

Дослідження штучних нейронних мереж для класифікації дорожніх знаків

Євтухов Сергій Костянтинович,
ІПСА, ДА-31

Об'єкт дослідження

штучні нейронні мережі

Предмет дослідження

модель, в основі якої закладена згорткова нейронна мережа, для вирішення задачі класифікації дорожніх знаків

Мета

створити модель на основі штучної нейронної мережі для класифікації дорожніх знаків;

дослідити, як залежить точність класифікації від особливостей обраної архітектури, передобробки набору даних і початкової ініціалізації гіперпараметрів.

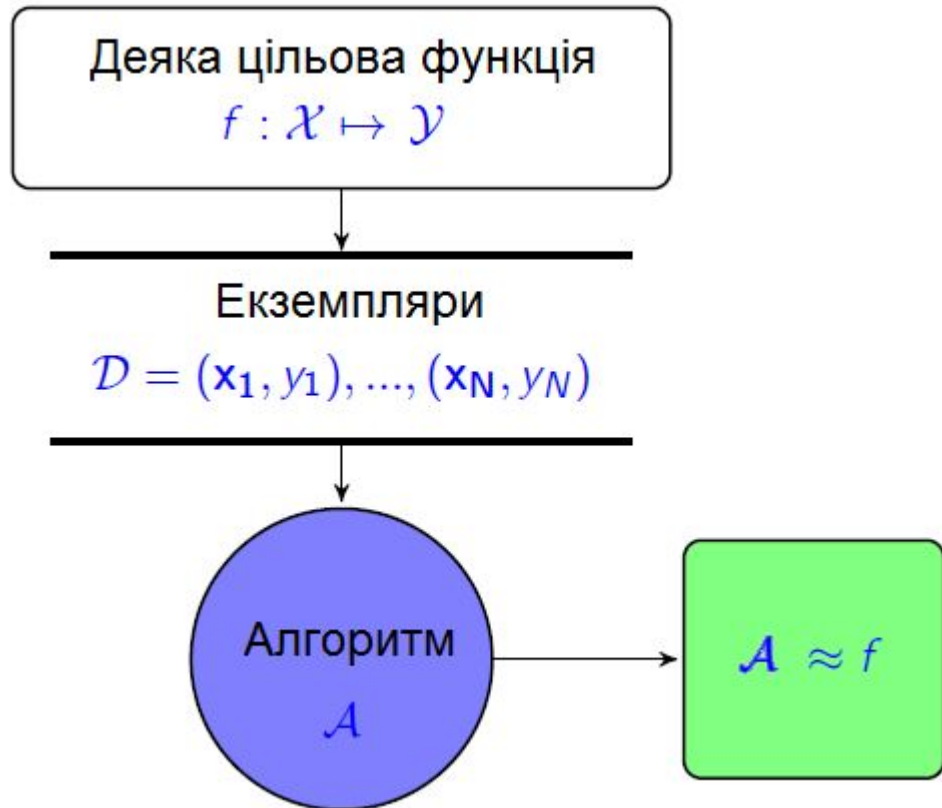
Задача класифікації

X - множина зображень
дорожніх знаків

Y - множина класів дорожніх
знаків

Вхід: розмічений тренувальний
набір даних

Вихід: модель, що здатна будь-
якому зображенню з X посавити
у відповідність клас з Y



Опис вхідних даних

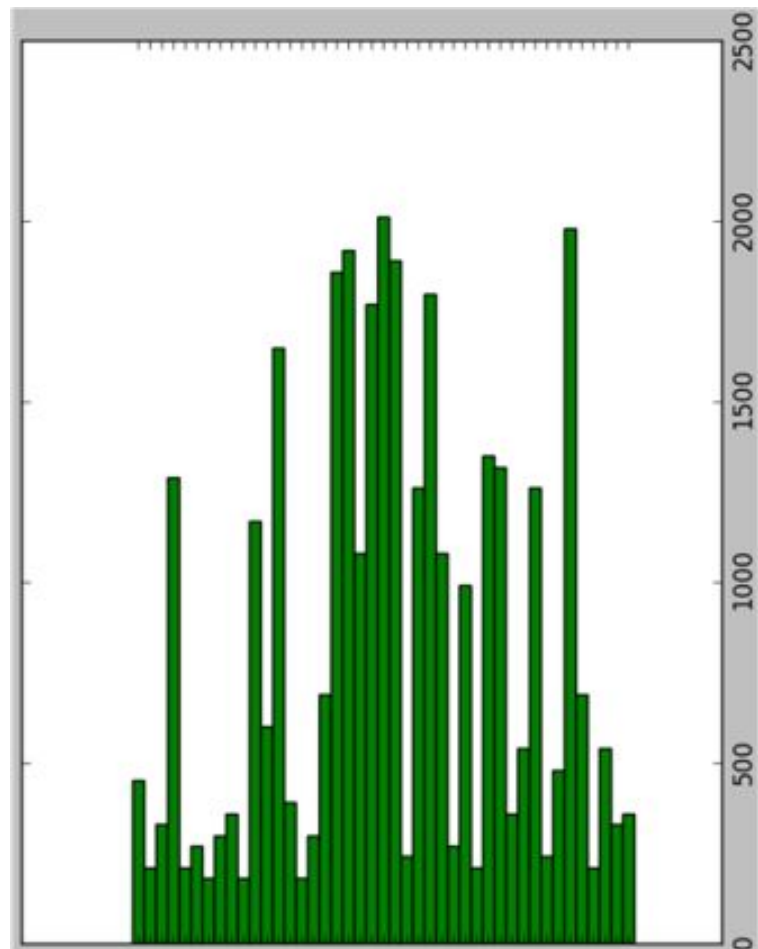
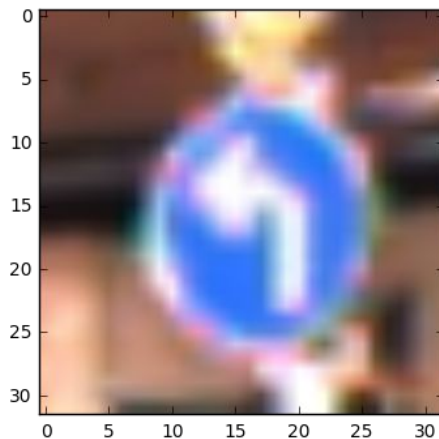
Об'єм навчальної вибірки: **34799**

Об'єм валідаційної вибірки: **4410**

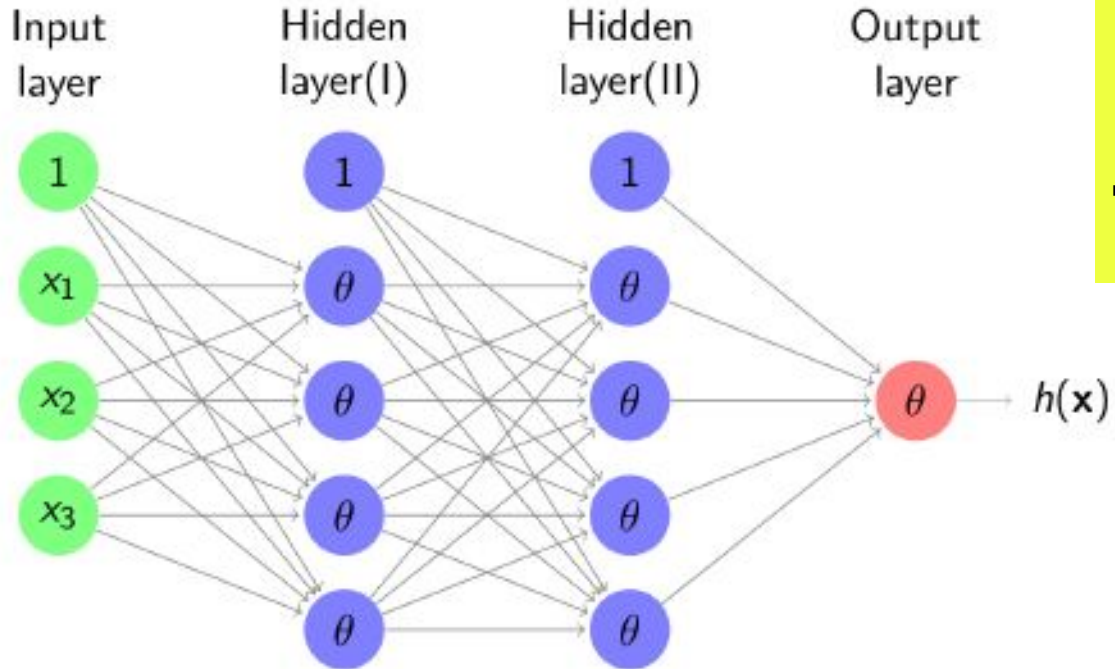
Об'єм тестувальної вибірки: **12630**

Розмірність екземпляра: **32x32x3**

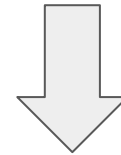
Кількість класів: **43**



Архітектура



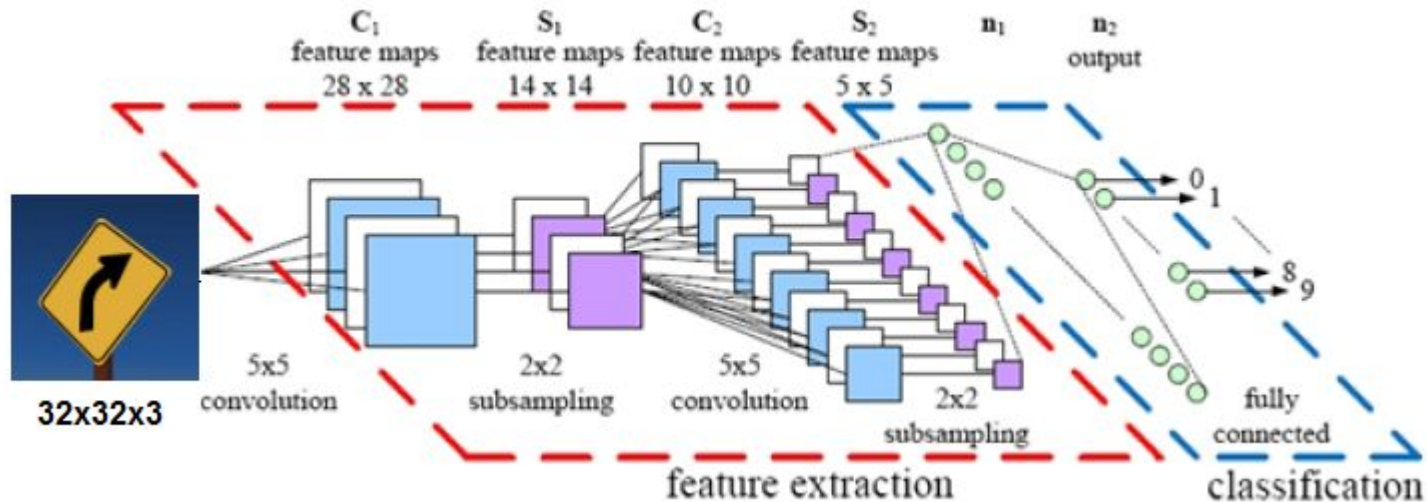
Вхідний шар: $32 \times 32 \times 3 = 900$
Ваги: $900 \times$ розмірність
другого шару = сотні тисяч
параметрів



Необхідно шукати
альтернативний підхід

Архітектура типового повнозв'язного перцептрона

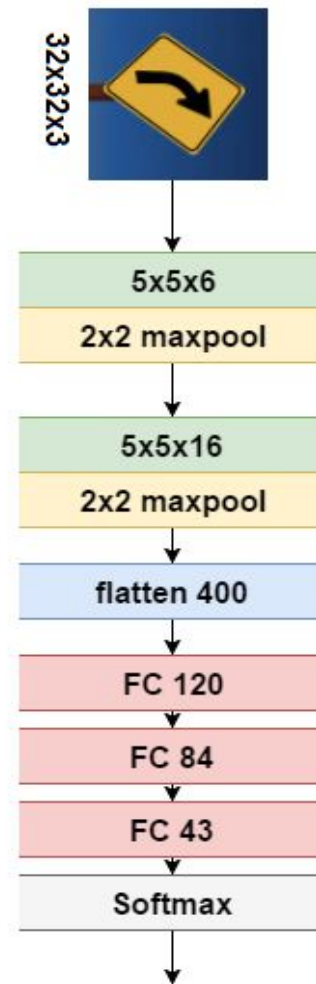
Архітектура. Згорткова нейронна мережа



За рахунок того, що ваги фільтра не змінюються при згортці зображення, маємо **набагато меншу кількість параметрів**, ніж у пласкій перцептронній архітектурі

Архітектура. Початкові припущення

| | |
|---------------------|--------------------------|
| Вхідне зображення | 32*32*3 |
| 1 шар згортки | 28*28*6 (фільтр 5x5x6) |
| 1 шар пулінгу | 14*14*6 (пулінг 2x2) |
| 2 шар згортки | 10*10*16 (фільтр 5x5x16) |
| 2 шар пулінгу | 5*5*16 (пулінг 2x2) |
| 1 повнозв'язний шар | 120 |
| 2 повнозв'язний шар | 84 |



Передобробка даних

1. Нормалізація зображень

$$(\text{pixel} - 128) / 128$$

2. Генерування нових даних:

a. Афінні перетворення

$$[0; 256] \Rightarrow [-1; 1]$$

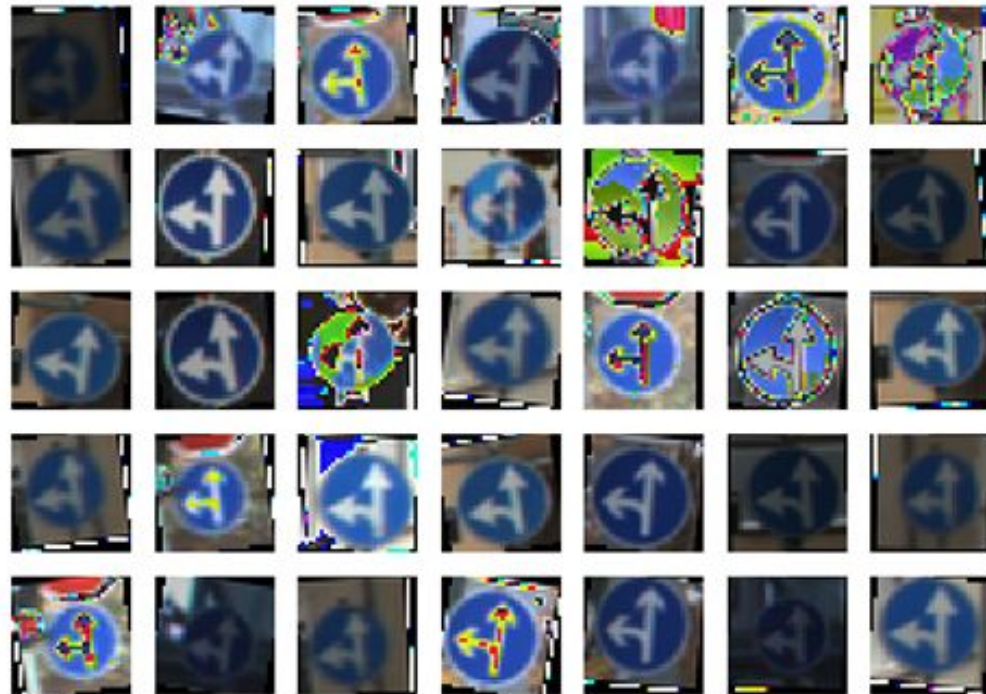
b. Зміна рівня яскравості

- пришвидшення збіжності методу оптимізації
- запобігання розмиванню градієнта
- зручність подальших маніпуляцій над даними

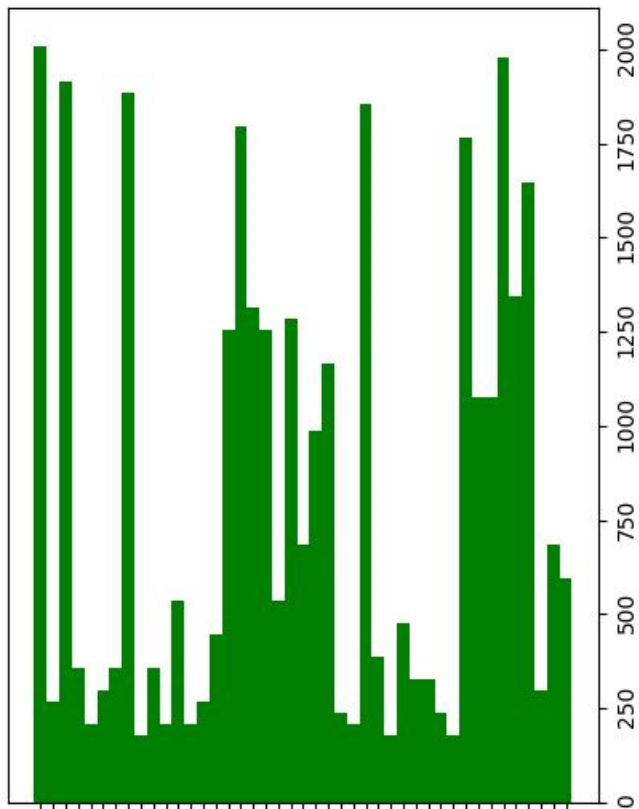
Передобробка даних

1. Нормалізація зображень
2. Генерування нових даних:
 - а. Афінні перетворення
 - б. Зміна рівня яскравості

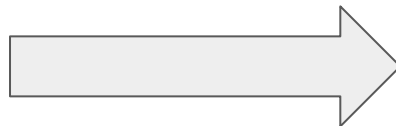
Зміщення: $[-3; 3]$ пікселі
Поворот: $[-10; 10]$ градусів
Яскравість: $\pm 25\%$



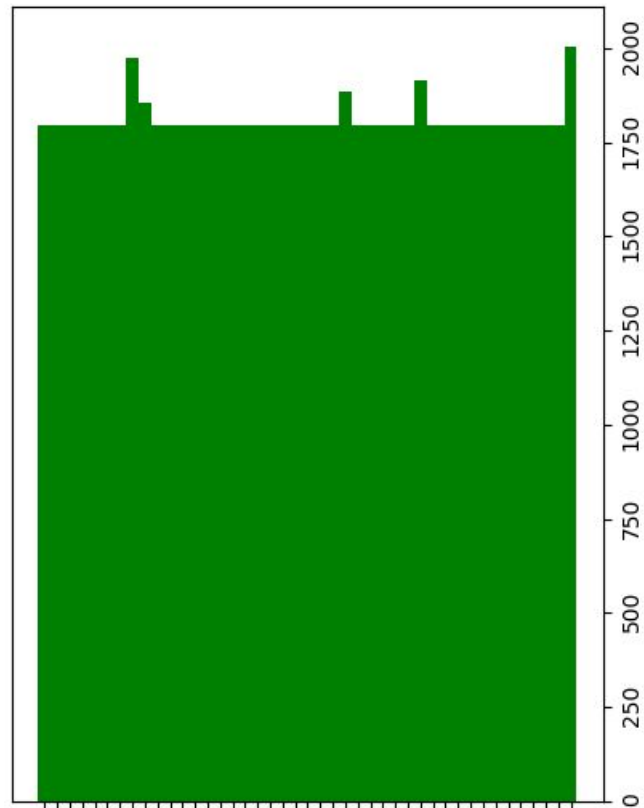
Передобработка данных



43261

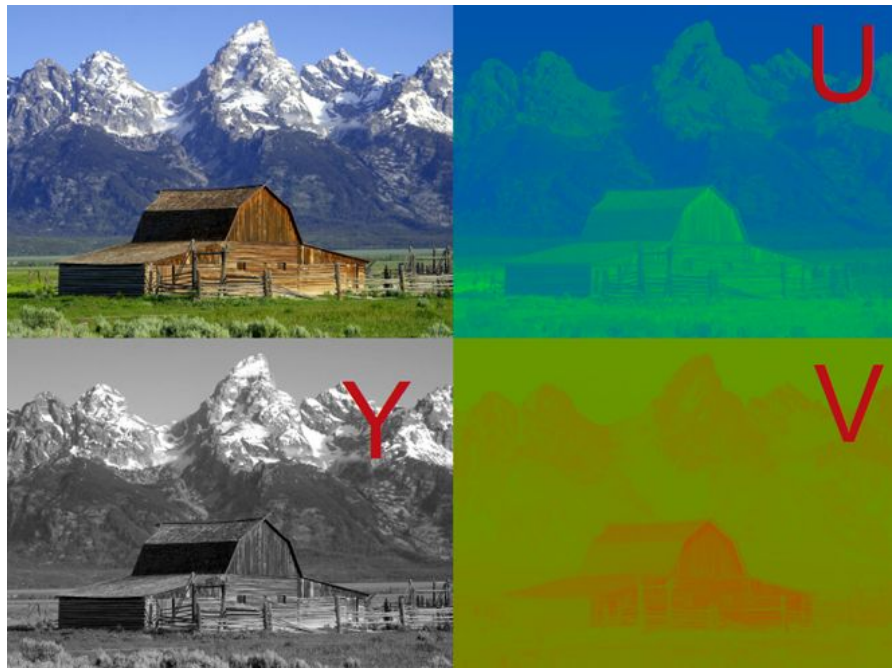


87% => 93.5%



Дослідження впливу особливостей архітектури нейронної мережі
на результат класифікації

Зміна розмірності вхідних даних. Перехід від RGB до напівтонів сірого



$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$

Вхідні зображення:

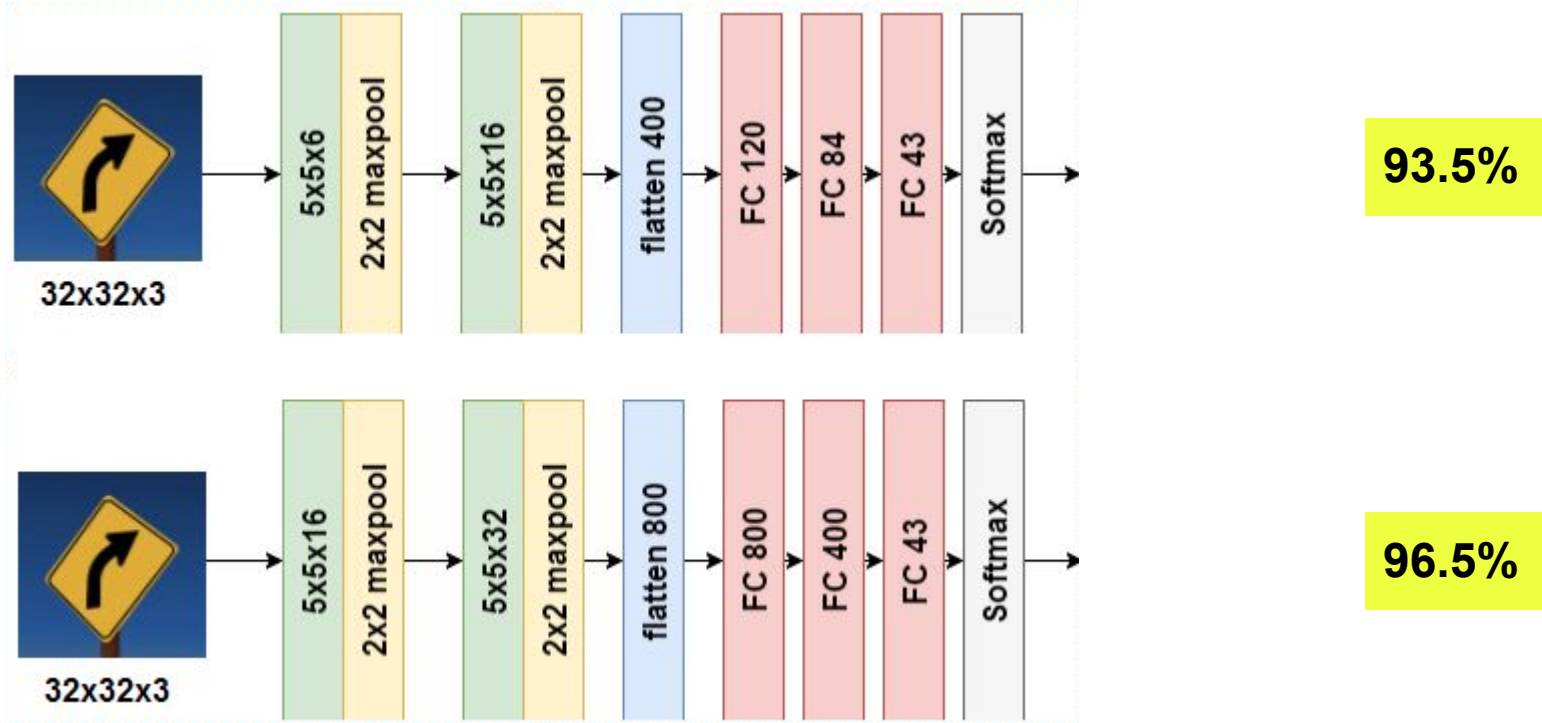
$32 \times 32 \times 3 \Rightarrow 32 \times 32 \times 1$

Фільтри першої згортки:

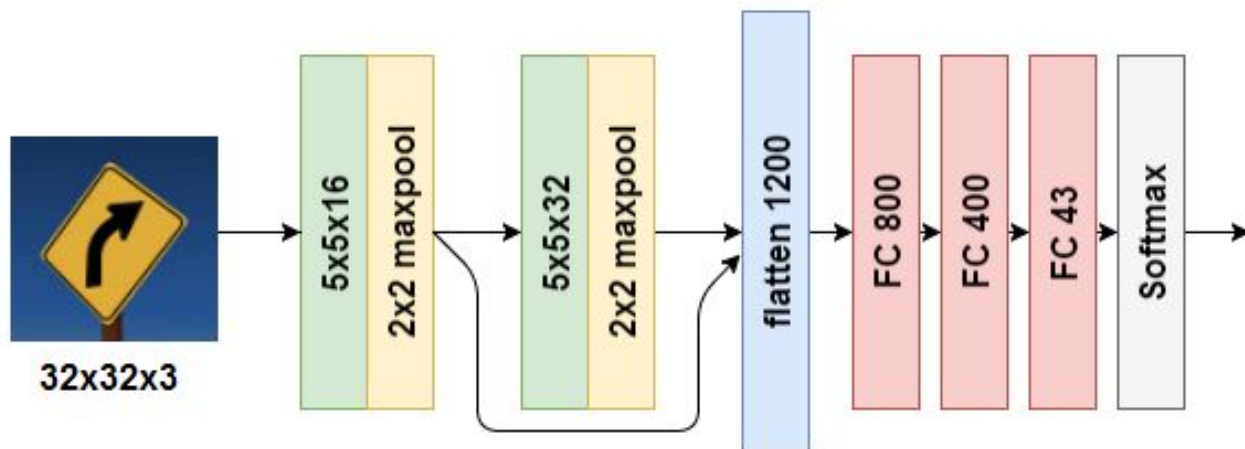
$5 \times 5 \times 3 \Rightarrow 5 \times 5 \times 1$

Точність класифікації: 92.7%

Масштабування шарів



Multi-scale архітектура



97.4%

Напрями майбутніх досліджень

дослідити точність класифікації моделі на максимально можливих обчислювальних ресурсах (AWS GPU)

інтегрувати отриману модель в онлайн-систему розпізнавання дорожніх знаків у відеопотоці, отриманому з камер автомобіля

