# Застосування штучного інтелекту для розмиття обличчя у відеопотоці реального часу

Підготував:
Студент групи КА-з82мп
Мостовий Дмитро Вікторович

Керівник:
професор кафедри ММСА,
д. т. н., проф., Данилов Валерій Якович

# Основні задачі

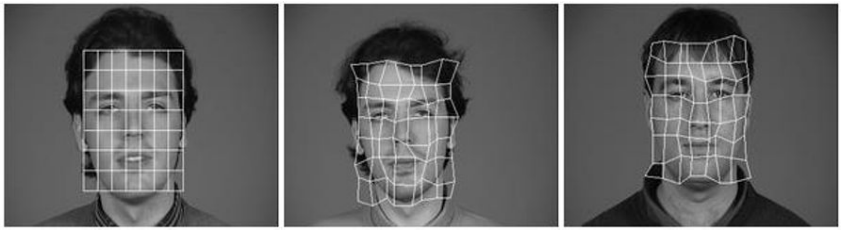
**Мета:** Створення автоматичного механізму розпізнавання та розмиття облич у потоковому відео.

**Об’єкт досліджень:** дані у відеопотоці та засоби розмиття облич.

**Предмет дослідження:** методи розмиття на основі нейронних мереж згорткового типу.

# Аналіз існуючих рішень

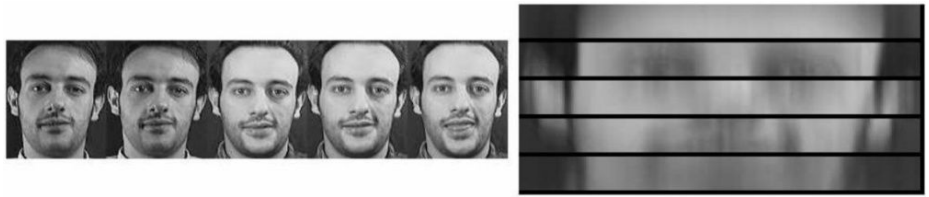
# Метод гнучкого порівняння на графах



Недоліки:

- Висока складність процедури розпізнавання.
- Низька адаптованість при запам'ятовуванні нових орієнтирів.
- Лінійна часова залежність від розміру бази даних осіб.

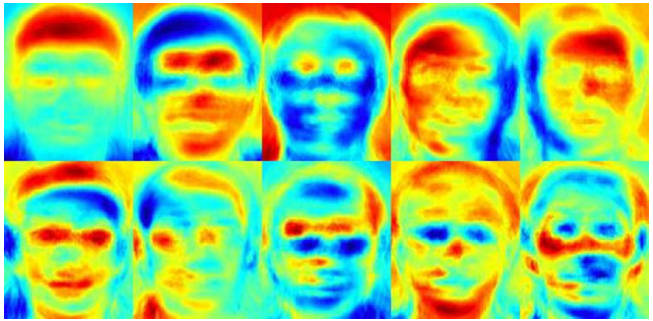
# Приховані моделі Маркова

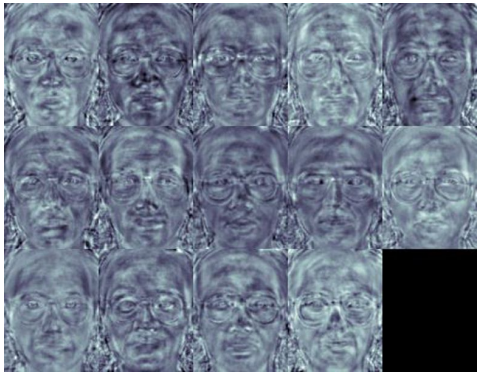


Недоліки:

- Необхідно вибрати параметри моделі для кожної бази даних.
- Немає прикладів комерційного застосування.

# OpenCV (Eigenface i FisherFaces)

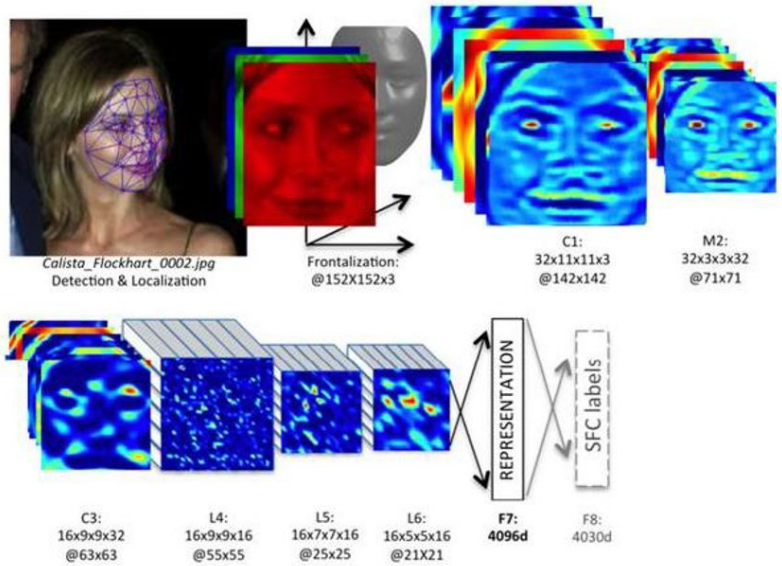




Недоліки:

- Відсутність стійкості до змін умов світла.
- Відсутність інваріантності афінним перетворенням.

# Нейронні мережі

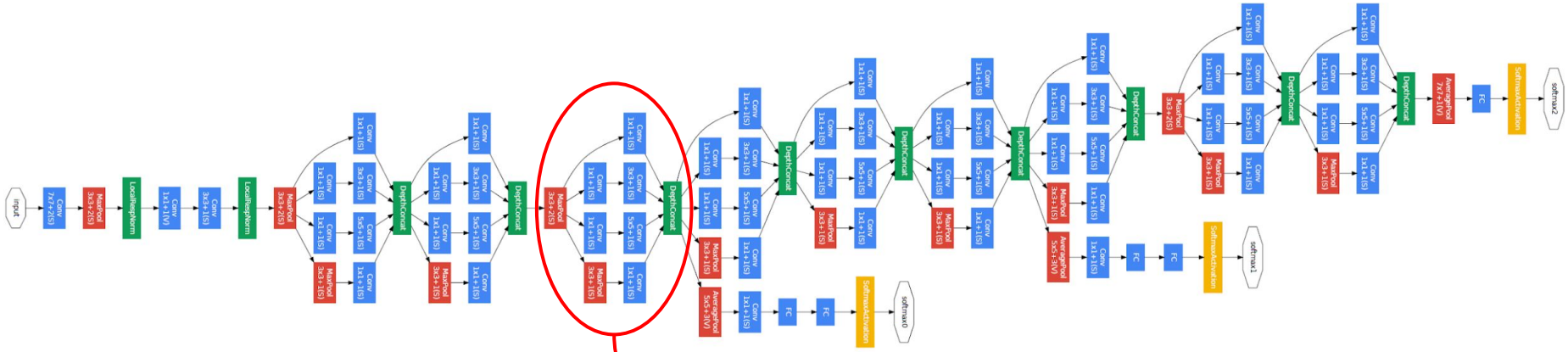



Недоліки:

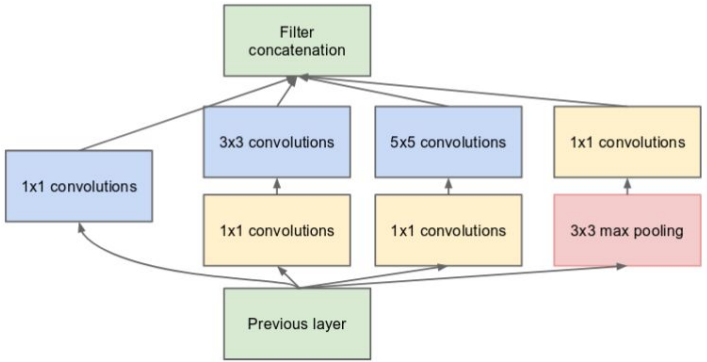
- Проблеми навчання: досягнення локального оптимуму, вибір оптимального кроку оптимізації, перепідготовка
- Тривалий процес навчання моделі (від 1 години до пару днів).

# Алгоритм розробленої системи

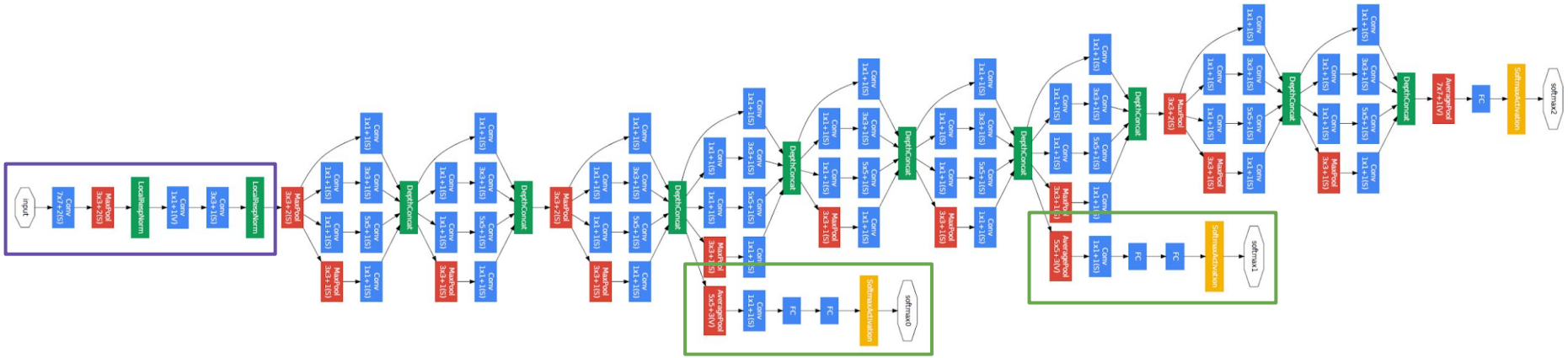
# Архітектура GoogLeNet (Inception v1)



- Складається з 9 Inception модулів, розміщених лінійно.
- Глибина 22 шари (27, включаючи шари об'єднання).
- Використовує середнє глобальне об'єднання в кінці останнього Inception модуля.



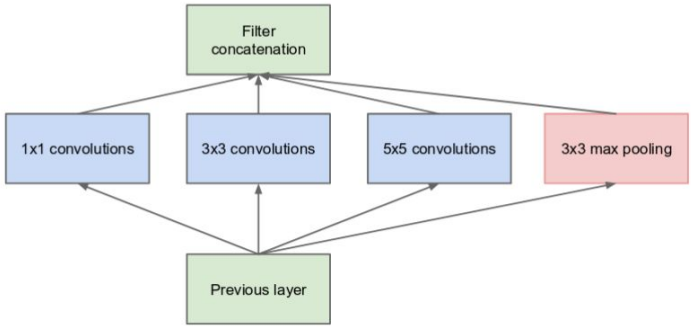
# Архітектура GoogLeNet (Inception v1)




Попередня згортка.

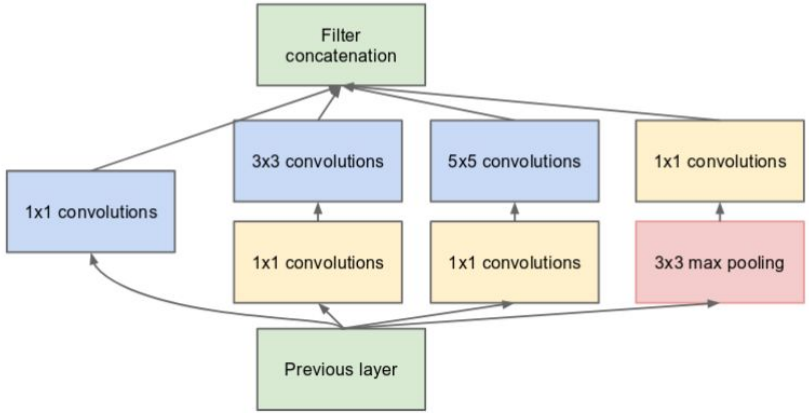
Допоміжні класифікатори. Вони застосовують softmax до виходів з двух Inception модулів, та обчислюють допоміжні втрати над тими ж мітками. Застосовують тільки для тренування моделі, під час виводу вони не застосовуються.

# Inception модуль



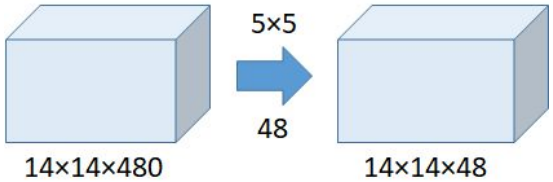

“Наївний” модуль виконує згортання на вході, з 3 різними розмірами фільтрів (1×1, 3×3, 5×5). Крім того, виконується також максимізаційне агрегування. Виходи об'єднуються і надсилаються до наступного Inception модуля.

Щоб зробити обчислення дешевшими, обмежується кількість вхідних каналів, додаючи додаткову згортку 1×1 перед 3×3 та 5×5. Хоча додавання додаткової операції може здатися не інтуїтивним, 1×1 згортки набагато дешевші, ніж 5×5 згортки, і також допомагає зменшити кількість вхідних каналів.

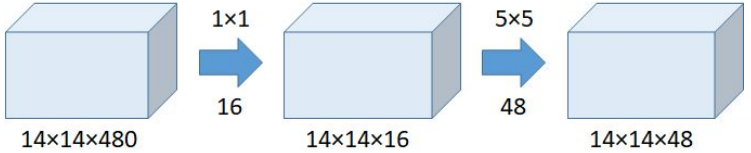


# Згортка 1×1

У GoogLeNet згортка 1×1 використовується як модуль зменшення розмірів для зменшення обчислень. 1×1 згортка може допомогти зменшити розмір моделі, що також може сприяти зменшенню проблеми з перенавчанням.




Кількість операцій = (14×14×48)×(5×5×480) = 112.9M




Кількість операцій для 1×1 = (14×14×16)×(1×1×480) = 1.5M

Кількість операцій для 5×5 = (14×14×48)×(5×5×16) = 3.8M

Загальна кількість операцій = 1.5M + 3.8M = 5.3M

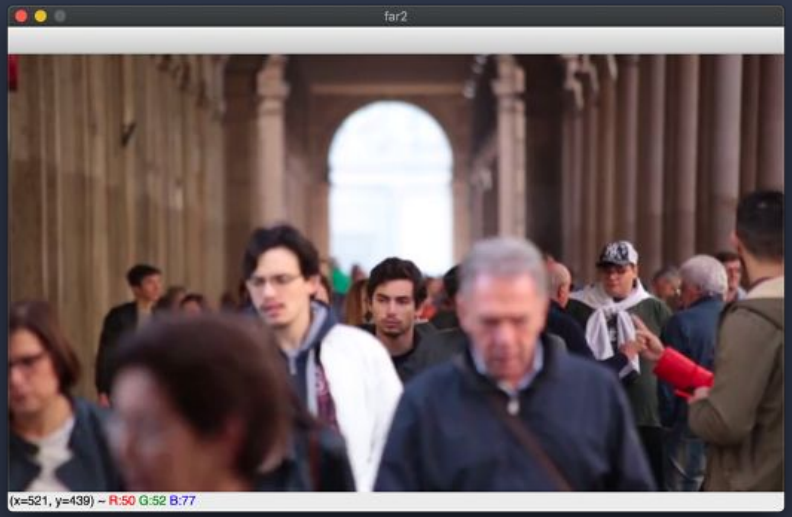
# Основні переваги GoogLeNet

- Висока точність роботи в порівнянні з іншими мережами.
- Низька потреба в обчислювальних потужностях.
- Потенціал розвитку мережі.

Еволюція GoogLeNet:

1. Inception-v1 (Inception Module)
2. Inception-v2 (Batch Normalization)
3. Inception-v3 (Factorization)
4. Inception-v4

# Відео на вході та результат на виході




# Висновки

- Нейронні мережі демонструють відмінні результати в задачах комп'ютерного зору
- Проблемою більшості алгоритмів є не універсальність та проблеми функціонування в умовах реального світу, а не лише на обраному наборі даних
- Розроблено додаток розмиття облич в потоковому відео
- Розроблена стратегія стартап-проекту, яка дозволить реалізувати представлену технологію в якості конкурентоспроможного продукту

Дякую за увагу!