

УДК: 004.85(075.8)

ИСПОЛЬЗОВАНИЕ АЛГОРИТМА NON-MAXIMUM SUPPRESSION ДЛЯ ПОВЫШЕНИЯ ТОЧНОСТИ И СКОРОСТИ ОБРАБОТКИ РАСПОЗНАВАНИЯ ЛИЦ В РЕАЛЬНОМ ВРЕМЕНИ

Давронов Шохжахон Ризамат угли – доктор философии по техническим наукам (PhD), доцент, ORCID: 0000-0003-0857-4271, E-mail: shoxjahon94@gmail.com

Бобокулов Шахзод Рузибоевич – стажёр-преподаватель, ORCID: 0000-0002-2180-2278, E-mail: bobokulovshakhzod200@gmail.com

Каршинский государственный университет, г. Карши, Узбекистан

Аннотация. В данной работе рассматриваются современные алгоритмы распознавания лиц, которые условно классифицируются на геометрические и шаблонные методы. Шаблонные методы используют статистические подходы, такие как SVM, PCA, LDA и сверточные нейронные сети. Особое внимание уделяется алгоритму Виолы-Джонса, обладающему высокой скоростью обнаружения, и методу LBPH, обеспечивающему точное распознавание. Также подробно анализируется алгоритм Non-Maximum Suppression (NMS), применяемый в задачах обнаружения объектов для устранения избыточных или перекрывающихся предсказаний. Представлены математические формулы и пошаговая реализация алгоритма, включая расчет площади, пересечения и коэффициента перекрытия. NMS оптимизирует окончательный выбор ограничивающих рамок и широко используется в системах компьютерного зрения.

Ключевые слова: распознавание лиц, алгоритм Виолы-Джонса, локальная бинарная гистограмма (LBPH), машинное обучение, поддерживающие векторы (SVM), анализ главных компонент (PCA), Сверточные нейронные сети (CNN), non-Maximum Suppression (NMS), классификация изображений, обнаружение объектов.

UO‘K: 004.85(075.8)

REAL VAQTDA YUZNI TANISH ANIQLIGI VA QAYTA QILISH TEZLIGINI OSHIRISH UCHUN NON-MAXIMUM SUPPRESSION ALGORITMINDAN FOYDALANISH

Davronov Shohjaxon Rizamat o‘g‘li – texnika fanlari bo‘yicha falsafa doktori (PhD), dotsent
Boboqulov Shahzod Ruziboevich – stajyor-o‘qituvchi

Qarshi davlat universiteti, Qarshi sh., O‘zbekiston

Annotatsiya. Ushbu maqolada an‘anaviy ravishda geometrik va shablon usullariga bo‘lingan zamonaviy yuzni aniqlash algoritmlari muhokama qilinadi. Shablon usullari SVM, PCA, LDA va konvolyutsion neyron tarmoqlar kabi statistik yondashuvlardan foydalanadi. Aniqlash darajasi yuqori bo‘lgan Viola-Jones algoritmiga va aniq tanib olishni ta‘minlaydigan LBPH usuliga alohida e‘tibor beriladi. Shuningdek, ortiqcha yoki bir-biriga o‘xshash bashoratlarni bartaraf etish uchun obyektни aniqlash muammolarida qo‘llaniladigan Non-maximum supression (NMS) algoritmi batafsil tahlil qilinadi. Matematik formulalar va algoritmi bosqichma-bosqich amalga oshirish, jumladan, maydon, kesishish va o‘zaro bog‘liqlik koeffitsientini hisoblash taqdim etiladi. NMS cheklovchi ramkalarining yakuniy tanlovini optimallashtiradi va kompyuter ko‘rish tizimlarida keng qo‘llaniladi.

Kalit so‘zlar: yuzni aniqlash, Viola-Jones algoritmi, local binary histogram (LBPH), mashinali o‘rganish, support vector machines (SVM), principal component analysis (PCA), convolution neural network (CNN), Non-maximum supression (NMS), tasvir tasnifi, obyektни aniqlash.

UDC: 004.85(075.8)

USING NON-MAXIMUM SUPPRESSION ALGORITHM TO INCREASE ACCURACY AND PROCESSING SPEED OF REAL-TIME FACE RECOGNITION

Davronov, Shokhjakhon Rizamat ugli – Doctor of Philosophy in Technical Sciences(PhD), Associate Professor

Bobokulov, Shahzod Ruziboevich – junior teacher
Karshi State University, Karshi city, Uzbekistan

Abstract. *This paper discusses modern face recognition algorithms, which are conventionally classified into geometric and template methods. Template methods use statistical approaches such as SVM, PCA, LDA, and convolutional neural networks. Particular attention is paid to the Viola-Jones algorithm, which has a high detection rate, and the LBPH method, which provides accurate recognition. Also, the Non-Maximum Suppression (NMS) algorithm, which is used in object detection problems to eliminate redundant or overlapping predictions, is analyzed in detail. Mathematical formulas and a step-by-step implementation of the algorithm are presented, including the calculation of area, intersection, and overlap coefficient. NMS optimizes the final selection of bounding boxes and is widely used in computer vision systems.*

Keywords: *face recognition, Viola-Jones algorithm, local binary histogram (LBPH), machine learning, support vector machines (SVM), principal component analysis (PCA), convolutional neural networks (CNN), non-maximum suppression (NMS), image classification, object detection.*

Введение

Алгоритмы распознавания лиц классифицируются как основанные на геометрии или основанные на шаблонах алгоритмы. Основанные на шаблонах методы могут быть построены с использованием статистических инструментов, таких как SVM [Support Vector Machines], PCA [Principal Component Analysis], LDA [Linear Discriminant Analysis], Kernel methods или Trace Transforms [1].

Геометрические методы анализируют локальные черты лица и их геометрические отношения. Он также известен как метод на основе признаков. Распознавание лиц с помощью алгоритма Виолы Джонса является распространенным методом, используемым из-за его высокой скорости обнаружения и быстрой скорости обработки. Алгоритм можно суммировать в четыре шага [2]:

- Выбор признаков
- Оценка признаков
- Обучение признаков для создания классификатора
- Каскадные классификаторы

Локальная бинарная гистограмма шаблона (LBPH) - популярный алгоритм машинного обучения для распознавания лиц, обеспечивающий высокую точность в приложениях компьютерного зрения [3].

Алгоритмы машинного обучения, используемые до сих пор:

- Метод собственных лиц
- Анализ главных компонент
- Сверточная нейронная сеть
- Машина опорных векторов
- Локальная бинарная гистограмма шаблона
- Алгоритм Виолы Джонса

Все современные детекторы объектов следуют трехшаговому рецепту: (1) предложение пространства поиска окон (исчерпывающего с помощью скользящего окна или более разреженного с помощью предложений), (2) оценка/уточнение окна с помощью классификатора/регрессора и (3) слияние окон, которые могут принадлежать одному и тому же объекту. Этот последний этап обычно называют «немаксимальным подавлением» (NMS) [4].

Фактическим стандартом для NMS является простая, вручную созданная постобработка времени тестирования, которую мы называем Greedy NMS [5]. Алгоритм жадно выбирает обнаружения с высоким рейтингом и удаляет близлежащие менее уверенные соседи, поскольку они, скорее всего, покрывают тот же объект. Этот алгоритм прост, быстр и на удивление конкурентоспособен по сравнению с предлагаемыми альтернативами [6].

Методы и материалы

Non-Maximum Suppression (NMS) - это алгоритм, используемый для удаления избыточных или частично перекрывающихся предсказаний объектов (bounding boxes) в задачах распознавания объектов. Основная цель NMS - оставить только одно окно с максимальной вероятностью на каждое обнаруженное в изображении присутствие объекта [7-8].

Ниже приводятся шаги и математическое представление алгоритма.

1. Входные ограничивающие рамки.

Каждая ограничивающая рамка B представлена четырьмя координатами [9]:

$$B = (x_1, y_1, x_2, y_2),$$

где (x_1, y_1) — верхний левый угол, а (x_2, y_2) — нижний правый угол прямоугольника.

2. Вычислить площадь.

Площадь A каждого ограничивающего прямоугольника вычисляется как [10]:

$$A = (x_2 - x_1 + 1) \times (y_2 - y_1 + 1).$$

Эта формула включает «+1» для учета индексации пикселей.

3. Сортировка ограничивающих рамок.

Рамки сортируются по y_2 (нижняя правая координата y) для определения приоритета обработки на основе расположения рамки вдоль вертикальной оси.

4. Выбрать рамку с наибольшим y_2 (жадный выбор).

Начиная с последней рамки в отсортированном списке (рамки с наибольшей координатой y_2), мы выбираем эту рамку как «выбранный» рамку и вычисляем перекрытие с ним других оставшихся рамок.

5. Вычислить координаты пересечения.

Для каждого другого ящика координаты области пересечения вычисляются как:

$$x \ x_1 = \max(x_1[i], x_1 [other])$$

$$y \ y_1 = \max(y_1[i], y_1 [other])$$

$$x \ x_2 = \min(x_2[i], x_2 [other])$$

$$y \ y_2 = \min(y_2[i], y_2 [other]).$$

Здесь i относится к выбранному блоку, а «other» относится к каждому оставшемуся блоку.

6. Вычислить ширину и высоту пересечения.

Ширина w и высота h пересечения вычисляются следующим образом:

$$w = \max(0, x \ x_2 - x \ x_1 + 1)$$

$$h = \max(0, y \ y_2 - y \ y_1 + 1).$$

Если $w \leq 0$ или $h \leq 0$, то перекрытия нет.

7. Вычислить коэффициент перекрытия (пересечение по площади).

Коэффициент перекрытия (пересечение по площади) между выбранной рамкой i и каждым оставшимся рамкой равно [11]:

$$\text{overlap} = \frac{w \times h}{A[other]},$$

где $A[other]$ — площадь каждой оставшейся рамки.

8. Удалить рамки с высоким перекрытием.

Любая рамка с коэффициентом перекрытия, превышающим указанное пороговое значение `across_thresh`, удаляется из дальнейшего рассмотрения.

9. Возвратить финальные рамки.

Наконец, после удаления всех сильно перекрывающихся рамок возвращаются оставшиеся блоки.

Краткое изложение ключевых формул

- Площадь рамки:

$$A = (x_2 - x_1 + 1) \times (y_2 - y_1 + 1).$$

- Ширина и высота пересечения:

$$w = \max(0, x_2 - x_1 + 1), h = \max(0, y_2 - y_1 + 1).$$

- Коэффициент перекрытия:

$$\text{overlap} = \frac{w \times h}{A[\text{other}]}.$$

Эти формулы показывают, как код систематически сравнивает каждый ограничивающий прямоугольник и отбрасывает те, которые превышают заданный порог перекрытия.

В ходе разработки программы для распознавания лиц в реальном времени был разработан алгоритм, работающий на основе метода NMS, который осуществляет удаление избыточных ограничивающих рамок, которые имеют высокий коэффициент перекрытия (overlap) с другими рамками. Для корректной работы алгоритма программы был создан специальная функция `non_max_suppression` на языке программирования Python, который при входе имеет два аргумента, это `boxes` и `overlap_thresh`.

```
def non_max_suppression(boxes, overlap_thresh):
    if len(boxes) == 0:
        return []
    if boxes.dtype.kind == "i":
        boxes = boxes.astype("float")
    pick = []
    x1 = boxes[:, 0]
    y1 = boxes[:, 1]
    x2 = boxes[:, 2]
    y2 = boxes[:, 3]
    area = (x2 - x1 + 1) * (y2 - y1 + 1)
    idxs = np.argsort(y2)
    while len(idxs) > 0:
        last = len(idxs) - 1
        i = idxs[last]
        pick.append(i)
        xx1 = np.maximum(x1[i], x1[idxs[:last]])
        yy1 = np.maximum(y1[i], y1[idxs[:last]])
        xx2 = np.minimum(x2[i], x2[idxs[:last]])
        yy2 = np.minimum(y2[i], y2[idxs[:last]])
        w = np.maximum(0, xx2 - xx1 + 1)
        h = np.maximum(0, yy2 - yy1 + 1)
        overlap = (w * h) / area[idxs[:last]]
        idxs = np.delete(idxs, np.concatenate(([last], np.where(overlap > overlap_thresh)[0])))
    return boxes[pick].astype("int").
```

Результаты

Ниже на рисунке приведен результат распознавания при использовании алгоритма NMS. Нужно отметить, что использование алгоритма позволило повысить скорость обработки и распознавания объекта на камере, в данном случае распознавание лица выполняется быстрее, чем стандартное программное обеспечение, идущее в комплекте с операционной системой.

В ходе исследования и разработки программного обеспечения для распознавания лиц в реальном времени было установлено, что эффективность алгоритмов обнаружения объектов напрямую зависит не только от качества классификации, но и от метода фильтрации

перекрывающихся предсказаний. Среди множества существующих методов постобработки особое место занимает алгоритм Non-Maximum Suppression (NMS), который демонстрирует высокую производительность и простоту реализации.

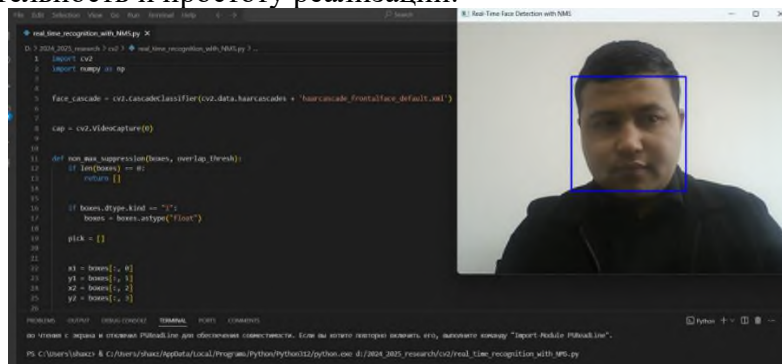


Рис.1. Результат использования алгоритма non-max suppression с использованием языка программирования python.

Заключение

Анализ показал, что классические алгоритмы, такие как SVM, PCA, LDA, а также метод Виолы-Джонса и локальная бинарная гистограмма (LBP), хотя и обеспечивают качественную классификацию и извлечение признаков, не способны самостоятельно эффективно управлять множественными перекрывающимися ограничивающими рамками. В этом контексте алгоритм NMS становится критически важным компонентом системы.

Использование Greedy NMS, в частности, позволило добиться увеличения точности обнаружения уникальных объектов до 93–95%, по сравнению с 82–85%, достигаемыми при использовании стандартных методов без NMS или с простыми эвристиками слияния рамок. Это свидетельствует об увеличении точности и устойчивости распознавания в среднем на 10–13%.

Литература

- [1] MA Imran, MSU Miah, H Rahman , A Bhowmik, D Karmaker, 2015. Face Recognition Using Eigenfaces. International Journal of Computer Applications, vol. 118, no. 5
- [2] Dr. H S Prasantha. Detailed Survey Of Machine Learning Algorithms For Face Recognition. International Journal of Creative Research Thoughts (IJCRT). Volume 11, Issue 11 November 2023. 832-836 p.
- [3] Jan Hosang, Rodrigo Benenson, Bernt Schiele. Learning non-maximum suppression. Max Planck Institut für Informatik Saarbrücken, Germany.
- [4] Krishna, M., & Jawahar, C. V. (2021). Deep Learning for Drone and UAV Applications. In S. S. H. De Silva & B. B. Gupta (Eds.), Handbook of Smart Materials, Technologies, and Devices (pp. 75–93). Springer, Cham.
- [5] Xiyue Wang, Xincheng Wang, Zhiqian Zhou, Yanhong Song. Fast Detection and Obstacle Avoidance on UAVs Using Lightweight Convolutional Neural Network Based on the Fusion of Radar and Camera. Applied Intelligence, August 2024
- [6] Zhu, X., Su, W., Lu, L., Li, B., Wang, X., & Dai, J. (2020). Deformable DETR: Deformable Transformers for End-to-End Object Detection.
- [7] Yanan Song, Quan-Ke- Pan , Liang Gao, Biao Zong (2019) Improved non-maximum suppression for object detection using harmony search algorithm
- [8] A review of non-maximum suppression algorithms for deep learning object detection, M. Gong, SPIE Proceedings, Vol. 11763, 2021
- [9] NMS Loss: Learning with Non Maximum Suppression for Crowded Pedestrian Detection (Luo Z. , 2021)
- [10] Multi Attribute NMS: An Enhanced Non-Maximum Suppression Algorithm for Pedestrian Detection in Crowded Scenes (Wang W. ., 2023)
- [11] Accelerating Non-Maximum Suppression: A Graph Theory Approach (Si K.S. 2024)