



Дипломна робота на тему:

# Згорткові нейронні мережі та їх застосування до розпізнавання дорожніх об'єктів в умовах зашумленості

**Виконав:**

Олексій Женчук, студент групи КА-51

**Науковий керівник:**

к.ф.-м.н., доц. Яковлева А.П.

Київ, 2019

# Вступ

**Мета роботи:** дослідження можливості та властивостей застосування згорткових нейронних мереж до задач розпізнавання та класифікації дорожніх об'єктів в умовах наявності цифрового шуму на зображеннях, а також перевірка ефективності та порівняння алгоритмів зменшення рівня шуму в застосуванні до вказаних типів задач.

**Об'єкт дослідження:** набори цифрових зображень, які містять дорожні об'єкти різних класів і які можуть мати певний рівень цифрового шуму, який необхідно оцінити та відновити зображення за відповідними методами для подальшого вирішення задачі їх розпізнавання та класифікації.

# Вступ

**Предмет дослідження:** моделі згорткових нейронних мереж з оптимізованим використанням параметрів та алгоритми зменшення шуму в зображеннях в їх застосуванні до задачі розпізнавання та класифікації дорожніх об'єктів на прикладі наборів реальних даних.

**Методи дослідження:** застосовані моделі згорткових нейронних мереж, методи зменшення рівня шуму в зображеннях, алгоритми навчання нейронних мереж, виконані за допомогою мови програмування Python.

# Актуальність

**Актуальність дослідження:** розпізнавання дорожніх об'єктів є ключовим для багатьох напрямків, які швидко розвиваються, зокрема, для безпілотних автомобілів, при цьому рішення на основі задач розпізнавання та класифікації мають прийматися в реальному часі та не тільки на основі отриманих даних, в яких може бути високий рівень цифрового шуму, що значно ускладнює роботу систем. Тому необхідні моделі, здатні визначати наявність шуму та мати відповідні механізми усунення його впливу.

# Постановка задачі

В роботі поставлено для виконання наступну задачу:

- провести огляд та аналіз наявних підходів до застосування згорткових нейронних мереж (ЗНМ) в задачах розпізнавання та класифікації зображень;
- розглянути механізми виникнення цифрового шуму в зображеннях та методи боротьби з ним з метою дослідити вплив алгоритмів зниження шуму на роботу згорткових нейронних мереж;

# Постановка задачі

- визначити моделі ЗНМ, придатні для застосування до заданого класу задач, та проаналізувати їх роботу для чистих, зашумлених та відновлених даних та ефективність їх застосування прикладі вибраних наборів даних;
- на основі отриманих результатів проаналізувати вплив алгоритмів зниження рівня шуму на роботу згорткових нейронних мереж та зробити висновки щодо особливостей та оптимальності їх застосування.

# Критерії знаходження оптимальних моделей

Критерієм оцінки ефективності роботи нейронної мережі є найменше значення функції втрат після кожної ітерації навчання моделі:

$$L(p, y) = - \sum_n y_n \ln p_n, n \in [1, N]$$

Але функція втрат вказує на величину розбіжності між отриманими та справжніми даними, але не кількість даних, для яких мережа робить правильні висновки щодо класу. Тому обраною метрикою оцінки результатів задачі класифікації є точність:

$$a = \frac{\sum_i c_i}{\sum_i n_i}, i \in [1, N]$$

Обраним критерієм оцінки оптимальності роботи моделей є точність на перевіірчній вибірці  $a_{val}$ , яка є і критерієм для оптимального методу зменшення рівня шуму.

# Моделі для згорткових нейронних мереж

Швидкість прийняття рішень для розпізнавання та класифікації дорожніх об'єктів на основі вирішення задачі класифікації має надзвичайно високе значення.

Застосування згорткових нейронних мереж з великою кількістю рівнів, які вимагають значної кількості операцій для такої сфери є неефективним.

Тому при виборі напряму розробки архітектур було вирішено зробити акцент на розгляді моделей, які пропонують ефективні моделі рівнів з меншою кількістю параметрів, при цьому зменшуючи складність моделі. Було обрано для розробки дві моделі з такого класу моделей.



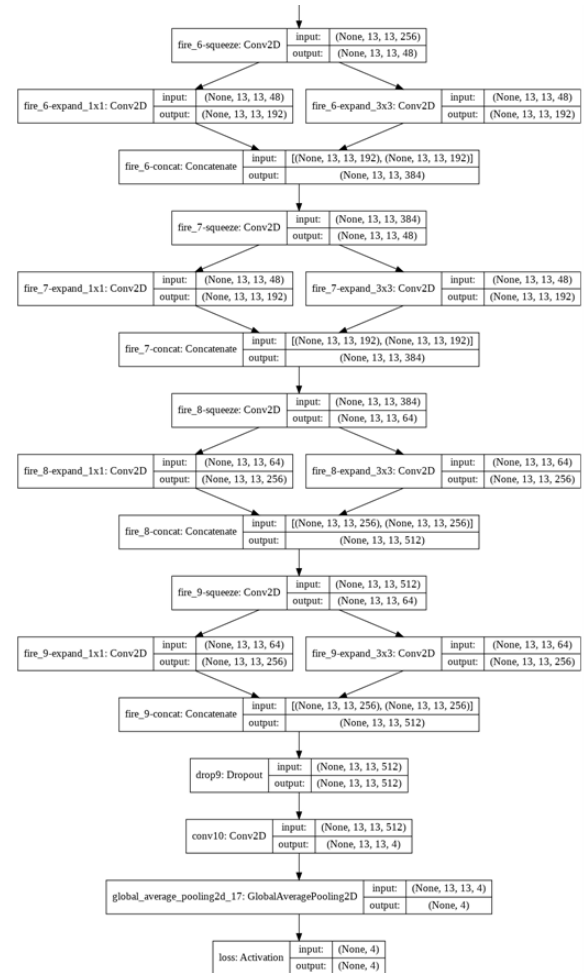
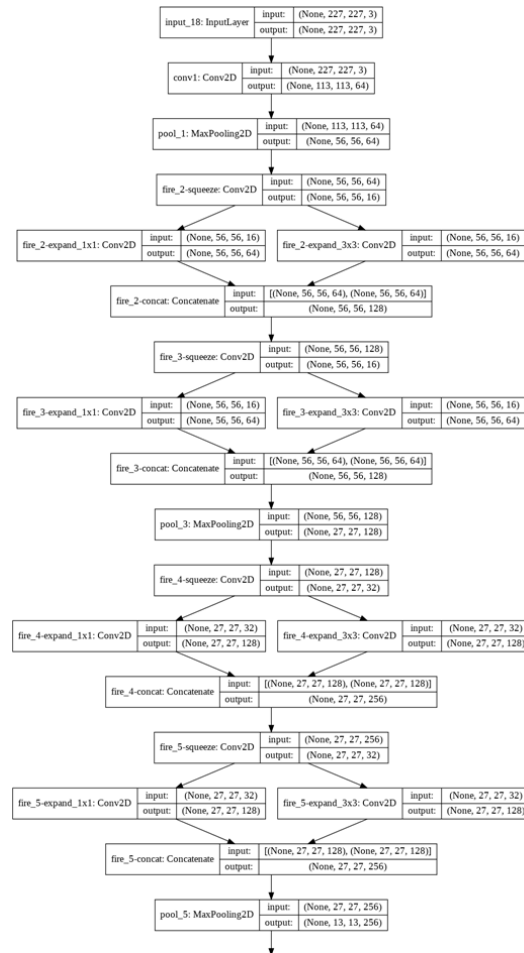
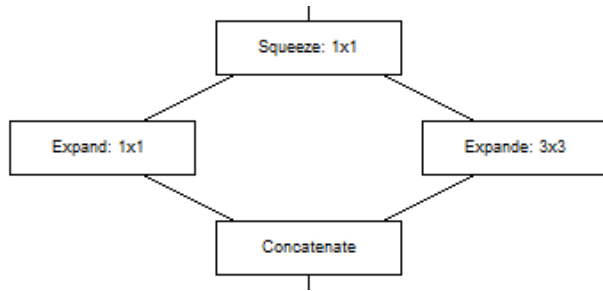
# Моделі для згорткових нейронних мереж

**Перша модель** збудована на основі ідеї мережі SqueezeNet. Головним принципом такої мережі є відмова від використання фільтрів великої розмірності, натомість використання фільтрів розмірності  $1 \times 1$ , які об'єднуються в окремі модулі, конфігурацією яких згодом і будується мережа.

Така мережа містить в кілька десятків разів менше параметрів, ніж моделі схожого класу.

Головним структурним елементом SqueezeNet є Fire Module, з варіантів якого будується вся система.

# Моделі для згорткових нейронних мереж



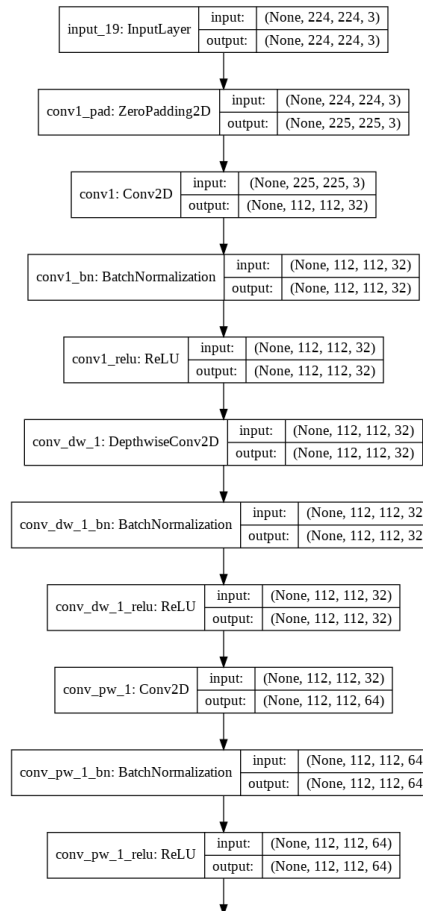
Структура мережі: на рисунку зліва схема елементарного модуля, посередині та справа структура першої моделі

# Моделі для згорткових нейронних мереж

**Друга модель** збудована на основі класу мереж MobileNet, які були запропоновані у 2017 р. для використання на пристроях з малими можливостями. Для зменшення кількості параметрів використовується ідея поглиблених розділених згорток. Ці згортки поєднують в собі послідовні операції поглибленої та поточної згортки, які відповідно є горизонтальними та вертикальними підвидами згорток  $1 \times 1$ . Як наслідок, кількість параметрів на окремому рівні зменшується в кілька разів без втрати характеристик.

Модель також дозволяє задавати параметри скорочення кількості каналів та розмірності на кожному рівні мережі.

# Моделі для згорткових нейронних мереж



Структура мережі: через її великий розмір вказаний елементарний модуль мережі з поглибленими та поточковими рівнями

# Набір даних Traffic Signs Dataset

Набір даних Traffic Signs Dataset було створено шляхом встановлення на автомобіль цифрової камери та запису даних на 350 кілометрах автодоріг, як в населеній зоні, так і поза нею, та на автодорогах різного значення. Зображення для знаків робилися з різної відстані як кадри з відеозапису, зробленого камерою.

Загалом, отриманий набір даних для задачі класифікації містить 2193 зображень, розділених на 7 класів найпопулярніших дорожніх знаків: «Пішохідний перехід», «Об'їзд перешкоди з правого боку», «Зупинку заборонено», «Обмеження максимальної швидкості 50 км/год», «Обмеження максимальної швидкості 70 км/год», «Головна дорога» та «Дати дорогу».

# Набір даних Graz02

Набір даних під назвою Graz02 було створено як набір зображень зі складними реальними характеристиками, де об'єкти класів з'являються на зображеннях в великій кількості кутів огляду та ситуацій.

Всього набір складається з 1476 зображень, які належать до чотирьох категорій основних учасників дорожнього руху: автомобілів, велосипедів, пішоходів та четвертого класу, який представляє всі інші об'єкти.

Особливостями набору даних є велика відмінність зображень для кожного представленого класу, тобто, наприклад, велика кількість різних положень та моделей автомобілів на зображеннях, а також різноманітність у фонових деталях зображення.

# Набір даних Graz02



Приклад зображень з чотирьох класів для набору реальних даних

# Моделювання цифрового шуму

Шум є випадковими коливаннями показників елементів зображення, наприклад, значень кольорових каналів чи яскравості, які не належать до реального зображення та є його спотвореннями. Сенсори, що сприймають зображення, завжди мають певний рівень шуму через те, що рівень освітленості для них визначається кількістю фотонів, які на них потрапили, яка є випадковим числом.

Найчастішим типом цифрового просторово незалежного шуму є гаусівський шум – такий шум, функція щільності імовірності якого відповідає нормальному (гаусівському) розподілу. Гаусівський шум зустрічається в незначній кількості в будь-якому сигналі та широко використовується для моделювання та наближення інших типів шуму. Всі практично значимі види шуму для цифрового зображення або моделюються як гаусівський, або наближуються ним та моделюються ним на практиці, або мають занадто мале поширення.



# Моделювання цифрового шуму

Оцінити точність наближення, тобто різницю між зображеннями, без наявності початкового в реальних умовах неможливо, крім того, вона не характеризує придатність алгоритму до роботи з моделлю розпізнавання. Тому в якості критерія точності результатів зниження рівня шуму використовуються остаточні результати роботи моделей та відповідні метрики для різних видів шуму.

Тому для моделювання зменшення рівня шуму та його механізму використовується модель адитивного гаусівського шуму, зображена на рисунку.



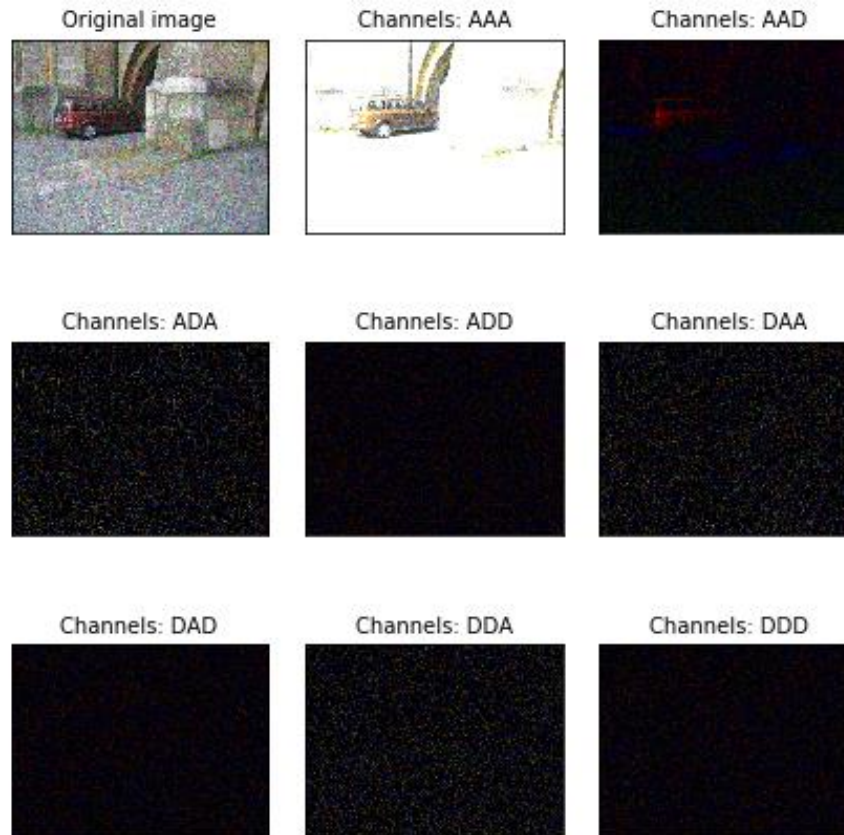
# Оцінка рівня шуму

Оцінка рівня шуму є необхідною для правильної побудови та роботи алгоритмів зменшення рівня шуму. В реальних випадках, чисте (реальне) зображення невідоме, тому оцінка рівня шуму надзвичайно важлива.

Оцінити рівень для шуму, який наближено моделюється гаусівським розподілом, можна за допомогою оцінки середнього квадратичного відхилення для зображення.

Метод оцінки рівня шуму, обраний для реалізації, засновується на знаходженні деталізуючих коефіцієнтів вейвлет-перетворення.

# Оцінка рівня шуму



Результати виділення шуму на  
зображенні з набору даних за  
допомогою вейвлет-  
перетворення

# Зменшення рівня шуму

Зменшення рівня шуму має на меті відновити чисте зображення у випадку, якщо на ньому присутні зашумлені компоненти і зазвичай засновується на відомостях або припущеннях щодо того, який саме тип шуму має зображення. Це як правило зводиться до наближення екстремальних значень точок з великою різницею з оточенням до менших наближених значень в різних формах.

Як ціль знаходження методу зниження шуму було обрано такі методи, які б зберігали більшість характеристик зображення. Зазвичай, такі методи є нелінійними та враховують структурні особливості.

# Зменшення рівня шуму: метод повної варіації

Серед таких методів зменшення рівня шуму було виділено:

- **метод повної варіації**, заснований на диференціальних рівняннях та полягає в мінімізації його варіації за ітеративними методами. Метод реалізований за допомогою двох алгоритмів:

- Метод повної варіації з алгоритмом оптимізації Брегмана зводиться до пошуку найменшого можливого значення повної варіації відновленого зображення  $u$  через суму абсолютних значень наближення градієнта  $d = \nabla u$  для всіх елементів зображення:

$$\operatorname{argmin}_d \sum_{i,j} |d_{i,j}| + \frac{\lambda}{2} \sum_{i,j} (f_{i,j} - u_{i,j})^2 + \frac{\gamma}{2} \sum_{i,j} |d_{i,j} - \nabla u_{i,j} - b_{i,j}|^2,$$

- Ітеративний метод повної варіації робить пошук відновленого зображення  $u$  за формулою:

$$\|u - f\|^2 = N\sigma^2,$$

де  $N$  – кількість точок. Після задання початкових значень, оцінка проводиться ітеративно шляхом пошуку проекції  $\|\pi_{\lambda K}(f)\|^2 = N\sigma^2$ :

$$\lambda_{n+1} = \frac{\sqrt{N}\sigma}{\|\pi_{\lambda_n K}(f)\|} \lambda_n.$$

# Зменшення рівня шуму: вейвлет-перетворення

- метод зменшення рівня шуму за допомогою **вейвлет-перетворення** з використанням порогових значень, який базується на властивості переводу ним білого шуму в білий шум та неінформативних коефіцієнтів без шуму до нуля. Порогові значення обираються за одним із двох алгоритмів:

- За алгоритмом VisuShrink, порогове значення  $\lambda$  для всіх елементів зображення обирається як

$$\lambda = \sigma\sqrt{2 \ln N},$$

де  $N$  – розмірність вхідних даних.

- За алгоритмом BayesShrink, порогове значення шукається із застосуванням попередньо оціненого  $\sigma^2$  та його різниці зі стандартним відхиленням для зашумленого зображення  $Y$ :

$$\lambda = \frac{\sigma^2}{\sqrt{\max\left(\frac{\sum_{i,j=1}^n Y_{ij}^2}{n^2} - \sigma^2, 0\right)}}$$

# Зменшення рівня шуму: метод двостороннього фільтра

- метод на основі **двостороннього фільтра**, який попри порівняну простоту використовує для відновлення нелінійну комбінацію точок з врахуванням не тільки їх просторової близькості, а ще й різницю між їх характеристиками та знайшов завдяки цьому значне застосування. Знаходження відновленого зображення може бути записане так:

$$I_{new}(p) = \frac{\sum_{q \in S} I(q) c_{\sigma_s}(\|p - q\|) s_{\sigma_r}(\|I(p) - I(q)\|)}{\sum_{q \in S} c_{\sigma_s}(\|p - q\|) s_{\sigma_r}(\|I(p) - I(q)\|)},$$

де  $I$  – початкове зашумлене зображення,  $p$  – поточна точка,  $q$  – інша точка в межах вікна  $S$ ,  $s_{\sigma_r}$  та  $c_{\sigma_s}$  – функції, що визначають характеристичну та геометричну подібності. Для гаусівського шуму:

$$f_{\sigma_i}(p, q) = e^{-\frac{1}{2} \left( \frac{\|d(p) - d(q)\|}{\sigma_i} \right)^2},$$

де  $d$  – координати точок.

# Зменшення рівня шуму

Зашумлене зображення



Відновлене зображення, BayesShrink



Відновлене зображення, VisuShrink



Відновлене зображення, метод Брегмана



Відновлене зображення, ітеративн



Зашумлене зображення



Відновлене зображення, BayesShrink



Відновлене зображення, VisuShrink



Відновлене зображення, метод Брегмана



Відновлене зображення, ітеративн



Зашумлене зображення



Відновлене зображення, двосторонній фільтр



Зашумлене зображення



Відновлене зображення, двосторонній фільтр



Приклади  
результатів  
застосування  
методів зниження  
рівня шуму до  
зображень з  
наборів даних



# Результати для набору Traffic Signs Dataset

Для набору даних Traffic Signs Dataset було реалізовано обидві з запропонованих моделей згорткових нейронних мереж. Для кожної з мереж було проведено навчання з використанням алгоритму Adam. Була застосована технологія запам'ятовування найкращої точки, коли зберігається версія моделі з найкращою метрикою (тестовою точністю). За результатами роботи було отримано такі результати:

Модель	Ориг. набір	Зашу- млені дані	ВПВ	ВРВ	МПВБ	МПВд	ДФ
SqueezeNet	0,8269	0,5490	0,5421	0,5444	0,6720	0,7608	0,7130
MobileNet	0,7350	0,6036	0,6014	0,6082	0,6492	0,6970	0,6720

# Результати для набору Graz02

Так само, як і для набору Traffic Signs Dataset, було реалізовано обидві з запропонованих моделей та проведено процес навчання за визначеною схемою з застосуванням технології запам'ятовування найкращої точки.

Для навчання моделі на основі SqueezeNet було пройдено 130 епох оновлення параметрів до появи перенавчання, тоді як для навчання моделі на основі MobileNet знадобилося 90 епох. Результати роботи наведено в таблиці:

Модель	Ориг. набір	Зашумлені дані	ВРВ	ВРВ	МПВб	МПВд	ДФ
SqueezeNet	0,9728	0,9348	0,9511	0,9538	0,9592	0,9647	0,9457
MobileNet	0,875	0,8179	0,8668	0,8668	0,875	0,8697	0,8697

# Висновки: моделі

Розроблені методи та моделі було застосовано до реальних наборів даних дорожніх об'єктів, де моделі показали високий рівень точності прогнозування, максимальний знайдений рівень точності для чистих даних склав 97,28 %.

Роботу згорткових нейронних мереж було перевірено для різних видів даних та різних методів зменшення рівня шуму, внаслідок чого виявлено, що одна з реалізованих моделей за нормальних умов вирішує задачу краще за другу, проте менш стійка до появи шуму на зображеннях.

# Висновки: цифровий шум

Було проаналізовано вплив алгоритмів зменшення рівня шуму на результати роботи мереж та виявлено, що їх застосування в загальному випадку покращує роботу алгоритмів, інколи доводячи рівень точності близько до початкового.

Для одного з наборів даних методи з використанням вейвлет-перетворення відчутно не покращували результати, але найоптимальнішим методом зменшення рівня шуму було визначено **ітеративний метод повної варіації**.

Методи повної варіації Брегмана та метод з двостороннім фільтром також показали гарні результати.

**Дуже дякую  
за увагу!**

