

## ДОДАТОК А

Графічний матеріал кваліфікаційної роботи

# МОДЕЛЬ ПРОГНОЗУВАННЯ НА ОСНОВІ ГЛИБОКОГО НАВЧАННЯ ДЛЯ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ

СТ. ГРУПИ СПм-21-2

БЕВЗЮК Р.Г.

КЕРІВНИК

доц. ФЕДОРЧЕНКО В.М.

## АКТУАЛЬНІСТЬ

- Однією з основних причин швидкого зростання глобального ринку автоматизації складів є значне зростання електронної комерції та вимоги до виконання замовлень, які створюються в результаті (швидка доставка, відстеження тощо). Існує підвищена потреба в системах, здатних обробляти великі обсяги замовлень із кількох позицій.
- Найбільша проблема в реалізації повністю автономних систем виконання замовлень полягає в ідентифікації та обробці сторонніх предметів. Сучасні методи сегментації зображення забезпечують достатню точність для обробки сторонніх об'єктів для операцій вибору та розміщення.
- Ці методи сегментації зображення дозволяють ідентифікувати та розрізнити геометричні форми, краї, поверхні та інші характеристики, достатні для надійної обробки сторонніх об'єктів. Однак цих ознак недостатньо для ідентифікації сторонніх предметів. Це обмеження технології зумовлює необхідність присутності фізичного працівника, який забезпечує ідентифікацію сторонніх предметів у процесі виконання замовлення. Для реалізації повної автоматизації система виконання замовлень повинна мати можливість автономно розпізнавати об'єкт, який обробляється, щоб об'єкт можна було транспортувати до відповідного місця призначення.

## МЕТА ТА ЗАДАЧІ КВАЛІФІКАЦІЙНОЇ РОБОТИ

**Мета кваліфікаційної роботи** побудувати модель прогнозування для класифікації об'єктів для системи яка використовується в складських підприємствах.

### Задачі:

- Розглянути правила розміщення штрих-кодів на товарі
- Обрати інструментарій для реалізації моделі класифікації товарів в яких може бути скритий штрих-код
- Описати вимоги до системи
- Провести моделювання системи та отримати результати

## ПРАВИЛА РОЗМІЩЕННЯ ШТРИХ-КОДУ НА ОБ'ЄКТАХ НА СКЛАДАХ

- Існує багато програм, які можуть знаходити та зчитувати штрих-коди зі складних сцен за допомогою камери та використовувати їх для ідентифікації об'єктів. Різниця в освітленні, позах і спотворенні перспективи може ускладнити ідентифікацію штрих-коду, але існують алгоритми, які пропонують адаптовані рішення для цих складних сцен. Найцікавішим є випадок, коли штрих-код приховано в результаті пози продукту, і його неможливо ідентифікувати в результаті цього без зміни положення об'єкта. Складський контейнер складається з предметів, на яких видно штрих-код, а також інших предметів зі штрих-кодами, прихованих за іншими поверхнями.
- Існують різні вимоги для розміщення штрих-кодів, але на жаль не всі виробники дотримуються цих вимог.



## ВИБІР ІНСТРУМЕНТАРІЮ ДЛЯ ПОБУДОВИ МОДЕЛІ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ

- В ході досліджень було визначено структуру повної системи розпізнавання об'єктів. Було вирішено, що систему буде створено з використанням згорткової нейронної мережі (ЗНМ). Тоді для цього методу локалізації штрих-коду потрібно побудувати схему розміщення штрих-коду на продуктах за допомогою згорткової нейронної мережі, щоб визначити, на якій поверхні товару лежить штрих-код.
- Щоб автоматизувати фазу повторного генерування системи виконання замовлень, нещодавно розроблена структура розпізнавання об'єктів має задовольнити двом важливим вимогам:
  1. Система повинна досягати 40% точності передбачення локалізації штрих-коду: оскільки системи виконання замовлень обробляють величезну кількість продуктів, нові системи бачення повинні прагнути до принаймні 40% точності передбачення. У подальшому можна покращити, додавши більше даних для навчання, або запровадивши більш оптимальні методи навчання.
  2. Система повинна бути масштабованою до дуже великих наборів продуктів: спроектована система бачення повинна мати можливість обробляти поточні набори продуктів, які складаються в порядку 107 різних продуктів.

## ВИЗНАЧЕННЯ ПАРАМЕТРІВ

Створений набір даних складається із зображень елемента категорії коробки, взятих із вигляду зверху та збоку. Приклади цих зображень з окремими предметами в різних позах бачимо на рисунках.



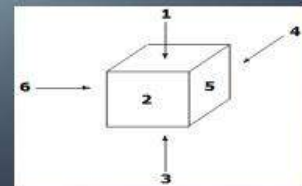
Зображення будуть імпортовані у застосунок як дані RGB з роздільною здатністю 400x300. Таким чином, вхідні дані складаються з матриці  $X \in \mathbb{R}^{400 \times 300 \times 3}$ , де «N» представляє кількість зразків у наборі даних.

Оскільки згорткові нейронні мережі є типом контрольованого навчання, для етапу навчання потрібні дані з мітками. Для ефективного маркування всіх даних вибрано наступне визначення класу:

$c_x = (1, 2, 3, 4, 5, 6) ==$  (верх, спереду, зліва, зверху, справа, зліва)



Пропонована структура розпізнавання об'єктів із визначенням класу



## МОДЕЛЬ ПРОГНОЗУВАННЯ З ВИКОРИСТАННЯМ ЗНМ

Вихідні дані застосунку складаються з вектора розподілу ймовірностей, який визначає ймовірність наявності штрих-коду на певній грані. Можливий результат можна визначити за допомогою:

$$c_i = \{1, 2, 3, 4, 5, 6\} == \{0.8, 0.1, 0.05, 0.02, 0.02, 0.01\}$$



Визначення класу для об'єкта у формі коробки

## ЗГОРТКОВА НЕЙРОННА МЕРЕЖА

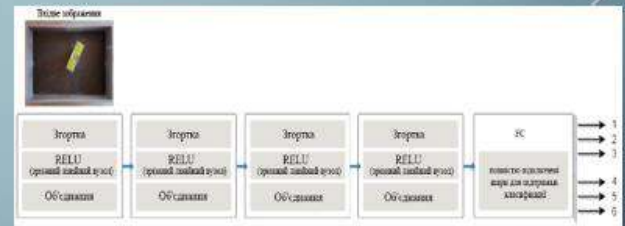
**Згорткові нейронні мережі** наразі є одним із найвідоміших алгоритмів глибокого навчання із зображеннями. У той час як для традиційного машинного навчання відповідні функції потрібно витягувати вручну, глибоке навчання використовує необроблені зображення, як вхідні дані для вивчення певних функцій. ЗНМ складаються з вхідного та вихідного рівня та кількох прихованих рівнів між входом і виходом.

Конфігурації ЗНМ складаються з безлічі прихованих шарів. У кожному шарі обсяги активації змінюються за допомогою диференційованих функцій. Існують чотири основні типи шарів, які використовуються для побудови конфігурацій ЗНМ:

1. Згортковий рівень – згорткові фільтри використовуються для отримання карти активації з вхідних даних.
2. Rectified Linear Unit Layer (ReLU) – фільтрує від'ємні значення, щоб отримати лише додатні значення для набагато швидшого часу навчання.
3. Рівень об'єднання – виконує нелінійну низьку дискретизацію та скорочує кількість параметрів для простішого виведення.
4. Повністю зв'язаний рівень – обчислює оцінки ймовірності класу шляхом виведення вектора розмірів  $C$ , де  $C$  є кількістю класів. Усі нейрони підключені до цього шару.

## КОНФІГУРАЦІЯ ЗНМ

При визначенні конфігурації мережевого рівня важливо зазначити, що не існує точної формули для оптимальної конфігурації рівня. Замість цього найкращий підхід – це метод проб і помилок, коли кілька конфігурацій шарів досліджуються та порівнюються, щоб побачити, наскільки добре вони працюють. Попередньо навчені мережі, можна використовувати як початкову конфігурацію рівня.

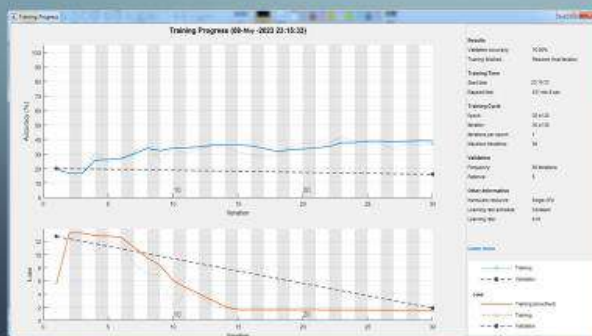


Навчання глибокої згорткової мережі, може зайняти різну кількість часу, залежно від розміру набору даних і доступної потужності обробки. Для навчання ЗНМ доступні три варіанти обчислень, і вибір найбільш оптимального є вирішальним залежно від кількості доступного часу для вирішення конкретного завдання:

1. Обчислення на основі CPU
2. Обчислення на основі GPU
3. Хмарні обчислення GPU

## НАВЧАННЯ ЗНМ (1)

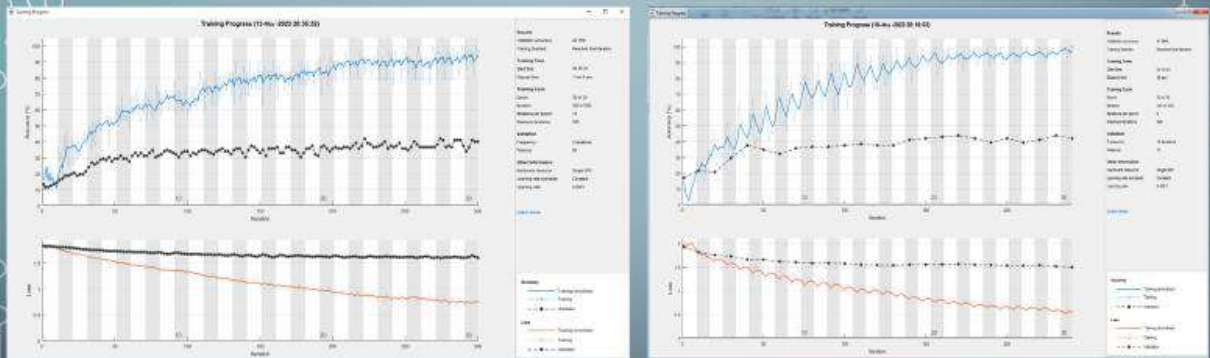
- Під час імпорту позначеного набору даних 75%/25% розподіляються між даними навчання та даними перевірки. Це означає, що 75% даних використовується для навчання мережі, а 25% даних використовується для перевірки мережі. Після досягнення достатньої точності перевірки можна використати додатковий набір тестів, щоб побачити, наскільки добре працює мережа.
- Перша конфігурація виконується на основі обчислювальної моделі ЦП. Використовується процесор Intel Core i7 720QM з чотирма фізичними та чотирма віртуальними ядрами.



ЗНМ навчальний запуск №1, з використанням центрального процесора. Точність перевірки 16% досягається за час навчання приблизно 431 хвилину

## НАВЧАННЯ ЗНМ (2)

Для наступних прогонів використовується графічний процесор NVidia GTX1060 і процесор Intel i7-6700HQ.



- Тренування ЗНМ №2 з використанням GPU. За рахунок зменшення розміру фільтра та кількості фільтрів точність перевірки 40,18% досягається з часом навчання 69 секунд

Навчальний запуск ЗНМ №3 (з GPU) зі зміненими темпами навчання та розмірами фільтрів шарів. Точність перевірки 41,96% досягається за час навчання 48 секунд.

## НАВЧАННЯ ЗНМ (3)



Навчальний запуск ЗНМ №4 (з GPU) з використанням детальних налаштувань. Точність перевірки 44,64% досягається за час навчання 63 секунди.

## ВИСНОВКИ

В ході кваліфікаційної роботи було:

- Розглянуто запропоновану структуру розпізнавання об'єктів у світлі визначених вимог і обмежень, можна зробити висновок, що досягнуто достатньої точності 44,64%, що перевищує вимогу 40%. Логічний висновок, який можна зробити з результатів навчання, полягає в тому, що набір даних має бути різко збільшений. Крім того, набори даних усіх інших категорій продуктів повинні бути додані до ЗНМ, щоб структура була реалізована на складі. Хоча генерація та збір цих даних може зайняти багато часу, з результатів, досягнутих на сьогоднішній день, можна побачити, що навіть з невеликими наборами даних показані багатообіцяючі результати.
- Під час навчання ЗНМ необхідна реалізація GPU. Навчання ЗНМ лише за допомогою процесора може зайняти місяці для повного набору даних категорії продуктів. Логічним є висновок, що тонке налаштування системи під час цього процесу неможливе в цьому часовому режимі. Масштабованість, не є проблемою при реалізації обчислень GPU.

По результатам досліджень було написано публікація "Система розпізнавання об'єктів для автоматизації складів з використанням згорткової нейронної мережі" та подано до фахового видання України категорії Б – Системи управління навігації та зв'язку