Видобування кобальту має надзвичайно негативний вплив на довкілля. Альтернативою кобальту для катодів може бути оксид літію нікелю, проте за звичайних умов він дуже нестабільний та його структура слабо досліджена на даний момент. Проте симуляції даної сполуки за допомогою квантового комп'ютера зможуть допомогти вивчити його краще та розробити методи для виготовлення стабільної версії даної сполуки [24]. Це дозволить позбутись від використання кобальту та замінити його менш екологічно шкідливим матеріалом та відповідно позитивно вплинути на стан довкілля в країнах, які зараз активно залучені у виробництво літій-іонних батарей та їх компонентів.

Проте наразі, ймовірно, найбільш практичним використанням сучасних квантових комп'ютерів є робота з квантовими сенсорами [25]. Квантові сенсори це наразі найбільш практичне використання квантових обчислень навіть з теперішнім станом квантового апаратного забезпечення. Квантові сенсори це сучасний SOTA в сфері високоточних сенсорів [26], їх наразі використовують

для виявлення та заміру гравітаційних хвиль для різноманітних астрономічних та фізичних досліджень, до прикладу в LIGO [27]. Окрім детекторів гравітаційних хвиль, ведуться дослідження з розробки інших квантових сенсорів для різноманітних задач: від задач сканування мозку [28] до розробки високоточних квантових радарів для точнішої навігації [29]. Квантові сенсори при проведенні заміру заміряють стан квантових частинок, які лежать в основі роботи сенсору, та перетворюють квантовий стан в класичний результат заміру, що вимагає великої кількості операцій та займає певний час. Квантові комп'ютери спроможні експоненційно оптимізувати даний процес за рахунок того, що зникає потреба перетворення квантового стану в класичний, що в свою чергу може суттєво прискорити весь процес та зробити квантові сенсори ще більш точними та швидкими [30].

1.2. Гібридні квантово-класичні нейронні мережі

Загалом, квантове машинне навчання і гібридні квантово-класичні нейронні мережі є перспективними та актуальними напрямками досліджень. Оскільки квантові обчислення на даний момент це дуже молода технологія та вона буквально знаходиться на зародкових етапах, їх застосування в області машинного навчання є теоретичним і знаходиться на етапі досліджень, і не має наразі практичного впровадження. Однак низка останніх досліджень свідчить про можливість застосування гібридних нейронних мереж для вирішення як теоретичних, так і практичних задач у різноманітних галузях досліджень.

Медичні застосування ШІ є однією з найважливіших сфер досліджень, оскільки вони спрямовані на прями покращення стану пацієнтів, допомагаючи лікарям аналізувати результати різноманітних досліджень та обстежень. У цій області проводиться велика кількість досліджень, що зокрема і залучають гібридні методи, такі як дослідження Месута Тогакара [31] щодо виявлення захворювань сітківки на основі зображень, отриманих за допомогою оптичної когерентної томографії (ОКТ). Медицина це одна з галузей, де гібридні

квантово-класичні нейронні мережі виявилися життєздатною альтернативою класичним моделям у виявленні захворювань на основі результатів обстеження пацієнтів. Одним з найновіших успішних прикладів є дослідження Агал'ї Р.К. [32] з виявлення ревматоїдного артрити на основі теплових зображень рук пацієнтів. Це дослідження показало, що точність роботи гібридних квантово-класичних нейронних мереж є достатньо високою для виконання даної задачі та вони є ефективним інструментом. Інші дослідження в медичній галузі, проведені Уллахом [33], щодо класифікації ішемічної кардіопатії та дослідження Юмін Дона [34] щодо класифікації пухлин головного мозку на основі результатів МРТ-сканів голови показали, що гібридні квантово-класичні нейронні мережі можуть досягати навіть вищої точності, ніж їхні класичні аналоги, що доводить можливість застосування гібридних систем ШІ для вирішення практичних завдань класифікації медичних зображень.

Іншими галузями досліджень, де квантово-класичні нейронні мережі довели свою ефективність, є хімія та біологія. Дослідження, проведене Грегорі В. Кіро [35], показало, що гібридні нейронні мережі можуть досягати SOTA точності у передбаченні спорідненості зв'язування білка з лігандом.

Алгоритми та моделі з залученням квантового машинного навчання для обробки зображень значно розвинулись та просунулись за останні роки [36]. Дослідники працюють над розкриттям потенціалу квантових обчислень для удосконалення традиційних методів глибокого навчання. Останні дослідження в цьому напрямі включають алгоритми квантового виділення ознак [37], квантову обробку зображень [36], квантово-натхненні згорткові нейронні мережі [38], що не будуються на квантовому апаратному забезпеченні, проте використовують ідеї та підходи, які були натхненні квантовими обчисленнями, та гібридні квантово-класичні архітектури [39, 40].

Квантове виділення ознак є процесом виділення характерних ознак або характеристик з квантових наборів даних за допомогою технік, що залучають квантові обчислення. Квантове виділення ознак передбачає використання принципів квантової механіки для аналізу та ідентифікації шаблонів або властивостей у квантових даних, які потім можна використовувати для різних

задач машинного навчання, наприклад для подальшої класифікації даних, їх кластеризації або зменшення розмірності сету даних без втрати їх ключових характеристик. Алгоритми квантового виділення ознак базуються на використанні можливостей квантових обчислень для підвищення ефективності та дієвості алгоритмів виділення ознак, особливо в сценаріях, де класичні методи можуть бути недоцільними або неефективними, зважаючи на величину або складність наборів даних, це особливо актуально при роботі з квантовими наборами даних [36].

Дослідження в сфері квантової обробки зображень зосереджується на використанні квантових обчислень для кодування, маніпулювання та відтворення квантових зображень у різних форматах для різноманітних задач [37]. Активно ведуться дослідження націлені на розробку схем кодування, які б використовували властивості квантової механіки квантового апаратного забезпечення.

Квантово-натхненні згорткові нейронні мережі використовують уявні подання та операції, де вхідний простір реальних чисел спочатку перетворюється на комплексний простір. У цьому комплексному просторі відбуваються маніпуляції параметрів за допомогою операцій, натхнених принципами квантових обчислень. Останні дослідження підкреслюють переваги використання комплексних параметрів для глибокого машинного навчання, що дозволяє досягти переваг у обчислювальній, біологічній сферах та сфері обробки сигналів [38]. Комплексні числа надають можливості покращеного представлення та відповідно потенційно полегшують оптимізацію даних, та пропонують ефективніше узагальнення та пришвидшення процесу навчання моделі.

Гібридні квантово-класичні архітектури [39, 40] набирають популярності в дослідженнях, направлених на ефективне виділення високорівневих ознак із даних, як приклад з зображень, для задач класифікаційних даних. Наприклад, у багатьох варіаціях гібридних квантово-класичних згорткових нейронних мереж квантова частина є параметризованою схемою для виділення важливих ознак зображень, а класична частина відповідає за класифікацію, базуючись на

ознаках, отриманих з квантової частини мережі. Крім того, існують і моделі квантово-класичних згорткових мереж, що використовують техніку кодування амплітуди для завдань класифікації зображень, що вимагає менше кубітів, порівнюючи з використанням кодування на основі обчислень [39]. Інша група архітектур квантово-класичних згорткових нейронних мереж – неглибокі гібридні квантово-класичні згорткові нейронні мережі. Даний підхід базується на доповненій варіаційній квантовій схемі та демонструє помітну ефективність у завданнях класифікації зображень [40]. Цей підхід використовує згортковий шар як вхідний шар нейронної мережі та спеціально розроблену варіаційну квантову схему у якості прихованого шару мережі, таким чином, зменшуючи глибину та складність моделі.

Останнім часом квантові згорткові нейронні мережі стали об'єктом великого дослідницького інтересу та багато нещодавніх досліджень спрямовані на вивчення даного підходу для вирішення завдань класифікації зображень різного рівня складності, здебільшого дослідження проводились на основі більш класичних, ніж практичних, задачах класифікації зображень, таких як класифікація зображень рукописних символів [41, 42] або класифікація квітів [43]. Деякі з досліджень були проведені на суттєво більш складних проблемах, таких як дослідження класифікації зображень ImageNet із доданим шумом, проведене Дебанджаном Коуаром [44], яке продемонструвало, що гібридні квантово-класичні нейронні мережі можуть перевершити класичні моделі, такі як ResNet-18 [45], у точності при вирішенні складної задачі класифікації зображень.

1.3. Сучасний стан використання квантових комп'ютерів

Не зважаючи на надзвичайно бурхливий та стрімкий розвиток квантового апаратного забезпечення, рівень доступності квантових обчислень залишається низьким (на момент проведення даного дослідження). Наразі є кілька компаній, що займаються розвитком квантових обчислень та надають доступ до своїх

продуктів у форматі хмарних обчислень, в даному розділі буде розглянуто найбільш поширені та перспективні платформи, що надають такі послуги.

Однією з найбільших платформ, що надають доступ до квантових обчислень для широкого загалу у вигляді хмарних обчислень, є IBM Quantum [46]. Ця платформа надає доступ до виконання задач на квантовому комп'ютері, що має 127 кубітів та базується на квантовому процесорі IBM Eagle [47]. Платформа IBM Quantum надає 10 хвилин квантових обчислень на місяць безкоштовного та з можливістю платного використання за форматом оплати за використаний час обчислень понад 10 хвилин з вартістю \$96 за хвилину обчислень. Для виконання задач на квантових процесорах, що надаються IBM, використовується бібліотека з відкритим кодом Qiskit [48], яка також надає широкий набір інструментарію з розробки та виконання квантових схем, дана бібліотека описана в більших деталях у розділі 2. Проте, не зважаючи на наявність, хоч і обмеженого безкоштовного доступу до квантових обчислень, використання даної платформи для проведення експериментів в межах дослідження виявилось проблематичним, адже на разі IBM забороняє доступ до даної платформи користувачам на території України.

Іншою великою та поширеною платформою, що надає доступ до квантових обчислень є Microsoft Azure Quantum [49]. Дана платформа є своєрідним агрегатором та надає доступ до квантового апаратного забезпечення від інших компаній, таких як IoniQ [50], Pasqal [51], Quantinuum [52] та Rigetti [53]. Таким чином, дана платформа надає дуже широкий вибір квантових комп'ютерів які базуються як на впійманих іонах, так і на технології надпровідності та з різноманітною кількістю кубітів (9-100 кубітів). Платформа Azure Quantum надає доступ до квантових чіпів IonQ, що базуються на технології впійманих іонів, стан яких маніпулюється за допомогою лазерів, та мають 25 та 32 кубіти. Доступ до даного квантового апаратного забезпечення здійснюється на основі оплати за кожну квантову операцію виконану на квантовому комп'ютері та вартість однієї операції не може бути нижча ніж \$12. Іншим квантовим апаратним забезпеченням, доступним на даній платформі є квантові комп'ютери від компанії Pasqal, що будуються на технології нейтральних атомів, стан яких

маніпулюється за допомогою лазерів, та які мають 100 кубітів. Вартість доступу до даного апаратного забезпечення становить \$3000 на годину. Іншим доступним квантовим апаратним забезпеченням на платформі Azure Quantum є квантові комп'ютери від компанії Quantinuum, які базуються на технології впійманих іонів та мають 20 та 56 кубітів. Доступ до даного квантового апаратного забезпечення надається у форматі підписки, вартість якої варіюється у межах \$135000 – \$185000 на місяць. Останнім типом доступного квантового апаратного забезпечення є квантові комп'ютери від Rigetti, що базуються на технології надпровідності та мають 9 та 84 кубіти. Вартість виконання квантових операцій на даному апаратному забезпеченні становить \$78 на хвилину.

Також однією з найбільших хмарних платформ, що надають послуги доступу до квантових обчислень є Amazon Web Services (AWS) Braket від компанії Amazon [54]. Дана платформа також є агрегатором пропозицій квантових обчислень наданих іншими компаніями, а саме IonQ, Rigetti, IQM [55] та QuEra [56]. Таким чином, платформа AWS Braket, так само як і вже розглянута Azure Quantum, надає широкий вибір квантового апаратного забезпечення, побудованого на різних технологіях та різного рівня потужності. Пропозиції доступу до квантового апаратного забезпечення побудованого на основі квантових чіпів від IonQ та Rigetti дублює пропозиції доступні на платформі Azure Quantum, проте має дещо іншу структуру вартості, адже AWS Braket надає доступ до потужностей усіх типів на основі оплати за кожну квантову операцію, що виконується на квантовому апаратному забезпеченні, на відміну від Azure Quantum, який для певних видів квантового апаратного забезпечення обчислює вартість на основі часу виконання квантових задач. Таким чином, запуск квантової задачі, що має 1000 «шотів» (shots) на квантових чіпах від IonQ буде коштувати \$30.3 – \$80.3, залежно від конкретного типу квантового процесора. В той же час, запуск аналогічної задачі на квантовому чіпі від Rigetti буде коштувати \$1.2. Запуск аналогічної задачі на квантовому чіпі від IQM, який будується на основі технології суперпровідності та має 20 кубітів, буде коштувати \$1.75. В той же час, запуск даної задачі на квантовому чіпі від QuEra, який будується на технології нейтральних атомів та має 256 кубітів, обійдеться

в \$10.3. Також за необхідності інтенсивного використання квантових обчислень, AWS Braket надає можливість погодинної оренди квантових комп'ютерів, вартість даної пропозиції варіюється в межах \$2500 – \$7000 за годину використання.

З даного короткого огляду та з зібраних даних можна зробити висновок, що у наш час квантові обчислення залишаються доволі обмеженими в доступності та мають надзвичайно високу вартість. Та в той же час сучасний стан квантового апаратного забезпечення диктує вимогу багаторазового запуску кожної задачі та агрегації результатів цих запусків для отримання усередненого результату для того, щоб мінімізувати вплив зовнішніх факторів, що спричиняють нестабільність кубітів на результат та отримати більш стабільний результат обчислень. Дану техніку називають «шотами» (shots) задачі, де кількість «шотів» відповідає кількості запусків кожної задачі. Таким чином, збільшуючи кількість «шотів» задачі, збільшується достовірність отриманого результату, проте в той же час збільшується і час її виконання та її вартість.

Проте дуже важливо зауважити, що з кожним роком квантове апаратне забезпечення удосконалюється, розробляються нові чіпи, що містять більше кубітів, кубіти стають якіснішими та проводиться величезна кількість досліджень, направлених на розвиток квантових комп'ютерів. Яскравим прикладом цього є випуск нового революційного квантового чіпу Willow від компанії Google в 2024 році, який спроможний експоненційно зменшити кількість помилок зі збільшенням кількості кубітів чіпу [57]. Також прикладом тенденції активного розвитку розмірності квантових чіпів є створення квантового чіпу IBM Osprey у 2022 році, який містить 433 кубіти [58]. Та в 2023 році IBM змогли досягти позначки у вражаючі 1121 кубітів в процесорі IBM Condor [59]. Дані тренди розвитку також підтверджуються опублікованими цілями компаній Google та IBM з розвитку квантових комп'ютерів до 2030-х років [60, 61]. Таким чином можна зробити прогноз, що актуальність досліджень у сферах дотичних до квантових обчислень є високою та буде лише збільшуватись разом з подальшим розвитком та майбутніми удосконаленнями квантового апаратного забезпечення.

Для проведення досліджень у більш стриманому бюджеті, що не дозволяє орендувати квантові апаратні потужності, або також на первинних етапах досліджень для апробації гіпотез, використовують симулятори квантових обчислень, що дозволяють запускати симуляції квантових процесів на класичному апаратному забезпеченні, таким чином значно зменшуючи вартість проведення експериментів, що залучають відносно нескладні квантові обчислення. Просунуті квантові симулятори, такі як Qiskit [48] та PennyLane [62] доступні у вигляді бібліотек з відкритим програмним кодом та доступні до використання безкоштовно. В той же час, їх розробники запевняють, що симулятори продукують результати достатньо наближені до результатів, які б були отримані у результаті виконання реальних квантових обчислень. Проте найбільшою проблемою використання симуляторів квантових процесів є дуже висока складність та кількість класичних операцій, що необхідно виконати для досягнення симуляції, що в свою чергу робить їх повільними та вимогливими до класичного апаратного забезпечення. Квантові симулятори розглянуті детальніше в розділі 2.

Але не зважаючи на активний розвиток квантового апаратного забезпечення, розвиток алгоритмічної та програмної частини помітно повільніший та менш динамічний. Тому постає необхідність дослідження і розробки нових методів гібридизації квантових та класичних обчислень для збільшення ефективності систем штучного інтелекту завдяки залученню квантового апаратного забезпечення.

ВИСНОВКИ ДО РОЗДІЛУ 1

В першому розділі було розглянуто сучасний стан предметної області дослідження. Важливим елементом огляду предметної області є огляд сучасного стану квантового апаратного забезпечення: проблеми, що виникають при роботі з ним та при виготовленні стабільних кубітів, переваги квантових обчислень та перспективи використання квантових обчислень для вирішення широкого ряду важливих проблем, вирішення яких за допомогою класичного апаратного забезпечення просто неможливе, або занадто неефективне, щоб бути раціональним.

Іншим важливим аспектом, що був розглянутий в межах першого розділу дисертаційної роботи, є сучасний стан систем ШІ. Було розглянуто тенденції розвитку сфери ШІ, її важливість в усіх сферах сучасного життя. Також було розглянуто і виклики та обмеження, що постають перед сучасними системами ШІ, такими як їх висока обчислювальна складність, високе споживання енергії сучасними системами ШІ та рух сучасних моделей в сторону ускладнення архітектур та збільшення вимог до апаратного забезпечення, що необхідне для забезпечення функціонування найсучасніших моделей, що базуються на нейронних мережах.

Гібридизація класичних та квантових обчислень може бути потужним інструментом для поєднання сильних сторін класичних глибоких нейронних мереж та квантових обчислень. Оскільки квантові обчислення надають дуже суттєве прискорення операцій, порівнюючи з класичним апаратним забезпеченням, та є значно більш енергоефективними, залучення квантових обчислень потенційно дозволяє збільшити швидкодію системи та зменшити споживання енергії системою, що адресує одні з найбільш гострих проблем в сфері ШІ. Також використання гібридизації нейронних мереж дозволяє залучати квантові обчислення в обмежених масштабах (на відміну від квантового машинного навчання), що в свою чергу дозволяє зменшити вимоги до квантового апаратного забезпечення, яке наразі є доволі нестабільним, складним у виготовленні, має дуже обмежене масштабування (на даний момент

найбільший квантовий комп'ютер оперує 1121-ма кубітами) та є недовговічним (кубіти втрачають когерентний стан за 20-100 мікросекунд), що накладає великі обмеження на об'єми обчислень, що можуть бути виконані на сучасному квантовому апаратному забезпеченні. Отож, використання гібридизації для побудови систем ШІ дозволяє як залучати переваги зі світу квантових обчислень та використовувати різноманітні техніки притаманні класичним нейронним мережам, так і компенсувати слабкі сторони квантового апаратного забезпечення завдяки виділення лише частини обчислень на квантовий пристрій, що дозволяє використовувати менші потужності квантового комп'ютера, та забезпечити збільшення ефективності роботи класичних систем ШІ завдяки зменшення їх складності та перенесення частини обчислень на квантове апаратне забезпечення.

Отже, розробка методу гібридизації квантових та класичних обчислень є доцільною та дозволить збільшити ефективність критичних систем ШІ завдяки залученню квантового апаратного забезпечення.

РОЗДІЛ 2

ОСОБЛИВОСТІ ГІБРИДИЗАЦІЇ КЛАСИЧНИХ ТА НЕКЛАСИЧНИХ ОБЧИСЛЕНЬ

2.1. Вибір задачі

Сфера ШІ неймовірно багатогранна та кількість проблем, які вирішуються системами ШІ майже безкінечна. Дане дослідження орієнтується на один з найбільш поширених класів проблем, для вирішення яких застосовуються системи на базі ШІ – задача класифікації зображень. Класифікація зображень є однією з класичних задач, що вирішується за допомогою ШІ і саме штучний інтелект є основним інструментом вирішення даної задачі.

Задача класифікації зображень була обрана як ключова проблема в межах даної дисертації з огляду на низку причин. Перш за все, це універсальність, адже класифікація зображень зустрічається майже у всіх сферах людського життя та існує надзвичайно широка кількість проблем, які можна звести до задачі класифікації різноманітних зображень. Яскравим прикладом є велика кількість проблем у сфері медицини, які за своєю суттю можуть бути зведені саме до оцінки тих чи інших типів зображень, наприклад така високорівнева задача як виявлення ревматоїдного артриту у пацієнтів може бути зведена до задачі бінарної класифікації зображень, що містять результати досліджень пацієнтів, як наприклад на основі теплових зображень рук пацієнтів [32]. Іншим прикладом з геть відмінної від медицини сфери може слугувати задача оцінки наслідків природних або техногенних катастроф, що також може бути зведено до задачі сегментації та бінарної класифікації супутникових знімків регіону, що постраждав від катастрофи [63]. Іншою важливою причиною вибору саме задачі роботи над зображеннями, як ключової проблеми дисертаційного дослідження, є той факт, що дана проблема часто зустрічається в системах критичної важливості, де залучено величезні масиви даних. Гарним прикладом цього є попередньо згадана задача аналізу супутникових знімків. У контексті критичних задач, що залучають великі об'єми даних, прискорення операцій завдяки

залученню квантових обчислень потенційно може значно підвищити рівень ефективності роботи згаданих систем.

Отже, прогрес та збільшення швидкодії систем ШІ для класифікації зображень завдяки залученню технологій, що базуються на квантових обчисленнях, спроможний виказати позитивний вплив на велику кількість сфер людського життя, зокрема, включаючи критично важливі для добробуту людей застосунки.

2.2. Вибір наборів даних для експериментів

Для проведення досліджень та експериментів в межах дисертаційного дослідження з гібридизації класичних та квантових обчислень для завдань ШІ було обрано задачу багатокласової класифікації зображень. Тому для експериментів було обрано кілька теоретичних та більш класичних наборів даних для задачі класифікації як чорно-білих, так і кольорових зображень та один набір даних, що представляє більш практичну задачу, для вирішення якої квантові обчислення можуть бути корисними, зважаючи на переваги, які вони надають.

Класичні набори даних, що було обрано для проведення експериментів:

- MNIST [64,65]
- MNIST-Fashion [66,67]
- CIFAR10 та CIFAR100 [68,69]

Набір даних, що представляє більш практичну задачу класифікації зображень, що було обрано для проведення експериментів – «Satellite Imagery of Hurricane Damage» (SIHD) [70].

Нижче розглянемо детально кожен з обраних наборів даних разом з обґрунтуванням їх вибору.

2.2.1. Набір даних MNIST

Набір даних MNIST є одним з найпопулярніших наборів даних в області комп'ютерного зору. Він часто слугує першим кроком в оцінці продуктивності різних алгоритмів класифікації зображень.

Набір даних MNIST складається з 70000 зображень рукописних цифр у чорно-білому кольорі. Кожне зображення відповідає одному з десяти класів, що представляють цифри від 0 до 9. Усі елементи набору даних мають уніфікований формат та представляють собою зображення розміром 28x28 пікселів. Набір даних розділений на дві частини: 60000 зображень для навчання і 10000 зображень для тестування. Кожен піксель на зображенні представлений 8-бітовим числом у діапазоні від 0 до 255, що позначає градацію сірого, де 0 позначає білий піксель (фон), а 255 – чорний піксель. Проміжні значення представляють різні рівні сірого, що дозволяє передавати дрібні деталі рукописних штрихів. Кожне зображення в наборі даних анотоване та мітка відповідає цифрі, яка міститься на зображенні. Рисунок 2.1 зображає приклад зображень з набору даних MNIST.

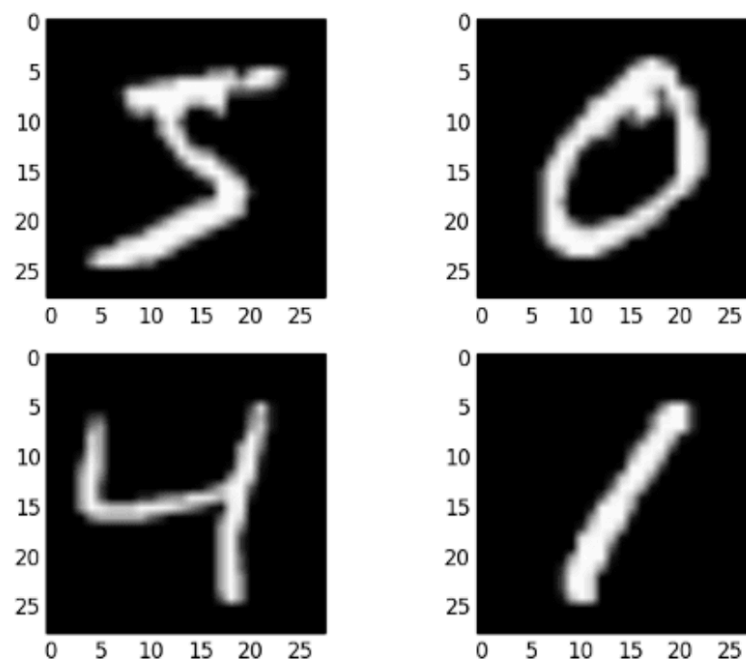


Рисунок 2.1. Приклад зображень набору даних MNIST

Набір даних MNIST є де-факто «hello-world» задачею класифікації зображень. Досягнення високої точності на MNIST часто розглядається як базова вимога для будь-якої нової техніки класифікації зображень. Тому його було обрано як перший крок у дослідженнях методів гібридизації класичних та неklasичних обчислень для дослідження та демонстрації життєздатності обраних методів та підходів на найпростішій задачі класифікації зображень.

2.2.2. Набір даних MNIST-Fashion

Набір даних MNIST-Fashion є дещо складнішим, але схожим за своїми характеристиками на оригінальний MNIST. Він був створений як більш складна та реалістична його альтернатива та широко використовується для оцінки ефективності алгоритмів класифікації зображень на більш складній задачі, ніж оригінальний набір даних MNIST.

Набір даних MNIST-Fashion складається з 70000 зображень, кожне з яких зображає предмет одягу або аксесуар. Всього MNIST-Fashion містить зображення, що відповідають десяти різним класам. Елементи набору даних мають уніфікований формат та представляють собою зображення розміром 28x28 пікселів у градації сірого. Набір даних розділений на дві частини: 60000 зображень для навчання і 10000 зображень для тестування. Кожен піксель на зображенні представлений 8-бітовим числом у діапазоні від 0 до 255, що позначає градацію сірого, де 0 позначає білий піксель (фон), а 255 – чорний піксель. Проміжні значення представляють різні рівні сірого, які відповідають за представлення різноманітних текстур та шаблонів зображених предметів. Також усі елементи набору даних анотовано мітками, що відповідають одному з десяти класів предметів моди, що представлені у наборі даних. Рисунок 2.2 наводить приклад зображень, що містяться в наборі даних MNIST-Fashion.

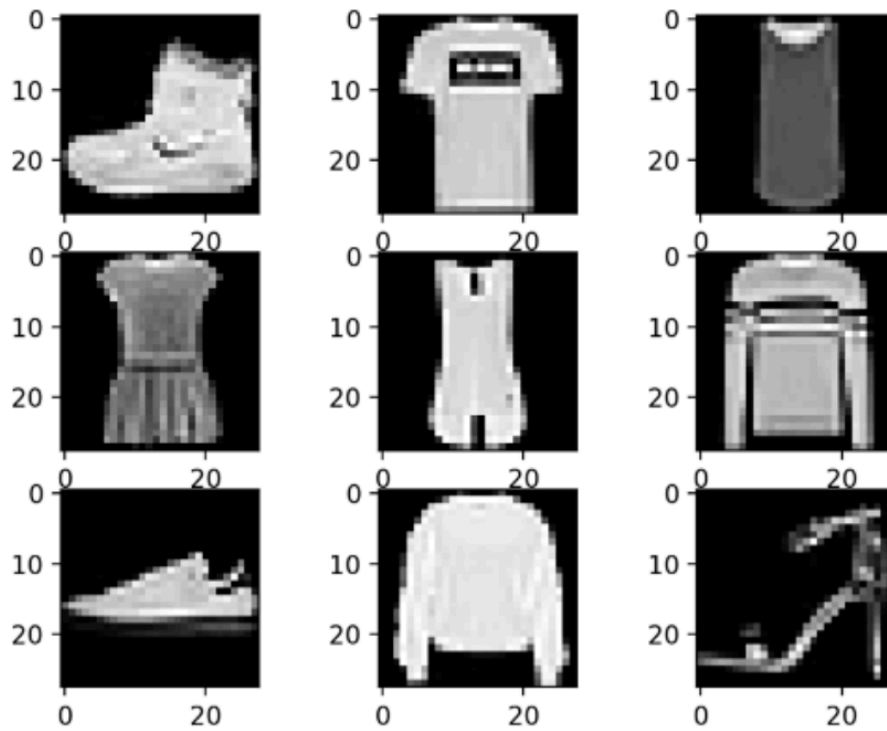


Рисунок 2.2. Приклад зображень набору даних MNIST-Fashion

Отже, можна побачити, що усі технічні характеристики даного набору даних повністю повторюють характеристики та структуру оригінального набору даних MNIST. Відповідність формату зображень та основних характеристик набору даних між MNIST та MNSIT-Fashion дозволяє переносити модель між даними наборами даних без будь-яких змін в структурі та архітектурі моделі. Це важлива особливість даного набору даних, адже він був створений для надання більш складної задачі класифікації зображень для оцінки продуктивності алгоритмів та моделей. На відміну від оригінального набору даних MNIST, який складається з простих зображень цифр, MNIST-Fashion містить більш різноманітні форми та шаблони, знайдені в одязі та аксесуарах, що представляє більш реалістичну та складну проблему для сучасних моделей машинного навчання. Він вводить варіативність у формах, текстурах та схожості між класами, що більшою мірою відповідає завданням, які виникають у практичному застосуванні, у порівнянні з простішими, більш однорідними шаблонами цифр з оригінального MNIST. Більшу складність задачі класифікації зображень з набору даних MNIST-Fashion порівняно з базовим MNIST яскраво видно, якщо порівняти Рисунок 3.1 та Рисунок 3.2, адже, для прикладу, зображення сандалів

помітно складніше за своєю структурою та текстурою ніж будь-яке з наведених зображень рукописних цифр.

Набір даних MNIST-Fashion було обрано для проведення експериментів у вигляді проміжного кроку між простим MNIST та більш складними наборами даних, такими як CIFAR10 та CIFAR100, адже він все ще є доволі простим, порівнюючи з вищезгаданими наборами даних, проте водночас складнішим, ніж базовий MNIST, що дозволяє краще оцінити життєздатність підходів, перш ніж переходити до вирішення більш серйозних та практичних проблем класифікації зображень.

2.2.3. Набори даних CIFAR10 та CIFAR100

Набори даних CIFAR10 та CIFAR100 є основними еталонами в галузі машинного навчання, особливо у сферах класифікації зображень та комп'ютерного зору. Розроблені Канадським інститутом передових досліджень (CIFAR) та опубліковані Алексом Кріжевським, Вінодом Наїром та Джеффри Хінтоном [69] ці набори на даний момент вважаються одним з стандартних інструментів для оцінки продуктивності різноманітних алгоритмів класифікації зображень, особливо методів ШІ, що базуються на згорткових нейронних мережах.

Набір даних CIFAR10 складається з 60000 кольорових зображень, кожне з яких є RGB картинкою розміром 32x32 пікселів. Всі зображення в наборі даних є анотованими та відносяться до одного з 10 класів. Кожен клас містить 6000 зображень. Набір даних розділений на дві частини: 50000 зображень для навчання та 10000 для тестування. Кожен з класів, що представлений в наборі даних, містить 5000 тренувальних зображень та 1000 тестових зображень, отже можна стверджувати, що всі класи представлені в однаковому співвідношенні в обох саб-сетах даних. Кожен з класів, що представлений в наборі даних CIFAR10, представляє реальний об'єкт, такий як, наприклад, літак, автомобіль, птах і т.ін. Усі класи є взаємовиключними, тобто усі зображення мають лише

одну мітку та жоден з елементів набору даних не відноситься одночасно до кількох різних класів. Рисунок 2.3 наводить приклад зображень, що містяться в наборі даних CIFAR10.

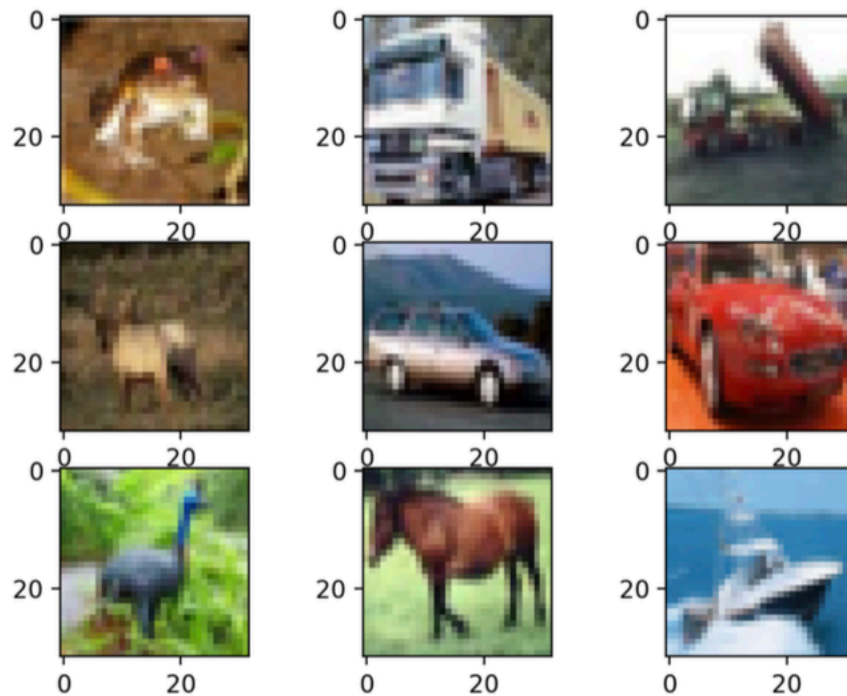


Рисунок 2.3. Приклад зображень набору даних CIFAR10.

На відміну від наборів даних у градаціях сірого, таких як MNIST та MNIST-Fashion, зображення з CIFAR10 представлені в кольоровому вигляді у форматі RGB, де кожен піксель представлений трьома 8-бітовими цілими числами, що відповідають червоному, зеленому та синьому каналам. Це робить набір даних помітно більш складним і репрезентативним для оцінки моделей, які працюють з кольоровими зображеннями. CIFAR10 вважається помірно складним набором даних через відносно малий розмір зображень (32x32 пікселів) та складність візуального контенту. Невелика роздільна здатність обмежує кількість деталей, які можуть бути захоплені, що ускладнює для моделей відмінність між класами, які мають подібні візуальні характеристики, для прикладу «кінь» та «олень», що можна помітити на Рисунку 2.3.

Набір даних CIFAR100 схожий на CIFAR10 як за кількістю зображень, що налічує 60000 кольорових зображень, так і за форматом зображень, кожне з яких є RGB зображенням розміром 32x32 пікселів. Важливою відмінністю CIFAR100

від CIFAR10 є кількість класів, на які поділено набір даних: CIFAR100 містить зображення 100 класів та кожному з класів відповідає по 600 зображень. Даний набір даних має ієрархічну структуру, де 100 класів згруповані в 20 супер-класів. Кожен супер-клас містить 5 пов'язаних класів. Наприклад, супер-клас «комахи» включає класи, такі як бджола, жук, метелик, гусениця та тарган. Отже, всі класи та супер-класи є збалансованими та рівномірно представлені у наборі даних. Рисунок 2.4 наводить приклад зображень з набору даних CIFAR100.

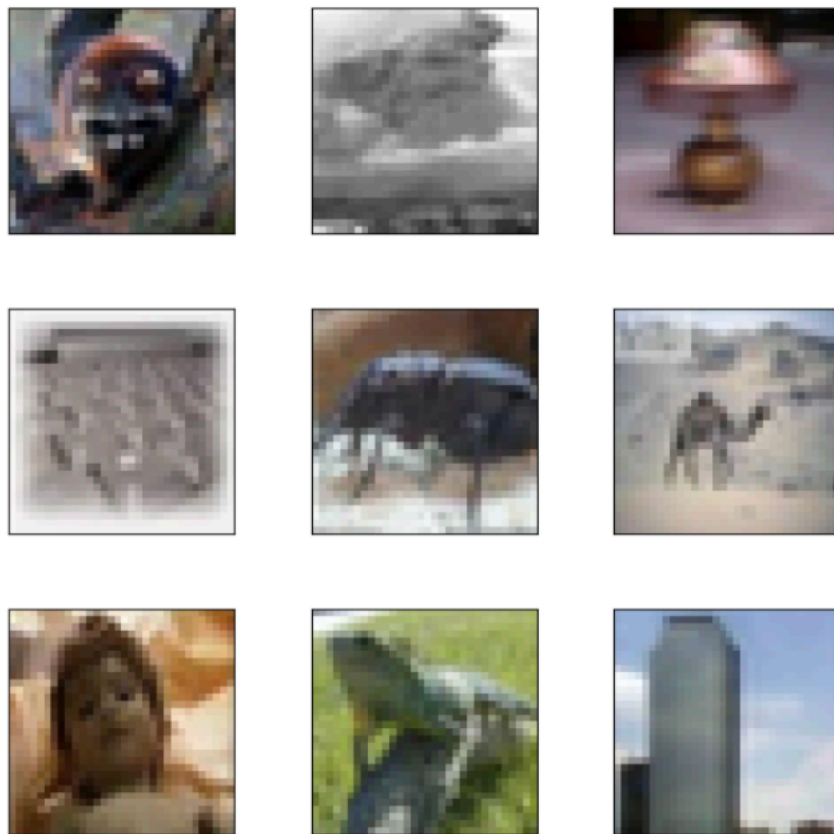


Рисунок 2.4. Приклад зображень набору даних CIFAR100.

CIFAR100 є значно складнішим для класифікації, ніж CIFAR10 через значно більшу кількість класів, меншу кількість зображень, що відповідають кожному з класів та ієрархічну структуру набору даних. Збільшена деталізація та різноманітність класів забезпечують більш суворе тестування здатності моделі до узагальнення. Тому набори даних CIFAR10 та CIFAR100 були використані у дослідженні через їхню популярність та поширеність у ролі еталонних наборів даних для сучасних моделей класифікації зображень.

2.2.4. Набір даних SIHD

Як приклад реального сценарію можливого застосування гібридних квантово-класичних нейронних мереж було обрано набір даних SIHD.

Набір даних SIHD містить 23000 кольорових (RGB) зображень розміром 256x256 пікселів, що містять супутникові знімки пошкоджених та непошкоджених будівель. Цей набір даних складається із зображень, зроблених супутником у районі Великого Х'юстона після урагану Харві, який вразив цей район у 2017 році. Рисунок 2.5 зображує приклад зображень, що містяться у даному наборі даних.

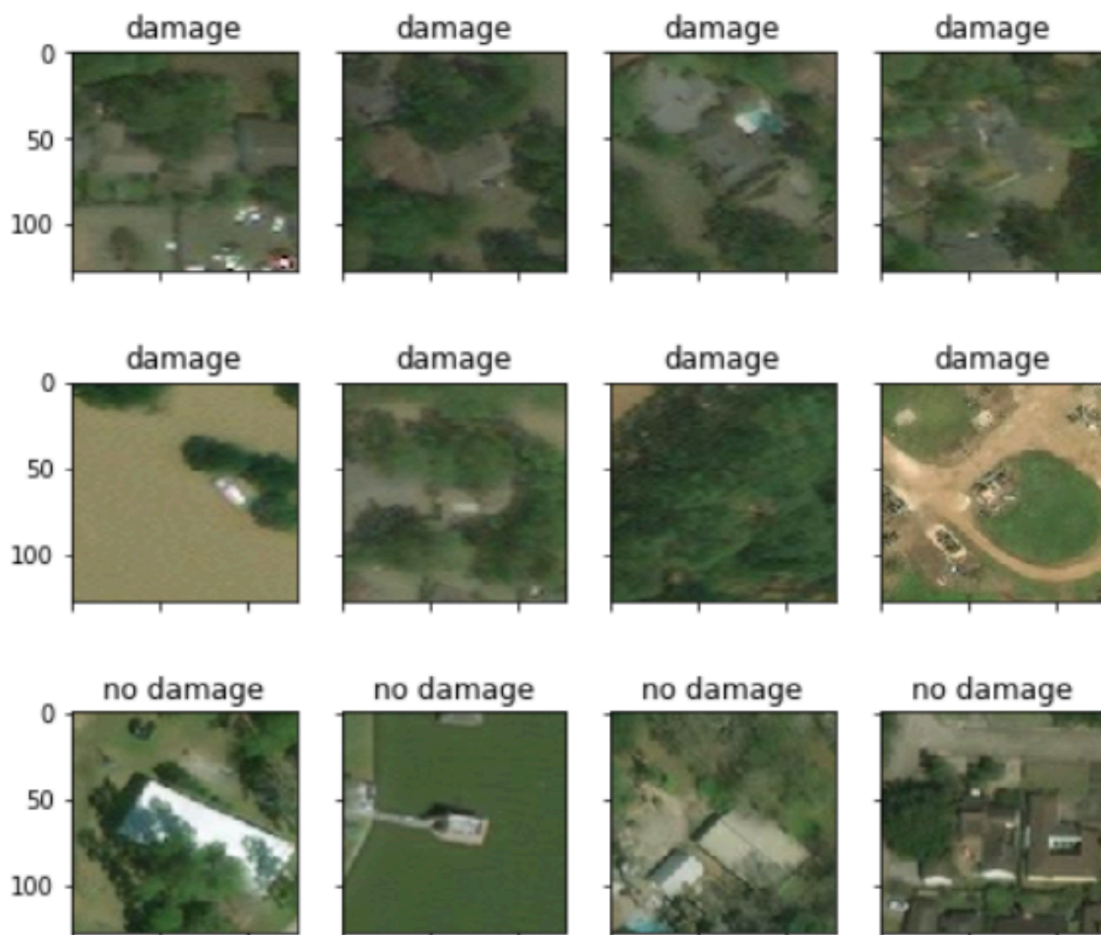


Рисунок 2.5. Приклад зображень набору даних
«Satellite Imagery of Hurricane Damage»

Даний набір даних не є збалансований і кількість зображень непошкоджених будівель значно більша, ніж кількість зображень, що містять

пошкоджені будівлі. У дослідженні було прийнято рішення мінімізувати потенційну упередженість набору даних в сторону непошкоджених будівель. Отже, було використано підмножину набору даних, яка містить однакову кількість зображень для кожного класу — 5000 зображень пошкоджених і 5000 зображень непошкоджених будівель.

Основні характеристики цього набору даних різко відрізняються від інших наборів даних, що були використані в межах дослідження, в першу чергу тим, що зображення є значно складнішими за своїми характеристиками:

- Їх розмір в рази більший, ніж розмір зображень з CIFAR10 чи CIFAR100.
- Форма будівель може дуже різко відрізнятися.
- Ушкодження кожної з будівель є високою мірою доволі унікальними.

Тож можна зробити висновок, що даний набір даних є як практичним, так і нетривіальним для класифікації та відносно складним, зважаючи на свою варіативність.

Задачу класифікації супутникових знімків було обрано у якості прикладу практичної задачі в межах дослідження, оскільки одним із провідних напрямків, де можуть використовуватися квантові обчислення, є аналіз супутникових зображень. Це особливо актуально та важливо в контексті того, як наразі зазвичай виконуються такі завдання, як оцінка збитків після природних або техногенних катастроф. Основним методом, що використовується для таких завдань, є огляд територій, які було зачеплено лихом, шляхом проїзду по даним територіям автомобілем та оцінка пошкоджень. Цей метод покладається на волонтерів і бригади екстреного реагування, які об'їжджають постраждалу територію та візуально оглядають ділянки. Це як трудомістка, так і небезпечна діяльність, яка в деяких випадках може бути неможливою, наприклад, у різних сценаріях, що пов'язані з бойовими діями та війною в цілому. Замість згаданого методу, або на додаток до нього, проблеми оцінки збитків можуть бути вирішені за допомогою аналізу супутникових зображень постраждалої території. Наша гіпотеза щодо користі та ефективності використання квантових обчислень для

цього завдання базується на тому, що супутники генерують величезну кількість даних, і завдання аналізу таких масивних наборів даних може отримати переваги від прискорення обчислень, яке можуть забезпечити квантові обчислення.

2.3. Визначення показників ефективності та метрик для експериментів

Визначення показників ефективності гібридних квантово-класичних нейронних мереж для експериментів є надзвичайно важливим кроком дослідження, адже, щоб зробити висновок про можливість застосування тієї чи іншої техніки, потрібно знати на які саме метрики потрібно звертати увагу та які з них є релевантними, а які ні.

Серед метрик, що часто використовуються для оцінки якості роботи різноманітних алгоритмів класифікації, можна виділити наступні:

- Максимальна точність класифікації зображень з тестового набору даних, досягнута моделлю.
- Значення функції втрат під час процесу навчання.
- Відсоток помилок при класифікації.
- Швидкодія моделі.
- Об'єм та складність обчислень, необхідних для роботи алгоритму.

Усі вищезгадані параметри є надзвичайно важливими для оцінки ефективності роботи моделі та усі вони впливають на можливість її використання у різних контекстах. Для прикладу, навіть якщо модель показує нижчу точність, ніж SOTA моделі, її використання все одно може бути можливим та ефективним, якщо вона виграє за швидкодією або залучає менше обчислювальних ресурсів для своєї роботи, такий компроміс може бути дуже корисним для цілого ряду різноманітних задач, які вимагають надзвичайно високої швидкості обробки даних або повинні працювати у середовищі, що

обмежене за доступними потужностями (для прикладу, так звані «периферійні обчислення»).

Експерименти, проведені в межах дослідження, з оцінки розглянутих методів гібридизації класичних та некласичних обчислень залучають симуляцію квантових обчислень у зв'язку з низькою доступністю квантового апаратного забезпечення та надзвичайно високою вартістю виконання квантових операцій. У зв'язку з використанням симуляції квантових обчислень на класичному апаратному забезпеченні було неможливо оцінити справжню швидкість гібридних моделей, тому оцінка даного аспекту та, відповідно, метрики швидкодії методів гібридизації класичних та квантових обчислень залишились поза межами дисертаційного дослідження та є важливим аспектом для подальших експериментів та досліджень.

Іншим нюансом є те, що ключовим аспектом дослідження є використання саме квантових обчислень, тому додатковою метрикою, яку було взято до уваги у межах дослідження є об'єм квантових операцій, що необхідні для підготовки гібридної моделі та для використання вже навченої моделі.

Отже, в межах даного дослідження було використано наступні метрики ефективності роботи гібридних моделей:

- Точність класифікації зображень моделлю.
- Значення функції втрат під час процесу навчання моделі.
- Відсоток помилок класифікації.
- Об'єм квантових обчислень, що необхідні для підготовки моделі.
- Об'єм квантових обчислень, що необхідні при використанні моделі.

Точність класифікації моделей була оцінена за допомогою наступних підходів. Перший підхід є простішим та більш прямолінійним та полягає в розбитті оригінального набору даних на тренувальний набір даних та тестовий набір даних, та точністю вважається відсоток правильно класифікованих зображень тестового набору даних. У якості другого та більш ґрунтовного підходу з оцінки точності роботи моделі було обрано підхід перехресної перевірки (cross-validation) [71]. Він полягає в тому, що тренувальний та

тестовий набір даних об'єднуються та обирається випадковим чином 20% з об'єданого набору даних як тестові дані, а решта 80% – тренувальні дані. Далі відбувається процес навчання моделі та оцінка метрик її роботи. Та після цього відбувається ротація тренувальних та тестових даних та весь процес повторюється знову, аж поки кожен окремий елемент набору даних не побував у складі як тренувального, так і тестового набору даних. Фінальним результатом вважається середнє значення точності серед усіх таких «згорток». Підхід перехресної перевірки дозволяє не лише краще оцінити точність роботи моделі, а й дає уявлення про статистичні метрики роботи моделі та її стійкість до зміни даних та усуває потенційно викривлені результати, що можуть виникнути, якщо тестовий набір даних буде містити «простіші» та «більш вдалі» для класифікації зображення.

2.4. Огляд доступних архітектур квантового пристрою

Квантовий пристрій є ключовим елементом гібридної нейронної мережі. Квантовий пристрій, котрий також називають квантовою схемою – це набір кубітів, що перебувають в когерентному стані, та певна кількість інструкцій та операцій, що застосовуються до кубітів з метою маніпуляції їх станом. Існує величезна кількість операцій, що можуть входити в склад квантових схем, проте в межах дисертаційного дослідження було прийнято рішення зупинитись на кількох широко розповсюджених базових операціях над кубітами.

- Перша з них це «Ry» операція, яка по суті є ротацією кубіта навколо осі «y» на певний заданий кут « θ » [72]. Дану операцію можна описати наступною формулою:

$$RY(\theta) = \exp\left(-i\frac{\theta}{2}Y\right) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{pmatrix} \quad (2.1)$$

- Також в дослідженні була використана «Rx» операція, яка є ротацією кубіта навколо осі «x» на певний заданий кут « θ » [73]. Дану операцію можна описати наступною формулою:

$$RX(\theta) = \exp\left(-i\frac{\theta}{2}X\right) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -i\sin\left(\frac{\theta}{2}\right) \\ -i\sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{pmatrix} \quad (2.2)$$

- «Rz» операція, яка є ротацією кубіта навколо осі «z» на певний заданий кут « θ » [74]. «Rz» операція може бути виконана за допомогою фізичної операції на апаратному забезпеченні, тобто за константний час та з нульовою помилкою. Дану операцію можна описати наступною формулою:

$$RZ(\theta) = \exp\left(-i\frac{\theta}{2}Z\right) = \begin{pmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{pmatrix} \quad (2.3)$$

- Операція Хадамарда (Hadamard Gate), яка є ротацією кубіта навколо осей «X» та «Z» на кут π [75]. Hadamard Gate є еквівалентом зміни бази обчислень з $|0\rangle, |1\rangle$ до $|-\rangle, |+\rangle$ та навпаки. Дану операцію можна описати наступною формулою:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad (2.4)$$

- Операція Паулі X (σ_x), яка є ротацією кубіта навколо осі «X» на π радіан [77]. Дана операція є еквівалентом класичної інверсії біту, що перетворює $|0\rangle$ в $|1\rangle$, а $|1\rangle$ в $|0\rangle$. Операцію σ_x можна описати наступною формулою:

$$\sigma_x = RX(\pi) = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = -iX \quad (2.5)$$

- Операція Паулі Y (σ_y), яка є ротацією кубіта навколо осі «Y» на π радіан [78]. Дана операція є еквівалентом інверсії біту та інверсії фази водночас, що перетворює $|0\rangle$ в $i|1\rangle$, а $|1\rangle$ в $-i|0\rangle$. Операцію σ_y можна описати наступною формулою:

$$\sigma_y = RY(\pi) = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = -iY \quad (2.6)$$

- Операція Паулі Z (σ_z), яка є ротацією кубіта навколо осі «Z» на π радіан [79]. Дана операція є еквівалентом інверсії фази, що перетворює $|0\rangle$ в $|0\rangle$, а $|1\rangle$ в $-|1\rangle$. Операцію σ_z можна описати наступною формулою:

$$\sigma_z = RZ(\pi) = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = -iZ \quad (2.7)$$

- Операція обміну станів кубітів, яка діє над двома кубітами «a» та «b» та встановлює стан кубіту «a» рівний поточному (до початку процесу обміну) стану кубіту «b» і навпаки [80]. Тобто цю операцію можна представити як перетворення $|a, b\rangle$ в $|b, a\rangle$.

Усі використані у дослідженні квантові схеми побудовано лише за допомогою найпростіших квантових операцій над кубітами. Проте навіть за допомогою вищезгаданих операцій можливо побудувати ефективну систему, що базується на основі квантових обчислень. У межах дисертаційного дослідження було використано квантові схеми, що містять відносно малу кількість кубітів, 2-9 кубітів. Частина алгоритмів побудована таким чином, щоб мінімізувати розмірність квантової схеми та, відповідно, кількість кубітів, що залучені до виконання обчислень. Це особливо важливо з огляду на сучасний етап розвитку квантового апаратного забезпечення і на складність процесу виготовлення якісних кубітів.

В дисертаційному дослідженні було використано кілька типів квантових схем для побудови гібридних квантово-класичних нейронних мереж.

RY квантові схеми, які є найпростішим типом квантових схем, використаних в дослідженні. Дані квантові схеми складаються лише з двох операцій, що виконуються над кожним з кубітів незалежно. Перша з операцій – Hadamard Gate, а друга – параметризована операція $RY(\theta)$ та після виконання схеми вимірюється стан усіх кубітів, що входять у її склад. Значення параметру θ індивідуальне для кожного кубіту та ці значення оновлюються в процесі тренування гібридної нейронної мережі. Діаграма RY квантової схеми, що містить 2 кубіти, наведена на Рисунку 2.6.

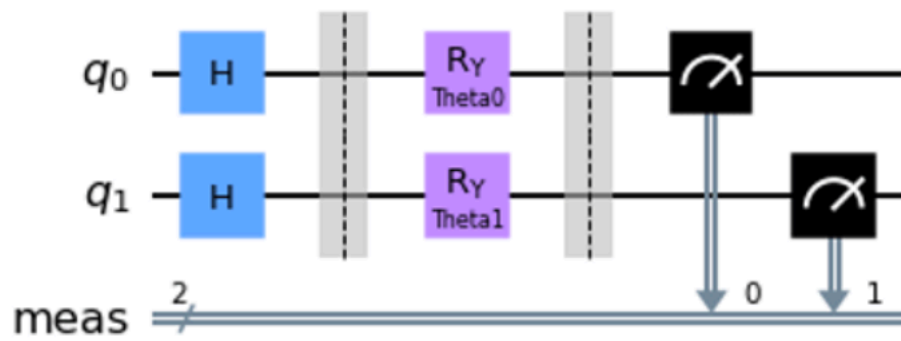


Рисунок 2.6. Діаграма RY квантової схеми з 2 кубітами

Іншим типом квантових схем, що були використані у дослідженні, були схеми, що реалізують алгоритм квантової апроксимації оптимізації (Quantum Approximation Optimization Algorithm, або скорочено QAOA). Даний тип квантових схем складніший, як порівняти з RY схемами, QAOA схеми складаються з більшої кількості операцій – певної кількості RX, RZ та AND операцій, кількість яких залежить від кількості кубітів в квантовій схемі. Аргументація за використанням даного квантового алгоритму для побудови гібридних квантово-класичних нейронних мереж базується на ідеї, що задача навчання нейронної мережі сама по собі є задачею оптимізації точності роботи мережі. Діаграма квантової схеми QAOA, що містить 2 кубіти, наведена на Рисунок 2.7.

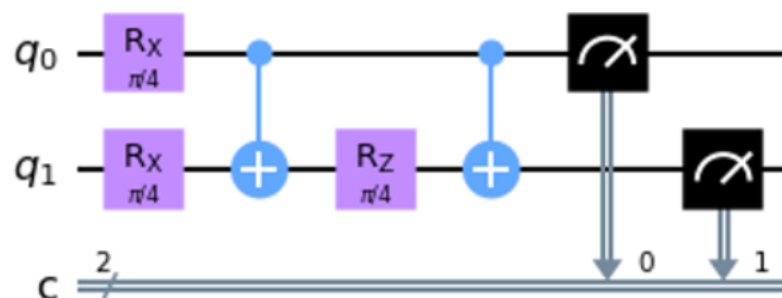


Рисунок 2.7. Діаграма QAOA квантової схеми з 2 кубітами

Третім типом квантових схем, що були використані у дослідженні, були схеми квантового перетворення Фур'є. Квантове перетворення Фур'є є не параметризованою схемою, тобто не містить конфігурованого параметру, а

результат роботи квантової схеми повністю залежить від вхідних даних (якщо не зважати на певний рівень помилок та неточностей, які наразі є невід'ємним складником квантових обчислень на практиці). Діаграма квантової схеми квантового перетворення Фур'є, що складається з 2 кубітів, наведена на Рисунку 2.8.

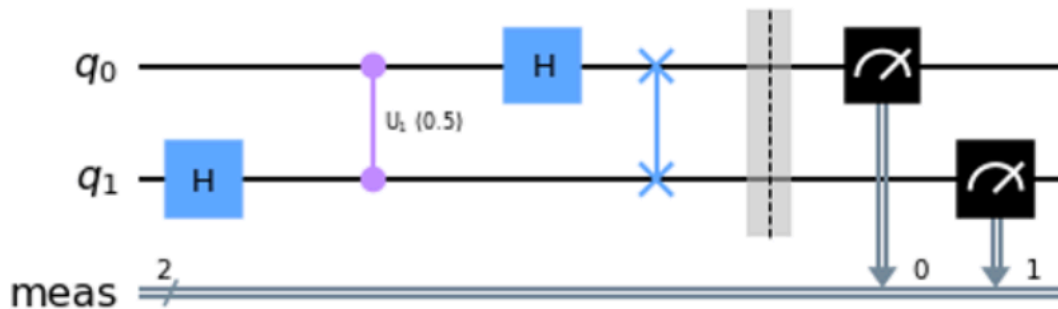


Рисунок 2.8. Діаграма квантового перетворення Фур'є з 2 кубітами

Четвертим типом квантових схем, що були використані в дослідженні, були квантові схеми, що реалізують алгоритм квантового стану Белла [81]. Стан Белла є одним з найпростіших варіантів реалізації квантової заплутаності (quantum entanglement) кількох кубітів. В дослідженні було використано варіацію квантової схеми, що містить додаткові параметризовані операції $R_Y(\theta)$, що дозволяє маніпулювати станом кубітів, змінюючи параметр θ під час процесу навчання гібридної квантово-класичної мережі, що побудована на основі даної схеми, таким способом залучаючи квантову частину мережі до навчання. Діаграма квантової схеми, що реалізовує даний алгоритм, наведена на Рисунку 2.9.

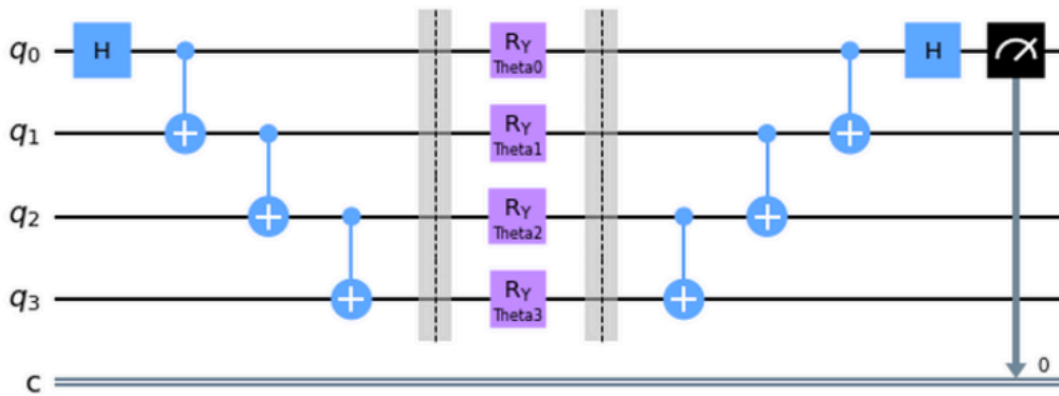


Рисунок 2.9. Діаграма реалізації квантового стану Белла

Також, окрім вищезазначених структурованих квантових схем, для частини експериментів було використано ортогональний підхід – замість дизайну квантових схем, що теоретично можуть бути ефективними для вирішення поставленої задачі, або використання загальновідомих квантових алгоритмів було використано підхід, що базується на генерації довільних квантових схем з певним рівнем квантової заплутаності та заданою кількістю кубітів, що необхідні для вирішення поставленої задачі. Даний підхід було використано з ідеї введення певного рівня квантового «шуму» в процес роботи нейронної мережі, таким способом спонукаючи мережу краще узагальнювати тренувальні дані, шукаючи ті чи інші шаблони, що присутні на зображеннях, що відносяться до одного класу, і менше саме «вивчати» та «запам'ятовувати» самі тренувальні дані. Оскільки генерація квантових схем даного типу відбувалась випадковим чином і часто генерувалось багато схем для одного експерименту, навести всі їх детальні діаграми неможливо, Рисунок 2.10. зображає приклад довільно згенерованої квантової схеми, що містить 9 кубітів.

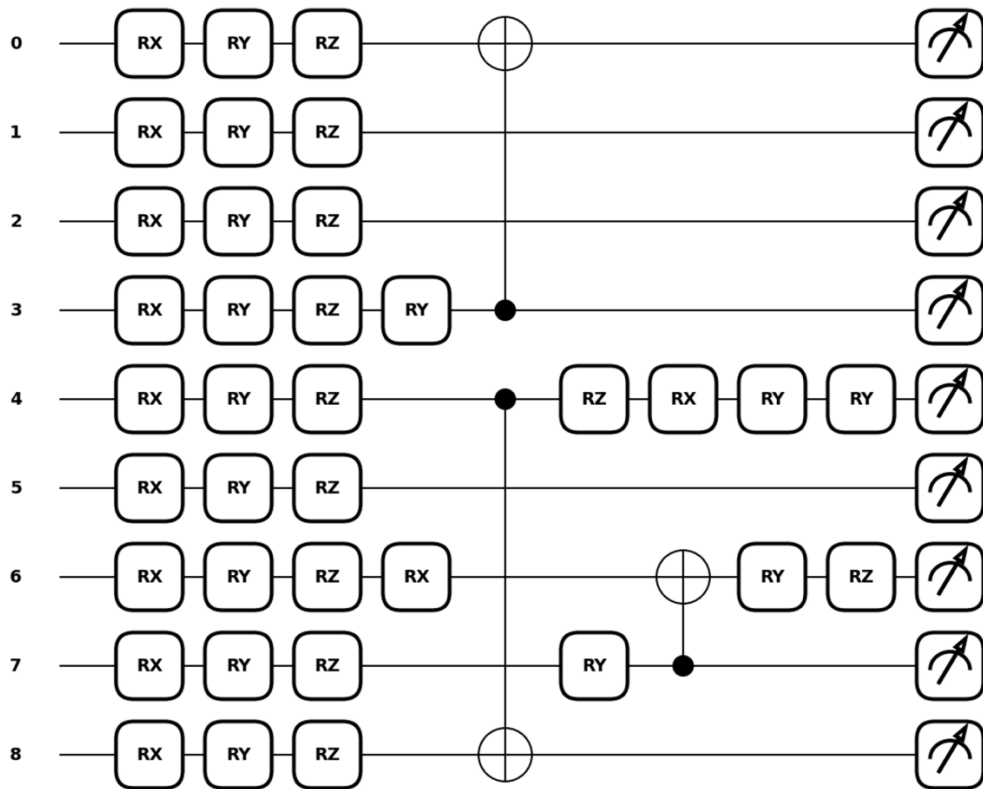


Рисунок 2.10. Діаграма випадково згенерованої квантової схеми, що містить 9 кубітів.

2.5. Вибір інструментів розробки та оточення

Незважаючи на те, що сфера ШІ надзвичайно динамічна, багатогранна та наповнена різноманітними новітніми дослідженнями, набір стандартних інструментів розробки, що використовуються для розробки технологій ШІ, не надто широкий і відносно сталий, що складається з 2-3 мов програмування, кількох фреймворків з інструментарієм для побудови нейронних мереж та кількох утилітарних бібліотек для спрощення різноманітних широко розповсюджених операцій.

В даному дослідженні використовується мова програмування Python [82]. Це інтерпретована динамічно типізована мова з доволі мінімалістичним синтаксисом, що є де-факто стандартною технологією «за замовчуванням» для будь-яких експериментів та розробок у сфері науки про дані в цілому та для нейронних мереж зокрема. Python було обрано через його поширеність в

індустрії та широкий набір готових та доступних інструментів та бібліотек, що спрощують реалізацію або й реалізують в цілому ті чи інші підходи, що були використані для побудови нейронних мереж в межах даного дослідження.

Всі нейронні мережі, використані в дослідженні, було побудовано та оперовано ними, використовуючи два фреймворки. Перший з них – це PyTorch [83]. PyTorch було використано для переважної більшості експериментів, проведених в межах даного дослідження. Це найпопулярніший інструмент у світі досліджень, пов'язаних з нейронними мережами та ШІ в цілому. Даний фреймворк надає неймовірно широкий набір інструментів для побудови та оперування моделями та, що найважливіше, надає інструменти для простого вбудовування будь-якого black-box рішення як частини моделі, що неймовірно важливо для побудови гібридних нейронних мереж. Іншим інструментом, який було використано для побудови та оперування нейронними мережами, є фреймворк Tensorflow [84]. Tensorflow це, як і PyTorch, бібліотека для побудови нейронних мереж різної складності, вона має зручний програмний інтерфейс та є широко поширеною у комерційних системах на базі ШІ. Tensorflow було використано для проведення одного з експериментів з огляду на те, що ця бібліотека надає дуже зручний функціонал для автоматичної генерації діаграм нейронних мереж, що було необхідно, зважаючи на природу експерименту, який вимагає тренування близько десятка моделей дуже схожої структури, для зручного представлення як суті експерименту, так і його результатів.

Всі експерименти було виконано у середовищі Jupyter Notebooks [85]. Це інтерактивне середовище розробки, яке представляє собою «ноутбук» з багатьма комітками з кодом, код з яких можна виконувати незалежно від інших коміток, проте маючи спільні між всіма комітками змінні та спільне середовище запуску. Jupyter Notebooks також являється стандартним інструментом у сфері науки про дані завдяки тому, що можливість незалежного виконання коміток є надзвичайно зручною для проведення різноманітних експериментів та надає можливість побачити результат виконання лише невеличкої частинки програми без необхідності її запуску цілком, адже весь експеримент може виконуватись протягом дуже значних проміжків часу (частина експериментів з дослідження

вимагають кількох тижнів безперервних обчислень). Також Jupyter Notebooks зберігають результат виконання кожної комірки та можуть виступати у вигляді інтерактивної презентації результатів експериментів та не вимагають створення додаткової документації для представлення результатів.

Іншим надзвичайно важливим аспектом оточення для проведення експериментів з побудови та використання гібридних квантово-класичних нейронних мереж є саме квантове апаратне забезпечення. Зважаючи на поточний етап розвитку квантових обчислень та фізичних квантових комп'ютерів, доступ до квантових обчислювальних ресурсів є дуже лімітованим та дорогим. Тому для обходу цієї проблеми для експериментів було використано симулятор квантових обчислень Qiskit [48]. Симулятори квантових обчислень дозволяють виконувати симуляції складних квантових процесів на класичному апаратному забезпеченні [86]. Результати симуляції максимально наближені до реальних результатів, які було б отримано при проведенні ідентичних операцій на фізичному квантовому комп'ютері. Проте симуляції квантових обчислень мають серйозні недоліки. Перший і найсерйозніший з недоліків – це складність самого симулятора та складність обчислень, які він виконує для забезпечення роботи симуляції. Це надзвичайно сповільнює процес проведення експериментів, адже частина кожної експериментальної гібридної мережі представляє собою квантовий пристрій, що виконує певні операції, та, відповідно, симуляції цих операцій відбуваються надзвичайно повільно. Іншим важливим недоліком симуляторів квантових процесів є їх недосконалість та певний рівень похибки та невідповідності результатів симуляції та реального виконання відповідної операції на фізичному квантовому комп'ютері. Проте, незважаючи на свої недоліки, симулятор квантових обчислень – це необхідна складова експериментів та все ж допомагає уникнути проблеми доступності квантового апаратного забезпечення. Окрім симулятора квантових обчислень, Qiskit надає вичерпний комплекс програмних інструментів для побудови та маніпуляції квантовими схемами, що виступають у якості квантового пристрою, на основі якого побудовані усі гібридні квантово-класичні нейронні мережі в межах даного дослідження. Також для експериментів, що залучали квантові

трансформації зображень, було використано бібліотеку PennyLane [62]. Вона надає більш високо-рівневий інтерфейс для взаємодії з квантовими схемами та, порівняно з Qiskit, має більший фокус на розробку застосунків, які поєднують квантові обчислення та машинне навчання за рахунок того, що надає широкий інструментарій інтеграцій з бібліотеками машинного навчання, такими як PyTorch та Tensorflow.

Усі експерименти було виконано на комп'ютері, що має наступні характеристики: процесор AMD Ryzen 9 5900X 12-Core CPU, 36GB RAM, відеокарта Nvidia GeForce 3080TI 16GB VRAM. Обчислення проводились із залученням графічного прискорювача на тих етапах, де це було можливо. Також частина обчислень проводилась на хмарних ресурсах Kaggle [87] та Google Colaboratory [88] із залученням графічних та тензорних прискорювачів.

ВИСНОВКИ ДО РОЗДІЛУ 2

В даному розділі було розглянуто основні конструкційні блоки квантової частини запропонованого методу гібридизації квантових та класичних обчислень для побудови глибоких гібридних нейронних мереж – квантові операції та основні архітектури квантового пристрою, що пропонуються до застосування при розробці методу гібридизації. Для реалізації квантової схеми для експериментів було обрано наступні квантові операції:

- Обороти кубіта навколо однієї з осей – R_x , R_y , R_z операції.
- Операція Хадамарда.
- Операція Паулі на основі різних осей (x , y , z).
- Операція обміну станів кубітів (swap).

Було розглянуто та обрано для проведення експериментів наступні архітектури квантових схем:

- Квантова схема на основі операції обороту кубіта (R_Y).
- Квантова схема, що реалізує алгоритм швидкої апроксимації (QAOA).
- Квантова схема, що реалізує алгоритм квантового перетворення Фур'є.
- Квантова схема, що реалізує квантовий стан Белла.
- Випадкові квантові схеми на основі операцій оберту кубітів та випадкового рівня квантової заплутаності кубітів.

Для вирішення проблеми доступності квантового апаратного забезпечення для проведення експериментів, було використано симулятори квантових обчислень для симулювання квантових процесів на класичному апаратному забезпеченні. У якості квантових симуляторів було обрано бібліотеки Qiskit та PennyLane, що надають широкий інструментарій моделювання квантових схем та реалістичні симуляції квантових процесів для запуску побудованих схем.

Для проведення експериментів з оцінки та апробації запропонованого методу гібридизації квантових та класичних обчислень було обрано задачу багатокласової класифікації зображень. Було обрано 4 класичні проблеми класифікації зображень та 1 практичну проблему для демонстрації практичності

та ефективності застосування гібридних моделей ШІ на основі запропонованого методу. У якості класичних проблем було обрано задачі класифікації зображень з наборів даних MNIST, MNIST-Fashion, CIFAR10 та CIFAR100 для оцінки ефективності методу на задачах різної складності та оцінки результатів роботи гібридних моделей, порівнюючи з референтними та сучасними SOTA моделями. У якості практичної задачі було обрано задачу класифікації супутникових знімків зони стихійного лиха для виявлення ділянок, що мають руйнування, спричинені стихійним лихом.

РОЗДІЛ 3

ПІДХОДИ ПОБУДОВИ ГІБРИДНИХ КВАНТОВО-КЛАСИЧНИХ НЕЙРОННИХ МЕРЕЖ

3.1. Вибір алгоритмів побудови архітектур гібридних квантово-класичних нейронних мереж

Наразі ведеться велика кількість досліджень багатьма різними командами, що намагаються створити різноманітні гнучкі гібридні інструменти, щоб зробити реальним використання квантових та класичних обчислень в тандемі для пошуку оптимального рішення для розв'язання різноманітних сучасних проблем, що потребують значної кількості обчислень. Зважаючи на це, різноманітні методи гібридизації класичних та некласичних обчислень мають великі перспективи для практичних застосунків, які є обчислювально-інтенсивними. Гарним прикладом сфер, що зацікавлені в розвитку методів гібридизації, є різноманітні системи ІІІ. В гібридних системах класичні апаратні потужності можуть бути використані для визначення найбільш високонавантаженої частини алгоритму та перенесення її обробки на квантове апаратне забезпечення або для розбиття великої проблеми на менші частини, які можуть бути вирішені за допомогою квантових обчислень, а потім знову об'єднані. Саме тому дослідження у сфері гібридних квантово-класичних нейронних мереж є надзвичайно перспективними та актуальними на даний момент.

У межах дисертаційного дослідження було розглянуто 3 різні підходи з гібридизації глибоких згорткових нейронних мереж для вирішення задачі багатокласової класифікації зображень.

1. Квантовий пристрій виконує роль одного з прихованих шарів глибокої згорткової нейронної мережі.

2. Квантовий пристрій виконує роль попередньої квантової обробки зображень до їх потрапляння у класичну частину гібридної нейронної мережі.
3. Квантовий пристрій виконує роль першого згорткового шару глибокої згорткової нейронної мережі.

Далі буде розглянуто кожен з вищезазначених підходів у деталях разом з оглядом ідей, що лежать в основі побудови кожного з зазначених підходів.

3.2. Квантовий пристрій як один з прихованих шарів нейронної мережі

Перший з методів гібридизації класичних та некласичних обчислень для завдань ШІ, що розглядається в межах дисертаційного дослідження, це підхід, що полягає в використанні квантового пристрою як частини гібридної нейронної мережі у вигляді одного з прихованих шарів мережі. Даний підхід спирається на припущення, що інкапсуляція частини необхідних обчислень в межах квантової схеми дозволить зробити класичну частину мережі менш глибокою та виконати частину необхідних обчислень на квантовому пристрої зі значним пришвидшенням їх виконання завдяки квантовим обчисленням, порівнюючи з класичною немодифікованою частиною мережі.

Даний підхід є відправною точкою проведених досліджень методів гібридизації. Теоретично він може дозволити не лише зробити класичну частину мережі менш глибокою, а й повністю винести обчислювально найскладнішу частину початкової класичної нейронної мережі в квантову схему, таким способом значно зменшуючи кількість обчислень, які повинні бути виконані на класичному апаратному забезпеченні. Проте, зважаючи на сучасний етап розвитку квантового апаратного забезпечення, важливим нюансом також є і оптимальний баланс складності класичної та квантової частин гібридної

нейронної мережі, адже використання надто складної квантової схеми або надто великої кількості кубітів може бути нераціональним з огляду на ряд чинників:

1. Найсучасніші квантові комп'ютери спроможні оперувати доволі обмеженою кількістю кубітів протягом дуже обмеженого проміжку часу.
2. Відсутність практичного алгоритму корекції помилок означає наявність доволі високого рівня помилок та неточностей квантових операцій, отже надмірний розмір та складність квантової схеми може спричинити низьку якість її роботи на сучасному квантовому апаратному забезпеченні [16, 17].

Оскільки квантовий пристрій слугує частиною нейронної мережі, він також повинен брати участь у процесі навчання, це означає, що повинні виконуватись наступні необхідні умови для навчання:

- Квантова схема повинна містити конфігуровані параметри, що використовуються, щоб змінити стан кубітів залежно від вхідних даних.
- Необхідно мати алгоритм «прямого» проходу (feedforward) через квантову схему та «зворотній» прохід для обчислення градієнту та зворотного поширення помилки (backpropagation), щоб зробити процес навчання можливим.

Прямий прохід (feedforward) досить простий:

1. Значення конфігурованих параметрів квантової схеми встановлюються рівними значенням вихідних ваг попереднього шару нейронної мережі.
2. Виконання квантової схеми та замір результуючого стану кубітів.
3. Збереження результатів виконання квантової схеми для подальшого їх використання для зворотного проходу.
4. Передача результатів до наступного шару нейронної мережі.

Зворотний прохід трохи складніший, оскільки він потребує обчислення градієнтів для оновлення ваг зв'язків нейронної мережі з метою коригування

прогнозів моделі відповідно до фактичних очікуваних результатів. У нашому випадку градієнт для квантової схеми реалізовано як градієнт для «чорного ящика». Це можна виразити наступною формулою:

$$qc_gradient = qc(\theta + \varepsilon) - qc(\theta - \varepsilon) \quad (3.1)$$

Де θ – це параметр квантової схеми, тобто вхідні значення попереднього прямого проходу, ε – це зміщення, яке є одним з гіперпараметрів гібридної квантово-класичної нейронної мережі, qc – це виконання квантової схеми, використовуючи передані параметри. Тобто градієнт квантового шару гібридної нейронної мережі є різницею між результатом виконання квантової схеми, обчисленими при правому зміщенні ($\theta + \varepsilon$) та лівому зміщенні ($\theta - \varepsilon$) [89, 90]. Таким чином отримано наступний алгоритм зворотного проходу:

1. Отримати значення градієнтів наступних шарів нейронної мережі.
2. Для кожного кубіту обчислити градієнт, використовуючи наведену вище формулу.
3. Результируючий градієнт рівня рівний векторному добутку попередніх градієнтів та вектору градієнтів кубітів.
4. Оновити ваги вхідних зв'язків нейронів відповідно до градієнтів.
5. Передача градієнтів попереднім шарам гібридної нейронної мережі.

Даний алгоритм дозволяє диференціювати квантову схему як частину алгоритму зворотного поширення помилки та, відповідно, використовувати квантовий пристрій як частину глибокої гібридної нейронної мережі.

Рисунок 3.1 зображує високорівневу діаграму гібридної квантово-класичної нейронної мережі, що використовує квантовий пристрій як один з прихованих рівнів.

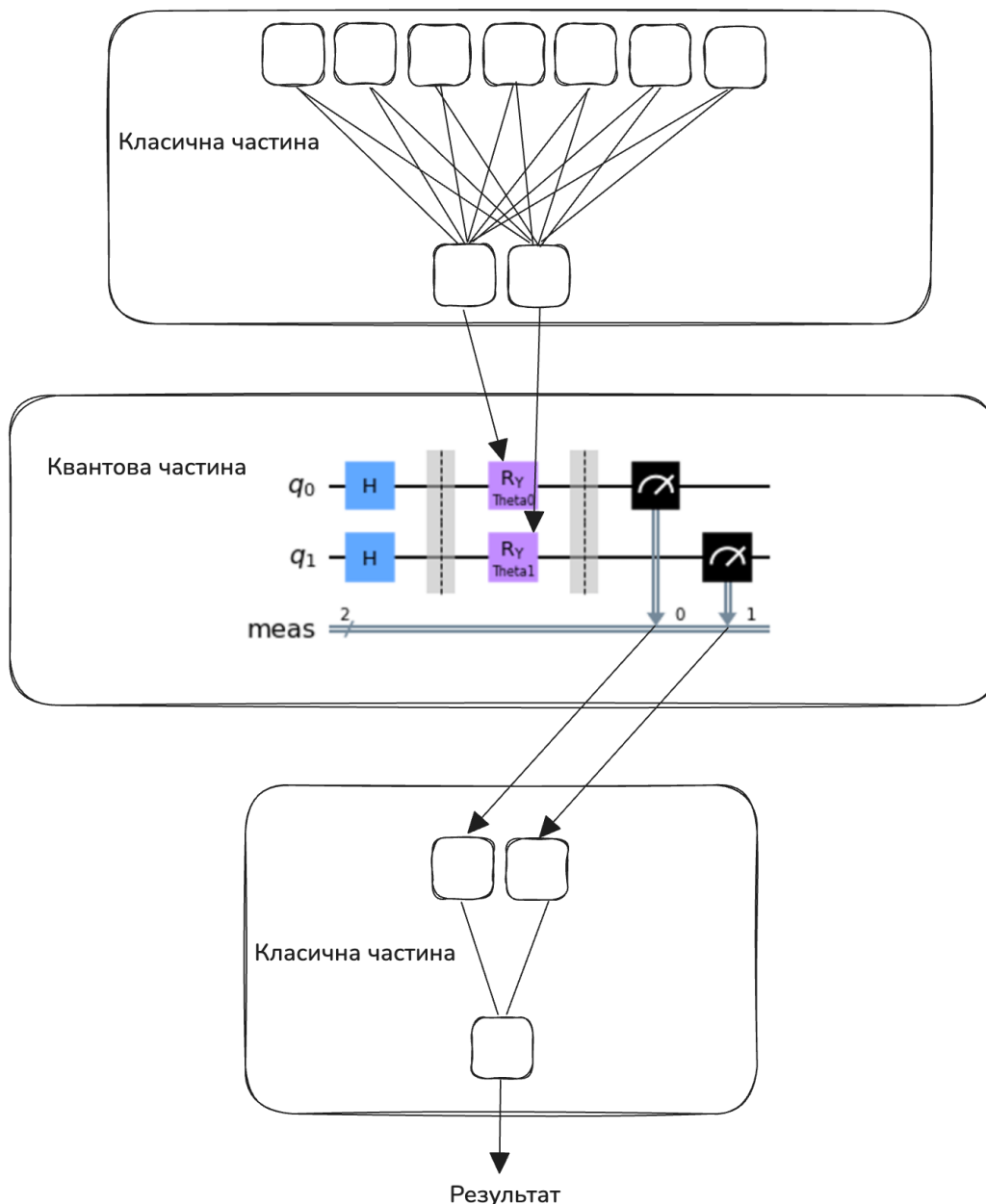


Рисунок 3.1. Діаграма гібридної нейронної мережі з квантовим пристроєм.

Отже, при використанні даного методу гібридизації квантовий рівень гібридної нейронної мережі бере безпосередню участь не лише у процесі використання результуючої моделі, а й у всьому процесі навчання, що вимагає значного залучення квантового апаратного забезпечення у процес підготовки моделі, що може бути одним з дуже серйозних обмежуючих факторів використання даного підходу на практиці на даному етапі розвитку технологій квантових обчислень, адже у сучасному світі доступ до квантових комп'ютерів є лімітованим та їх вартість залишається високою.

3.3. Попередня квантова обробка даних

Другим розглянутим в межах дисертаційного дослідження підходом гібридизації класичних та некласичних обчислень для завдань ШІ є використання квантового пристрою для попередньої квантової обробки даних. Даний підхід спирається на ідеї, що квантовий пристрій може бути використаний для внесення певного рівня квантового шуму у тренувальні дані, таким способом виконуючи роль DA кроку.

Квантове штучне розширення різноманітності даних дещо не відповідає класичному значенню даного підходу, адже його результат не є детерміністичним, не може бути оберненим та результуючі змінені зображення є дуже відмінними від оригінальних даних. Проте, концептуально, дану техніку все ж можна віднести до певного роду розширення різноманітності даних з огляду на те, що задача, яку переслідує квантове «зашумлення» даних, є ідентичною до задачі класичного процесу розширення різноманітності даних та переслідує мету збільшення ефективності процесу тренування моделі.

«Зашумлення» тренувальних даних це дуже поширена техніка попередньої обробки даних. Дана техніка особливо ефективна в сценаріях, коли об'єм тренувальних даних обмежений, адже кожна нова варіація опрацьованих вхідних даних може слугувати додатковим розширенням оригінального набору даних. Звісно, таке розширення даних менш якісне, проте це все одно дуже дієва техніка, коли немає можливості розширити набір даних новими оригінальними даними, а його початковий розмір не є оптимальним для тренування моделі. Також розширення різноманітності вхідних даних провокує кращу генералізацію нейронною мережею, адже модель не «запам'ятовує» вхідні дані, оскільки при кожній ітерації навчання вхідні дані дещо змінені та відрізняються від даних, що були представлені моделі на будь-якій з попередніх ітерацій навчання.

Підхід з попередньої квантової обробки вхідних даних реалізовано у вигляді квантової згорткової операції над зображеннями з оригінального набору даних. Квантово-згорткову операцію часто називають «кванволюційною»

операцією (quantvolutional). Кванволюційна операція була описана Максвелом Хендерсоном [91]. Обробку зображення за допомогою квантово-згорткових операцій можна описати наступним алгоритмом:

1. Обираються гіперпараметри обробки, такі як розмір квантово-згорткового ядра, розмір кроку квантово-згорткового ядра та «падінг» оригінальної картинки. Детальний опис гіперпараметрів та їх ефект на результат наведено нижче.
2. Квантово-згорткове ядро рухається по всій площі оригінальної картинки відповідно до обраних значень гіперпараметрів.
3. Для кожної позиції квантово-згорткового ядра значення кожного каналу пікселя (RGB) слугує вхідним значенням для операції квантової схеми.
4. Виконується квантова схема для кожного з каналів.
5. Результуючий стан кубітів заміряється та їх результат використовується як значення відповідного каналу результуючого зображення.

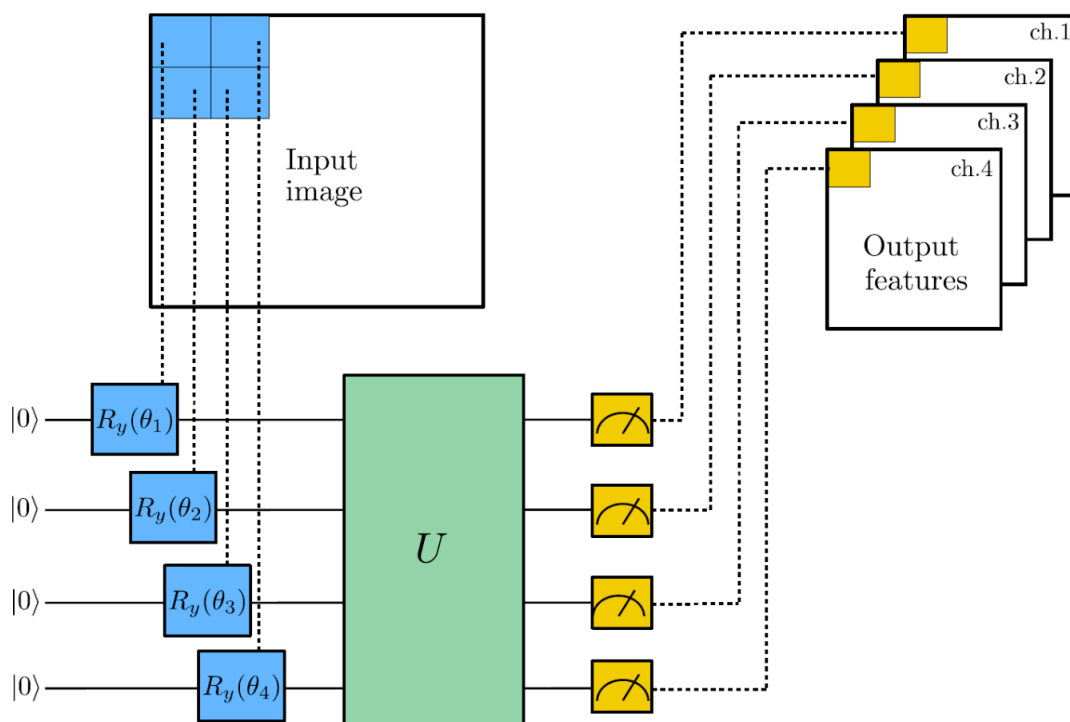


Рисунок 3.2. Діаграма квантово-згорткового процесу [91].

В алгоритмі згадуються наступні гіперпараметри обробки даних:

- Розмір квантово-згорткового ядра (kernel size) – гіперпараметр, що визначає кількість пікселів, що використовуються для кожного окремого кроку згорткової обробки зображення. Розмір ядра повинен бути рівним кількості кубітів квантової схеми, що використовується для операції. Також, оскільки ядро повинно мати вигляд квадрату для коректної роботи алгоритму, то, відповідно, значення даного гіперпараметру повинно бути квадратом цілого числа та бути більшим 1, адже згорткова операція з ядром розміру 1 не має жодного сенсу. Отже, зважаючи на обмеження, що їх накладає сучасний стан квантового апаратного забезпечення, оптимальними значеннями розміру квантово-згорткового ядра є 4 та 9 пікселів, у тому розмір квантової схеми, що використовується для квантово-згорткової операції, повинен також відповідно містити 4 або 9 кубітів.
- Розмір кроку квантово-згорткової операції (stride) – даний параметр є за своєю природою ідентичним відповідному параметру класичної згорткової операції та визначає розмір кроку, на який зміщується квантово-згорткове ядро на кожному кроці алгоритму. На практиці даний параметр впливає на розмір результуючого зображення та збільшення значення кроку робить результуюче зображення меншим за розміром та меншим за роздільною здатністю.
- Паддінг оригінального зображення (padding) – даний параметр також є ідентичним відповідному параметру класичної згорткової операції. Значення даного параметру регулює, чи буде додано до оригінального зображення рамку з пікселів розміром рівним значенню даного параметру та відповідно впливає на результуючий розмір опрацьованого зображення. Це особливо важливо, якщо результуюче зображення повинно мати певний заданий розмір або зберегти оригінальний розмір. За замовчуванням значення даного гіперпараметру рівне нулю.

- Тип квантових операцій, що використовуються в квантовій схемі. Даний гіперпараметр впливає на те, які саме операції виконуються в квантовому пристрої на квантовому апаратному забезпеченні. Набір операцій, що входять до складу квантової схеми, впливає на складність схеми та відповідно змінює і результат квантової обробки кожного пікселя. В межах даного дослідження було використано три типи квантових операцій: RX , RY , RZ у різноманітних їх комбінаціях.

Між класичною згортковою та квантово-згортковою операціями, не зважаючи на їх схожість, є дуже суттєва відмінність: результатом згорткової операції є одне зображення, результатом же квантово-згорткової операції є кілька зображень. Причиною цього є те, що класичний алгоритм згорткової операції продукує лише одне значення для кожної позиції згорткового ядра, в той же час як результатом квантово-згорткової операції є результат заміру стану кубітів квантової схеми після її виконання, а отже їх кількість співпадає з кількістю кубітів, які містить квантова схема. Тож значення кожного кубіту та, відповідно, кожне окреме результуюче значення квантової схеми було використано у вигляді окремого шару результату. Отже, для розміру квантово-згорткового ядра 2×2 ми використовуємо квантову схему, що містить 4 кубіти та результат квантово-згорткової операції буде містити 4 шари. Також кожен з результуючих шарів містить 3 канали, що відповідають RGB каналам зображення. Тому розмірність даних збільшується в n разів після виконання квантово-згорткової операції, де значення n відповідає розміру квантово-згорткового ядра.

Також важливим аспектом даного підходу є архітектура квантового пристрою, що бере участь у квантово-згортковій операції та обмеження, які накладає квантовий пристрій на решту параметрів операції. Перш за все потрібно зазначити, що для експериментів з використання методу гібридизації класичних та некласичних обчислень, який полягає у попередній квантовій обробці даних, було використано квантові схеми, згенеровані випадковим чином, що містять задані типи квантових операцій, що є одним з гіперпараметрів методу. Також

дуже важливим аспектом квантової схеми є певний рівень квантової заплутаності кубітів, що беруть участь у виконанні операцій на квантовому апаратному забезпеченні. Квантова заплутаність є важливим аспектом, адже її наявність гарантує, що значення кожного пікселя, що входить в квантово-згорткову операцію, впливають на багато каналів/шарів результуючих даних та, відповідно, значення в кожному з результуючих каналів/шарів залежить від багатьох пікселів, що беруть участь в кроці квантово-згорткової операції. Дана властивість впливає у те, що квантово-згорткова операція, що залучає квантову схему з квантовою заплутаністю, працює за схожим принципом з класичними згортковими операціями, де значення результуючого пікселя залежить від значень усіх пікселів, що входять у крок згорткової операції.

В межах дослідження було проведено експерименти, метою яких було дослідження раціональності використання додаткових каналів/шарів, отриманих в результаті квантово-згорткової операції, та розглянуто набір підходів до використання багатоканальності, яку надає попередня квантова обробка даних у вигляді різноманітних архітектур класичної частини гібридної мережі.

Загалом архітектура гібридної квантово-класичної моделі III, що залучає метод гібридизації з попередньою квантовою обробкою даних, виглядає наступним чином:

- Квантовий пристрій, що виконує попередню квантову обробку даних.
- Класична модель, що оперує над багатоканальними даними, які є результатом роботи квантового пристрою.

Даний підхід до побудови гібридних моделей III має дуже цікаву відмінність від методу, розглянутого в розділі 3.2. «Квантовий пристрій як один з прихованих шарів нейронної мережі»: обробка кожного окремого елементу оригінального набору даних відбувається лише один раз. Це не впливає на процес експлуатації вже навченої моделі, проте ця особливість дозволяє значно зменшити кількість квантових операцій при процесі навчання моделі, адже можна виконати обробку тренувальних даних лише один раз – перед початком процесу навчання моделі та використати оброблені дані у вигляді нового набору

даних, який вже буде використовуватись для навчання, адже процес квантової обробки даних не бере безпосередньої участі у процесі навчання моделі, що прибирає одне з дуже важливих обмежень, які має попередньо розглянутий метод – квантовий пристрій потрібен лише на початку процесу навчання, тому сам процес навчання на вже підготовлених даних немає ніяких залежностей на квантове апаратне забезпечення. Це може бути дуже корисно, зважаючи на обмежену доступність квантових обчислень у сучасному світі, тому спрощення процесу тренування моделі та мінімізація кількості квантових обчислень, потрібних для цього, може мати дуже позитивний ефект на практиці. Також можливим корисним наслідком є можливість проводити квантову обробку даних паралельно, якщо це, звісно, буде раціонально, зважаючи на доступність квантових обчислень.

Тож можна підсумувати, що описаний метод гібридизації класичних та неklasичних обчислень має наступні переваги:

- Можливість розділити в часі процес підготовки даних за допомогою квантових обчислень та процес навчання класичної частини моделі.
- Можливість проводити обробку даних паралельно.
- Процес навчання класичної частини моделі не вимагає доступу до квантових обчислень.
- Зашумлення вхідних даних за допомогою квантових обчислень, що може мати позитивний вплив на процес навчання моделі.
- Розширення розмірності оригінального набору даних за допомогою квантово-згорткової операції.

3.4. Квантовий пристрій як перший згортковий шар нейронної мережі

Третім розглянутим в межах дисертаційного дослідження підходом гібридизації класичних та неklasичних обчислень для завдань ШІ є

використання квантового пристрою у якості першого згорткового шару гібридної квантово-класичної глибокої згорткової нейронної мережі.

Даний підхід будується на ідеях та підходах методу, що був викладений в розділі 2.4.2. «Попередня квантова обробка даних» та є логічним розвитком вищезгаданої ідеї. Основною відмінністю даного підходу є те, що результат квантово-згорткової операції використовується не у вигляді єдиного багатоканального елемента даних, а у вигляді багатьох варіацій одного і того ж самого зображення.

Процес підготовки моделі можна описати наступним алгоритмом:

1. Квантова обробка оригінального набору даних за допомогою квантово-згорткових операцій.
2. Результати квантової обробки перетворюються у новий тренувальний набір даних, що містить в n разів більше зображень, порівнюючи з оригінальним набором даних, де n – рівний кількості каналів в результаті квантово-згорткової операції.
3. Новий тренувальний набір даних використовується для навчання класичної частини моделі.

Отже, в результаті виконання попередньої квантової обробки даних створюється новий тренувальний набір даних, що складається з n варіацій кожного з вхідних зображень, як вже було зазначено, значення n є рівним кількості каналів в результаті квантово-згорткової операції та зі свого боку відповідає гіперпараметру попередньої квантової обробки даних – розміру квантово-згорткового ядра та, відповідно, розміру квантової схеми, що використовується для попередньої квантової обробки зображень. Це дозволяє додати певної варіативності тренувальному набору даних, що може бути дуже корисно у сценаріях, де розмір оригінального набору даних є не оптимальним для навчання моделі, таким способом даний підхід може дозволити збільшити точність роботи моделі у вищезгаданому сценарію за рахунок розширення тренувального набору даних.

Алгоритм процесу тренування класичної частини гібридної моделі нічим особливим не вирізняється та є ідентичним процесу навчання будь-якої класичної моделі ШІ. Проте процес використання підготовленої моделі має ключову відмінність – зважаючи на те, що квантово-згортковий шар гібридної нейронної мережі, на відміну від класичної згорткової операції, продукує як результат не одне зображення, а набір, що містить n зображень, класична частина гібридної моделі повинна опрацювати та відповідно класифікувати n зображень. Отже, для отримання фінального передбачення гібридною квантово-класичною нейронною мережею потрібно останнім кроком узагальнити результати передбачень для кожної з варіацій опрацьованого зображення. Для цього можна використати багато різноманітних підходів та алгоритмів, що можуть залежати від контексту кожної окремої задачі, з таких підходів можна виділити алгоритм «голосування» за результат MV [92]. Це дозволяє ввести певний рівень гнучкості у процес визначення фінального передбачення гібридної мережі, що може бути корисно у багатьох різноманітних застосунках.

Отже, отримано наступний алгоритм використання гібридної квантово-класичної моделі ШІ, побудованою за допомогою підходу з використанням квантово-згорткової операції як першого шару гібридної нейронної мережі:

1. Квантова обробка вхідного зображення.
2. Для кожної з отриманих варіацій зображення у результаті квантової обробки:
 - a. Опрацювання зображення класичною частиною моделі.
 - b. Збереження передбачення класичної частини моделі.
3. Узагальнення усіх передбачень, отриманих за допомогою класичної частини моделі.

Також іншим важливим аспектом даного методу є можливість використання сучасних та ефективних архітектур згорткових нейронних мереж у якості класичної частини гібридної моделі. Та окрім використання SOTA архітектур, що не менш важливо, описана архітектура гібридної моделі дозволяє

залучати також різноманітні просунуті техніки, такі як трансферне навчання (transfer learning) [93].

Діаграму, що описує високорівневу архітектуру гібридних моделей, побудованих за допомогою даного методу гібридизації, наведено на Рисунку 3.3.

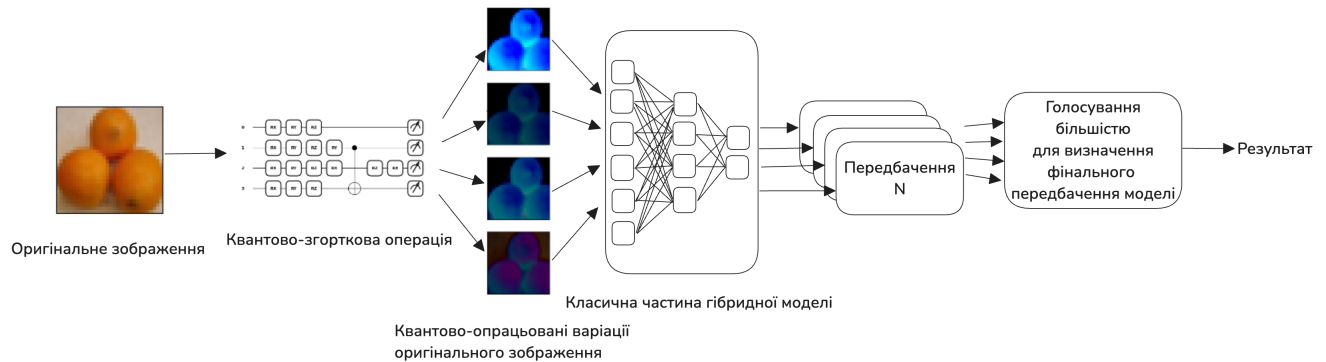


Рисунок 3.3. Архітектура гібридної нейронної мережі, що використовує квантово-згортковий шар [94]

Тож можна підсумувати, що гібридну модель ШІ, побудовану використовуючи даний метод гібридизації, також можна розділити на дві окремі частини:

- Квантова обробка вхідного зображення за допомогою квантового пристрою.
- Використання результатів квантової обробки зображення класичною частиною моделі.

Дане розділення приносить всі переваги, згадані в розділі 3.3. «Попередня квантова обробка даних» такі як можливість розділити в часі процес попередньої обробки даних, можливість провести обробку вхідних даних за допомогою квантового пристрою лише один раз та не повторювати її на кожній епісі навчання та можливість виконувати квантову обробку зображень з необмеженим паралелізмом. Проте використання кожного з каналів результатів квантової обробки даних як окремого RGB зображення приносить кілька додаткових серйозних переваг:

- Можливість залучати SOTA моделі як класичну частину гібридної нейронної мережі.
- Можливість ефективного залучення просунутих технік, що використовуються для покращення результатів роботи класичних нейронних мереж, яскравим прикладом такого підходу є трансферне навчання [93].
- Гнучкість інтерпретації результатів роботи моделі за рахунок використання підходів, таких як голосування за визначення фінального результату MV.

3.5. Багатоструктурна класична частина гібридної квантово-згорткової нейронної мережі

Даний підхід розширює попередньо описаний метод гібридизації, що базується на квантовому пристрої, що виконує роль першого квантового згорткового шару гібридної глибокої нейронної мережі. При використанні квантово-згорткової операції, на основі одного вхідного зображення формуються додаткові канали зображення та з огляду на те, що кожен з додаткових каналів квантово-опрацьованих варіацій оригінального зображення має специфічні відмінності та особливості, раціональним розвитком ідеї використання квантово-згорткової операції у якості першого шару глибокої нейронної мережі, є використання багатоструктурної моделі у якості класичної частини гібридної моделі та замість однієї класичної моделі, використати N нейронних мереж (де N рівний кількості каналів, створених в результаті квантово-згорткової операції), кожна з яких опрацьовує лише один з нових каналів зображення [95]. Архітектурну діаграму підходу, що використовує багатоструктурну архітектуру класичної частини гібридної квантово-згорткової нейронної мережі, наведено на Рисунку 3.4.

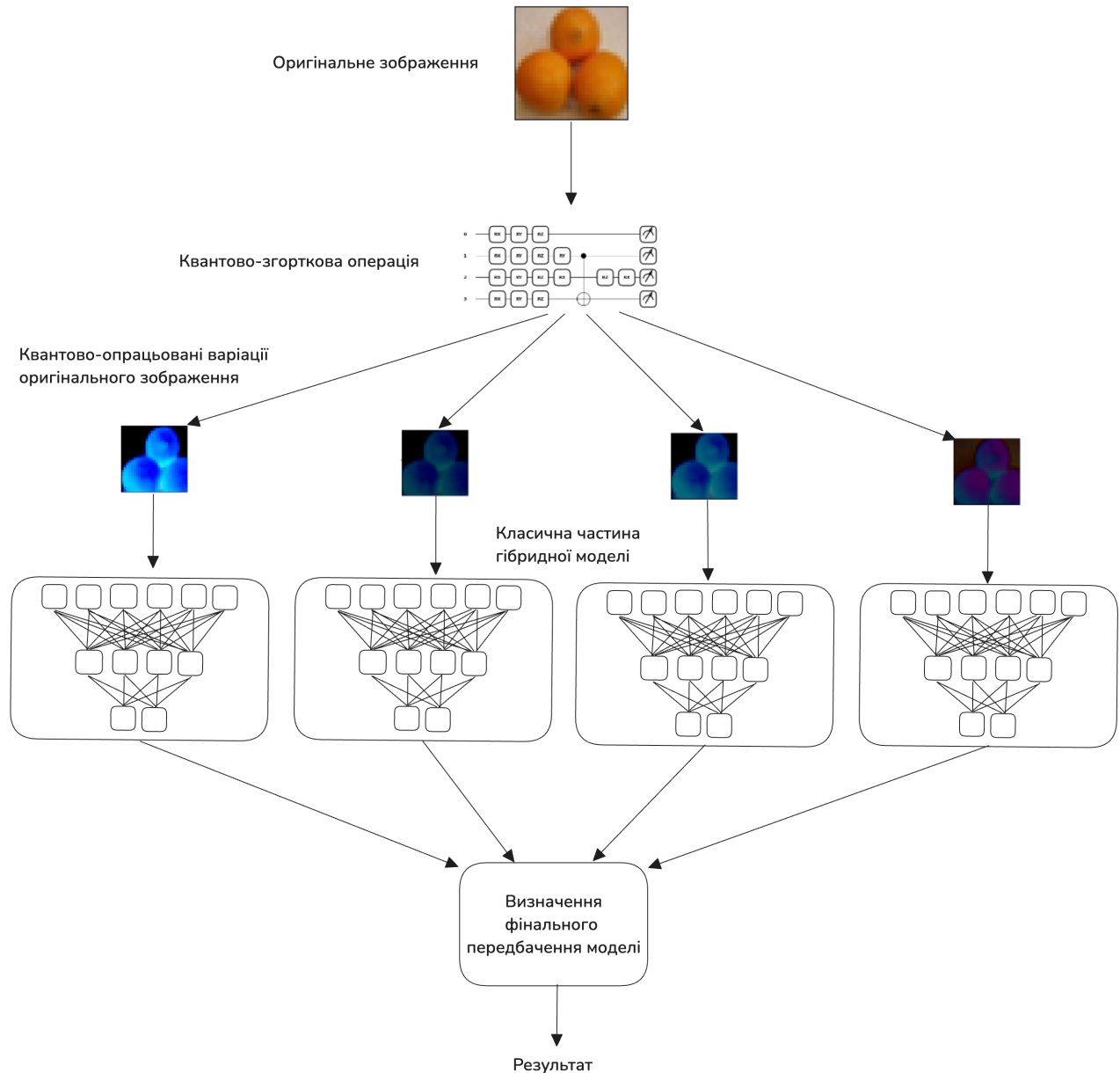


Рисунок 3.4. Архітектура гібридної нейронної мережі, що використовує квантово-згортковий шар та багатоструктурну класичну частину моделі.

Також дану ідею можна розвинути та додати ще один додатковий вхідний канал класичної частини гібридної моделі та передавати оригінальне зображення в даний додатковий канал. Архітектурну діаграму підходу, що містить $N+1$ базових мереж, наведено на Рисунку 3.5.

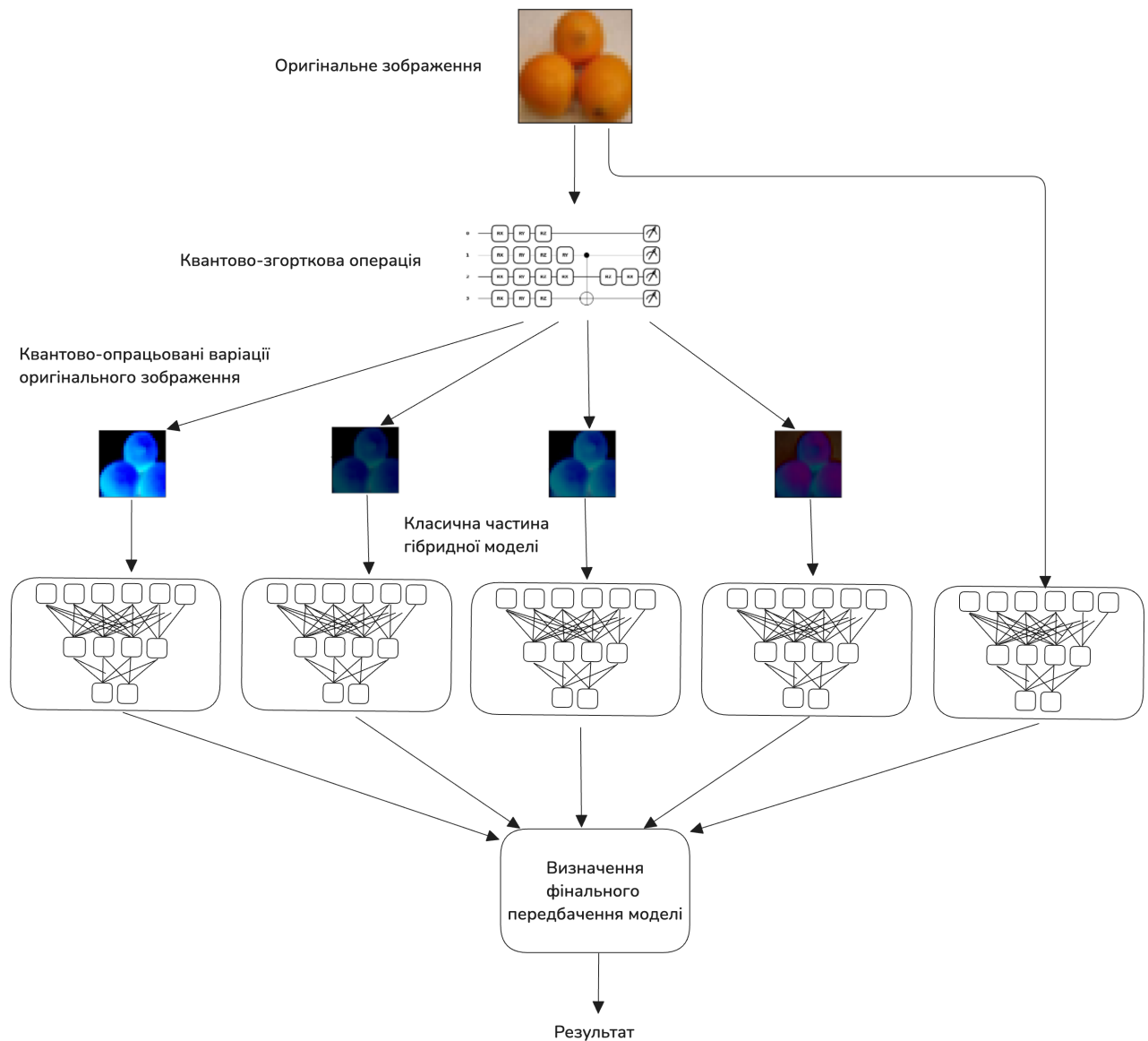


Рисунок 3.5. Архітектура гібридної нейронної мережі,
що використовує квантово-згортковий шар та
багоструктурну класичну частину моделі
з додатковим каналом для оригінального зображення.

Порівнявши діаграми наведені на Рисунках 3.4 та 3.5 з діаграмою, описаного вище підходу на Рисунку 3.3, можна помітити, що загальна ідея підходу повторюється, з єдиною важливою відмінністю – замість однієї моделі у якості класичної частини гібридної мережі використовується багато моделей, кожна з яких працює виключно з одним з каналів даних. Даний підхід успадковує всі переваги попереднього підходу, проте пропонує потенційно вищу точність роботи гібридної моделі за рахунок більш вузької спеціалізації кожної з частин

багатоструктурної класичної частини моделі та більшої гнучкості підходу, що дозволяє також розширення підходу за рахунок залучення додаткового вхідного каналу даних – оригінального не зміненого зображення. Проте порівняно з попереднім підходом, даний підхід також має суттєвий недолік – залучення кожної додаткової базової моделі до класичної багатоструктурної частини гібридної нейронної мережі збільшує об'єм пам'яті, що необхідний для опрацювання даних, що у свою чергу вимагає використання більш потужного апаратного забезпечення та збільшує вартість використання даного підходу.

Тож можна підсумувати, що гібридна модель ШІ, що побудована на основі квантового пристрою, що виступає у ролі першого згорткового шару глибокої нейронної мережі та має багатоструктурну класичну частину моделі, має всі переваги підходу гібридизації, що було розглянуто у розділі 3.4 «Квантовий пристрій як перший згортковий шар нейронної мережі», проте відрізняється тим, що:

- Дозволяє потенційно отримати вищу точність роботи моделі за рахунок використання багатоструктурності.
- Більша гнучкість підходу, що дозволяє залучати додаткові базові моделі, залежно від вимог конкретної задачі.
- Вимагає більш потужного класичного апаратного забезпечення, адже збільшуються вимоги до необхідного об'єму пам'яті.

3.6. Комплексний метод гібридизації класичних та квантових обчислень для завдань штучного інтелекту

На основі розглянутих методів гібридизації можна побудувати комплексний підхід до побудови гібридних квантово-класичних нейронних мереж. Ключовим фактором для побудови комплексного методу є вимоги конкретної задачі, яка вимагає залучення гібридних обчислень. Обидва основні розглянуті методи мають ряд своїх як переваг, так і суттєвих недоліків, тому

комплексний метод полягає у використанні того чи іншого підходу гібридизації, залежно від вимог задачі з ціллю максимального використання його переваг та можливості змиритись з недоліками.

Основною перевагою, що має метод гібридизації, що полягає в заміні частини прихованих шарів нейронів класичної глибокої нейронної мережі на квантовий пристрій, є прискорення обчислень за рахунок перенесення частини обчислень на квантове апаратне забезпечення. Проте глибше дослідження даного методу гібридизації показало, що цей підхід також знижує максимальну точність, яку модель може досягнути в результаті тренування. Процес дослідження та детальні його результати описано в межах розділу 4.

Перевагою методу гібридизації, що передбачає залучення квантово-згорткової операції є помітне збільшення точності роботи моделі, яке досягається за рахунок використання квантових обчислень для штучного розширення набору даних та створення додаткових каналів. Проте дослідження показало, що даний метод є відносно дорогим з точки зору кількості класичних обчислень, яка помітно зростає при використанні цього підходу, що в свою чергу спричиняє збільшення часу роботи моделі. Процес дослідження та детальні результати експериментів також наведено в межах розділу 4.

Таким чином, на основі сильних та слабких сторін кожного з розглянутих методів гібридизації можна зробити висновок, що для задач, що вимагають надвисокої швидкодії, проте можуть толерувати дещо нижчий рівень точності роботи моделі підходить використання методу гібридизації, що полягає в заміні частини прихованих шарів нейронів класичної глибокої нейронної мережі на квантовий пристрій. Для задач, що вимагають підвищеного рівня точності роботи моделі та толерантні до зниження швидкодії роботи моделі пропонується використання методу гібридизації нейронних мереж, що полягає в залученні квантово-згорткової операції в якості першої операції в межах роботи моделі і з можливістю залучення багатоструктурної класичної частини гібридної глибокої мережі, якщо цього дозволяють потужності класичної апаратної частини. Особливо ефективним даний метод є для задач, що мають субоптимальний розмір тренувального набору даних, адже штучне квантове розширення набору

тренувальних даних показало себе ефективним механізмом для часткового подолання даної проблеми.

ВИСНОВКИ ДО РОЗДІЛУ 3

В даному розділі дисертаційного дослідження було розглянуто 3 різні підходи з гібридизації глибоких згорткових нейронних мереж для вирішення задачі багато-класової класифікації зображень.

1. Квантовий пристрій виконує роль одного з прихованих шарів глибокої згорткової нейронної мережі.
2. Квантовий пристрій виконує роль попередньої квантової обробки зображень до їх потрапляння у класичну частину гібридної нейронної мережі.
3. Квантовий пристрій виконує роль першого квантово-згорткового шару глибокої згорткової нейронної мережі.

Перший з підходів наразі є найбільш поширеним підходом гібридизації та був розглянутий як частина комплексного методу гібридизації класичних та неklasичних обчислень. Даний підхід дозволяє збільшити швидкодію систем ШІ за рахунок зменшення глибини класичної частини гібридної нейронної мережі шляхом заміни частини прихованих шарів мережі на квантову схему. Проте потенційно даний підхід може призвести до зниження точності роботи гібридної моделі.

Підхід, що базується на попередній квантовій обробці зображень, є відомим методом покращення ефективності роботи гібридних квантово-класичних нейронних мереж та він отримав подальший розвиток в межах даної роботи. Даний підхід потенційно дозволяє розширити тренувальний набір даних та таким способом позитивно вплинути на стійкість моделі та збільшити точність роботи моделі для задач, що мають субоптимальний розмір тренувального набору даних.

Було запропоновано новітній підхід гібридизації глибоких нейронних мереж за рахунок використання квантового пристрою у якості першого квантово-згорткового шару глибокої гібридної квантово-класичної згорткової нейронної мережі та використання підходу MV для визначення фінального

передбачення моделі. Даний підхід дозволяє потенційно підвищити точність роботи моделей ШІ за рахунок створення додаткових каналів вхідних даних та гнучкого механізму конфігурації механізму визначення фінального передбачення моделі за рахунок використання технік голосування для встановлення результату роботи моделі (для прикладу, MV).

Було запропоновано розширення підходу гібридизації глибоких нейронних мереж, що будуються на основі квантово-згорткової операції, за рахунок використання багатоструктурної класичної частини гібридної моделі. Дане розширення дозволяє потенційно досягати вищих рівнів точності роботи моделі за рахунок залучення додаткових класичних обчислювальних ресурсів.

Також було запропоновано комплексний метод гібридизації квантових та класичних обчислень для збільшення ефективності роботи систем ШІ, що базується на використанні одного з підходів залежно від вимоги задачі. Якщо задача потребує надвисокої швидкодії та є толерантною до зниження точності роботи моделі – пропонується застосування підходу із заміною частини прихованих шарів глибокої нейронної мережі на квантовий пристрій. Якщо ж задача потребує підвищення точності класифікації, є толерантною до зниження швидкодії та має субоптимальний розмір тренувального набору даних – пропонується застосування квантового пристрою у якості першого квантово-згорткового шару гібридної моделі.

РОЗДІЛ 4

АНАЛІЗ РЕЗУЛЬТАТІВ МЕТОДІВ ГІБРИДИЗАЦІЇ КЛАСИЧНИХ ТА НЕКЛАСИЧНИХ ОБЧИСЛЕНЬ ДЛЯ ПОБУДОВИ ГІБРИДНИХ НЕЙРОННИХ МЕРЕЖ

4.1. Аналіз впливу архітектури квантового пристрою на ефективність роботи гібридних квантово-класичних нейронних мереж

Експерименти з дослідження впливу архітектури квантового пристрою на метрики роботи гібридних квантово-класичних нейронних мереж, побудованих за допомогою підходу гібридизації, що включає квантовий пристрій як один з прихованих шарів глибокої нейронної мережі, були проведені, базуючись на двох найпростіших наборах даних, що були використані у дисертаційному дослідженні – MNIST та MNIST-Fashion. Для даних експериментів було обрано прості набори даних, адже дані експерименти можна вважати першими кроками дослідження, оскільки від їх результатів залежить подальший хід дослідження та архітектури гібридних нейронних мереж, що будуть використані для вирішення більш складних та комплексних задач.

Спочатку розглянемо архітектуру глибокої гібридної квантово-класичної згорткової нейронної мережі. Експериментальна мережа має наступну структуру шарів:

1. Два згорткові шари.
2. Шар випадкових виключень (dropout).
3. Два лінійних шари.
4. Квантовий пристрій.
5. Вихідний лінійний шар.

Згорткові шари здійснюють перетворення вхідного зображення у вектор із 320 елементів, обидва шари використовують функцію активації ReLU [96]. Шар випадкових виключень застосовується для боротьби з перенавчанням моделі, виключаючи деякі випадкові зв'язки між згортковими та лінійними шарами. Два

лінійні шари виконують зменшення розмірності даних з вектору розміром 320 елементів до вектору, що містить кількість елементів, що відповідає кількості кубітів у квантовому пристрої. Перший з лінійних шарів використовує функцію активації ReLU, а другий — функцію активації тангенс. Тангенс був обраний як функція активації через його властивість обмежувати значення параметрів в інтервалі $(-1; 1)$. Перед подачею даних у квантовий пристрій всі значення множаться на π , оскільки варіаційна частина квантової схеми є квантовими операціями, що виконують ротацію кубітів на певну задану кількість радіанів та навколо певної осі. Отже, виконується ротація кубітів у квантовій схемі у діапазоні $(-\pi; \pi)$ радіан. Завершальний шар нейронної мережі — це лінійний шар, який перетворює вихідні дані з квантового пристрою на результат класифікації. Цей шар використовує функцію активації softmax [97] для перетворення результатів заміру стану кубітів в фінальний результат класифікації [98].

Для інтерпретації результатів роботи квантового пристрою було використано два наступних підходи.

Перший з підходів класичний – значення заміру стану кожного кубіту відповідає одному результуючому класу і результатом класифікації є клас з порядковим номером кубіту з найбільшим значенням після виконання функції активації. Таким чином алгоритм обрання фінального результату класифікації наступний:

1. Виконання квантової схеми та замір результуючого стану кубітів.
2. Виконання функції активації над результатами заміру стану кубітів.
3. Обрати порядковий номер кубіта з найбільшою ймовірністю результату «1» та результуючий клас є відповідним за порядком.

Альтернативним підходом до визначення фінального передбачення класу гібридною квантово-класичною нейронною мережею було використано підхід з «двійковим кодуванням» результатів. В даному підході результати індивідуальних кубітів представляють собою число у двійковій системі числення. Алгоритм даного підходу можна описати наступним чином:

1. Виконання квантової схеми та замір результуючого стану кубітів.

2. Виконання функції активації над результатами заміру стану кубітів.
3. Обрати порядковий номер класу залежно від найбільш ймовірного двійкового числа, яке є результатом роботи квантової схеми.

Другий підхід було обрано з огляду на те, що потужності сучасних квантових комп'ютерів та кількість кубітів, які можливо виготовити на даному етапі розвитку квантових комп'ютерів, є обмеженими. Тому другий підхід хоч і є складнішим, проте дозволяє сильно заощадити на кількості кубітів, що необхідні для вирішення задачі класифікації. Розмірність задачі класифікації, яку можна вирішити, використовуючи квантову схему, що містить задану кількість кубітів, можна виразити наступною формулою:

$$N = 2^{n-q} \quad (4.1)$$

Де N – максимальна розмірність задачі класифікації (кількість класів класифікації), $n-q$ – кількість кубітів, що використовується в квантовій схемі.

Даний експеримент було проведено для наступних варіантів квантових схем:

- RY;
- Квантовий алгоритм оптимізації апроксимації (QAOA);
- Квантове перетворення Фур'є;
- Квантова схема стану Белла, що реалізовує квантову заплутаність.

Та для наступних варіацій кількості класів класифікації:

- 2 класи;
- 4 класи;
- 6 класів;
- 8 класів;
- 10 класів.

Даний вибір параметрів експерименту не випадковий, адже вибір розмірності квантової схеми повністю залежить від кількості класів, що потрібно класифікувати за допомогою гібридної нейронної мережі, що включає дану квантову схему.

Для оцінки ефективності різноманітних архітектур квантового пристрою та для отримання референтного значення було використано також модель ШІ класичної архітектури. Архітектура референтної моделі повністю імітує архітектуру гібридних квантово-класичних моделей з єдиною важливою відмінністю – квантовий пристрій замінено на шар класичних штучних нейронів.

Рисунок 4.1 демонструє графік функції втрат та рисунок 4.2 демонструє графік точності класифікації 4-х класів зображень з MNIST-Fashion набору даних гібридної квантово-класичної моделі, що містить RY квантову схему, що складається з двох кубітів. З даних графіків можна помітити, що процес навчання гібридної моделі сходиться та алгоритм зворотного поширення помилки працює вірно для квантової схеми, та, відповідно, процес навчання гібридної мережі також працює вірно.

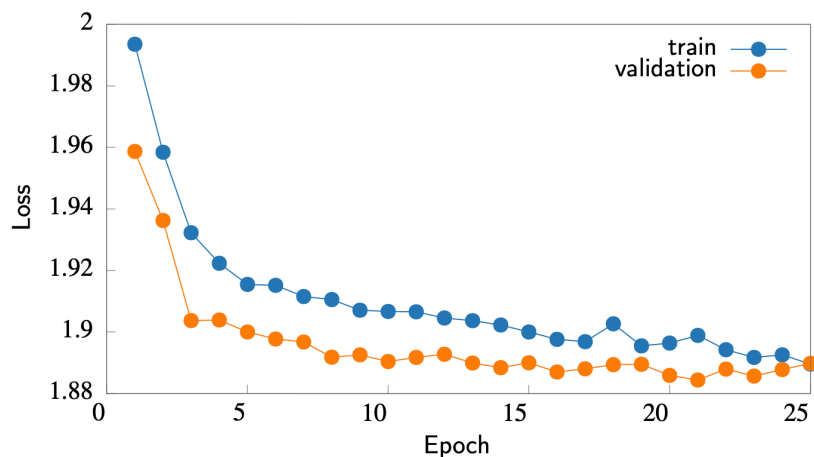


Рисунок 4.1. Значення функції втрат під час тренування гібридної моделі на наборі даних MNIST-Fashion [98].

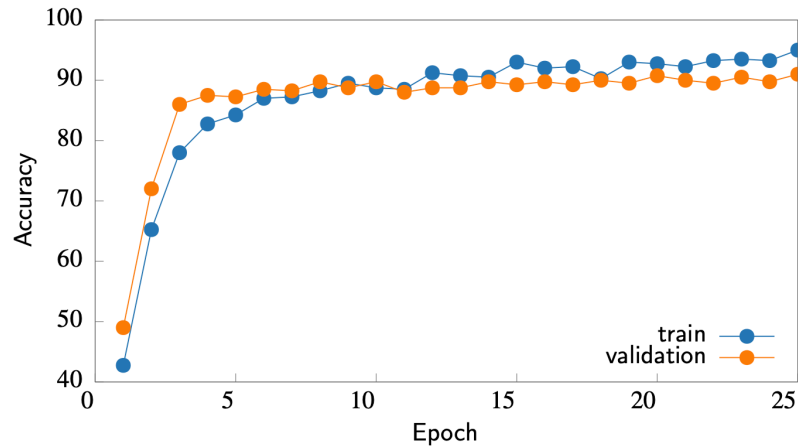


Рисунок 4.2. Точність класифікації під час тренування гібридної моделі на наборі даних MNIST-Fashion [98].

Результати експерименту з вирішення задачі класифікації 2-х класів набору даних MNIST за допомогою гібридних квантово-класичних нейронних мереж з різними архітектурами квантового рівня наведено в таблиці 4.1. З результатів можна помітити, що наступні архітектури квантового пристрою є ефективними для їх використання в контексті вирішення задачі класифікації: RY, QAOA та квантова схема, що реалізує стан Белла. В той же час, архітектура QFT виявилась неефективною. З огляду на результати експерименту було прийнято рішення проводити подальші дослідження та експерименти, використовуючи найефективнішу архітектуру – RY квантову схему, яка є не лише найефективнішою, але і найпростішою.

Таблиця 4.1. Точність класифікації 2 класів набору даних MNIST [98].

Квантова схема	Точність класифікації
RY, 2 кубіти	100%
QAOA, 2 кубіти	98.89%
QFT, 1 кубіт	57%
Квантова схема Белла	99.41%

Таблиця 4.2 ілюструє результати експерименту з вирішення задач класифікації різної розмірності за допомогою гібридних квантово-класичних

нейронних мереж, що побудовані на основі квантової схеми RY архітектури. З результатів експерименту можна зробити висновок, що даний підхід може бути робочим та, теоретично, практичним в різних сценаріях, які вимагають переваги швидкості над точністю роботи мережі.

Таблиця 4.2. Точність класифікації MNIST та MNIST-Fashion [98].

Квантова схема	Кількість класів	Точність на наборі даних MNIST	Точність на наборі даних MNIST-Fashion
RY, 2 кубіти	4	96%	95.6%
RY, 3 кубіти	6	91.67%	59.48%
RY, 3 кубіти	8	87.75%	73.92%
RY, 4 кубіти	10	65.13%	65.48%

Таблиця 4.3 доповнює попередні результати та ілюструє метрики перехресної перевірки та результати експерименту демонструють задовільну стійкість моделі до зміни тренувальних та тестових даних.

Таблиця 4.3. Метрики перехресної перевірки MNIST-Fashion [98].

Квантова схема	Кількість класів	Середня точність класифікації	Стандартне відхилення точності класифікації
RY, 2 кубіти	4	95.6%	0.87%
RY, 3 кубіти	6	59.48%	9.37%
RY, 3 кубіти	8	73.92%	3.01%
RY, 4 кубіти	10	65.48%	6.57%

4.2. Аналіз методу гібридизації нейронних мереж, використовуючи квантовий пристрій як один з прихованих шарів нейронної мережі

Для експерименту, що досліджує ефективність та практичність використання методу гібридизації класичних та некласичних обчислень, що полягає в використанні квантового пристрою як одного з прихованих шарів глибокої гібридної згорткової нейронної мережі, було використано набори даних CIFAR10 та SIHD.

В даному експерименті CIFAR10 використано як приклад класичної задачі класифікації зображень помірної складності, адже набір даних містить кольорові зображення відносно низької роздільної здатності, що робить їх проблематичними для класифікації між класами, що містять візуально схожі риси.

В той же час, набір даних SIHD представляє практичну задачу класифікації зображень та є дещо простішим набором даних за своїми характеристиками, адже він містить лише 2 класи (пошкоджені будівлі та непошкоджені будівлі) та зображення мають достатньо високу роздільну здатність.

Для проведення даного експерименту було побудовано глибоку гібридну нейронну мережу, що має архітектуру майже ідентичну до гібридної моделі, що була використана в розділі 4.1, а саме:

1. Два згорткові шари.
2. Шар випадкових виключень (dropout).
3. Два лінійних шари.
4. Квантовий пристрій.
5. Вихідний лінійний шар.

Функції активації шарів ідентичні до відповідних частин моделі, використаної в розділі 4.1. Проте дещо змінена конфігурація перших згорткових

шарів моделей, для того щоб адаптувати їх до формату зображень з використаних наборів даних.

У якості квантового пристрою було використано квантову схему, що проводить конфігуровані RY операції над кубітами та містить 1-4 кубіти в цілому, залежно від конкретної кількості класів, що були класифіковані в тому чи іншому конкретному експерименті. Архітектура квантового пристрою була обрана з огляду на результати попереднього експерименту, в межах яких гібридні моделі, що будуються на квантових пристроях RY архітектури, показали найкращі результати точності класифікації [98].

У якості референтних моделей було використано класичні глибокі нейронні мережі, що мають ідентичну архітектуру, порівняно з гібридними квантово-класичними моделями, але з шаром класичних штучних нейронів, що замінюють квантовий пристрій.

Також додатково для оцінки реальної практичності підходу результати роботи гібридної моделі на наборі даних SIHD було також порівняно з SOTA моделями для цього набору даних [99]. У цій якості було обрано VGG-16 [100] та ResNet-50 [101]. Дані моделі було обрано, оскільки вони показали найкращу точність класифікації на даному наборі даних та, отже, являються поточними SOTA моделями для набору даних SIHD.

VGG-16 це класична глибока згорткова нейронна мережа, вона має 13 згорткових шарів, 3 лінійні шари нейронів та має відносно просту «пірамідальну» архітектуру.

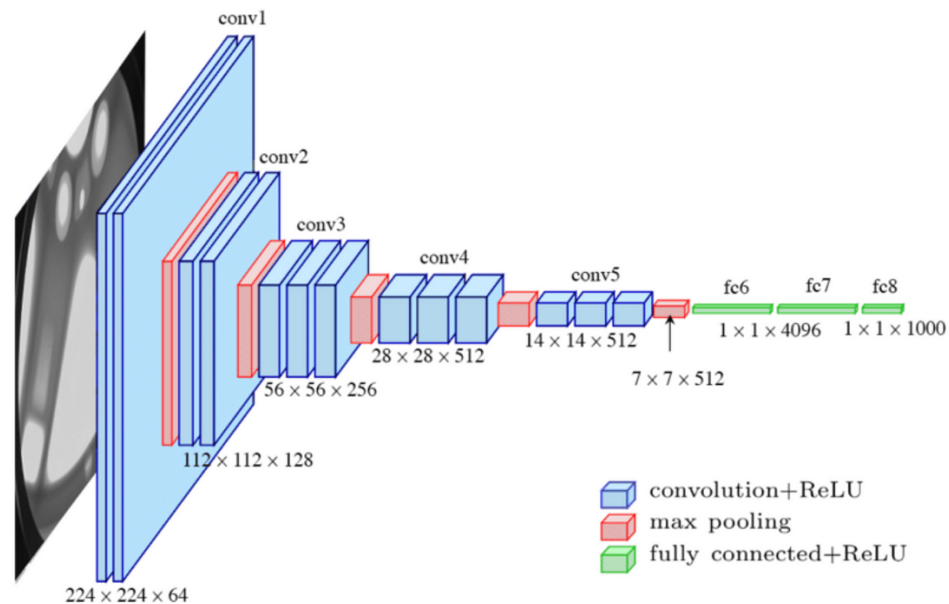


Рисунок 4.3. Діаграма архітектури моделі VGG-16 [100].

ResNet-50 – це також широковідома класична модель, архітектура якої базується на «залишкових блоках» (residual block). Рисунок 4.4 наводить діаграму залишкового блоку. Модель ResNet-50 складається з 46 згорткових шарів, які в свою чергу є залишковими блоками, та з одного лінійного шару нейронів.

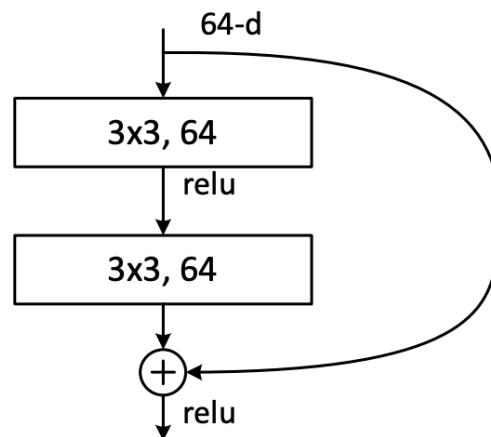


Рисунок 4.4. Діаграма залишкового блоку моделі ResNet-50 [101].

Отже, в межах експерименту було проведено тренування гібридних квантово-класичних нейронних мереж та референтних класичних моделей на наборах даних CIFAR10 та SIHD та оцінено ефективність їх роботи з огляду на максимальний рівень точності, який змогли досягнути моделі. Точність

класифікації було оцінено, використовуючи перехресну перевірку (cross-validation), тому, окрім самої метрики точності, вдалось дослідити також і її стійкість та залежність від конкретного під-набору вхідних даних, що були використані для тренування.

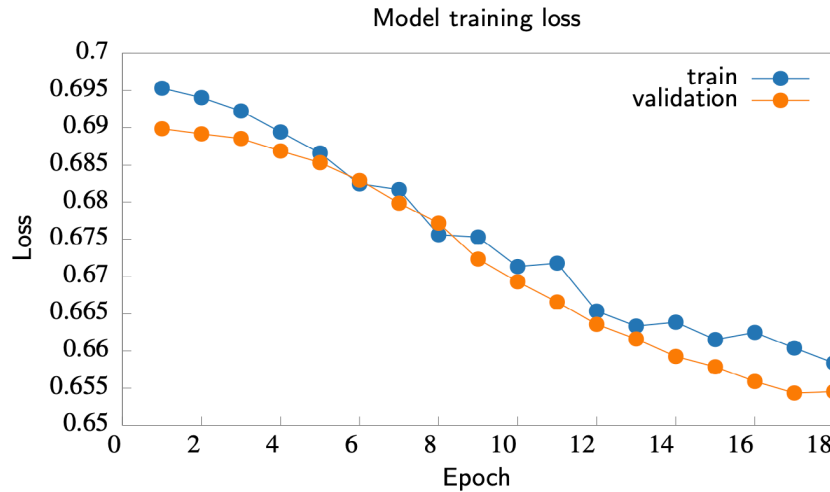


Рисунок 4.5. Графік функції втрат під час тренування гібридної моделі на наборі даних SIHD [102].

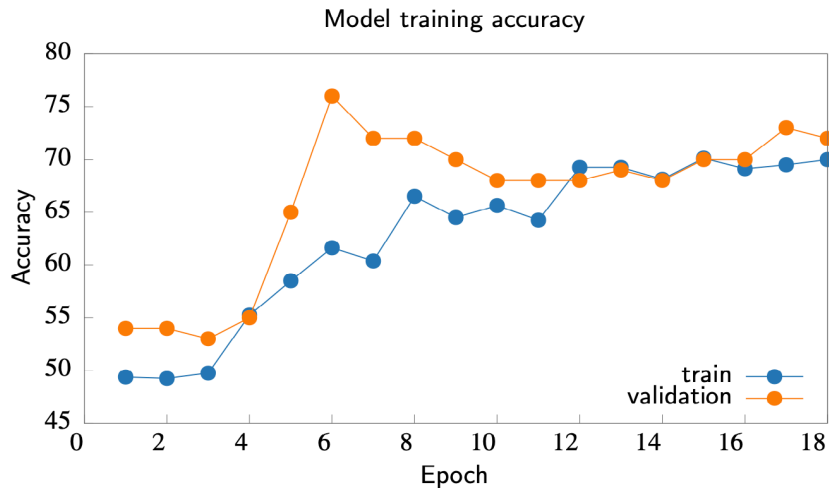


Рисунок 4.6. Графік точності гібридної моделі під час тренування на наборі даних SIHD [102].

Рисунки 4.5 та 4.6 наводять графіки процесу навчання гібридної квантово-класичної моделі. На графіках можна побачити, що процес навчання відбувається та модель навчається, покращуючи свої метрики точності. Таблиця

4.4 наводить результати експерименту класифікації зображень набору даних CIFAR10 та порівняно результати гібридної моделі та референтної моделі.

Таблиця 4.4. Точність класифікації зображень набору даних CIFAR10 [103].

Кількість класів	Точність гібридної моделі	Точність референтної моделі
4	79.6%	67.4%
6	52%	76.4%
8	56.58%	71.2%
10	45.13%	70.1%

Результати експерименту на наборі даних CIFAR10 показали, що точність роботи гібридної моделі помітно нижча, ніж точність референтної моделі та точність значно знижується зі збільшенням розмірності задачі.

Таблиця 4.5 наводить результати експерименту з бінарної класифікації зображень з набору даних SIHD за допомогою гібридної моделі та рядом референтних моделей. Точність роботи гібридної моделі очікувано нижча, порівнюючи з референтними моделями. Проте відставання складає лише 3% від референтної моделі.

Таблиця 4.5. Точність класифікації зображень набору даних «Hurricane Damage» [102].

Модель	Точність
Гібридна модель	72.75%
Референтна класична модель	76%
VGG-16	89%
ResNet-50	88%

З результатів експериментів можна зробити висновок про практичність використання методу гібридизації класичних та некласичних обчислень на

основі заміни прихованих рівнів глибокої нейронної мережі на квантовий пристрій. Даний підхід може бути корисним у контексті задач, що вимагають високої швидкодії та є толерантними до дещо нижчої точності роботи моделі.

4.3. Аналіз результатів попередньої квантової обробки даних

Для оцінки ефективності попередньої квантової обробки даних було проведено 4 експерименти, до кожного з яких було залучено одну гібридну квантово-класичну модель та 2 референтні класичні моделі. Всі експерименти були проведені на наборі даних CIFAR100.

Гібридна модель складається з двох частин:

1. Квантовий пристрій, що виконує попередню квантову обробку даних.
2. Класична нейронна мережа, що опрацьовує результати квантової обробки даних.

Квантова частина гібридної моделі являє собою квантово-згорткову операцію над вхідними даними, що залучає випадкову квантову схему, що складається з 4 кубітів та містить RX операції над кубітами та випадковий рівень квантової заплутаності між кубітами.

Рисунок 4.7 демонструє результати обробки даних за допомогою квантово-згорткової операції, що залучає випадкову квантову схему, що містить RX операції та має 4 кубіти. Після квантової попередньої обробки даних розмір набору даних збільшується в 4 рази при використанні квантового пристрою з 4 кубітами і в 9 разів при використанні схеми, що містить 9 кубітів. З проілюстрованого прикладу видно, що основні особливості об'єкта на зображенні залишаються збереженими після квантової попередньої обробки на всіх вихідних шарах, але загалом зображення зазнає значних змін.

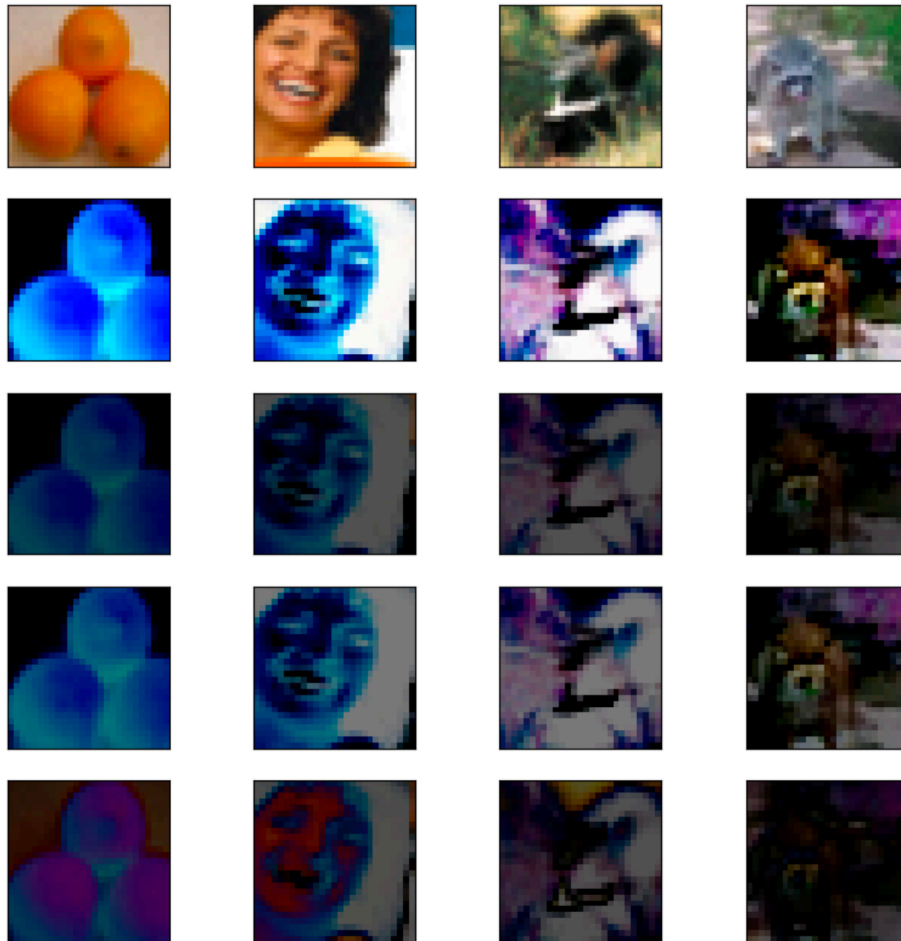


Рисунок 4.7. Приклад квантово-опрацьованих зображень набору даних CIFAR100 [104].

Рисунок 4.8 наводить діаграму класичної частини гібридної моделі. Як можна помітити, її архітектура надзвичайно проста та вона складається лише з двох частин:

1. Перетворення вхідного зображення в плоский вектор.
2. Один шар нейронів, що має кількість вхідних значень рівну розмірності вектору з кроку 1 та кількість вихідних значень рівну кількості класів для класифікації.

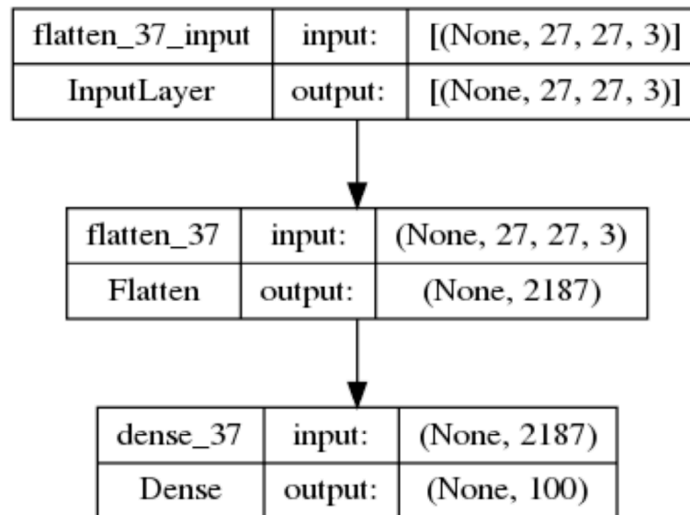


Рисунок 4.8. Діаграма класичної частини гібридної моделі [104].

У якості референтних моделей було використано дві класичні моделі. Архітектура першої з них майже повністю повторює архітектуру класичної частини гібридної моделі, проте з ключовою відмінністю – референтна модель працює не з квантово-опрацьованими даними, а з оригінальним зображенням. Рисунок 4.9 наводить архітектурну діаграму першої з референтних моделей.

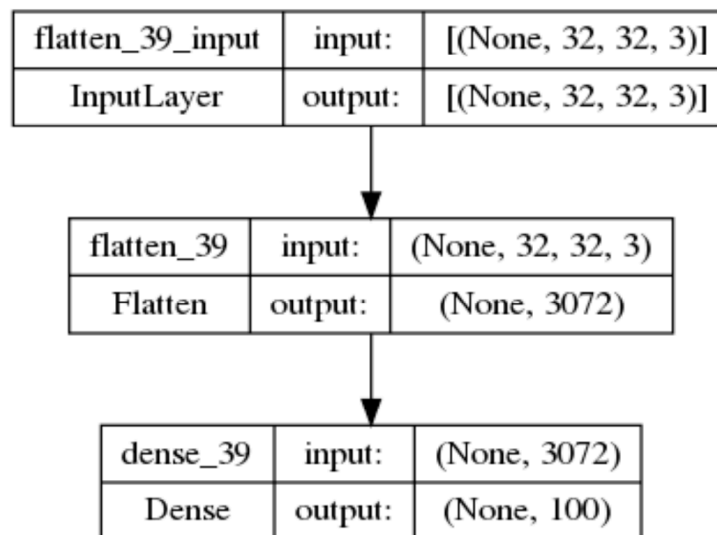


Рисунок 4.9. Діаграма першої референтної моделі [104].

У якості другої референтної моделі було обрано просту згорткову нейронну мережу, що складається з одного згорткового шару та одного лінійного шару нейронів. Проста згорткова нейронна мережа була обрана у якості референтної моделі, тому що згорткові мережі є де-факто стандартним рішенням

для вирішення різноманітних задач, пов'язаних з роботою із зображеннями, та демонструють значно вищу точність, порівнюючи зі звичайними глибокими нейронними мережами. Рисунок 4.10 наводить діаграму референтної згорткової нейронної мережі, що була використана в експериментах.

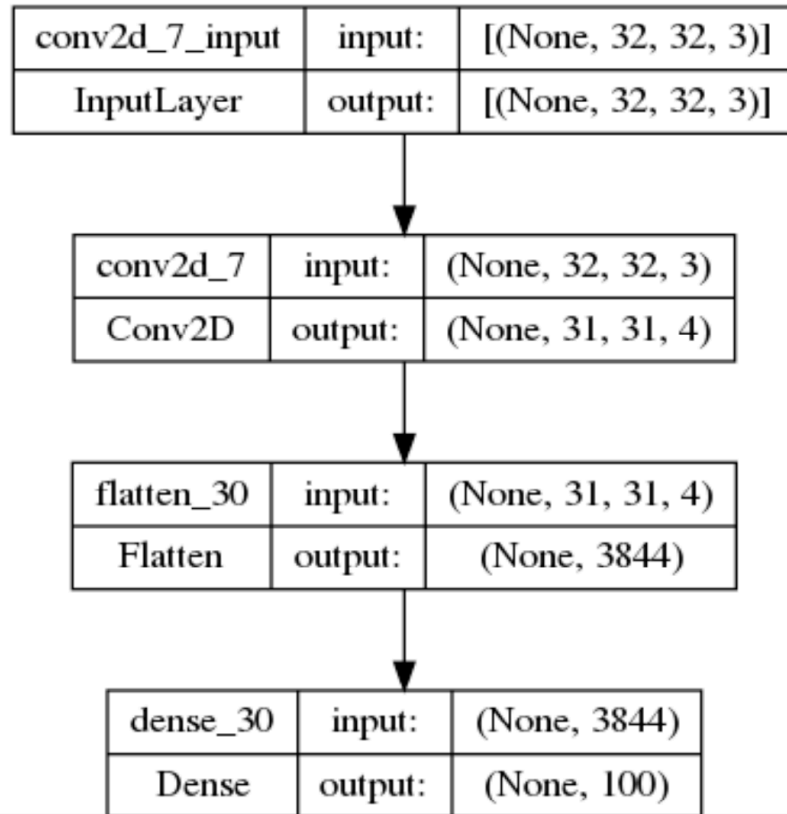


Рисунок 4.10. Діаграма другої референтної моделі [104].

У першому з експериментів було проаналізовано ефективність гібридної квантово-класичної моделі та проведено її порівняння з референтними моделями у відносно простому варіанті її застосування. Квантова частина опрацьовує кожне зображення з оригінального набору даних, і кожен вихідний канал квантово-згорткової операції розглядається як окрема картинка у новому наборі даних. Отже, кінцевий розмір набору даних для класичної частини гібридної моделі збільшується до 200000 картинок (оригінальний розмір набору даних CIFAR100 – 50000 картинок, помножених на 4, що є кількістю вихідних каналів із квантової частини моделі). Для референтних моделей жодних модифікацій не було застосовано, їх навчання проводилось на незмінених зображеннях з оригінального набору даних CIFAR100.

Для досягнення процесу класифікації в першому експерименті було використано наступні кроки:

1. Кожне вихідне зображення попередньо обробляється за допомогою квантової схеми.
2. Кожен канал з виходу квантового пристрою розглядається як окреме RGB-зображення, що призводить до створення нового набору даних із виходів квантової схеми. Отже, розмір результуючого набору даних у n разів більший за початковий набір даних, де n — це кількість кубітів у квантовій схемі.
3. Кожне зображення з попередньо обробленого набору даних використовується як вхід для класичної частини гібридної моделі.
4. Класична частина гібридної модулі генерує результат, який є прогнозом класу для кожного зображення.

Графіки метрик точності та втрат для даного експерименту наведені на Рисунку 4.11. З цих метрик можна помітити, що процес навчання гібридної моделі сходиться та гібридна модель досягає насичення навчання приблизно після 8 епох. В той же час, процес навчання референтних моделей є повільнішим, потребує більше епох для досягнення насичення, але, зрештою, забезпечує значно вищу точність. Цей покращений результат, ймовірно, пояснюється тим, що референтні моделі мають більшу кількість параметрів, що навчаються, що дозволяє їм вивчати складніші особливості та досягати кращої точності класифікації.

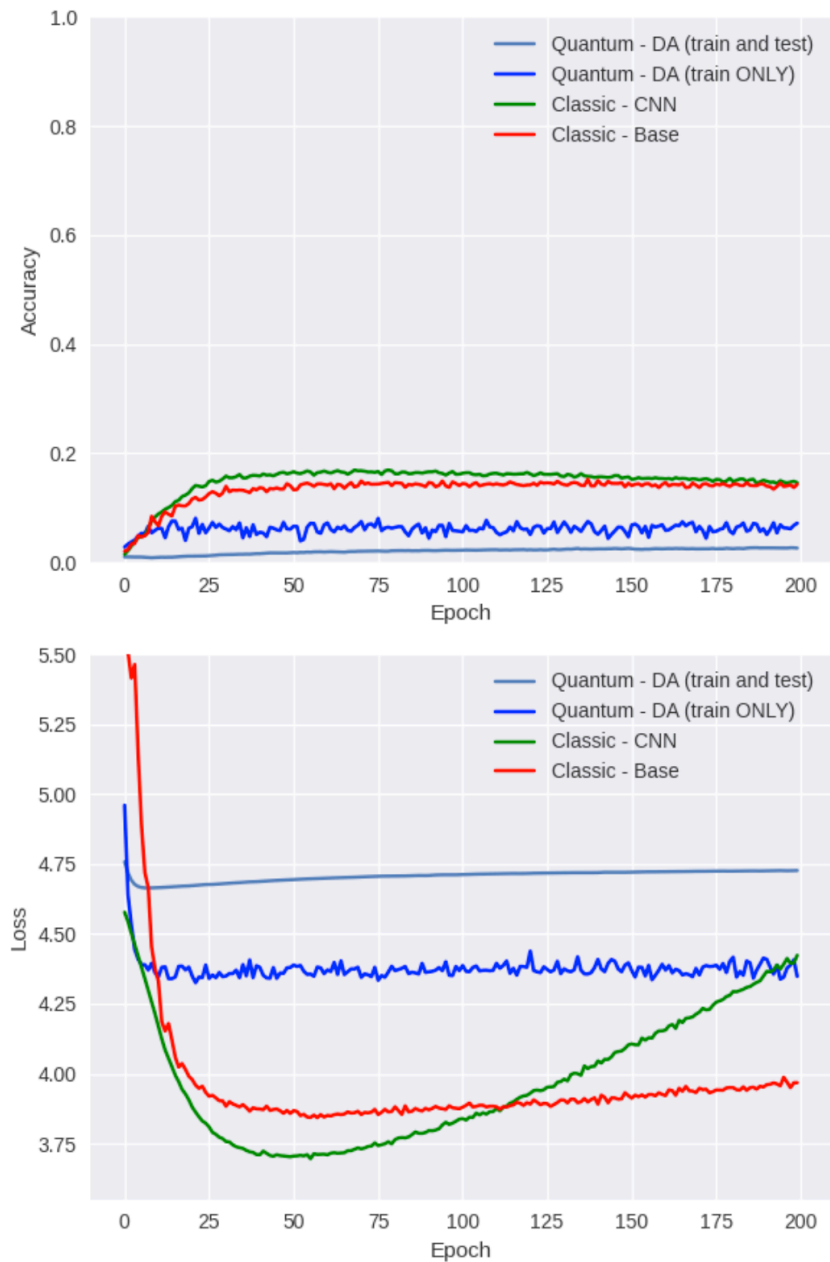


Рисунок 4.11. Графік результатів першого експерименту [104].

У другому експерименті глибину гібридної нейронної мережі та референтної згорткової нейронної мережі було збільшено шляхом додавання двох додаткових згорткових шарів напочатку. У такий спосіб, класична частина гібридної моделі містила два класичні шари, що обробляли квантово-опрацьовані дані, при цьому кожен додатковий канал зображення, створений квантовим пристроєм, розглядався як окремий елемент набору даних. Архітектуру референтної згорткової нейронної мережі також було розширено,

додавши два додаткові згорткові шари, таким способом отримавши три згорткових шари загалом, за якими слідує один лінійний повнозв'язний шар.

Метрики точності та втрат для другого експерименту представлені на Рисунку 4.12. Результати показують, що гібридна нейронна мережа значно перевершила референтну плоску нейронну мережу як за точністю, так і за показниками функції втрат. Крім того, гібридна нейронна мережа з другого експерименту продемонструвала значне покращення точності, порівнюючи з референтною класичною згортковою моделлю без додаткових згорткових шарів, що була використана у попередньому експерименті.

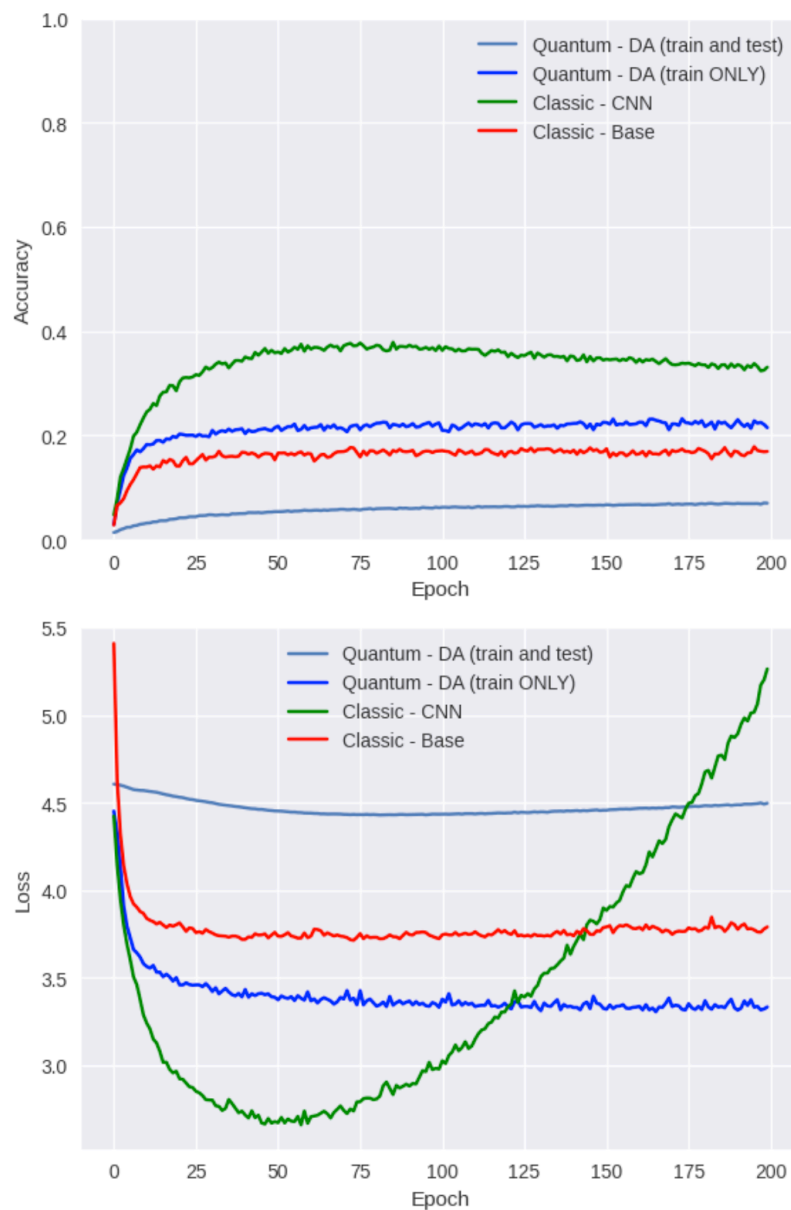


Рисунок 4.12. Графік результатів експерименту з складнішими архітектурами [104].

В межах третього експерименту згортковий шар було замінено на перетворення квантово опрацьованих варіантів кожного зображення в плоский вектор перед його подачею до повнозв'язного лінійного шару гібридної нейронної мережі. Ця модифікація збільшує кількість каналів з 3 (RGB) до 12 (4 виходи з QC, кожен з яких має 3 канали (RGB)). Референтні моделі залишилися незмінними.

Алгоритм роботи гібридної квантово-класичної нейронної мережі для третього експерименту:

1. Кожне вихідне зображення попередньо обробляється за допомогою квантового пристрою.
2. Усі результуючі канали з квантового пристрою перетворюються в єдиний вектор і отримані дані розглядаються як окремий елемент набору даних. Розмір набору даних зберігається.
3. Кожен елемент з попередньо обробленого набору даних використовується як вхід для класичної частини гібридної мережі.
4. Класична частина гібридної мережі продукує результат, що є прогнозом класу для вхідного зображення.

Результати третього експерименту, представлені на Рисунку 4.13, свідчать про те, що цей підхід є неефективним, оскільки процес навчання моделі не сходиться.

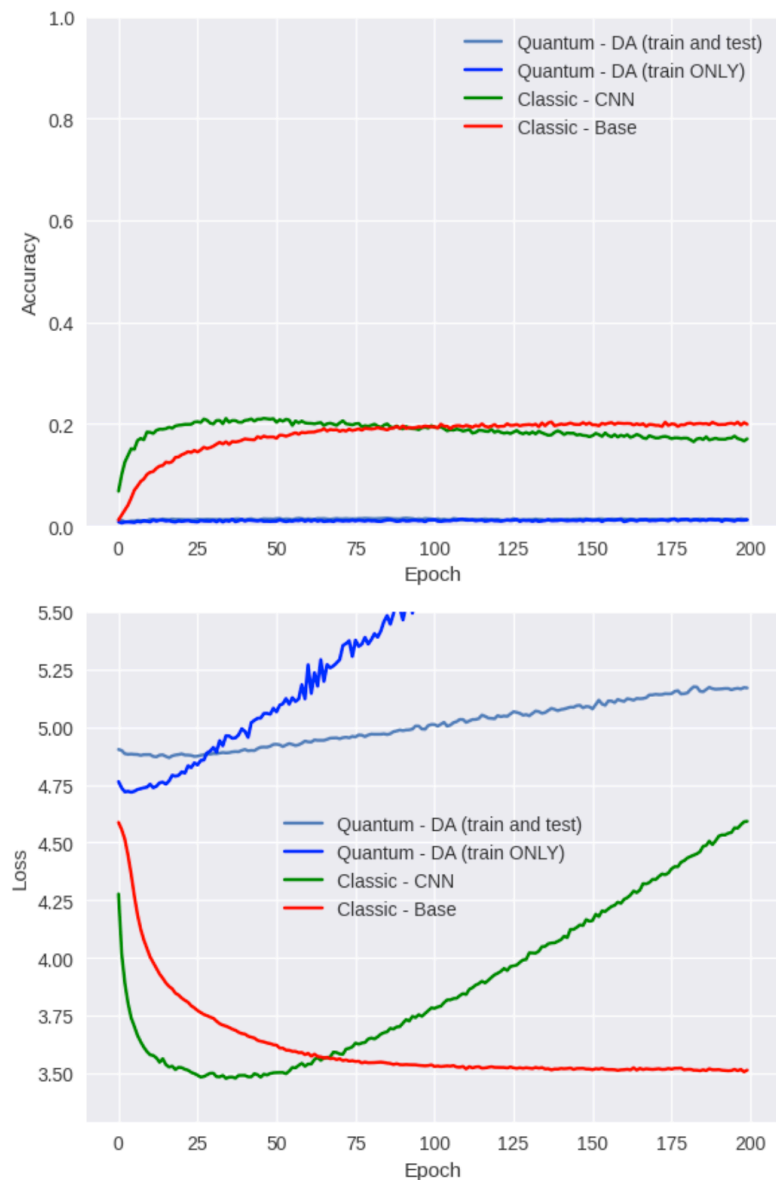


Рисунок 4.13. Графік результатів експерименту з 12-ма каналами [104].

В межах четвертого експерименту всі шари з результатів квантової операції були подані до повнозв'язного лінійного шару класичної частини гібридної нейронної мережі як паралельні входні канали. Референтні моделі залишились без змін.

Алгоритм роботи гібридної квантово-класичної нейронної мережі для четвертого експерименту:

1. Кожне вихідне зображення попередньо обробляється за допомогою квантового пристрою.

2. Усі результуючі канали з квантового пристрою розглядаються як додаткові канали нової точки даних. Розмір набору даних зберігається, проте кожна точка даних містить більше каналів, ніж у початковому наборі.
3. Кожен елемент з попередньо обробленого набору даних використовується як вхід для класичної частини гібридної мережі, при цьому кожен канал з попередньо обробленого набору даних використовується як паралельний вхід.
4. Класична частина гібридної мережі продукує результат, що є прогнозом класу для вхідного зображення.

Результати четвертого експерименту, представлені на рисунку 4.14, показали, що цей підхід є неефективним, оскільки процес навчання моделі не сходиться.

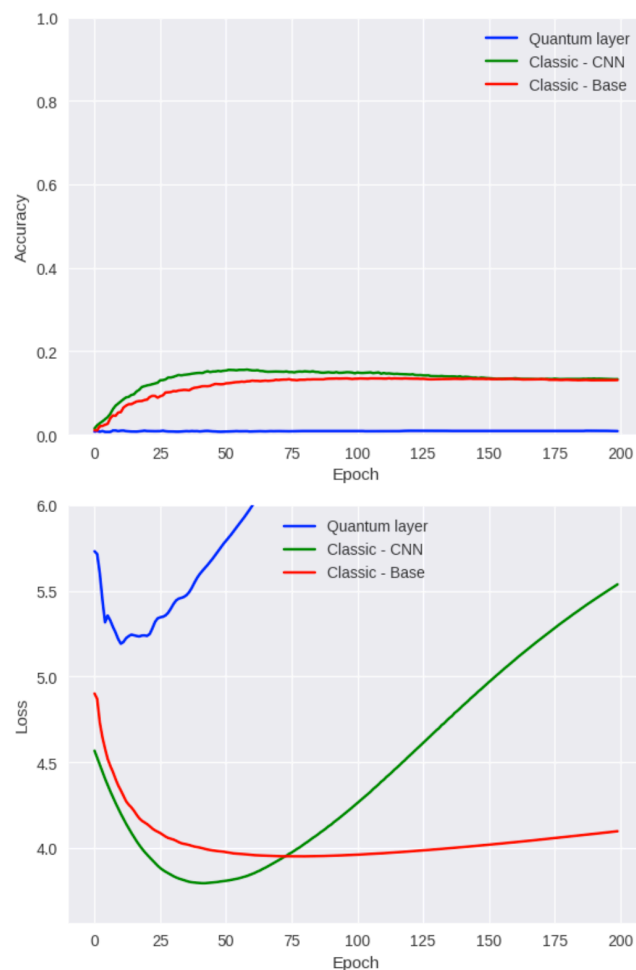


Рисунок 4.14. Графік результатів експерименту з додатковими паралельними каналами [104].

4.4. Аналіз методу гібридизації нейронних мереж, використовуючи квантовий пристрій як перший рівень нейронної мережі

В даному розділі буде розглянуто метод гібридизації, заснований на наступній ідеї: якщо використовувати квантовий пристрій як перший шар гібридної нейронної мережі, то для процесу навчання можна обробити кожен точку даних у навчальному наборі лише один раз. Це дозволить ефективніше використовувати ресурси квантового комп'ютера під час навчання. Такий підхід значно оптимізує процес навчання моделі, а також дає змогу застосовувати техніку трансферного навчання для гібридних нейронних мереж, вбудовуючи квантовий пристрій як перший шар у передові моделі. Це підвищує ефективність гібридних моделей і дає можливість використовувати перевірені архітектури класичних нейронних мереж для їх побудови. Ще однією важливою перевагою запропонованого методу гібридизації є те, що результати, отримані з квантового шару, можуть бути повторно використані багатьма різними архітектурами класичної частини гібридних моделей. Оскільки ці результати можна розглядати як еквівалент процесам попередньої обробки та збільшення даних з оригінального набору, їх можна використовувати як новий набір даних для повторного використання [94].

Для оцінки впливу гібридизації на відомі архітектури нейронних мереж було обрано та перевірено моделі ResNet [101], EfficientNet [105] і VGG-16 [100]. Вибір цих архітектур базувався на підході, що передбачає використання двох передових моделей для кожного набору даних, застосованого в цьому дослідженні, з подальшим порівняльним аналізом їхньої продуктивності та ефективності в контексті гібридних нейронних мереж, створених на основі цих архітектур.

Набір даних CIFAR100 є класичним тестовим набором для задач класифікації зображень, широко використовуваним у численних дослідженнях. Найсучаснішою моделлю для розв'язання задачі класифікації зображень на цьому наборі є EfficientNet, яка досягла класифікаційної точності 91,7% [105].

EfficientNet використовує метод сполучених коефіцієнтів для масштабування моделей та автоматизоване машинне навчання (AutoML). Другою широко застосовуваною моделлю для набору даних CIFAR100 є ResNet, яка продемонструвала точність 77,3%. ResNet є глибокою згортковою нейронною мережею, що використовує залишкові блоки для підвищення точності, порівнюючи зі стандартними згортковими нейронними мережами.

Набір даних SIHD значно рідше використовується у наукових дослідженнях у порівнянні з CIFAR100, і кількість моделей, оцінених на цьому наборі, є значно меншою. Найвищу класифікаційну точність на наборі даних HD продемонструвала модель EfficientNet, досягнувши вражаючих 99%. Інша високоефективна модель для цього набору даних — VGG-16, яка досягла точності 97%. VGG-16 є глибокою згортковою нейронною мережею з 16 шарами та пірамідальною архітектурою.

Рисунок 4.15 демонструє результати обробки набору даних SIHD за допомогою квантово-згорткової операції, що залучає випадкову квантову схему, що містить RX операції та має 4 кубіти. З проілюстрованого прикладу видно, що основні особливості об'єкта на зображенні залишаються збереженими після квантової попередньої обробки на всіх вихідних шарах, але загалом зображення зазнає значних змін.

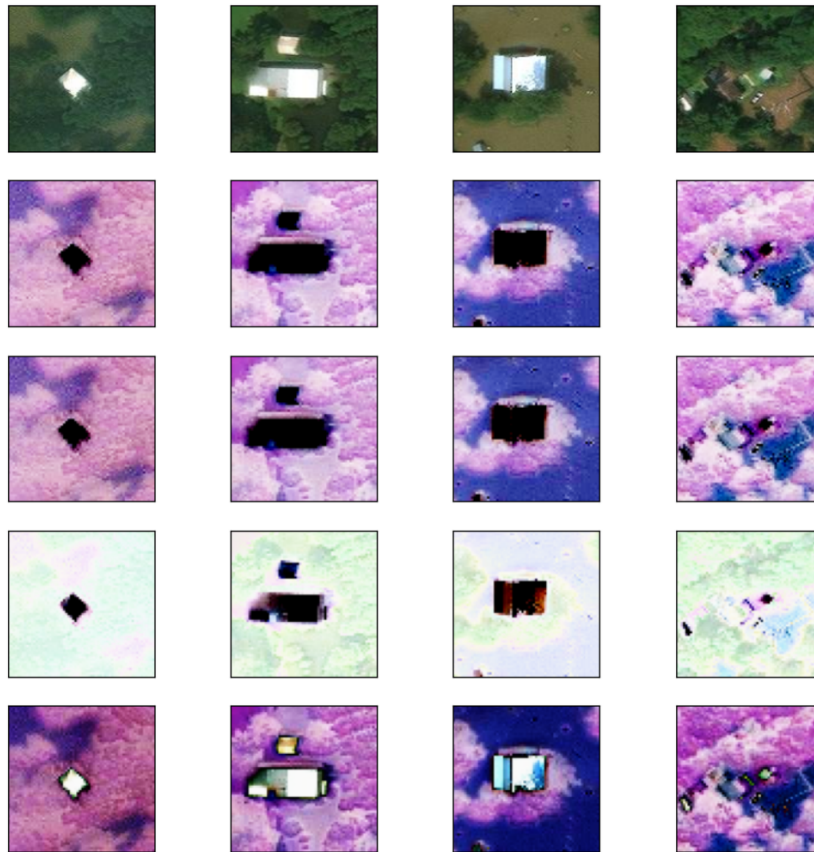


Рисунок 4.15. Приклад квантово-опрацьованих зображень набору даних SIHD [94].

Гібридні квантово-класичні мережі, розглянуті в даному підрозділі, використовують підхід та архітектуру, що були описані в розділі 3.4. «Квантовий пристрій як перший згортковий шар нейронної мережі» та базуються на підході, що полягає в використанні квантової схеми у квантово-згортковій операції, що слугує першим шаром гібридної квантово-класичної нейронної мережі. Отже, усі експерименти в межах даного підрозділу були направлені на дослідження ефективності даного підходу та мають на меті оцінку його ефективності та перспективності для застосування у вирішенні практико-орієнтованих задач.

Для розуміння ролі квантово-згорткових операцій були досліджені прості гібридні нейронні мережі у наступних конфігураціях:

- Класична згорткова нейронна мережа — вхідний шар згорткової нейронної мережі та вихідний шар для цільової кількості класів.

- Класична щільна нейронна мережа — вхідний шар з операцією перетворення зображень у плоский вектор (flattening) для оригінальних вхідних зображень та вихідний шар для цільової кількості класів.
- «Квантово-згортковий шар + CNN» — вхідний шар з квантово-згортковою операцією для попередньої обробки вхідних зображень, шар згорткової нейронної мережі та вихідний шар для цільової кількості класів.
- «Квантово-згортковий шар + щільна нейронна мережа» — вхідний шар з квантово-згортковою операцією для попередньої обробки вхідних зображень, операція перетворення оброблених зображень у плоский вектор (flattening) та вихідний шар для цільової кількості класів.

Для експериментів з залученням простих гібридних нейронних мереж, описаних вище, квантова операція мала наступну комбінацію гіперпараметрів квантово-згорткової операції: розмір квантово-згорткової матриці 2×2 , згортковий крок (stride) рівний 1 та випадкова квантова схема, що залучає операції RX обертання кубіта.

Набір даних CIFAR100 був використаний у двох режимах класифікації:

- Класифікація всіх 100 класів набору даних CIFAR100.
- Класифікація зображень набору даних CIFAR100 відповідно до їх супер-класів – 20 супер-класів.

Для дослідження впливу розміру тренувального набору даних на ефективність роботи гібридних моделей було проведено експеримент з порівняння ефективності вищезазначених моделей на підмножині набору даних CIFAR100 розміром 10000 елементів та на повному наборі даних, який містить 50000 елементів. Результати експерименту наведено на Рисунку 4.16, де ліва частина графіків демонструє процес навчання на 10000 зображень, а права частина – повний набір даних – 50000 зображень з задачею класифікації всіх 100 класів набору даних CIFAR100.

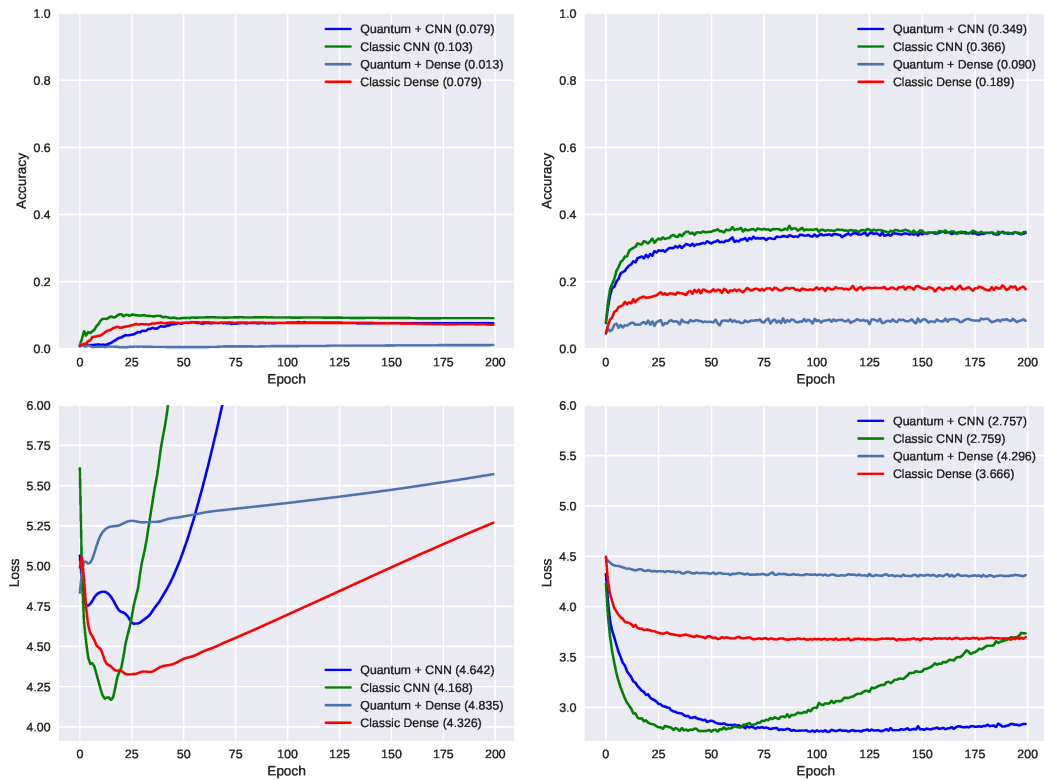


Рисунок 4.16. Результати експерименту з дослідження впливу розміру набору даних на ефективність гібридних моделей [94].

Рисунки 4.17 та 4.18 зображують результати експерименту з різними розмірами набору даних та різними рівнями складності моделей. Рисунок 4.17 демонструє результати експерименту з моделями, що містять 14 згорткових фільтрів, в той час як Рисунок 4.18 зображає результати експериментів з моделями, що містять 256 згорткових фільтрів. Розміри наборів даних, використаних для експериментів, знаходяться від 1000 до 50000 зображень.

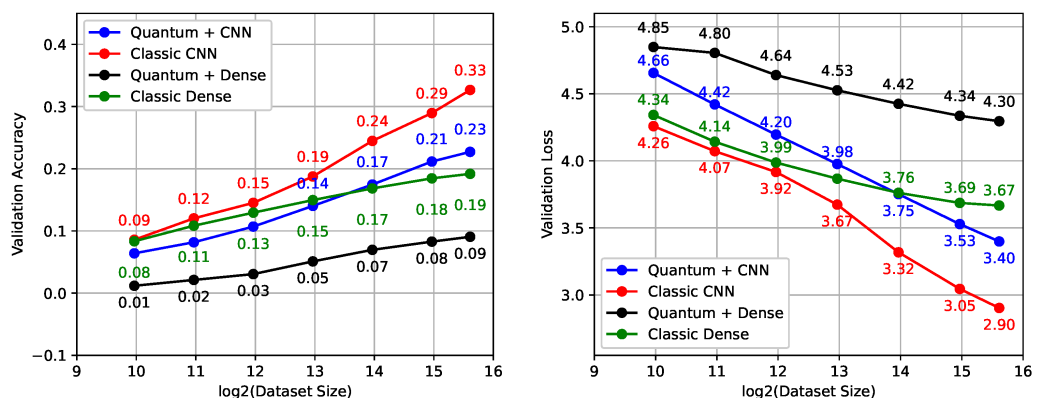


Рисунок 4.17. Результати експерименту з різними розмірами набору даних з моделями, що містять 16 згорткових фільтрів [94].

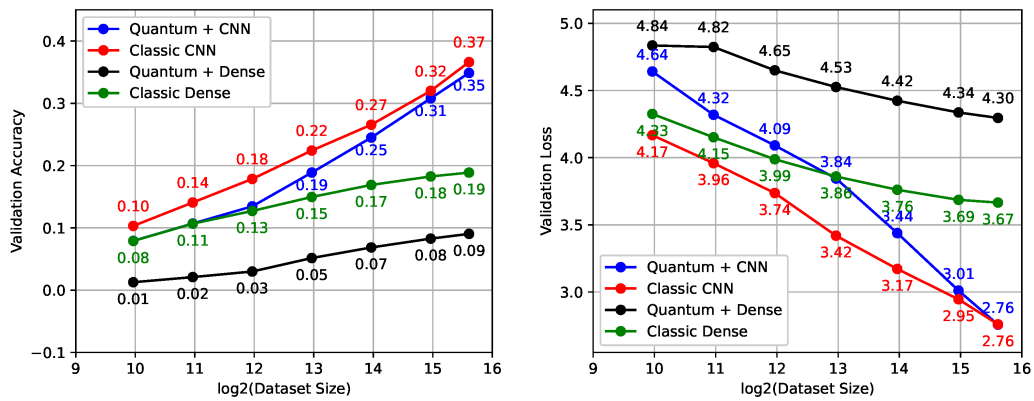


Рисунок 4.18. Результати експерименту з різними розмірами набору даних з моделями, що містять 256 згорткових фільтрів [94].

З отриманих результатів можна зробити висновок, що конфігурація «квантово-згортковий шар + CNN» демонструє найкращі результати на більших розмірах тренувальних наборів даних та за умов використання більш складної класичної частини нейронної мережі (збільшена кількість фільтрів у згортковому шарі). Можна стверджувати, що гібридна нейронна мережа є більш чутливою до параметрів тренувальних даних та складності моделі, ніж референтна згорткова нейронна мережа.

Ці результати пояснюються тим фактом, що квантово-згорткова операція створює більш складну версію вхідних даних, оскільки вона генерує майже вчетверо більше інформації на виході (чотири шари, де кожен є зображенням у форматі RGB, яке трохи менше за розміром, порівнюючи з вхідним), і тому потребує більш складної класичної частини мережі для ефективної генералізації шаблонів, що відповідають кожному окремому класу. Крім того, потрібен більший обсяг тренувальних даних для досягнення вищого рівня точності та досягнення показників, порівнянних з референтною моделлю.

Варто також зазначити, що для моделі гібридної конфігурації «Квантово-згортковий шар + CNN» із 256 фільтрами у згортковому шарі (Рисунок 3.23) збільшення розміру набору даних призводить до кардинально іншої динаміки покращення точності та втрат. Для розмірів набору даних у діапазоні від 1000 до 4000 початкових тренувальних зображень крива точності (перші три точки на Рисунку 4.18, ліворуч) демонструє повільний темп покращення, подібний до

темтів моделі класичної щільної нейронної мережі. Проте для розмірів набору даних у діапазоні від 4000 до 50000 зображень крива точності (наступні п'ять точок на Рисунку 4.18, ліворуч) демонструє швидкий темп покращення, навіть перевищуючи темп моделі класичної згорткової нейронної мережі.

Ще одним висновком, який можна зробити з цих експериментів, є те, що конфігурація «Квантово-згортковий шар + щільна нейронна мережа» в рамках експерименту демонструє нижчі показники точності незалежно від розміру набору даних і значно поступається за метриками іншим конфігураціям, описаним у цьому дослідженні. Єдиною перевагою цієї конфігурації є теоретично більша швидкість обчислень завдяки простій архітектурі та тому, що найскладніша частина даної гібридної моделі — квантово-згортковий шар — обчислюється на квантовому обладнанні. Однак для завдань, що вимагають значної точності, застосування такої конфігурації без додаткового згорткового шару не є доцільним.

Виявлена різниця у продуктивності при різній кількості згорткових фільтрів у шарі згорткової нейронної мережі, показана на Рисунках 4.17 і 4.18, підкреслює необхідність детальнішого аналізу залежності від кількості згорткових фільтрів у згорткових шарах нейронної мережі. Були проведені дослідження максимальної точності роботи моделі (Рисунок 4.19, ліворуч) і мінімальних значень функції втрат (Рисунок 4.19, праворуч) як функції кількості згорткових фільтрів у шарі згорткової нейронної мережі для набору даних, що містить 50000 зображень у навчальному наборі і 5000 зображень у тестовому наборі. Результати подано на Рисунку 4.19.

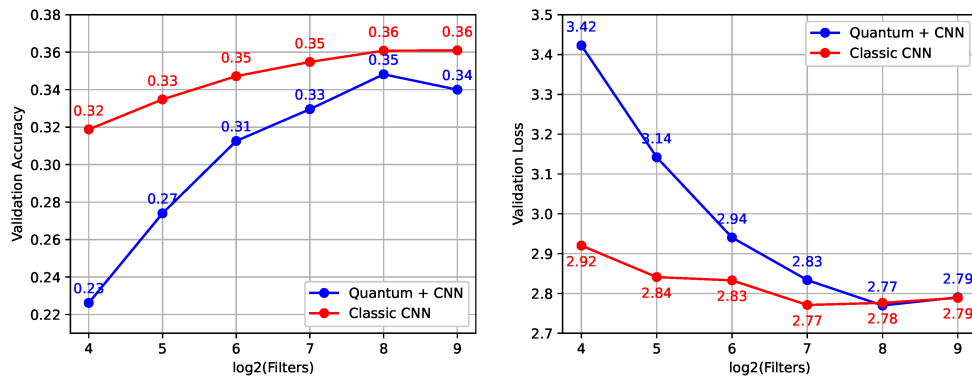


Рисунок 4.19. Максимальна точність (ліворуч) і мінімальна втрата (праворуч) як функція кількості згорткових фільтрів [94].

На Рисунку 4.19 відображено, що гібридна нейронна мережа демонструє підвищену чутливість до змін складності моделі, що супроводжується суттєвими коливаннями у точності при збільшенні її складності. Для досягнення оптимальної продуктивності гібридна нейронна мережа вимагає ретельного налаштування архітектури класичної моделі, де кількість згорткових шарів є ключовим гіперпараметром, що потребує точного налаштування. Перевищення складності мережі за певний поріг (у нашому випадку понад 256) призводить до погіршення продуктивності гібридної нейронної мережі, ймовірно, через перенавчання. Варто також зазначити, що збільшення кількості згорткових фільтрів у шарі згорткової нейронної мережі забезпечує значно швидше покращення квантово-згорткової нейронної мережі у порівнянні з класичною згортковою нейронною мережею.

Аналіз впливу додаткових фільтрів на ефективність моделі поставив питання про роль додаткових шарів згорткової нейронної мережі. Було проведено дослідження максимальної точності (ліворуч) та мінімального значення функції втрат (праворуч) у залежності від кількості згорткових фільтрів у двох шарах згорткової нейронної мережі для набору даних, що складається з 50000 зображень у навчальному наборі та 5000 зображень у тестовому наборі. Результати наведено на Рисунку 4.20.

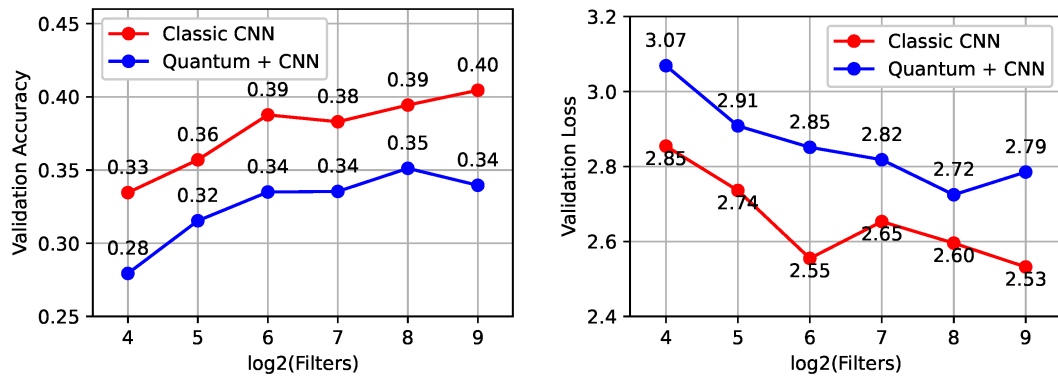


Рисунок 4.20. Максимальна точність (ліворуч) і мінімальна втрата (праворуч) як функція кількості згорткових фільтрів [94].

Загалом, впровадження другого шару згорткової нейронної мережі призвело до значного підвищення ефективності класичної моделі згорткової нейронної мережі в значно більшій мірі, ніж спостерігалось у випадку моделі гібридної нейронної мережі.

У межах наступного кроку було досліджено подібні залежності різних параметрів від продуктивності простих гібридних нейронних мереж у порівнянні з іншими типами нейронних мереж для набору даних CIFAR100 з грубою класифікацією, що включає 20 грубих класів (на відміну від тонкої класифікації, що складається з 100 класів, як це було раніше).

Вивчався вплив різних розмірів навчального набору даних (при цьому розмір тестового набору залишався незмінним і становив 5000 зображень), максимальні значення точності та мінімальні значення втрат, результати яких узагальнені на Рисунку 4.21 для 16 фільтрів і на Рисунку 4.22 для 256 фільтрів у шарі згорткової нейронної мережі.

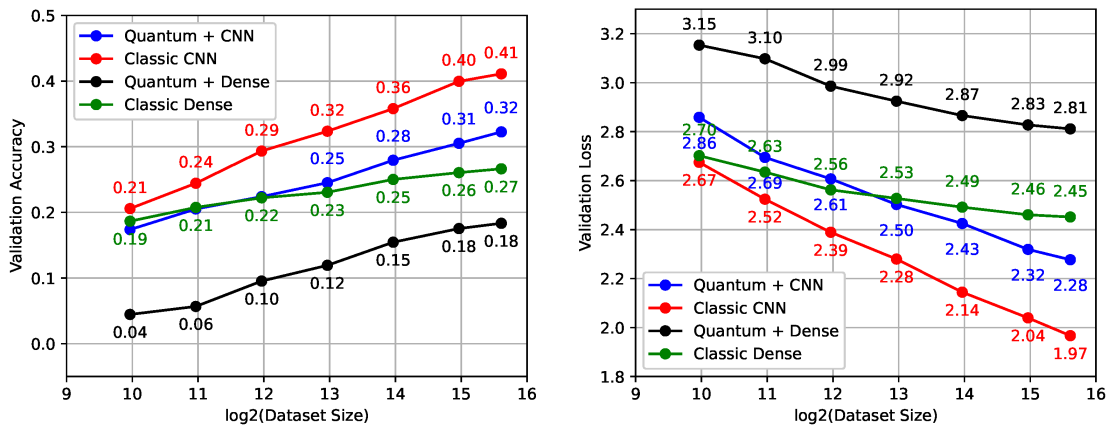


Рисунок 4.21. Максимальна точність (ліворуч) і мінімальна втрата (праворуч) як функція розміру навчального набору даних (від 1000 до 50,000 зображень), виміряна на тому ж розмірі тестового набору даних у 5000 зображень для 16 фільтрів у шарі згорткової нейронної мережі для CIFAR100 з грубою класифікацією [94].

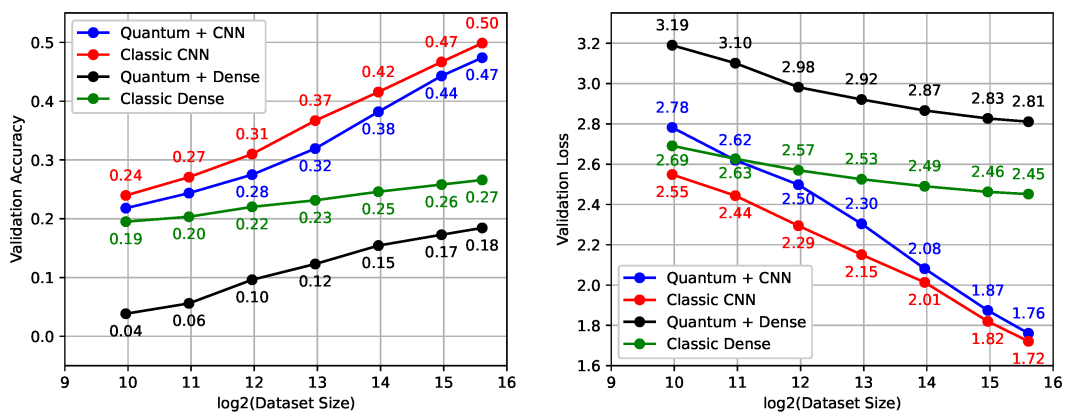


Рисунок 4.22. Максимальна точність (ліворуч) і мінімальна втрата (праворуч) як функція розміру навчального набору даних для 256 фільтрів у шарі згорткової нейронної мережі для CIFAR100 з грубою класифікацією [94].

Результати проведеного експерименту демонструють аналогічну закономірність, що спостерігається у відповідних результатах більш складної задачі (задача класифікації з 100 класами), підкреслюючи, що спостережені ефекти є схожими та потенційно універсальними для різних кількостей класів.

Різні показники продуктивності, що характеризують вплив різної кількості згорткових фільтрів у шарі згорткової нейронної мережі, представлені на Рисунках 4.21 і 4.22, вказують на необхідність детальнішого дослідження

залежності продуктивності від кількості згорткових фільтрів у шарі згорткової нейронної мережі. Було проведено аналіз максимальної точності (ліворуч) та мінімальних втрат (праворуч) як функцій від кількості згорткових фільтрів у шарі згорткової нейронної мережі (для 50000 зображень в навчальному наборі даних і 5000 зображень в тестовому наборі даних), результати якого наведені на Рисунку 4.23.

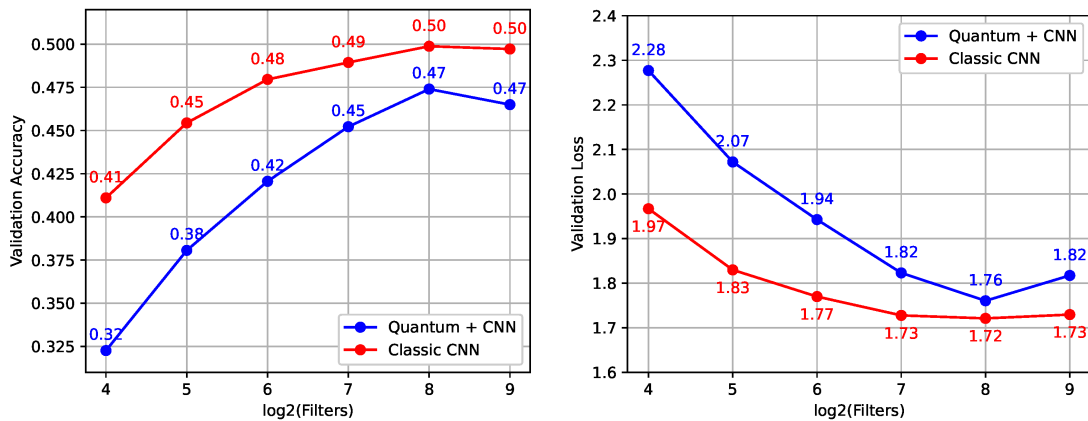


Рисунок 4.23. Максимальна точність (ліворуч) і мінімальна втрата (праворуч) як функція кількості згорткових фільтрів у шарі згорткової нейронної мережі для CIFAR100 з грубою класифікацією [94].

Незважаючи на нижчі максимальні показники точності та більшу кількість епох, необхідних для гібридних нейронних мереж для досягнення оптимальної точності класифікації в порівнянні з класичними референтними моделями, як показано на Рисунках 4.16-4.23, гібридні нейронні мережі все ще демонструють значні переваги, порівнюючи з класичними нейронними мережами. Основними перевагами гібридних нейронних мереж є підвищена енергетична ефективність і значно вища швидкість обчислень, що забезпечується квантовим апаратним забезпеченням. Ці переваги можуть бути надзвичайно корисними і надавати значну цінність, навіть в умовах нижчих показників точності в широкому спектрі застосувань.

Попередні експерименти були націлені на дослідження впливу розміру набору даних, складності набору даних та структури гібридних нейронних мереж на показник точності, який порівнюється з показниками класичних референтних моделей. Головною метою цієї частини дослідження є не виявлення конфігурації

гібридних нейронних мереж, яка перевищує точність класичної нейронної мережі, а вивчення впливу різних параметрів на продуктивність гібридних нейронних мереж.

Після проведення експериментів на простих конфігураціях гібридних квантово-згорткових моделей, було вирішено на наступному кроці дослідження залучити для експериментів складніші конфігурації гібридних моделей з ціллю дослідження практичності та ефективності даного підходу для вирішення задачі багатокласової класифікації зображень.

Рисунки 4.24 та 4.25 ілюструють процес навчання гібридних нейронних мереж, що базуються на складних сучасних класичних архітектурах і референтних моделях на наборі даних CIFAR100. З графіків очевидно, що процес навчання сходиться. Спостерігається, що гібридні нейронні мережі демонструють нижчу точність, порівнюючи з класичними аналогами. Крім того, важливо зазначити, що більш складні архітектури гібридних нейронних мереж перевершують простіші гібридні нейронні мережі, а також прості класичні референтні моделі; наприклад, гібридні нейронні мережі на базі EfficientNet демонструють значно вищу точність, ніж більш проста модель VGG-16.

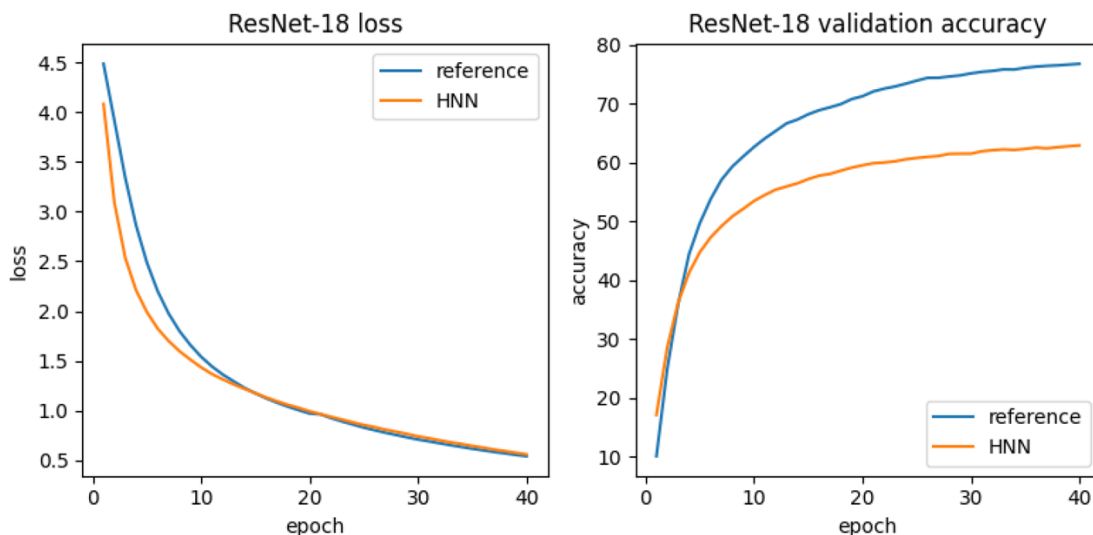


Рисунок 4.24. Значення функції втрат та точність гібридної нейронної мережі ResNet-18 та референтної моделі на наборі даних CIFAR100 [94].

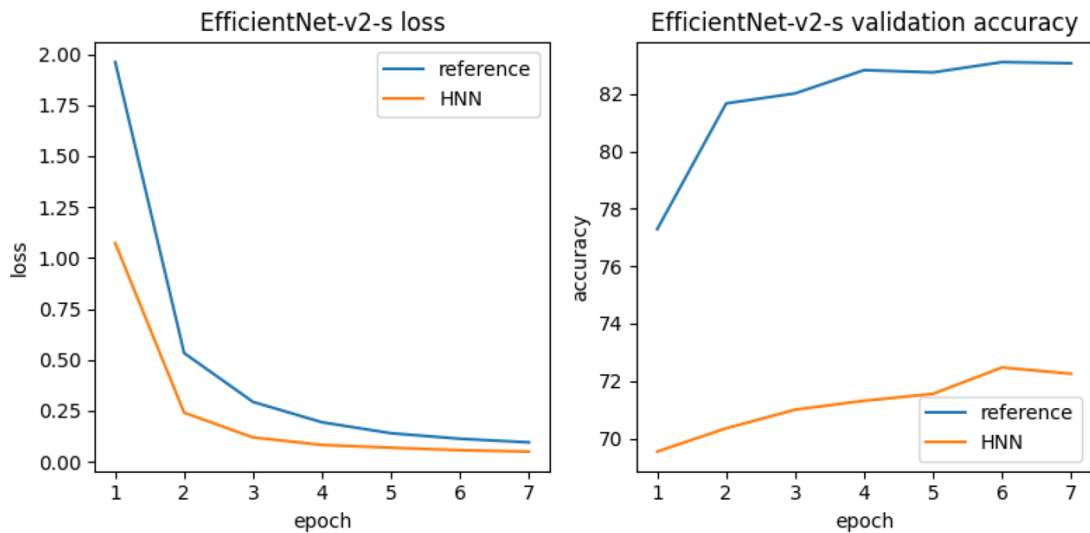


Рисунок 4.25. Значення функції втрат та точність гібридної нейронної мережі EfficientNet-v2-s та референтної моделі на наборі даних CIFAR100 [94].

Рисунки 4.26 і 4.27 ілюструють процес навчання гібридних нейронних мереж та референтних моделей на наборі даних SIHD. Аналіз графіків підтверджує, що процес навчання сходиться. На наборі даних SIHD спостерігається, що певні гібридні нейронні мережі демонструють кращі результати, порівняно з референтними моделями, за показниками точності та швидкості навчання (кількість епох, необхідних для досягнення максимальної точності). Це явище можна пояснити тим, що завдяки квантовому штучному розширенню даних нам вдалося значно збільшити обсяг навчального набору даних, що сприяло кращому засвоєнню моделями шаблонів, які вказують на пошкодження будівель.

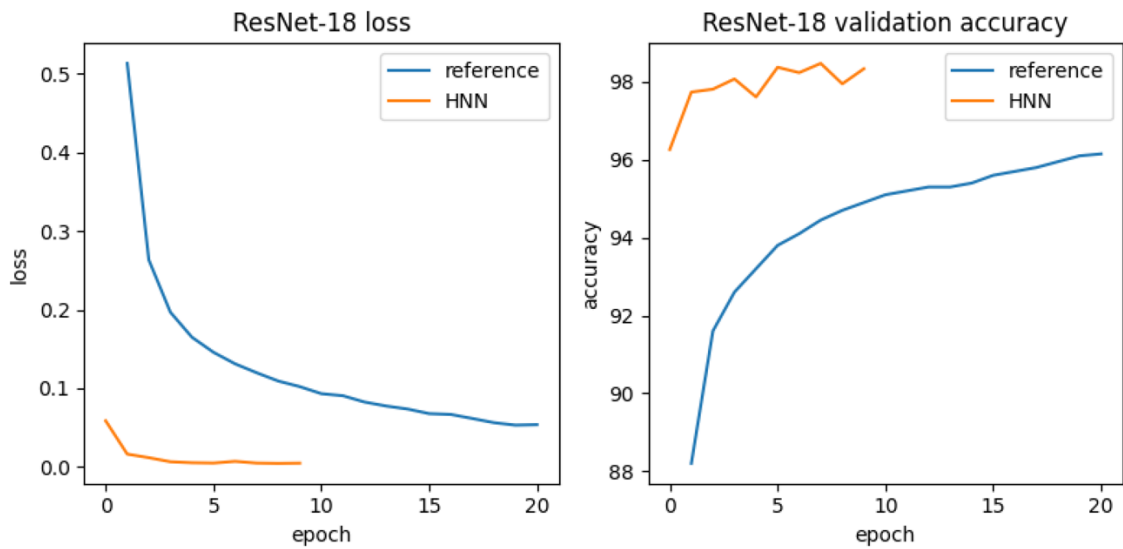


Рисунок 4.26. Значення функції втрат та точність гібридної нейронної мережі ResNet-18 і референтної моделі на наборі даних SIHD [94].

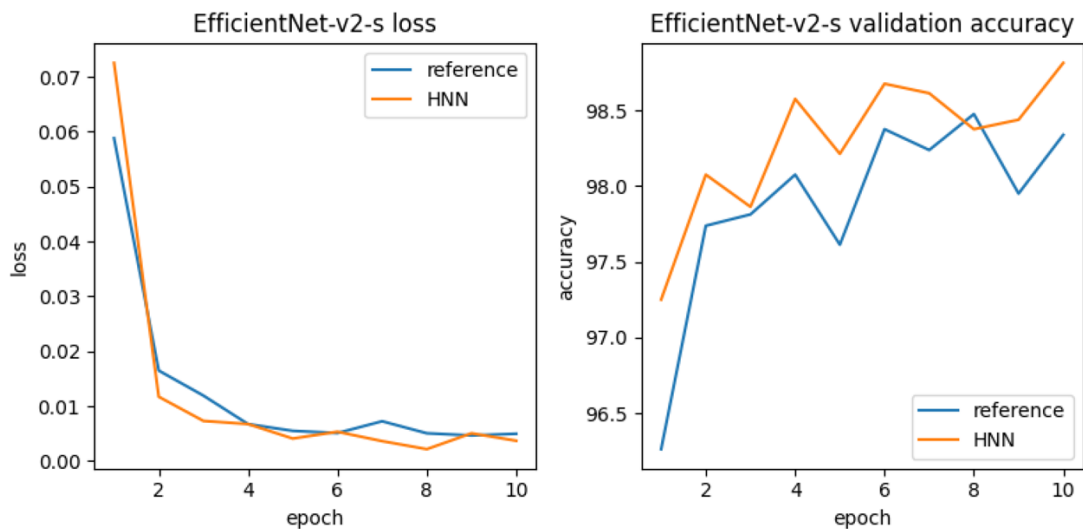


Рисунок 4.27. Значення функції втрат та точність гібридної нейронної мережі EfficientNet-v2-s і референтної моделі на наборі даних SIHD [94].

Таблиця 4.5 ілюструє метрики точності, на тестовому наборі даних, гібридних нейронних мереж та референтних класичних нейронних мереж, на яких базуються гібридні нейронні мережі. Аналіз результатів для набору даних CIFAR100 показує, що найвищу точність досягнуто класичною нейронною мережею EfficientNet. Ситуація на наборі даних SIHD є більш захопливою, оскільки найкраща точність була продемонстрована гібридною нейронною мережею, що ґрунтується на архітектурі EfficientNet. Ці результати

підтверджують доцільність застосування гібридних нейронних мереж для вирішення реальних проблем та навіть свідчать про те, що в певних сценаріях гібридні нейронні мережі можуть невеликою мірою перевершувати класичні моделі.

Таблиця 4.5. Точність гібридних нейронних мереж та референтних моделей (найвища точність виділена жирним шрифтом) [94].

Модель	CIFAR100	SIHD
ResNet-18 референтна	76.74%	96.15%
EfficientNet-v2-s референтна	78.76%	98.5%
VGG-16 референтна	62.19%	96.5%
ResNet-18 HNN	62.88%	98.62%
EfficientNet-v2-s HNN	72.26%	98.81%
VGG-16 HNN	54.95%	85.94%

Таблиця 4.6 ілюструє вплив техніки голосування більшістю (MV) на точність гібридних нейронних мереж на тестовій підмножині. Порівняння статистичних даних між Таблицею 1 та Таблицею 2 виявляє, що впровадження техніки голосування більшістю забезпечує значне підвищення метрик точності гібридних нейронних мереж (4,83% для гібридної нейронної мережі на основі EfficientNet та 5,17% для гібридної нейронної мережі на основі VGG-16) при вирішенні більш складної задачі класифікації зображень CIFAR100. Гібридна EfficientNet майже досягла продуктивності референтної моделі, а гібридна VGG-16 навіть перевершила свою немодифіковану версію в цій задачі.

Таблиця 4.6. Тестова точність моделей гібридних нейронних мереж із використанням техніки голосування більшістю [94].

Модель	CIFAR100	SIHD
EfficientNet-v2-s HNN	77.09%	97.75%
VGG-16 HNN	65.12%	98.1%

На наборі даних SIHD спостерігається інша картина. Застосування техніки голосування більшістю не призвело до покращення точності для гібридної нейронної мережі EfficientNet і навіть виявилось дещо менш точним (на 1,06%). Натомість гібридна нейронна мережа VGG-16 з технікою голосування більшістю перевершила свою класичну версію, продемонструвавши підвищення точності на 1,6%. До того ж техніка голосування більшістю для обчислення точності показала вищу метрику точності, ніж альтернативна методика обчислення.

4.5. Аналіз ефективності використання багатоструктурної класичної частини гібридної квантово-згорткової моделі

Підхід із залученням багатоструктурної класичної частини гібридної квантово-згорткової нейронної мережі є розширенням підходу гібридизації що залучає квантовий пристрій у якості першого згорткового рівня гібридної моделі, аналіз якого було проведено в межах розділу 4.4 «Аналіз методу гібридизації нейронних мереж, використовуючи квантовий пристрій як перший рівень нейронної мережі». Для проведення експериментів з аналізу ефективності залучення багатоструктурної класичної частини гібридних нейронних мереж у якості базової моделі було обрано класичну архітектуру LNet050 [106]. LNet050 було обрано як приклад невеликої, але досить ефективної групи згорткових нейронних мереж, яка демонструє чудову продуктивність, незважаючи на невеликий розмір.

Для проведення експериментів було використано кілька стандартних метрик для оцінки моделей. До них відносяться значення функції втрат під час

навчання та перевірки на тестовому наборі даних, що обчислювалася за допомогою функції втрат категоричної крос-ентропії, що відображає помилку між прогнозами моделі та фактичними мітками. Також відслідковувалась точність підчас навчання та точність класифікації тестового набору даних. Крім цих стандартних метрик, було використано площу під кривою (area under the curve, або ж AUC) для кривої робочих характеристик приймача (receiver operating characteristics, або ж ROC) [107]. Дана метрика використовується для оцінки здатності моделі розрізняти класи та полягає в порівнянні кількості істинно позитивних результатів класифікації та хибно позитивних (1-специфічність) для різних порогів класифікації. Мікросереднє (micro-average) значення AUC узагальнює метрики глобально, враховуючи всі класи класифікації та всі елементи набору даних, тоді як макросереднє (macro-average) значення AUC обчислює метрики для кожного класу окремо та потім усереднює їх.

Експерименти проводились на наступних конфігураціях моделей:

- Референтна класична нейронна мережа, яка є незміненою моделлю LCNet050.
- Гібридна квантово-згорткова нейронна мережа, що залучає квантовий пристрій, що містить 4 кубіти та класична частина якої складається з багатоструктурної моделі на основі базової моделі LCNet050 та яка містить 4 базові моделі (HNN-QC4). Діаграма класичної частини даної моделі наведена на Рисунку 4.28.
- Гібридна квантово-згорткова нейронна мережа, що залучає квантовий пристрій, що містить 4 кубіти та класична частина якої складається з багатоструктурної моделі на основі базової моделі LCNet050 та яка містить 5 базових моделей (HNN-QC5), 4 з яких опрацьовують результати квантово-згорткової операції, а одна – оригінальне вхідне зображення. Діаграма класичної частини даної моделі наведена на Рисунку 4.29.

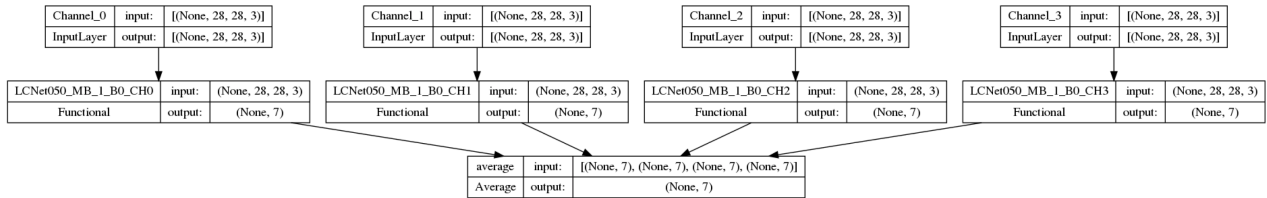


Рисунок 4.28. Діаграма класичної частини моделі HNN-QC4 [95].

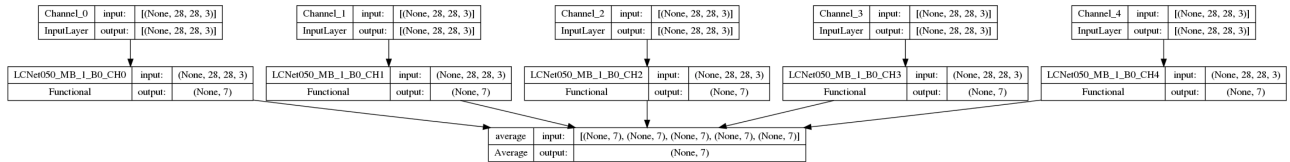


Рисунок 4.29. Діаграма класичної частини моделі HNN-QC5 [95].

Усі експерименти для оцінки ефективності даного підходу було проведено на наборі даних CIFAR10 та у якості квантового пристрою було залучено квантову схему, що містить 4 кубіти та базується на RY операціях. Навчання проводилось протягом 200 ітерацій. Також результати всіх експериментів базуються на результатах перехресної перевірки (cross-validation) для отримання більш надійних результатів, що визначаються на основі повного набору даних.

Рисунок 4.30 демонструє значення AUC метрик референтної моделі залежно від розмірів пакету (batch size). Рисунок 4.31 демонструє залежність точності роботи референтної моделі залежно від розміру пакету. На основі отриманих результатів, можна зробити спостереження, що референтна модель демонструє тенденцію до зниження як AUC (мікросередньої та макросередньої), так і точності зі збільшенням розміру пакета від 16 до 8192. Помітно, що мікросередня AUC зменшується з 0.961 для розміру пакета 16 до 0.917 для розміру пакета 8192, а точність аналогічно знижується з 72.7% до 61%.

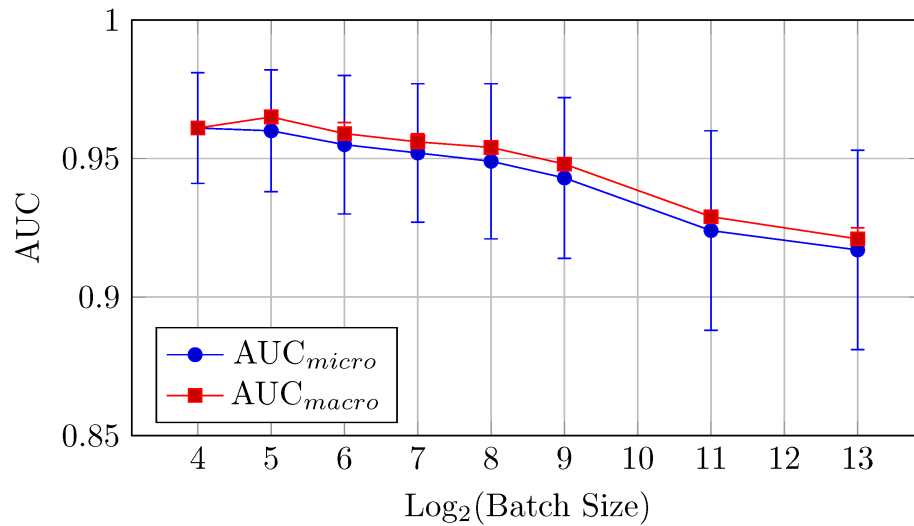


Рисунок 4.30. Залежність AUC від розміру пакету референтної моделі [95].

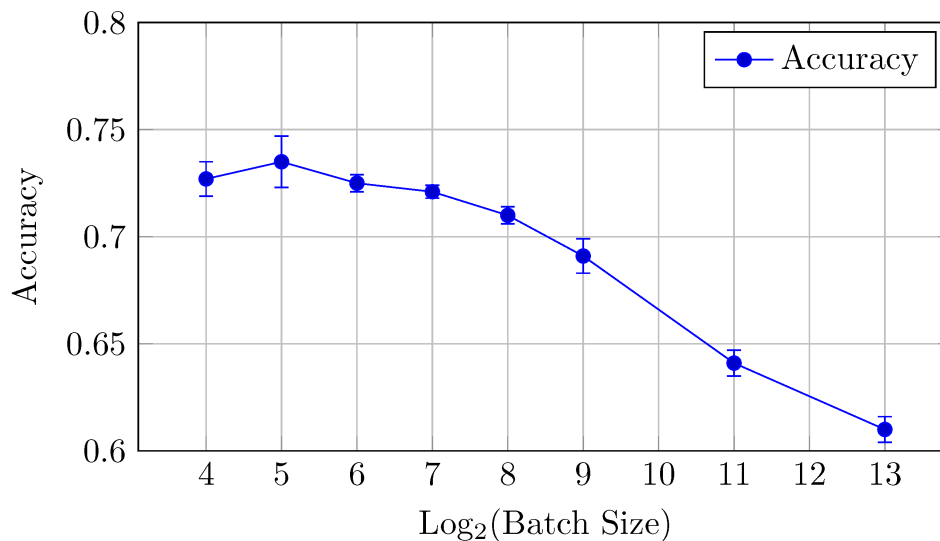


Рисунок 4.31. Залежність точності від розміру пакету референтної моделі [95].

Спостерігається тенденція до зниження AUC зі збільшенням розміру пакету, причому найвищі значення мікросередньої та макросередньої AUC спостерігаються для менших розмірів пакетів (16 та 32). Обидві метрики падають зі збільшенням розміру пакета понад 64, але різниця між середніми значеннями залишається в межах стандартного відхилення до розміру пакета 512. Це вказує на те, що менші розміри пакетів є ефективнішими для оптимізації AUC референтної моделі. Значення мікросередньої та макросередньої AUC виявилися дуже близькими для всіх розмірів пакетів, особливо для розміру пакета 16, що свідчить про стабільність продуктивності моделі для всіх класів набору даних.

Точність референтної моделі починається з 72.7% для розміру пакета 16 і поступово знижується зі збільшенням розміру пакета, але різниця між середніми значеннями залишається в межах стандартного відхилення до розміру пакета 128. Найкраща продуктивність досягається при розмірі пакета 32 з точністю 73.5%.

Загалом, невеликі розміри пакетів (16 або 32) забезпечують кращу продуктивність як для AUC, так і для точності референтної моделі.

Рисунок 4.32 демонструє значення AUC метрик моделі HNN-QC4 залежно від розмірів пакету. Рисунок 4.33 демонструє залежність точності роботи моделі HNN-QC4 залежно від розміру пакету. На основі отриманих результатів можна зробити спостереження, що значення AUC та точності демонструють покращення порівняно з референтною моделлю, особливо для середніх розмірів пакетів (від 32 до 256). Продуктивність залишається відносно стабільною для розмірів пакетів між 64 та 256 з незначними коливаннями.

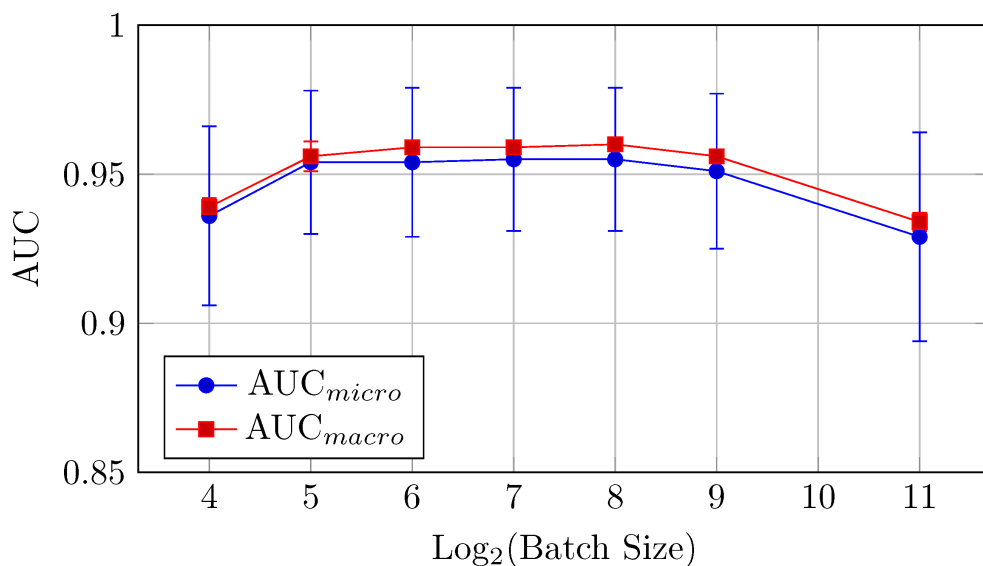


Рисунок 4.32. Залежність AUC від розміру пакету моделі HNN-QC4 [95].

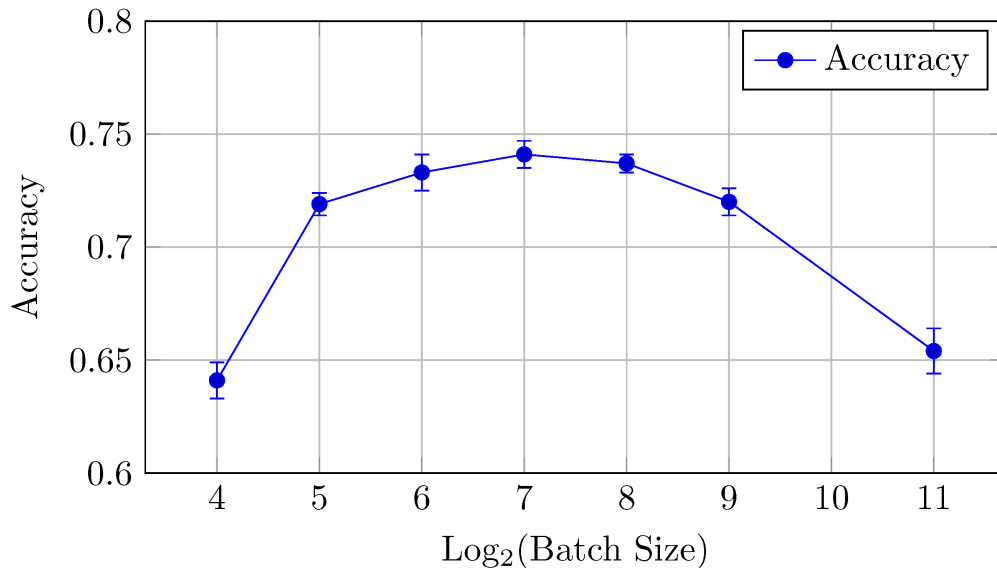


Рисунок 4.33. Залежність точності від розміру пакету моделі HNN-QC4 [95].

Значення AUC досить стабільні для розмірів пакетів між 64 та 256, причому мікросередня AUC тримається на рівні близько 0.955. Ця стабільність AUC свідчить про те, що модель HNN-QC4 менш чутлива до змін розміру пакета в межах цього діапазону. Також дуже мала різниця між значеннями мікросередньою та макросередньою AUC для всіх досліджуваних розмірів пакетів вказує на стабільну продуктивність моделі для всіх класів набору даних.

Найвища точність досягається при розмірі пакета 128 та рівна 74.1%, що свідчить про те, що середні розміри пакетів є найкращими для даної варіації гібридної моделі. Варто також зауважити, що точність починає знижуватися при розмірах пакетів як менше 64, так і більше 256.

Рисунок 4.34 демонструє значення AUC метрик моделі HNN-QC5 залежно від розмірів пакету. Рисунок 4.35 демонструє залежність точності роботи моделі HNN-QC5 залежно від розміру пакету.

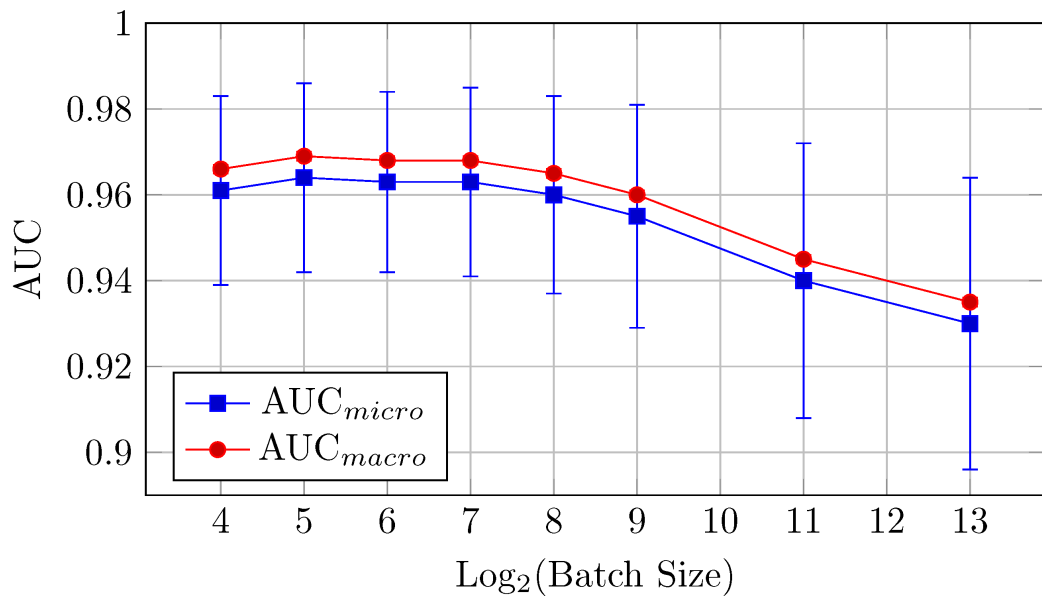


Рисунок 4.34. Залежність AUC від розміру пакету моделі HNN-QC5 [95].

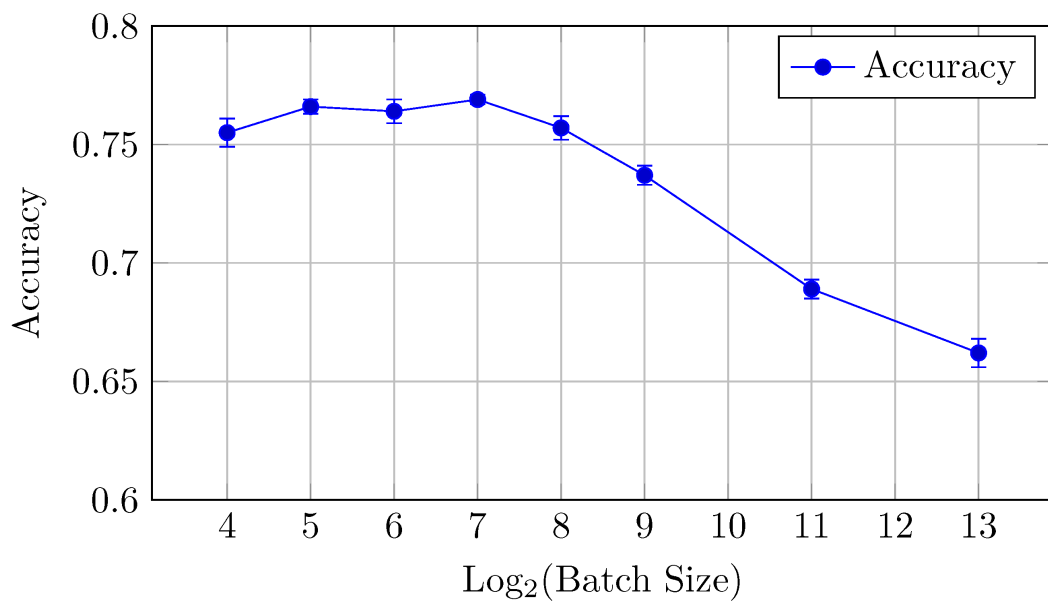


Рисунок 4.35. Залежність точності від розміру пакету моделі HNN-QC5 [95].

Найвищі значення AUC спостерігаються при розмірах пакетів 16 та 32, де мікросередня AUC досягає 0.964, а макросередня – 0.969. Дана варіація гібридної моделі демонструє кращу продуктивність при невеликих розмірах пакетів, але різниця між середніми значеннями залишається в межах стандартного відхилення до розміру пакета 512. Близькість значень мікросередньої та

макросередньої AUC для всіх зразків також вказує на стабільність продуктивності моделі для всіх класів у наборі даних.

Модель HNN-QC5 перевершує HNN-QC4 та референтну моделі за точністю, з найвищим значенням 76.9% при розмірі пакета 128. Також варто зауважити, що модель демонструє стабільну продуктивність для середніх розмірів пакетів (між 32 та 256).

Дана варіація моделі показує, що невеликі та середні розміри пакетів (від 16 до 128) забезпечують найкращі результати для AUC та точності. Продуктивність знижується при більших розмірах пакетів (512, 2048), слідуючи аналогічній тенденції з іншими розглянутими моделями.

Загалом, порівнюючи всі моделі на наборі даних CIFAR10, можна спостерігати, що модель HNN-QC5 демонструє найкращу загальну продуктивність серед усіх моделей. Як значення AUC, так і точність стабільно вищі, порівняно з референтною моделлю та моделлю HNN-QC4.

Також, найкращі значення площі під кривою та точності для різних розмірів пакетів для моделей HNN-QC_n на наборі даних CIFAR10, що наведені в Таблиці 4.7, можна відзначити тенденцію до підвищення середніх значень та зменшення стандартного відхилення точності зі збільшенням кількості основних моделей та квантових каналів. Також середні значення точності підвищуються зі збільшенням кількості основних моделей та квантових каналів для розмірів пакетів 64 і більше. Це демонструє Рисунок 4.36, який показує графіки залежності точності від розміру пакета для моделей HNN-QC4, HNN-QC5 та референтної моделі на наборі даних CIFAR10.

Таблиця 4.7. Найкращі значення метрик AUC та точності роботи розглянутих моделей HNN-QC_n та референтної моделі [95].

Модель	AUC _{micro}	AUC _{macro}	Точність
Референтна модель	0.961 ± 0.020	0.965 ± 0.002	73.5% ± 1.2%
HNN-QC4	0.955 ± 0.024	0.960 ± 0.001	74.1% ± 0.6%
HNN-QC5	0.964 ± 0.022	0.969 ± 0.001	76.9% ± 0.2%

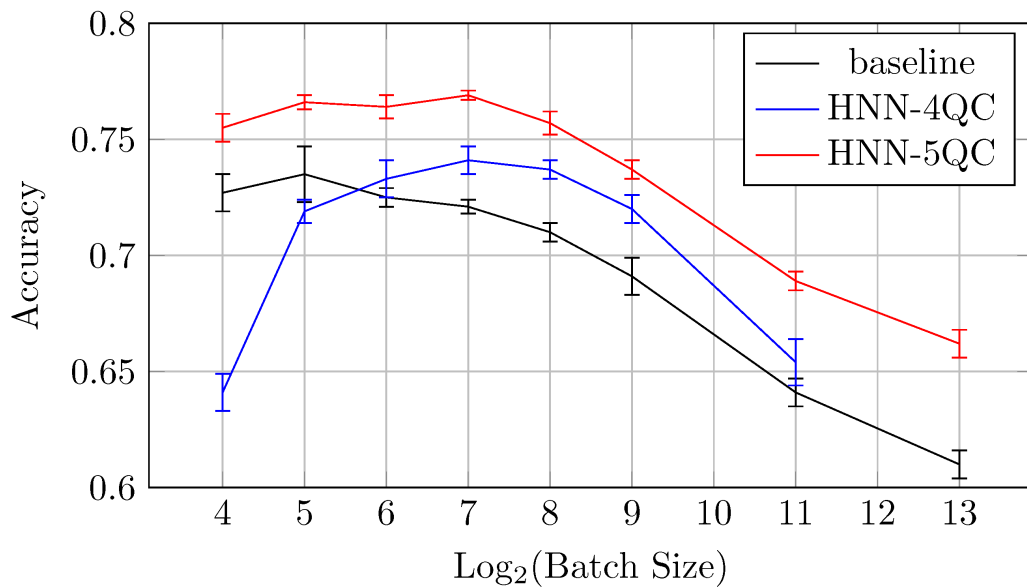


Рисунок 4.36. Залежність точності розглянутих моделей від розміру пакету [95].

На основі наведених даних видно, що точність, продемонстрована моделлю HNN-QC5, стабільно вища порівняно з точністю моделі HNN-QC4, а точність HNN-QC4 стабільно вища порівняно з референтною моделлю для розмірів пакетів 64 і більше.

4.6. Порівняльний аналіз ефективності розглянутих методів гібридизації нейронних мереж

Для побудови ефективного комплексного методу гібридизації класичних та некласичних обчислень необхідно розглянути та оцінити ефективність, сильні та слабкі сторони кожного з розглянутих підходів. Для оцінки їх ефективності важливо визначити метрики, які будуть використовуватися для порівняння цих підходів. Оскільки експерименти були проведені на класичному апаратному забезпеченні та для емуляції квантових процесів використовувався симулятор квантових обчислень, визначити часову складність кожного підходу неможливо. Тому замість вимірювання часу обчислень було вирішено вимірювати кількість

квантових операцій, необхідних для навчання та роботи моделі, а також оцінити порівняльну складність класичної частини гібридних моделей.

Кількість необхідних квантових операцій є важливою метрикою, оскільки квантове апаратне забезпечення на даний момент значно дорожче, а доступ до нього значно більш обмежений, порівнюючи з класичним апаратним забезпеченням.

Крім того, було використано набір стандартних метрик, таких як максимальна точність моделі, досягнута на тестовому наборі даних, кількість епох, необхідних для навчання моделі, та мінімальне значення функції втрат під час навчання моделі. Отже, повний перелік метрик ефективності, використаних у експериментах, є наступним:

- максимальна точність, продемонстрована моделлю на тестовому наборі даних;
- кількість епох, необхідних для навчання моделі;
- кількість квантових операцій, необхідних для навчання моделі;
- кількість квантових операцій, необхідних для отримання одного прогнозу моделі;
- порівняна складність класичної частини моделі.

В межах даного експерименту квантово-згорткову модель було побудовано на основі архітектури VGG-16. Гібридну нейронну мережу з квантовим пристроєм, що заміняє лінійний шар нейронів, було побудовано на основі простої архітектури згорткової моделі, що є еквівалентом гібридних моделей, що використовувались в експериментах в розділах 4.2 та 4.4. Також було використано глибоку згорткову нейронну мережу, що має архітектуру VGG-16, у якості референтної моделі.

Рисунок 4.37 демонструє графік точності моделей під час навчання до моменту досягнення максимальних значень точності на наборі даних CIFAR10. На графіку результати гібридної нейронної мережі, що містить квантовий пристрій як один з прихованих шарів нейронної мережі, позначено синім

кольором, результати квантово-згорткової мережі – червоним, а результати референтної моделі – коричневим.

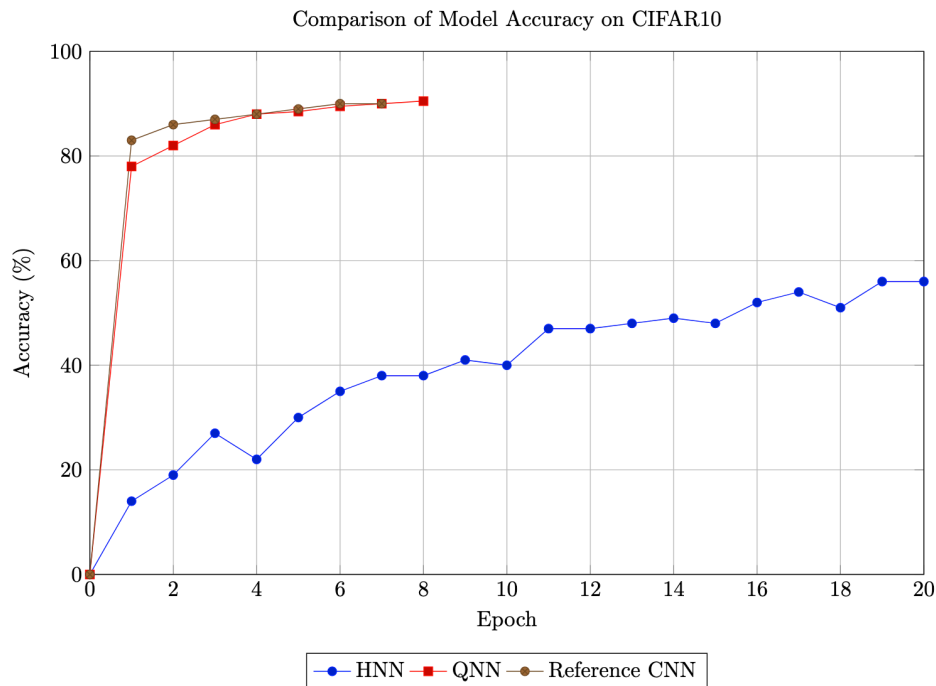


Рисунок 4.37. Точність роботи моделей на наборі даних CIFAR10.

На основі результатів експерименту можна зробити висновок, що квантово-згорткова модель продемонструвала найкращі результати точності роботи, досягнувши точності у 90.5% на наборі даних CIFAR10 та перевершивши референтну модель на незначні 0.5%. Гібридна мережа з квантовим пристроєм в якості прихованого шару показала точність класифікації у 56%, що є на 34.5% гіршим результатом, порівнюючи з квантово-згортковою моделлю, та на 34% гіршим результатом, порівнюючи з референтною моделлю.

Також помітно, що гібридна модель на основі квантового пристрою в якості прихованого шару нейронної мережі потребує більшої кількості епох для завершення процесу навчання.

Рисунок 4.38 демонструє графік точності моделей під час навчання до моменту досягнення максимальних значень точності на наборі даних SIHD. На графіку результати гібридної нейронної мережі, що містить квантовий пристрій як один з прихованих шарів нейронної мережі, позначено синім кольором,

результати квантово-згорткової мережі – червоним, а результати референтної моделі – коричневим.

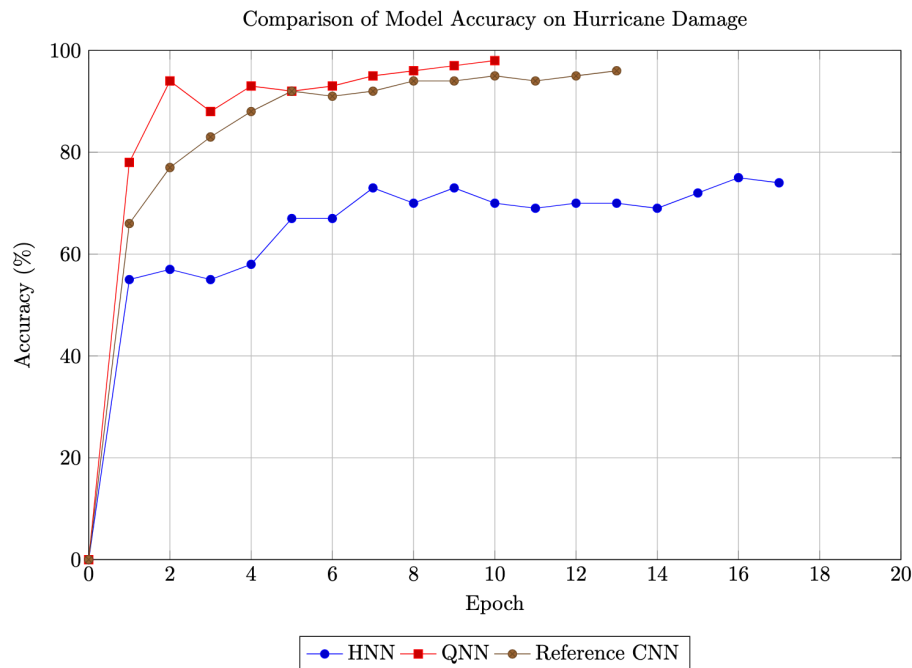


Рисунок 4.38. Точність роботи моделей на наборі даних Hurricane Damage [108].

На основі результатів експерименту можна зробити висновок, що квантово-згорткова модель продемонструвала найкращі результати точності роботи, досягнувши точності у 98.1% на наборі даних SHD та перевершивши референтну модель на 1.6%. Гібридна мережа з квантовим пристроєм в якості прихованого шару показала точність у 75%, що є на 23.1% гіршим результатом, порівнюючи з квантово-згортковою моделлю, та на 21.5% гіршим результатом, порівнюючи з референтною моделлю [108].

Також помітно, що гібридна модель на основі квантового пристрою в середині нейронної мережі потребує більшої кількості епох для завершення процесу навчання. В той же час, процес навчання квантово-згорткової моделі є найкоротшим, випереджаючи на 3 ітерації навчання референтну модель та на 7 ітерацій гібридну модель.

Кількість квантових операцій, необхідних для підготовки моделі з квантовим пристроєм в якості прихованого шару нейронної мережі, можна описати за допомогою наступної формули:

$$N_{QC} = epoch * N_{train} \quad (4.2)$$

Де N_{QC} – кількість квантових операцій, необхідних для підготовки моделі, $epoch$ – кількість епох навчання моделі, а N_{train} – кількість елементів тренувального набору даних. Що складає 34000 квантових операцій, необхідних для підготовки моделі на наборі даних Hurricane Damage, та 60000 квантових операцій для підготовки моделі на наборі даних CIFAR10 або CIFAR100.

Для отримання одного прогнозу гібридна модель даної архітектури потребує лише 1 квантової операції.

Для квантово-згорткової моделі кількість квантових операцій, необхідних для підготовки моделі, не залежить від кількості ітерацій навчання і може бути описана наступною формулою:

$$N_{QC} = image_h * image_w * (qubits * 3 + 1) * N_{train} \quad (4.3)$$

Де N_{QC} – кількість квантових операцій, необхідних для підготовки моделі, $image_h$ та $image_w$ – висота та ширина зображення у пікселях, $qubits$ – кількість кубітів у квантовому пристрої, а N_{train} – кількість елементів тренувального набору даних. Що складає 39936000 квантових операцій, необхідних для підготовки моделі на наборі даних CIFAR10 або CIFAR100, та 425984000 квантових операцій для підготовки моделі на наборі даних Hurricane Damage, зважаючи, що розмір зображень у наборі даних було зменшено до 128x128 пікселів.

Для отримання одного прогнозу модель квантово-згорткової архітектури потребує кількості квантових операцій, що може бути описана наступною формулою:

$$N_{QC} = image_h * image_w * (qubits * 3 + 1) \quad (4.4)$$

Де N_{QC} – кількість квантових операцій, необхідних для підготовки моделі, $image_h$ та $image_w$ – висота та ширина зображення у пікселях, а $qubits$ –

кількість кубітів у квантовому пристрої. Що складає 13312 квантових операцій, необхідних для обробки моделлю одного зображення з набору даних CIFAR10 або CIFAR100, та 212992 квантових операцій для обробки моделлю одного зображення з набору даних Hurricane Damage, зважаючи, що розмір зображень у наборі даних було зменшено до 128x128 пікселів.

Таким чином, з огляду на результати обчислень кількості квантових операцій, необхідних для підготовки, можна зробити висновок, що незважаючи на те, що квантово-згортова модель обробляє кожне зображення навчального набору даних лише один раз, для підготовки моделі їй потрібно значно більше квантових операцій – 39.49 мільйона операцій для навчання моделі на наборі даних CIFAR10 і 425.984 мільйона операцій для навчання моделі на наборі даних SIHD, що на кілька порядків більше ніж вимагає модель альтернативної гібридної архітектури – 60000 та 34000 квантових операцій відповідно.

Кількість квантових операцій, що необхідні для функціонування моделі також відповідає тому ж шаблону та квантово-згортова модель вимагає на порядки більше квантових операцій для обчислення одного прогнозу моделі. Кількість квантових операцій, необхідна для роботи квантово-згортової моделі лінійно залежить від розміру зображень у наборі даних та складає 13312 квантових операцій для обробки одного зображення з набору даних CIFAR10 та 212992 квантових операцій для обробки одного, зменшеного до розміру 128×128 пікселів, зображення з набору даних SIHD. В той же час кількість квантових операцій, необхідних для обчислення одного передбачення моделі, що базується на основі квантового пристрою, який виконує роль одного з прихованих шарів глибокої нейронної мережі, є константною та завжди складає 1 операцію. Таким чином можна зробити висновок, що збільшення розміру зображень набору даних значно підвищує вартість застосування квантово-згортової архітектури для вирішення задачі.

Для комплексного порівняння ефективності роботи розглянутих методів важливо також розглянути і порівняльну складність класичної частини розглянутих гібридних моделей. Складність класичної частини квантово-згортової моделі у 4 рази вища порівняно з еталонною моделлю та більш ніж у

4 рази вища порівняно з гібридною моделлю альтернативної архітектури. Це пояснюється тим, що кожне зображення в наборі даних потребує обробки чотирьох варіацій, створених квантовим шаром, що безпосередньо збільшує час обробки класичною частиною гібридної нейронної мережі в 4 рази. В той же час, складність класичної частини гібридної квантово-класичної нейронної мережі, що містить квантовий пристрій у якості одного з прихованих шарів нейронів, є дещо нижчою порівняно з еталонною моделлю, адже частина шарів (в нашому експериментів 1 шар) замінений на квантовий пристрій, що дещо зменшує складність класичної частини моделі.

ВИСНОВКИ ДО РОЗДІЛУ 4

Фаза попередньої обробки даних включала розширення вхідних наборів даних шляхом генерації кількох версій зображень на основі різних комбінацій гіперпараметрів. Зразки попередньо оброблених наборів даних CIFAR100 та SIHD показали впізнавані візерунки з оригінальних зображень, що підкреслює ефективність процесу розширення даних. Зокрема, кожне зображення, створене квантово-згортковим шаром, мало унікальні характеристики, що підкреслює різноманітність, введено квантовими операціями.

Дослідження простих конфігурацій гібридних нейронних мереж, включаючи варіації з квантово попередньо обробленими вхідними даними та класичними шарами нейронної мережі, надало цінні порівняння з традиційними згортковими нейронними мережами. Аналіз залежності продуктивності гібридних нейронних мереж від розміру навчального набору даних показав, що продуктивність гібридних нейронних мереж, зокрема конфігурації «Квантовий пристрій + згорткова нейронна мережа», покращується зі збільшенням розміру навчальних даних та складності архітектури мережі. Ця чутливість до параметрів навчальних даних і складності моделей підкреслює важливість ретельного проектування моделей і обробки наборів даних для максимізації продуктивності гібридних нейронних мереж.

Крім того, результати дослідження впливу кількості згорткових фільтрів і кількості шарів згорткових нейронних мереж на продуктивність гібридних нейронних мереж показали, що збільшення складності моделі може підвищити продуктивність до певного порогу, але надмірна складність може призвести до зниження результатів або перенавчання. Зокрема, додавання другого згорткового шару значно покращило продуктивність традиційних згорткових нейронних мереж, підкреслюючи складну взаємодію між архітектурою моделі та її продуктивністю.

Дослідження продуктивності гібридних нейронних мереж у схемах грубої класифікації підтвердили виявлені тенденції, припускаючи можливу універсальність цих ефектів у задачах різної складності класифікації.

Експерименти з простою конфігурацією гібридних нейронних мереж показали, що квантово-згорткова операція створює більш складну версію вхідних даних, тому для досягнення максимальної точності необхідна складніша класична частина гібридної нейронної мережі. Це стосується як структури класичної частини (згорткова нейронна мережа показала значно вищу точність, ніж щільна нейронна мережа), так і її складності (кількість фільтрів у шарі згорткової нейронної мережі). Крім того, експерименти показали, що конфігурація квантово-згорткової гібридної нейронної мережі є більш чутливою до розміру навчального набору даних, ніж її класичний аналог.

Експерименти, проведені в рамках дослідження на більш складних конфігураціях гібридних нейронних мереж, показали доцільність застосування квантово-згорткових гібридних нейронних мереж для вирішення практичних завдань на прикладі задач класифікації зображень. Найкращі результати у багатокласовій класифікації на наборі даних CIFAR100 показала класична модель EfficientNet. Усі гібридні нейронні мережі продемонстрували точність на 5-10% нижчу, порівнюючи з класичними моделями. Експерименти на наборі даних SIHD показали іншу картину — гібридні нейронні мережі, побудовані на основі ResNet і EfficientNet, перевершили свої класичні аналоги з незначною перевагою (0,3–2,2%). Це можна пояснити тим, що операція квантової згортки значно збільшує розмір навчального набору даних, що виявилось корисним у ситуаціях, коли початковий набір даних був оптимального розміру.

Техніка більшості голосів була застосована для прогнозування фінальних результатів, що відображає реальний сценарій використання, оскільки квантово-згортковий шар гібридних нейронних мереж множить розмір набору даних на n , де n є кількістю кубітів у квантовій схемі. Крім того, коли модель використовується для нового елемента даних, ми очікуємо отримати прогноз для одного класу вхідних даних; отже, класична частина моделі обробляє кожен шар, створений квантовим пристроєм, а потім алгоритм голосування визначає найбільш поширений клас, зафіксований класичною частиною гібридних нейронних мереж, як остаточне рішення. Ця методика довела свою високу ефективність для гібридних нейронних мереж у складніших задачах, як це було

продемонстровано на наборі даних CIFAR100. Для менш складних задач класифікації зображень, таких як бінарна класифікація на наборі даних SIHD, де гібридні нейронні мережі продемонстрували високий рівень точності (98,81%), техніка голосування виявилася менш ефективною. Проте ця техніка також забезпечує гнучкість конфігурації для вирішення ситуацій, коли модель не може надійно класифікувати зображення; наприклад, якщо техніка голосування призводить до нічийного результату в питанні, чи була будівля пошкоджена, ми можемо автоматично класифікувати будівлю як пошкоджену, щоб зменшити кількість хибних негативних результатів, що є більш небажаними, ніж хибні позитивні результати в цьому випадку.

Проте застосування гібридних нейронних мереж також має певні суттєві обмеження. Найбільшим з них є обмежена доступність ресурсів квантових обчислень на даний момент. Цей недолік частково компенсується запропонованою архітектурою гібридної нейронної мережі, де квантовий пристрій слугує першим шаром нейронної мережі, що дозволяє виконувати квантову трансформацію навчальних даних лише один раз. Однак навіть у цьому випадку забезпечення квантовими обчислювальними ресурсами все ще є основним вузьким місцем для масштабованості гібридних нейронних мереж.

Ще одним можливим обмеженням є необхідність передачі даних між квантовою та класичною частинами нейронної мережі. Оскільки квантові та класичні частини виконуються на різному обладнанні, потрібна передача даних через мережу, що може бути досить повільною, якщо фізична відстань між квантовим і класичним обладнанням є значною. Однак ми вважаємо, що прискорення обчислень завдяки використанню квантових обчислень буде більшим за часові витрати на передачу даних через мережу.

Порівняння розглянутих методів гібридизації класичних та квантових обчислень для завдання класифікації зображень підтвердили, що гібридні квантово-згорткові нейронні мережі, здатні досягати вищого рівня точності порівняно з референтною класичною моделлю та порівняно з альтернативним розглянутим підходом гібридизації. Однак водночас квантово-згорткові нейронні мережі вимагають значної кількості квантових операцій як для

підготовки, так і для функціонування моделі, а також потребують у 4 рази більше обчислювального часу класичного обладнання через необхідність обробки у 4 рази більшого обсягу даних порівняно з альтернативними підходами, розглянутими в цьому дослідженні та порівняно з референтними моделями.

Результати експериментів також показали, що гібридні нейронні мережі, побудовані з використанням квантового пристрою, який виконує роль одного із прихованих шарів нейронної мережі, можуть бути життєздатним підходом, незважаючи на значно нижчу точність порівняно з альтернативними методами гібридизації. Доцільність цього підходу пояснюється відносно низькою вартістю в термінах кількості квантових операцій та зменшеною складністю класичної частини нейронної мережі, що може підвищити швидкість обчислень моделі завдяки зменшенню глибини класичної частини моделі та використанню квантового пристрою замість видалених шарів нейронної мережі.

Метод гібридизації на основі квантово-згорткових нейронних мереж показав себе доцільним та потенційно ефективним рішенням для критично важливих практичних завдань, які вимагають високої точності моделі та, водночас, толерантні до збільшення часу обробки і значно вищої вартості через велику кількість необхідних квантових та класичних операцій.

Метод гібридизації на основі глибоких гібридних мереж, що містять квантовий пристрій у якості одного з прихованих шарів нейронної мережі, довели свою доцільність для практичних завдань, критичних до часу, які потребують підвищеної швидкості обробки даних та можуть допустити зниження точності моделі.

Отже, отримані результати можна узагальнити, зробивши висновок, що використання запропонованого комплексного методу гібридизації квантових та класичних обчислень дозволяє підвищити ефективність роботи моделей ШІ в ряді випадків, залежно від вимог конкретної задачі.

ВИСНОВКИ

В дисертаційній роботі розв'язано актуальне наукове завдання гібридизації квантових та класичних обчислень для створення гібридних квантово-класичних глибоких нейронних мереж. В результаті дисертаційної роботи отримано наступні наукові та практичні результати:

- Розроблено комплексний метод гібридизації класичних та квантових обчислень, який, на відміну від існуючих рішень, може бути направлений на збільшення швидкодії моделей ШІ або на збільшення точності роботи моделей залежно від вимог задачі за рахунок залучення квантових обчислень у різних частинах гібридних нейронних мереж (у якості квантово-згорткового (quanvolutional) шару або ж у якості заміни частини прихованих повнозв'язних шарів глибокої нейронної мережі).
- Розроблено метод побудови гібридних систем ШІ, який, на відміну від існуючих рішень, дозволяє залучати незмінні класичні SOTA моделі та дозволяє досягнути вищих значень точності роботи моделей, порівнюючи з альтернативними підходами, за рахунок залучення квантових обчислень у якості першого квантово-згорткового (quanvolutional) шару гібридних моделей.
- Удосконалено метод використання трансферного навчання у контексті гібридних квантово-класичних нейронних мереж, що дозволяє зменшити обсяг ресурсів (час обчислень, енергія), потрібних для підготовки моделі, та підвищити точність роботи гібридних квантово-класичних моделей ШІ.
- Розроблено удосконалення методу використання багатоканальності (multi-channel) даних для вирішення задачі класифікації зображень, що дозволяє підвищити точність роботи гібридної моделі ШІ за рахунок використання квантового пристрою для створення додаткових каналів на основі вхідних даних.
- Запропоновано розширення методу штучного розширення різноманітності даних (data augmentation) для вирішення задачі

класифікації зображень, що дозволяє підвищити стійкість по відношенню до невідомих даних (generalization) гібридної моделі III за рахунок використання квантово-згорткового штучного розширення різноманітності даних (quantvolutional data augmentation) і використання результатів роботи квантового пристрою в якості додаткових каналів вхідних даних.

- Розроблено розширення методу використання багатоструктурної (multi-backbone) моделі у якості класичної частини гібридної квантово-згорткової (quantvolutional) моделі для вирішення задачі класифікації зображень, що дозволяє підвищити точність роботи гібридної моделі III за рахунок незалежного паралельного опрацювання додаткових каналів, що були створені за допомогою квантового пристрою, спеціалізованими базовими моделями, що входять до складу багатоструктурної класичної частини моделі.
- Проведено аналіз впливу архітектури квантового пристрою на ефективність роботи гібридної моделі, що будується на його основі. Найкращий результат показала квантова схема з RY архітектурою.
- Проведено аналіз ефективності запропонованого методу гібридизації квантових та класичних обчислень.

Результати було продемонстровано на прикладі практичної задачі класифікації пошкоджень в зонах стихійних лих та в контексті використання систем III загального застосування, особливо для задач, що вимагають надвисокої швидкодії.

Дослідження дотримується основних принципів Open Science [109] тому вихідний код всіх експериментів, проведених в межах дисертаційного дослідження, надається у вільному доступі на платформах GitHub та Kaggle у вигляді Jupyter Notebooks [110-115]. Квантово-опрацьовані набори даних CIFAR100 та SIHD також надаються у вільному доступі на платформі Kaggle [116, 117].

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. What Kind of Problems Can Quantum Computers Solve? [Електронний ресурс]. — Режим доступу: <https://levelup.gitconnected.com/what-kind-of-problems-can-quantum-computers-solve-63234eb9fd76>
2. Y. Kanamori, S.-M. Yoo, W.D. Pan, F.T. Sheldon: A short survey on quantum computers. *International Journal of Computers and Applications* Vol. 28, No. 3, (2006)
3. Biamonte, J.; Wittek, P.; Pancotti, N.; Rebentrost, P.; Wiebe, N.; Lloyd, S. Quantum machine learning. *Nature* 2017, 549, 195–202.
4. García, D.P.; Cruz-Benito, J.; García-Peñalvo, F.J. Systematic Literature Review: Quantum Machine Learning and Its Applications. *arXiv* 2023, arXiv:2201.04093.
5. Brown, Tom B.; Mann, Benjamin; Ryder, Nick; Subbiah, Melanie; Kaplan, Jared; Dhariwal, Prafulla; Neelakantan, Arvind; Shyam, Pranav; Sastry, Girish; Askell, Amanda; Agarwal, Sandhini; Herbert-Voss, Ariel; Krueger, Gretchen; Henighan, Tom; Child, Rewon; Ramesh, Aditya; Ziegler, Daniel M.; Wu, Jeffrey; Winter, Clemens; Hesse, Christopher; Chen, Mark; Sigler, Eric; Litwin, Mateusz; Gray, Scott; Chess, Benjamin; Clark, Jack; Berner, Christopher; McCandlish, Sam; Radford, Alec; Sutskever, Ilya; Amodei, Dario: Language Models are Few-Shot Learners. *arXiv*:2005.14165, (2020).
6. Number of Parameters in GPT-4 (Latest Data) [Електронний ресурс]. — Режим доступу: <https://explodingtopics.com/blog/gpt-parameters>.
7. Anne-Laure Ligozat, Alex De Vries, “Generative AI: energy consumption soars.” [Електронний ресурс]. — Режим доступу: <https://www.polytechnique-insights.com/en/columns/energy/generative-ai-energy-consumption-soars/>.
8. Ragel, R. G., and Parameswaran, S.: A hybrid hardware–software technique to improve reliability in embedded processors. *ACM Transactions on Embedded Computing Systems (TECS)* 10(3), 1–16 (2011).

9. Pellizzoni, R., and Caccamo, M., “Hybrid hardware-software architecture for reconfigurable real-time systems.” In: 2008 IEEE Real-Time and Embedded Technology and Applications Symposium, pp. 273–284. IEEE (2008) <https://doi.org/10.1109/RTAS.2008.14>.
10. Z. Yang, M. Zolanvari and R. Jain, “A Survey of Important Issues in Quantum Computing and Communications,” in IEEE Communications Surveys & Tutorials, vol. 25, no. 2, pp. 1059-1094 (2023), doi: 10.1109/COMST.2023.3254481.
11. Nielsen, Michael A., and Isaac L. Chuang. 2010. “Quantum Computation and Quantum Information.” 10th anniversary edition. Cambridge: Cambridge University Press.
12. Moses, S. A.; Baldwin, C. H.; Allman, M. S.; Ancona, R.; Ascarrunz, L.; Barnes, C.; Bartolotta, J.; Bjork, B.; Blanchard, P.; Bohn, M.; Bohnet, J. G.; Brown, N. C.; Burdick, N. Q.; Burton, W. C.; Campbell, S. L. (2023-12-18). “A Race-Track Trapped-Ion Quantum Processor.” Physical Review X. 13 (4). arXiv:2305.03828
13. Castelvechi, Davide (5 January 2017). “Quantum computers ready to leap out of the lab in 2017.” Nature. 541 (7635): 9–10. Bibcode:2017Natur.541....9C. doi:10.1038/541009a. PMID 28054624. S2CID 4447373.
14. Shen, L. Y. L. (1972-02-01). “Superconductivity of Tantalum, Niobium and Lanthanum Studied by Electron Tunneling: Problems of Surface Contamination.” AIP Conference Proceedings. 4 (1): 31–44. Bibcode:1972AIPC....4...31S. doi:10.1063/1.2946195. ISSN 0094-243X
15. Arute, Frank; Arya, Kunal; Babbush, Ryan; Bacon, Dave; Bardin, Joseph C.; Barends, Rami; Biswas, Rupak; Boixo, Sergio; Brandao, Fernando G. S. L.; Buell, David A.; Burkett, Brian; Chen, Yu; Chen, Zijun; Chiaro, Ben; Collins, Roberto; Courtney, William; Dunsworth, Andrew; Farhi, Edward; Foxen, Brooks; Fowler, Austin; Gidney, Craig; Giustina, Marissa; Graff, Rob; Guerin, Keith; Habegger, Steve; Harrigan, Matthew P.; Hartmann, Michael J.; Ho, Alan; Hoffmann, Markus; Huang, Trent; Humble, Travis S.;

- Isakov, Sergei V.; Jeffrey, Evan; Jiang, Zhang; Kafri, Dvir; Kechedzhi, Kostyantyn; Kelly, Julian; Klimov, Paul V.; Knysh, Sergey; Korotkov, Alexander; Kostritsa, Fedor; Landhuis, David; Lindmark, Mike; Lucero, Erik; Lyakh, Dmitry; Mandrà, Salvatore; McClean, Jarrod R.; McEwen, Matthew; Megrant, Anthony; Mi, Xiao; Michielsen, Kristel; Mohseni, Masoud; Mutus, Josh; Naaman, Ofer; Neeley, Matthew; Neill, Charles; Niu, Murphy Yuezhen; Ostby, Eric; Petukhov, Andre; Platt, John C.; Quintana, Chris; Rieffel, Eleanor G.; Roushan, Pedram; Rubin, Nicholas C.; Sank, Daniel; Satzinger, Kevin J.; Smelyanskiy, Vadim; Sung, Kevin J.; Trevithick, Matthew D.; Vainsencher, Amit; Villalonga, Benjamin; White, Theodore; Yao, Z. Jamie; Yeh, Ping; Zalcman, Adam; Neven, Hartmut; Martinis, John M. (October 2019). “Quantum supremacy using a programmable superconducting processor.” *Nature*. 574 (7779): 505–510. arXiv:1910.11333. Bibcode:2019Natur.574..505A. doi:10.1038/s41586-019-1666-5. PMID 31645734
16. Shor, Peter W. (1995). “Scheme for reducing decoherence in quantum computer memory.” *Physical Review A*. 52 (4): R2493–R2496. Bibcode:1995PhRvA..52.2493S. doi:10.1103/PhysRevA.52.R2493. PMID 9912632
17. Dyakonov, M. I. (14 October 2006). S. Luryi; Xu, J.; Zaslavsky, A. (eds.). “Is Fault-Tolerant Quantum Computation Really Possible?” *Future Trends in Microelectronics. Up the Nano Creek*: 4–18. arXiv:quant-ph/0610117. Bibcode:2006quant.ph.10117D
18. Bacon, D. (2001). “Decoherence, control, and symmetry in quantum computers” in arXiv:quant-ph/0305025
19. Torsten Hoefler; Thomas Häner; Matthias Troyer (May 2023). “Disentangling Hype from Practicality: On Realistically Achieving Quantum Advantage.” *Communications of the ACM*.
20. Travis L. Scholten and Carl J. Williams and Dustin Moody and Michele Mosca and William Hurley and William J. Zeng and Matthias Troyer and Jay

- M. Gambetta (2024). “Assessing the Benefits and Risks of Quantum Computers”. arXiv:2401.16317. <https://doi.org/10.48550/arXiv.2401.16317>
21. Joshua J. Goings, Alec White, Joonho Lee, Christofer S. Tautermann, Matthias Degroote, Craig Gidney, Toru Shiozaki, Ryan Babbush, and Nicholas C. Rubin (2022). “Reliably assessing the electronic structure of cytochrome P450 on today’s classical computers and tomorrow’s quantum computers”. <https://doi.org/10.1073/pnas.2203533119>
22. “What is carbon sequestration?” [Электронный ресурс]. — Режим доступа: <https://www.usgs.gov/faqs/what-carbon-sequestration>
23. “We need to decarbonize software” [Электронный ресурс]. — Режим доступа: <https://spectrum.ieee.org/green-software>
24. Nicholas C. Rubin, Dominic W. Berry, Fionn D. Malone, Alec F. White, Tanuj Khattar, A. Eugene DePrince, III, Sabrina Sicolo, Michael Kühn, Michael Kaicher, Joonho Lee, and Ryan Babbush (2023). “Fault-Tolerant Quantum Simulation of Materials Using Bloch Orbitals”. PRX Quantum 4, 040303 — Published 6 October 2023. <https://doi.org/10.1103/PRXQuantum.4.040303>
25. Degen, C. L. and Reinhard, F. and Cappellaro, P. “Quantum sensing”. Reviews of Modern Physics. 3 (89). DOI: 10.1103/revmodphys.89.035002.
26. “The quantum enhanced LIGO detector sets new sensitivity record” [Электронный ресурс]. — Режим доступа: <https://www.ligo.org/science/Publication-SqueezedVacuum/index.php>
27. Tse, M.; Yu, Haocun; Kijbunchoo, N.; Fernandez-Galiana, A.; Dupej, P.; Barsotti, L.; Blair, C. D.; Brown, D. D.; Dwyer, S. E.; Effler, A.; Evans, M. (5 December 2019). “Quantum-Enhanced Advanced LIGO Detectors in the Era of Gravitational-Wave Astronomy”. Physical Review Letters. 123 (23): 231107. Bibcode:2019PhRvL.123w1107T. doi:10.1103/PhysRevLett.123.231107. hdl:1721.1/136579.2. PMID 31868462

28. “Researchers build first modular quantum brain sensor, record signal” [Электронный ресурс]. — Режим доступа: <https://phys.org/news/2021-06-modular-quantum-brain-sensor.html>
29. Lanzagorta, Marco (October 31, 2011). “Quantum Radar.” *Synthesis Lectures on Quantum Computing*. 3 (1): 1–139. doi:10.2200/S00384ED1V01Y201110QMC005. S2CID 27569963
30. Hsin-Yuan Huang et al., “Quantum advantage in learning from experiments”. *Science* 376, 1182-1186 (2022). DOI: 10.1126/science.abn7293
31. Toğaçar, M.; Ergen, B.; Tümen, V. “Use of dominant activations obtained by processing OCT images with the CNNs and slime mold method in retinal disease detection.” *Biocybern. Biomed. Eng.* 2022, 42, 646–666.
32. Ahalya, R.K.; Almutairi, F.M.; Snehalatha, U.; Dhanraj, V.; Aslam, S.M. “RANet: A custom CNN model and quantum convolutional neural network for the automated detection of rheumatoid arthritis in hand thermal images.” *Sci. Rep.* 2023, 13, 15638.
33. Ullah, U.; Jurado, A.G.O.; Gonzalez, I.D.; Garcia-Zapirain, B. “A Fully Connected Quantum Convolutional Neural Network for Classifying Ischemic Cardiopathy.” *IEEE Access* 2022, 10, 134592–134605.
34. Dong, Y.; Fu, Y.; Liu, H.; Che, X.; Sun, L.; Luo, Y. “An improved hybrid quantum-classical convolutional neural network for multi-class brain tumor MRI classification.” *J. Appl. Phys.* 2023, 133, 064401.
35. Kyro, G.W.; Brent, R.I.; Batista, V.S. “HAC-Net: A Hybrid Attention-Based Convolutional Neural Network for Highly Accurate Protein–Ligand Binding Affinity Prediction.” *J. Chem. Inf. Model.* 2023, 10, 1947–1960.
36. Yan, F.; Iliyasu, A.M.; Venegas-Andraca, S.E. “A survey of quantum image representations.” *Quantum Inf. Process* 2016, 15, 1–35.
37. Dou, T.; Zhang, G.; Cui, W. “Efficient quantum feature extraction for CNN-based learning.” *J. Frankl. Inst.* 2023, 360, 7438–7456.
38. Shi, S.; Wang, Z.; Cui, G.; Wang, S.; Shang, R.; Li, W.; Wei, Z.; Gu, Y. “Quantum-inspired complex convolutional neural networks.” *Appl. Intell.* 2022, 52, 17912–17921.

39. Liu, J.; Lim, K.H.; Wood, K.L.; Huang, W.; Guo, C.; Huang, H.L. “Hybrid quantum-classical convolutional neural networks.” *Sci. China Phys. Mech. Astron.* 2021, 64, 290311.
40. Wang, A.; Hu, J.; Zhang, S.; Li, L. “Shallow hybrid quantum-classical convolutional neural network model for image classification.” *Quantum Inf. Process* 2024, 23, 17.
41. Patil, A.P.; Pandey, S.; Das Kasat, N.; Modi, S.; Raj, S.; Kulkarni, R. “Implementation of Handwritten Character Recognition using Quantum Convolutional Neural Network.” In *Proceedings of the 2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon)*, Vijaypur, India, 20–21 November 2022; pp. 1–4.
42. Li, Y.C.; Zhou, R.-G.; Xu, R.Q.; Luo, J.; Hu, W.W. “A quantum deep convolutional neural network for image recognition.” *Quantum Sci. Technol.* 2020, 5, 044003.
43. Chalumuri, A.; Kune, R.; Manoj, B.S. “A hybrid classical-quantum approach for multi-class classification.” *Quantum Inf. Process* 2021, 20, 119.
44. Konar, D.; Sarma, A.D.; Bhandary, S.; Bhattacharyya, S.; Cangi, A.; Aggarwal, V. “A shallow hybrid classical–quantum spiking feedforward neural network for noise-robust image classification.” *Appl. Soft Comput.* 2023, 136, 110099.
45. He, K.; Zhang, X.; Ren, S.; Sun, J. “Identity Mappings in Deep Residual Networks.” *CoRR* 2016, arXiv:1603.05027.
46. “IBM Quantum.” [Электронный ресурс]. — Режим доступа: <https://www.ibm.com/quantum>.
47. “IBM Unveils Breakthrough 127-Qubit Quantum Processor.” [Электронный ресурс]. — Режим доступа: <https://newsroom.ibm.com/2021-11-16-IBM-Unveils-Breakthrough-127-Qubit-Quantum-Processor>.
48. “Qiskit” [Электронный ресурс]. — Режим доступа: <https://www.ibm.com/quantum/qiskit>.
49. “Azure Quantum cloud service.” [Электронный ресурс]. — Режим доступа: <https://azure.microsoft.com/en-us/products/quantum>.

50. “IonQ.” [Электронный ресурс]. — Режим доступа: <https://ionq.com/>.
51. “Pasqal.” [Электронный ресурс]. — Режим доступа: <https://www.pasqal.com/>.
52. “Quantinuum.” [Электронный ресурс]. — Режим доступа: <https://www.quantinuum.com/>.
53. “Rigetti.” [Электронный ресурс]. — Режим доступа: <https://www.rigetti.com/>.
54. “AWS Braket.” [Электронный ресурс]. — Режим доступа: <https://aws.amazon.com/braket/>.
55. “IQM.” [Электронный ресурс]. — Режим доступа: <https://www.meetiqm.com/>.
56. “QuEra.” [Электронный ресурс]. — Режим доступа: <https://www.quera.com/>.
57. Acharya, Rajeev; Abanin, Dmitry A.; Aghababaie-Beni, Laleh; Aleiner, Igor; Andersen, Trond I.; Ansmann, Markus; Arute, Frank; Arya, Kunal; Asfaw, Abraham; Astrakhantsev, Nikita; Atalaya, Juan; Babbush, Ryan; Bacon, Dave; Ballard, Brian; Bardin, Joseph C., “Quantum error correction below the surface code threshold”. *Nature*: 1–3. arXiv:2408.13687. doi:10.1038/s41586-024-08449-y (2024-12-09).
58. “IBM Unveils 400 Qubit-Plus Quantum Processor and Next-Generation IBM Quantum System Two.” [Электронный ресурс]. — Режим доступа: <https://newsroom.ibm.com/2022-11-09-IBM-Unveils-400-Qubit-Plus-Quantum-Processor-and-Next-Generation-IBM-Quantum-System-Two>.
59. “The hardware and software for the era of quantum utility is here.” [Электронный ресурс]. — Режим доступа: <https://www.ibm.com/quantum/blog/quantum-roadmap-2033>.
60. “Technology for the quantum future.” [Электронный ресурс]. — Режим доступа: <https://www.ibm.com/quantum/technology#roadmap>.
61. “Google quantum computing roadmap.” [Электронный ресурс]. — Режим доступа: <https://quantumai.google/roadmap>.

62. “PennyLane.” [Электронный ресурс]. — Режим доступа: <https://pennylane.ai/>.
63. Cheng C-S, Behzadan AH, Noshadravan A. “Deep learning for post-hurricane aerial damage assessment of buildings.” *Comput Aided Civ Inf.* 2021; 36: 695–710. <https://doi.org/10.1111/mice.12658>.
64. Yann LeCun, Corinna Cortes, Christopher J. C. Burges, “*MNIST handwritten digits database*” [Электронный ресурс]. — Режим доступа: <http://yann.lecun.com/exdb/mnist/>.
65. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. “Gradient-based learning applied to document recognition” in *Proceedings of the IEEE*, 86(11), 2278-2324, 1998.
66. “Fashion-MNIST” [Электронный ресурс]. — Режим доступа: <https://github.com/zalandoresearch/fashion-mnist>.
67. Xiao, H., Rasul, K., & Vollgraf, R. “Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms” in *arXiv preprint arXiv:1708.07747*, 2017.
68. “CIFAR10 and CIFAR100 datasets” [Электронный ресурс]. — Режим доступа: <https://www.cs.toronto.edu/~kriz/cifar.html>.
69. Krizhevsky, A., & Hinton, G. (2009). “Learning multiple layers of features from tiny images” [Электронный ресурс]. — Режим доступа: <https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf>.
70. Quoc Dung Cao, Youngjun Choe, “Building Damage Annotation on Post-Hurricane Satellite Imagery Based on Convolutional Neural Networks” in *arXiv:1807.01688*, 2018.
71. Stone, M (1977). “An Asymptotic Equivalence of Choice of Model by Cross-Validation and Akaike's Criterion”. *Journal of the Royal Statistical Society, Series B (Methodological)*. 39 (1): 44–47. doi:10.1111/j.2517-6161.1977.tb01603.x.
72. “RYGate” [Электронный ресурс]. — Режим доступа: <https://docs.quantum.ibm.com/api/qiskit/qiskit.circuit.library.RYGate>.

73. “RXGate” [Электронный ресурс]. — Режим доступа: <https://docs.quantum.ibm.com/api/qiskit/qiskit.circuit.library.RXGate>.
74. “RZGate” [Электронный ресурс]. — Режим доступа: <https://docs.quantum.ibm.com/api/qiskit/qiskit.circuit.library.RZGate>.
75. “HGate” [Электронный ресурс]. — Режим доступа: <https://docs.quantum.ibm.com/api/qiskit/qiskit.circuit.library.HGate>.
76. “The Pauli spin matrices” [Электронный ресурс]. — Режим доступа: https://www.feynmanlectures.caltech.edu/III_11.html#Ch11-S1.
77. “XGate” [Электронный ресурс]. — Режим доступа: <https://docs.quantum.ibm.com/api/qiskit/qiskit.circuit.library.XGate>.
78. “YGate” [Электронный ресурс]. — Режим доступа: <https://docs.quantum.ibm.com/api/qiskit/qiskit.circuit.library.YGate>.
79. “ZGate” [Электронный ресурс]. — Режим доступа: <https://docs.quantum.ibm.com/api/qiskit/qiskit.circuit.library.ZGate>.
80. “SwapGate” [Электронный ресурс]. — Режим доступа: <https://docs.quantum.ibm.com/api/qiskit/qiskit.circuit.library.SwapGate>.
81. Nielsen, M. A., & Chuang, I. L. (2010). “*Quantum Computation and Quantum Information: 10th Anniversary Edition.*” Cambridge: Cambridge University Press.
82. “Python” [Электронный ресурс]. — Режим доступа: <https://www.python.org/>.
83. “PyTorch” [Электронный ресурс]. — Режим доступа: <https://pytorch.org/>
84. “TensorFlow” [Электронный ресурс]. — Режим доступа: <https://www.tensorflow.org/>.
85. “Jupyter” [Электронный ресурс]. — Режим доступа: <https://jupyter.org/>.
86. I.M. Georgescu, S. Ashhab, Franco Nori, “Quantum simulation” in *Rev. Mod. Phys.* 86, 2014.
87. “Kaggle: Your Machine Learning and Data Science Community.” [Электронный ресурс]. — Режим доступа: <https://www.kaggle.com/>.
88. “Google Colaboratory.” [Электронный ресурс]. — Режим доступа: <https://colab.google/>.

89. “Hybrid quantum-classical Neural Networks with PyTorch and Qiskit” [Электронный ресурс]. — Режим доступа: <https://qiskit.org/textbook/ch-machine-learning/machine-learning-qiskit-pytorch.html>
90. Gavin E. Crooks, “Gradients of parameterized quantum gates using the parameter-shift rule and gate decomposition” in arxiv:1905.13311, 2019.
91. Maxwell Henderson, Samriddhi Shakya, Shashindra Pradhan, Tristan Cook: “Quantum Evolutional Neural Networks: Powering Image Recognition with Quantum Circuits.” arxiv:1904.04767, (2019)
92. Sina Aeeneh and Nikola Zlatanov and Jiangshan Yu, “New Bounds on the Accuracy of Majority Voting for Multi-Class Classification” in arXiv:2309.09564, (2023).
93. Fuzhen Zhuang and Zhiyuan Qi and Keyu Duan and Dongbo Xi and Yongchun Zhu and Hengshu Zhu and Hui Xiong and Qing He, “A Comprehensive Survey on Transfer Learning” in arXiv:1911.02685, (2019).
94. Yuri Gordienko, Yevhenii Trochun and Sergii Stirenko, “Multimodal Quantum Evolutional and Convolutional Neural Networks for Multi-class Image Classification”. *Big Data and Cognitive Computing*. 2024, 8(7), 75, ISSN: 2504-2289. DOI 10.3390/bdcc8070075.
95. Y. Gordienko, Y. Trochun, V. Taran, A. Khmelnytskyi, S. Stirenko, “HNN-QCn: Hybrid Neural Network with Multiple Backbones and Quantum Transformation as Data Augmentation Technique”. *AI* 2025, 6, 36, ISSN: 2673-2688. DOI 10.3390/ai6020036.
96. Casper Hansen, “ReLU” [Электронный ресурс]. — Режим доступа: <https://mlfromscratch.com/activation-functions-explained/#relu>.
97. Bala Priya C, “Softmax Activation Function: Everything You Need to Know” [Электронный ресурс]. — Режим доступа: <https://www.pinecone.io/learn/softmax-activation/>.
98. Yevhenii Trochun, Sergii Stirenko, Evgen Pavlov, Yuri Gordienko, “Impact of Hybrid Neural Network Structure on Performance of Multiclass Classification”, in *In Proceedings of the IEEE 19th International Conference on Smart Technologies, EUROCON'2021*, 2021.

99. Weining Hu, "Flooding Damage Detection from Post-Hurricane Satellite Imagery Based on Convolutional Neural Networks" [Электронный ресурс]. — Режим доступа: <https://github.com/weining20000/Flooding-Damage-Detection-from-Post-Hurricane-Satellite-Imagery-Based-on-CNN>
100. Xiangyu Zhang, Jianhua Zou, Kaiming He, Jian Sun, "Accelerating Very Deep Convolutional Networks for Classification and Detection" in *arXiv:1505.06798*, 2015.
101. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition" in *arXiv:1512.03385*, 2015.
102. Y. Trochun et al., "Hurricane Damage Detection by Classic and Hybrid Classic-Quantum Neural Networks," 2021 International Conference on Space-Air-Ground Computing (SAGC), Huizhou, China, 2021, pp. 152-156. IEEE. ISBN: 978-1-6654-0012-1. DOI 10.1109/SAGC52752.2021.00033.
103. Y. Trochun, S. Stirenko, O. Rokovyi, O. Alienin, E. Pavlov and Y. Gordienko, "Hybrid Classic-Quantum Neural Networks for Image Classification," 2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Cracow, Poland, 2021, pp. 968-972. IEEE. ISSN: 2770-4254. DOI 10.1109/IDAACS53288.2021.9661011.
104. Yevhenii Trochun, Sergii Stirenko and Yuri Gordienko, "Multichannel Quantum Data Preprocessing for Image Classification Problems", Springer Book Series, Algorithms for Intelligent Systems.
105. Tan, M.; Le, Q.V. "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks". CoRR 2019, arXiv:1905.11946.
106. Cui, C.; Gao, T.; Wei, S.; Du, Y.; Guo, R.; Dong, S.; Lu, B.; Zhou, Y.; Lv, X.; Liu, Q.; et al. "PP-LCNet: A lightweight CPU convolutional neural network" in arXiv:2109.15099, 2021.
107. Bradley, A.P., "The use of the area under the ROC curve in the evaluation of machine learning algorithms". Pattern Recognit. 1997, 30, 1145–1159.
108. Y. Trochun and Y. Gordienko, "Effectiveness of Hybrid Quantum-Classical and Quantum Convolutional Neural Networks for image classification", Inf.

Comput. and Intell. syst. j., no. 5, pp. 68–79, Dec. 2024. ISSN 2708-4930.
DOI: 10.20535/2786-8729.5.2024.318743.

109. UNICEF, “Understanding Open Science” [Электронный ресурс]. — Режим доступа: <https://unesdoc.unesco.org/ark:/48223/pf0000383323>.
110. Yevhenii Trochun, “Quanvolutional Neural Networks.” [Электронный ресурс]. — Режим доступа: <https://github.com/ZheniaTrochun/quantvolutional-neural-networks>.
111. Yevhenii Trochun, “Quantum Deep Learning.” [Электронный ресурс]. — Режим доступа: <https://github.com/ZheniaTrochun/quantum-deep-learning>.
112. Yevhenii Trochun, “vgg_cifar10_Y_w4_s1.” [Электронный ресурс]. — Режим доступа: <https://www.kaggle.com/code/yevheniitrochun/vgg-cifar10-y-w4-s1>.
113. Yevhenii Trochun and Yuri Gordienko, “CIFAR10_FULL_Q5_W4_LCNet050_MB1_BATCH_64_TRIAL11.” [Электронный ресурс]. — Режим доступа: <https://www.kaggle.com/code/yevheniitrochun/cifar10-full-q5-w4-lcnet050-mb1-batch-64-trial11>.
114. Yevhenii Trochun and Yuri Gordienko, “CIFAR10_FULL_Q4_W4_LCNet050_MB1_BATCH_64_TRIAL11.” [Электронный ресурс]. — Режим доступа: <https://www.kaggle.com/code/yevheniitrochun/cifar10-full-q4-w4-lcnet050-mb1-batch-64-trial11>.
115. Yevhenii Trochun and Yuri Gordienko, “CIFAR10_FULL_CLASSIC_LCNet050_MB1_TRIAL07_BATCH_64.” [Электронный ресурс]. — Режим доступа: <https://www.kaggle.com/code/yevheniitrochun/cifar10-full-classic-lcnet050-mb1-trial07-batch-64>.
116. Yevhenii Trochun, “Quantum-Augmented CIFAR100.” [Электронный ресурс]. — Режим доступа:

<https://www.kaggle.com/datasets/yevheniitrochun/quantum-augmented-cifar100>.

117. Yevhenii Trochun, “Quantum-Augmented Images of Hurricane Damage.” [Электронний ресурс]. — Режим доступу: <https://www.kaggle.com/datasets/yevheniitrochun/quantum-augmented-images-of-hurricane-damage-part1>.

ДОДАТОК А**Список публікацій здобувача**

1. Y. Trochun, E. Pavlov, S. Stirenko and Y. Gordienko, “Impact of Hybrid Neural Network Structure on Performance of Multiclass Classification,” IEEE EUROCON 2021 - 19th International Conference on Smart Technologies, Lviv, Ukraine, 2021, pp. 152-156. IEEE. ISBN 978-1-6654-3299-3. DOI 10.1109/EUROCON52738.2021.9535586 (Web of Science).
2. Y. Trochun et al., “Hurricane Damage Detection by Classic and Hybrid Classic-Quantum Neural Networks,” 2021 International Conference on Space-Air-Ground Computing (SAGC), Huizhou, China, 2021, pp. 152-156. IEEE. ISBN: 978-1-6654-0012-1. DOI 10.1109/SAGC52752.2021.00033 (Scopus).
3. Y. Trochun, S. Stirenko, O. Rokovyi, O. Alienin, E. Pavlov and Y. Gordienko, “Hybrid Classic-Quantum Neural Networks for Image Classification,” 2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Cracow, Poland, 2021, pp. 968-972. IEEE. ISSN: 2770-4254. DOI 10.1109/IDAACS53288.2021.9661011 (Scopus).
4. Y. Trochun, S. Stirenko and Y. Gordienko, “Multichannel Quantum Data Preprocessing for Image Classification Problems”, Springer Book Series, Algorithms for Intelligent Systems, ISSN: 2524-7565.
5. Yuri Gordienko, Yevhenii Trochun and Sergii Stirenko, “Multimodal Quantvolutional and Convolutional Neural Networks for Multi-class Image Classification”. Big Data and Cognitive Computing. 2024, 8(7), 75, ISSN: 2504-2289. DOI 10.3390/bdcc8070075 (Scopus Q2).
6. Y. Trochun and Y. Gordienko, “Effectiveness of Hybrid Quantum-Classical and Quantvolutional Neural Networks for image classification”, Inf. Comput. and Intell. syst. j., no. 5, pp. 68–79, Dec. 2024. ISSN 2708-4930. DOI: 10.20535/2786-8729.5.2024.318743 (Фахове видання категорії «Б»).

7. Y. Gordienko, **Y. Trochun**, V. Taran, A. Khmelnytskyi, S. Stirenko, “HNN-QCn: Hybrid Neural Network with Multiple Backbones and Quantum Transformation as Data Augmentation Technique”. *AI 2025*, 6, 36, ISSN: 2673-2688. DOI 10.3390/ai6020036 (Scopus Q2).