УДК 004.932

Очерклевич О.Б., Куриляк Д.Б.

Національний університет “Львівська політехніка”,

кафедра електронних обчислювальних машин

**ОСОБЛИВОСТІ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ВІДЕОПОТОКУ З ДОПОМОГОЮ ЗГОРТОЧНИХ НЕЙРОННИХ МЕРЕЖ**

© Очерклевич О.Б.,Куриляк Д.Б.., 2019

**У статті аналізуються особливості вибору архітектури згорткової нейронної мережі для системи інтелектуального аналізу відеопотоку. Розглядаються особливості побудови згорткових нейронних мереж.**

**Ключові слова: згортка, нейрон, нейромережі, відеопотік**

Ocherklevych O.B., Kurilyak D.B.

Lviv Polytechnic National University,

department of electronic computers

**FEATURES OF INTELLECTUAL ANALYSIS OF VIDEO FLOW USING СONVOLUTION NEURAL NETWORKS**

© Ocherklevych O.B., Kurilyak D.B.., 2019

# The article analyzes the features of choosing a convolutional neural network architecture for the video stream mining system. Features of construction of convolutions of neural networks are considered.

# Keywords: convolution, neuron, networks, video-stream

# Вступ

Побудова систем комп'ютерного зору та аналізу відеопотоку – актуальна задача сьогодення. Рішення у цій області знаходять широке застосування у найрізноманітніших сферах життя людини: відстеження появи машин на стоянці, забутого багажу в приміщеннях аеропортів та вокзалів, забезпечення захисту від несанкціонованого проникнення на певні об’єкти, розпізнавання облич людей, автомобільних номерів, написів, тощо. Розпізнавання об’єктів – одна з найважливіших задач в області комп’ютерного зору. Задача, з якою людина легко вправляється кожного дня, є неймовірно складною для обчислювальної техніки. Пошук відповідності між двома зображеннями деякої сцени чи об’єкту є складовою частиною багатьох систем аналізу візуальної інформації , зокрема відеопотоків .

На основі аналізу науково-технічної літератури можна виділити два основних напрями для вирішення проблеми розпізнавання образів: часткове (інваріантність щодо повороту, зсуву, масштабу) і загальне (інваріантність щодо форми, ракурсу, текстури). Для задач часткового інваріантного розпізнавання вже зараз існує ряд математичних методів, таких як нейронні мережі, неокогнітрон, згорткові нейронні мережі вищих порядків, статистичні моделі. Основна проблема полягає в практичних складнощах реалізації таких методів. Задача загального інваріантного розпізнавання не вирішена досі. Є окремі успіхи на якість розпізнавання для обмежених класів об’єктів, як правило, облич осіб, дорожніх знаків, банківських чеків, які демонструють точність розпізнавання навіть вище, ніж у людини.

Метою статті є дослідження особливостей вибору архітектури згорткової нейронної мережі для системи інтелектуального аналізу відео-потоку.

# Постановка задачі

Пропонується вирішити задачу вибору архітектури згорткової нейронної мережі для детекції об'єктів, щоб користувач мав можливість проводити моніторинг об'єктів, що знаходяться в зоні особливого контролю . При вирішенні даного завдання об'єктом послужать деталі технічних пристроїв. Завдання виявлення об'єктів включає в себе кілька підзадач, такі як: отримання даних для розробки моделі виявлення, вибір методів класифікації для навчання моделі, побудова моделі для ідентифікації об'єктів, розробка програмної складової для роботи інтелектуального аналізу відеопотоку.

**Вибір архітектури згорткової нейронної мережі**

Архітектура ЗНМ складається з багатьох шарів. Є два основні типи шарів: згорткові (Convolutional) і підвибіркові (Subsampling), Згорткові та підвибіркові шари чергуються один з одним. Але останнім часом є тенденція до скорочення кількості під вибіркових. Тому підвибіркових шарів часто менше, ніж згорткових або немає взагалі.

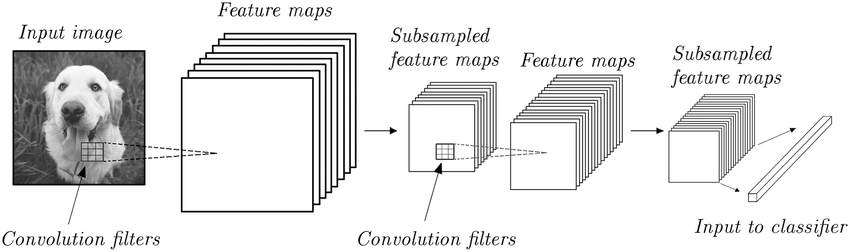


Рисунок 1 – Загальний вигляд згорткової нейронної мережі

У кожному загортковому шарі є набір з декількох фільтрів(ядер згортки), які мають невелике рецептивне поле, але займають усю глибину вхідного об’єму, причому нейрони одного фільтра мають однакові ваги, що застосовуються до всіх локальних ділянок попереднього шару (тим самим здійснюючи згортку). Зображення попереднього шару ніби сканується невеликим вікном і помножується на набір ваг, а результат зберігається на відповідному нейроні поточного шару. Таким чином набір площин є картою ознак (feature maps), де кожна площина знаходить «свої» ознаки зображення в будь-якому місці попереднього шару.

Наступний за згорточним шаром, підвиборочний шар зменшує масштаб площин даних, для полегшення виділення основних ознак в наступних шарах. Найефективнішою є «максимальна підвибірка». Остання частина згорткової мережі – це класифікатор ознак, як правило, представлений кількашаровим перцептроном або опорно-векторною машиною. Кількість нейронів у шарі-класифікаторі визначається кількістю класів, до яких слід віднести вхідне зображення.

Було застосовано та порівняно два найсучасніших алгоритму згорткових нейронних мереж SSD та YOLO, що мають найвищі швидкість та точність детекції.

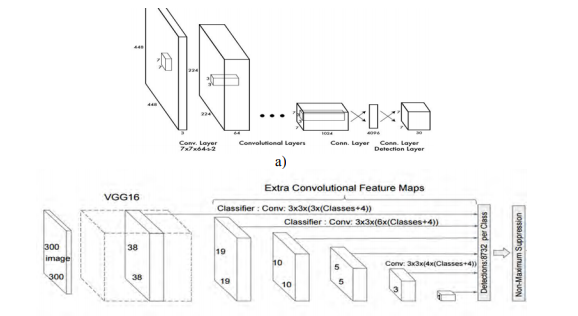


Рисунок 2 – Архітектура YOLO (a) та SSD (б)

YOLO дуже швидка. Під час тестування на одному з найпотужніших графічних процесорів (Nvidia Titan X GPU) базова мережа обробляє 45 кадрів за секунду. YOLO оцінює усе зображення, коли робить прогноз, на відміну від методів, що базуються на ковзаючому вікні або пропозиціях регіонів. Система поділяє зображення за допомогою сітки та для кожної клітинки передбачає декілька обмежувальних рамок і ймовірність наявності в них об’єкта певного класу.

SSD – є наступним поколінням ЗНМ для розпізнавання об’єктів, вона також розглядає зображення цілком, але розглядає сітки різних масштабів. Додавши ряд поліпшень, вдалося значно підвищити точність, порівняно з попереднім варіантом, не втративши про цьому у швидкості розпізнавання. Застосовуючи ці фільтри до різних карт ознак, отриманих з різних шарів на пізніших етапах, мережі дозволяють виконувати детекцію у різних масштабах. Це вирішує значну проблему попередніх мережі і дозволяє досягти високої точності виявлення об’єктів різних форм та розмірів на зображеннях з порівняно низькою роздільністю.

# Висновки

# У результаті аналізу літератури було виявлено, що найбільш ефективним з інтелектуальних методів розпізнавання об’єктів з відео потоку є згорткові нейронні мережі. Серед них найбільш удосконалені архітектури є YOLO та SSD. На основі архітектури SSD розроблено систему розпізнавання об’єкта на зображенні з камери системи .

# 

*1.Згорткова нейронна мережа [Електронний ресурс]..Режим доступу:* [*http://uk.wikipedia.org/wiki/Згорткова\_нейронна\_мережа*](http://uk.wikipedia.org/wiki/Згорткова_нейронна_мережа)*. 2. Сховище зображень для навчання нейронних мереж [Електронний ресурс]. «The CIFAR-10 dataset». Режим доступу:* [*https://www.cs.toronto.edu/~kriz/cifar.html*](https://www.cs.toronto.edu/~kriz/cifar.html)*. 3. Іван Гудфелов, Йоша Бенгіо, Арон Коурвілле. «Машинне навчання» MIT Press, 2016. Режим доступу:* [*http://www.deeplearningbook.org*](http://www.deeplearningbook.org)*. 4. Ліла В.Б. Алгоритм та программна реализація адаптивного метода навчання штучних нейронних мереж// Инженерний вестн. Дона, 2012. 5. Девід Е. Голберг «Дзен та мистецтво генетичного алгоритму». 3rd International Conference on Genetic Algorithms, pp. 80-85.*