

Звіт “Розпізнавання облич”

1. Огляд проблеми

Натренувати модель, здатну класифікувати зображення для вирішення проблеми ідентифікації обличчя в наявному наборі даних. Для вирішення цієї задачі буде використовуватись [база даних облич ORL](#), яка містить набір зображень облич, зроблених в період з квітня 1992 року по квітень 1994 року в лабораторії Кембриджського університету (AT&T Laboratories Cambridge)

Набір даних складається з 400 чорно-білих зображень. Кожне зображення належить одному з 40 суб'єктів. Кожен суб'єкт має 10 фотографій, при цьому знімки були зроблені в різний час, змінюючи освітлення, міміку (відкриті / закриті очі, посміхаючись / не посміхаючись) і деталі обличчя (окуляри / без окулярів). Розмір кожного зображення становить 112x92 пікселів.

Під час роботи з набором даних та тренування моделі, необхідно ознайомитись з різними методами покращити якість набору даних та точність і ефективність моделі, включаючи масштабування та нормалізацію даних, зменшення розмірності даних методом головних компонент (МГК, англ. principal component analysis, PCA) та попередню обробку зображень (нахил, збільшення, обрізання).

2. Існуючі математичні моделі для вирішення цієї проблеми

На даний час відомо та використовується велика кількість методів розпізнавання облич. Поширеними є методи та їх похідні такі як метод Віоли-Джонса, еластичних графів, метод головний компонент, методи принцип яких базується на геометричному методі розпізнавання, методи засновані на нейронних мережах такі як метод прихованої Маркової моделі, метод згорткової нейронної мережі, метод локальних бінарних шаблонів та ін.

Кожен із методів має свої переваги та недоліки, які проявляються при тих чи інших ситуаціях, що призводить до непередбачуваних випадків “пропустити чужого чи відмова у доступі для свого”. Фактори, що впливають на якість роботи методів зазвичай одні і ті ж, та все ж таки кожен із методів має свою стійкість до певних із них.

Такими факторами являється рівень та кут освітлення, відстань від камери, міміка, настрої, кут нахилу обличчя, кут самого обличчя відносно камери, стан здоров'я людини, вікові зміни, наявність бороди чи вус та ін.

Геометричний (антропометричний) метод розпізнавання обличчя є одним із перших серед використовуваних методів розпізнавання обличчя. Спочатку застосовувався в криміналістиці і був там детально розроблений. Потім з'явилися комп'ютерні реалізації цього методу.

В методах цього виду розпізнавання полягає у виділенні набору ключових точок (або областей) особи і наступному формуванні набору ознак. Серед ключових

точок можуть бути куточки очей, губ, кінчик носа, центр ока тощо. Даний метод задає високі вимоги до зйомки зображень і потребує надійного алгоритму знаходження ключових точок для загального випадку.

Метод гнучкого порівняння на графах, суть якого зводиться до порівнянні графів, що описують зображення обличчя особи. В окремих публікаціях вказується 95-97% ефективність розпізнавання навіть при наявності різних емоційних виразів і зміні ракурсу при формуванні зображення особи до 15 градусів. Має низьку чутливість до рівня освітленості обличчя.

Проте цей метод потребує багато часу та ресурсів ПК. Іншим недоліком такого підходу є низька технологічність при запам'ятовуванні нових еталонів, що загалом призводить до нелінійної залежності часу роботи від розміру бази даних осіб. Загалом цей підхід має нижчі показники за достовірністю розпізнавання, ніж методи, побудовані із використанням нейромереж.

Метод Віюлі-Джонса дозволяє виявляти об'єкти на зображеннях в реальному часі. Метод добре працює при спостереженні об'єкта під невеликим кутом, приблизно до 30°. Точність розпізнавання з використанням даного методу частково досягає понад 90%, що є хорошим результатом. Однак при куті відхилення понад 30° ймовірність розпізнавання різко падає. Враховуючи дану особливість унеможливорюється детектування особи під довільним кутом.

Використання нейронних мереж. Одні з найкращих результатів в області розпізнавання осіб досягається за допомогою використання згорткових нейронних мереж (ЗНМ). ЗНМ забезпечує часткову стійкість до змін масштабу, зсувам, поворотам, зміні ракурсу та іншим спотворень. До недоліків методів, які побудовані на основі нейронних мереж можна віднести додавання нового еталонного особи в базу даних, що вимагає повного перенавчання мережі на всьому наявному наборі, а це досить тривала процедура, яка залежно від розміру вибірки вимагає годин роботи а то і декількох днів. Також їм властиві проблеми математичного характеру, пов'язані з навчанням: попадання в локальний екстремум, вибір оптимального кроку оптимізації, перенавчання, тощо.

Локальні бінарні шаблони (ЛБШ) вперше були запропоновані в 1996 році для аналізу текстури піктонових зображень. Дослідження показали, що ЛБШ інваріантні до невеликих змін в умовах освітлення і невеликим поворотам зображення. Методи на основі ЛБШ добре працюють при використанні зображень облич із різною мімікою, різним освітленням, поворотами голови. Серед недоліків - необхідність якісної попередньої обробки зображень через високу чутливість до шуму, оскільки за його присутності зростає кількість помилкових бінарних кодів.

Приховані моделі Маркова. Прихована марковська модель — статистична модель, що імітує роботу процесу, схожого на марківський процес з невідомими параметрами. Згідно моделі ставиться задача знаходження невідомих параметрів на основі інших параметрів, за якими ведеться спостереження. Отримані параметри можуть бути використані в подальшому аналізі для розпізнавання облич. З точки зору розпізнавання – зображення це двовимірний дискретний сигнал. Важливу роль в побудові моделі зображення грає вектор спостереження. Для того, щоб уникнути

розбіжностей в описах, зазвичай використовують прямокутне вікно для розпізнавання. Щоб не втрачати області даних, прямокутні вікна мають перекривати одне інше.

Метод головних компонент (МГК) зводить процес розпізнавання чи класифікації до побудови для вхідного зображення певної кількості головних компонент зображень.

Метод k-найближчих сусідів (англ. k-nearest neighbor method) – простий непараметричний класифікаційний метод, де для класифікації об'єктів у рамках простору властивостей використовуються відстані (зазвичай евклідові), порашовані до усіх інших об'єктів. Відбираються об'єкти, до яких відстань найменша, і вони виділяються в окремий клас.

Ми будемо використовувати саме метод найближчих сусідів.

3. Роз'яснення алгоритму

Основним принципом методу найближчих сусідів є те, що об'єкт присвоюється тому класу, який є найбільш поширеним серед сусідів даного елемента. Сусіди беруться, виходячи з множини об'єктів, класи яких уже відомі, і, виходячи з ключового для даного методу значення k , вираховується, який клас є найчисельнішим серед них. Кожен об'єкт має кінцеву кількість атрибутів (розмірностей). Передбачається, що існує певний набір об'єктів з уже наявною класифікацією.

У нас є набір даних 400 зображень розміром 112x92 пікселів. Від так ми можемо представити кожне зображення матрицею з розмірністю 112x92 (глибину ігноруємо, оскільки чорно-білі зображення мають глибину 1).

Спочатку ми векторизуємо кожне зображення, таким замість матриці 112x92 ми отримаємо вектор довжиною 10304, а загалом весь набір даних буде матрицею розміром 400x10304, тобто 400 спостережень, кожне з яких має 10304 ознак. Кожна ознака має величину в діапазоні від 0 до 255 включно (насиченість пікселя від білого до чорного). Масштабуємо всі ознаки, так щоб їх величини були в діапазоні від 0 до 1.

Далі розіб'ємо наш датасет на тренувальну та тестові виборки (80% тренувальна і 20% тест). Так ми отримуємо 320 спостережень в тренувальному наборі, та 80 спостережень в тестовому наборі.

Далі для кожного прикладу в тестовому наборі даних знайдемо евклідові відстані (MSE) до усіх інших об'єктів в тренувальному наборі даних. Присвоїмо тестовим прикладам ті класи, які відповідають їх найближчим сусідам (найкоротша евклідова відстань) в тренувальній вибірці.

4. Результати

Порівняємо отримані мітки, з відомими мітками тестового набору. Я отримав точність (ассигасу) моделі 97.5%

Вихідний код на [GitHub](#)

4.1 Нормалізація

Спробуємо покращити точність моделі з допомогою нормалізації даних. Існує багато способів нормалізації набору даних, ми використаємо 2 з них:

- **за спостереженнями (прикладами)** - у випадку безперервно зростаючого набору даних, як наприклад у випадку онлайн-навчання, ми нормалізуємо кожне спостереження (в нашому випадку кожне зображення обличчя) окремо: обчислити середнє значення від введення даних, відняти середнє від вихідного зображення і розділити його на стандартне відхилення зображення;
- **за ознаками** - якщо у нас є весь набір даних, ми можемо для кожної ознаки в наборі даних обчислити середнє значення і відняти його від ознак.

Порівняємо результати.

Нормалізація	За ознаками	За прикладом	Без нормалізації
Точність	0.9875	0.95	0.9875

Як бачимо, нормалізація за прикладами погіршила точність, а нормалізація за ознаками не вплинула на точність класифікації.

4.2 Вирівнювання зображення

Інколи покращити точність класифікації можна за рахунок обробки зображень. Так, деякі моделі є дуже чутливі до нахилу зображень, і якщо відцентрувати зображення, можна покращити результат класифікації. Проте в моєму випадку мені не вдалось це зробити, змінюючи нахил голови (повертаючи зображення). На мою думку це пояснюється тим, що помилки в класифікації обумовлені не хибним нахилом голови (набір даних від початку є якісним і більш-менш відцентрованим), а наявністю на деяких фотографіях окулярів, нетипової міміки, або розворотом обличчя, що неможливо виправити зміною нахилу голови.

4.3 PCA

Інколи для збільшення точності роботи моделі або пришвидшення аналізу, використовують зменшення розмірності даних (зменшення кількості ознак). Це може значно зменшити час тренування та прогнозування Нейронної Мережі. Інколи зменшення розмірності може збільшити точність такого алгоритму машинного навчання, як Дерево ухвалення рішень. Крім того зменшення розмірності з багатовимірною простору, до, наприклад, тривимірною, дозволяє візуалізувати дані.

Сенс полягає в тому, що в деяких наборах даних далеко не всі ознаки об'єкта несуть корисну інформацію про об'єкт, дублюють одна одну, не мають суттєвий вплив на точність роботи моделі, а відтак кількість таких ознак може бути скорочена до найбільш важливих.

Одним їх алгоритмів зменшення розмірності є метод головних компонент (PCA).

Етапи алгоритму PCA:

- масштабувати дані в діапазоні від -1 до 1
- обчислити матрицю коваріацій. $C = X.T * X$
- обчислити власні значення λ_i та власні вектори v_i матриці коваріацій
- відсортувати власні вектори, за власними значеннями
- відібрати k головних компонент відповідно до k найбільших власних значень
- обчислити матрицю перетворення. $W = k_important_eigenvectors * X$
- проєціювання набору даних у підпростір PCA. $X_reduced = X * W$

Застосувавши PCA мені вдалося зменшити кількість корисних ознак з 10304 до 166 і при цьому зберегти точність класифікації моєї моделі на рівні 98.75% що відповідає класифікації без застосування PCA та нормалізації.

	PCA energy = 0.95	PCA energy = 0.98	PCA energy = 0.99	PCA energy = 1.00	Без нормалізації
Точність	0.9875	0.9875	0.9875	0.9875	0.9875
Кількість ознак	166	234	268	10304	10304

При цьому час, затрачений на проєціювання тестового набору даних у підпростір PCA та класифікацію зріс приблизно у 2 рази у порівнянні з часом затраченим на класифікацію оригінального тестового набору без застосування PCA.

Посилання та джерела

1. "[Методи Розпізнавання Облич: Короткий Огляд](#)" І. Голуб'як кафедра інформаційних технологій ДВНЗ «Прикарпатський національний університет.
2. "Обзор математических методов распознавания изображений лиц" С. А. Пучинин, аспирант кафедры «Прикладная математика и информатика» Ижевский государственный технический университет.
3. <https://habr.com/ru/post/304214/>
4. <https://m.habr.com/ru/company/ods/blog/325654/>
5. <http://qaru.site/questions/124527/principal-component-analysis-pca-in-python>
6. <https://www.programcreek.com/python/example/61552/numpy.cov>