

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ, СИСТЕМНИЙ АНАЛІЗ ТА КЕРУВАННЯ

УДК 004.62

DOI: 10.20535/1810-0546.2017.6.115181

С.С. Ніколаєв*, Ю.О. Тимошенко, К.Ю. Матвіїв

КПІ ім. Ігоря Сікорського, Київ, Україна

ЗАЛЕЖНІСТЬ ЯКОСТІ ДЕТЕКТОРА ОБЛИЧ НА ОЗНАКАХ ХААРА ВІД ВАРІАТИВНОСТІ НАВЧАЛЬНОЇ ВИБІРКИ

Background. When training generalized face detectors based on Haar cascades, there is a problem of long learning time of the resulting cascades and their poor quality. Therefore, in practice, frontal and profile face detectors are trained separately. Such approach makes recognition systems more complex.

Objective. The aim of the paper is to compare the impact of the training set composition with faces at different inclination angles on the quality of the trained detectors.

Methods. It is proposed to train a series of face detectors on sub-sets that cover different ranges of face angles. All other parameters of training are fixed. As the result, the learning time and the quality of the obtained cascades will be compared.

Results. The quality and the training time of face classifiers are evaluated depending on the composition of the training subsets. Also the quality of the frontal and side face classifiers is compared having the same sizes of training sets. It is shown that the *AUC* metric has a difference of 0.003 between the frontal and profile face detectors.

Conclusions. It has been shown experimentally that the more variations present in the object's dataset (the side-view of faces compared to the frontal positions), the longer and harder the Haar cascade learns, given the same amounts of the training samples. Using the proposed approach, the quality of the final classifier can be controlled by selecting the appropriate composition of the training samples.

Keywords: face detector; Haar cascades; training set; boosting; training set composition.

Вступ

Однією з актуальних проблем розпізнавання образів є задача детекції облич, яка набула широкого застосування в охоронних системах, медицині, соціальній сфері тощо. Для її розв'язання ключовим завданням є побудова швидких та якісних детекторів облич людини, які знаходять присутні обличчя на статичних зображеннях або кадрах у відео.

На сьогодні відомо чимало методів розпізнавання облич, зокрема дерева класифікації, машини опорних векторів, класичні та глибокі нейронні мережі тощо [1]. Для розв'язання задач реального часу, таких як знаходження обличчя у відеопотоці, необхідні методи, які одночасно мають задовольняти дві вимоги, а саме: бути достатньо швидкими, щоб до отримання наступного кадру встигнути обробити поточний, і в той же час бути достатньо точними. Ці вимоги є протилежними, адже чим більш точний метод, тим зазвичай більше обчислювальних ресурсів він вимагає.

Стандартом де факто за швидкодією функціонування та якістю роботи вважається детектор у вигляді каскаду на ознаках Хаара, вперше представлений П. Віолою та М. Джонсом [2, 3], який і був вибраний для нашого дослідження.

Ознаки Хаара (Haar-Like features) – ознаки цифрового зображення, які складаються із суміжних прямокутних областей і позиціонуються на зображенні. Вони використовуються в каскадах Хаара як слабкі класифікатори для розпізнавання образів. Такий підхід дає змогу швидко обчислювати суму інтенсивностей пікселів для будь-яких прямокутних областей завдяки використанню оптимізації, відомої як інтегральне перетворення зображення.

Ще один важливий момент у методі Віоли–Джонса – це використання каскадної моделі або виродженого дерева прийняття рішень: у кожному вузлі дерева приймається рішення, чи може на зображенні міститися шуканий об'єкт, чи ні. Навчання побудовано таким чином, щоб на початкових рівнях із найменшими витратами відкидати більшу частину вікон, у яких не може міститися об'єкт. Таким чином, отримується детектор шуканого об'єкта, який являє собою навчений каскад з ознаками Хаара.

“Універсальний” детектор облич повинен правильно класифікувати будь-яке обличчя при довільному ракурсі та освітленні у вхідному зображенні, що означає наявність варіативної навчальної вибірки об'єкта.

При спробі натренувати такий універсальний детектор може виявитися, що на практиці

* corresponding author: sergiynicolaev@gmail.com

після навчання він або робитиме багато хибних класифікацій, або узагалі не знайде закономірностей.

Відомо, що чим більш варіативний об'єкт намагається апроксимувати такий метод, тим більше часу і розмічених даних необхідно на навчання і тим повільніше падає крива похибки. При цьому тренування моделей може тривати кілька тижнів або місяців навіть на сучасних серверах.

Для зменшення варіативності вибірки, що описує об'єкт, на практиці застосовується її поділ на підвибірки за якоюсь ознакою. Наприклад, положення кожного обличчя може бути охарактеризоване трьома кутами нахилу – в горизонтальній площині, у вертикальній площині та нахилом голови вбік. Чим більші кути нахилу є у вибірці, тим більш повно описано об'єкт детекції, але водночас тим складнішою стає задача навчання детектора на сформованій підмножині форм обличчя.

Зазвичай рішення проблеми створення універсального детектора полягає у навчанні низки спеціалізованих детекторів, що відповідають виключно за певні підмножини форм об'єкта, наприклад детектор фронтальних облич, детектор облич у профіль і т.п. Такий підхід допомагає значно підвищити якість, але при цьому страждає швидкодія, бо до зображення доводиться застосовувати декілька спеціалізованих детекторів замість одного. Тому необхідно мінімізувати саму кількість спеціалізованих детекторів, кожен із яких повинен покривати підвибірку з якомога ширшою підмножиною форм об'єкта. На прикладі облич це означає, що кожна підвибірка повинна включати якомога ширші діапазони кутів нахилу голови.

Постановка задачі

Метою роботи є проведення порівнянь, яким чином формування тренувальної вибірки облич впливає на якість, швидкість роботи та процес навчання детекторів, адже кожна тренувальна вибірка описує об'єкт по-різному і встановити шукані параметри можливо тільки практичним навчанням детектора.

Для цього необхідно:

- побудувати ряд фронтальних і бокових класифікаторів через попереднє їх навчання за допомогою тренувальної вибірки;
- оцінити якість отриманих класифікаторів і час навчання;

- порівняти результати для різних класифікаторів залежно від величини вибірки та виду детекції (фронтальна чи бокова);

- вибрати серед побудованих класифікаторів найбільш якісний.

Збір і підготовка даних

Фронтальним класифікатором будемо вважати такий, що повинен відслідковувати обличчя, яке може бути повернуте в межах 15° по горизонтальній осі. Боковим класифікатором будемо вважати, відповідно, такий, що відслідковуватиме обличчя, повернуте на більш ніж 15° . Таке розділення є обґрунтованим з тої точки зору, що для успішного навчання класифікатора на навчання йому необхідно подавати якомога більш стандартизовані зображення облич.

В інтернеті є значна кількість готових баз даних із зображенням обличчя людини та відповідними файлами розмітки. Так, у [4] наведено 60 таких баз, які містять також опис файлів для розмітки. Пріоритетним для авторів є вибір такої бази, що не тільки містить репрезентативну вибірку, достатньо велику та різноманітну щодо положення обличчя, але й має необхідну кількість розмітного матеріалу. Таким чином, у роботі вибрано базу облич UMD Faces Dataset, яка містить 367920 зображень і 3 розмітних файли у форматі csv, що характеризують положення облич по 21 ключовій точці для кожного обличчя [5].

Навчальна вибірка сформована вирізанням і масштабуванням із оригінальних зображень облич по координатах із розмітного файлу. Далі зображення розділено на 9 груп залежно від кутів повороту обличчя у горизонтальній і вертикальній площинах. Виявлено, що в горизонтальній площині обличчя з вибірки знаходяться в інтервалі $[-65^\circ; 65^\circ]$, тому весь інтервал розділено на 3 підмножини:

$$A_1 = [-65^\circ; -30^\circ] \cup (30^\circ; 65^\circ];$$

$$A_2 = [-30^\circ; 15^\circ] \cup (15^\circ; 30^\circ];$$

$$A_3 = [-15^\circ; 15^\circ].$$

Для вертикальної площини виявлено, що обличчя із заданої вибірки знаходяться в інтервалі $[-35^\circ; 30^\circ]$, після чого виділено такі 3 підмножини:

$$B_1 = [-35^\circ; -10^\circ); B_2 = [-10^\circ; 10^\circ); B_3 = [-10^\circ; 30^\circ].$$

У деяких групах опинилось замало зображень для навчання, тому виділено найбільші

4 групи зображень. Найменша з цих груп містила вибірку в 10000 зображень. Далі з кожної групи випадковим чином було відібрано частину зображень для тестування. Відібрані зображення не брали участі в подальшому навчанні класифікаторів, щоб мати можливість перевірити якість навчених класифікаторів на незалежній вибірці, яку вони не бачили до цього. Для зображень, що залишились для навчання, формувався конфігураційний файл *Good.dat* за допомогою бібліотеки *OpenCV* такою командою:

```
opencv_createsamples.exe -info Good.dat
-vecsamples.vec -w 20 -h 20 -num 60621
```

На виході ця утиліта формувала бінарний файл із масштабованими зображеннями розміром 20×20 пікселів. На цьому формування позитивної вибірки завершується.

Наступним кроком є підготовка негативних зображень (зображень фону). Такими зображеннями можуть бути ті, які не містять людського обличчя та відповідають фонам, на яких можна зустріти людину. Це різноманітні зображення міста, природи, побуту тощо. Такі зображення відібрано з колекції баз *The PASCAL Object Recognition Database Collection* [6]. Таким чином, відібрано більше 300000 зображень, після чого сформовано конфігураційний файл *Bad.dat*.

Процедура побудови класифікаторів

Побудова класифікаторів здійснювалась за допомогою утиліти *opencv_traincascade* за допомогою команди

```
opencv_traincascade -data classifier
-vecsamples.vec -bg Bad.dat -numStages 20
-minHitRate 0.999 -maxFalseAlarmRate 0.3
-numPos 28000 -numNeg 9000 -w 20 -h 20
-mode All -precalcValBufSize 8192
-precalcIdxBufSize 8192
```

Було виділено чотири типи класифікаторів, два з яких фронтальні, а два інших – бокові. Для зручності будемо називати класифікатором 1 той, що має розпізнавати фронтальні обличчя, де кут нахилу обличчя в горизонтальній площині знаходиться в інтервалі $[-15^\circ; 15^\circ]$, а у вертикальній – в діапазоні $[-10^\circ; 10^\circ]$ (рис. 1).

Для класифікатора 2 кут нахилу у горизонтальній площині знаходиться в інтервалі $[-15^\circ; 15^\circ]$, а у вертикальній – у діапазоні $[-35^\circ; -10^\circ]$ (рис. 2).



Рис. 1. Приклади зображень для класифікатора 1



Рис. 2. Приклади зображень для класифікатора 2

Для класифікаторів 3 і 4 кут у горизонтальній площині знаходиться в інтервалі $[-30^\circ; 15^\circ] \cup (15^\circ; 30^\circ]$ та $[-65^\circ; -30^\circ] \cup (30^\circ; 65^\circ]$ (рис. 3 і 4 відповідно).

Приклад детекції обличч одним із натренованих фронтальних класифікаторів можна побачити на рис. 5.

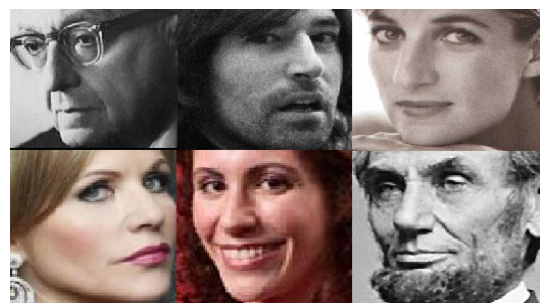


Рис. 3. Приклади зображень для класифікатора 3



Рис. 4. Приклади зображень для класифікатора 4

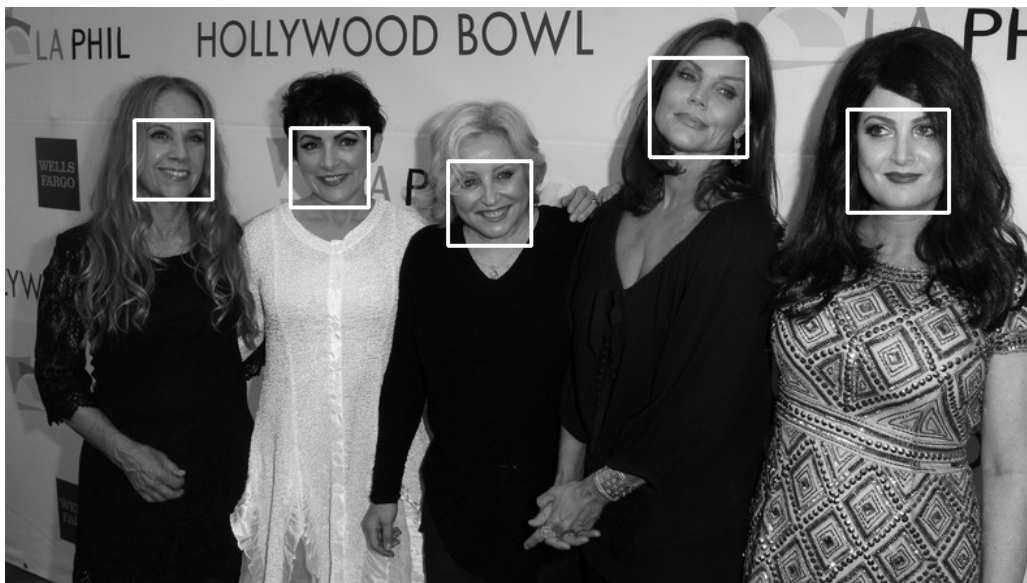


Рис. 5. Приклад детекції обличчя фронтальним класифікатором

Порівняння якості класифікаторів

Введемо такі позначення метрик якості:

- tp – кількість обличчя, розпізнаних детектором як обличчя (true positive);
- fp – кількість фонових зображень, розпізнаних детектором як обличчя (false positive);
- tn – кількість фонових зображень, розпізнаних детектором як фон (true negative);
- fn – кількість обличчя, розпізнаних детектором як фон (false negative);
- точність детекції обличчя:

$$precision = \frac{tp}{tp + fp};$$

- імовірність того, що позитивний зразок буде розпізнано:

$$recall = \frac{tp}{tp + fn};$$

- імовірність правильної класифікації зразка:

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn}.$$

Щоб розрахувати якість, для кожного типу класифікаторів виділено частину вибірки, на якій не проводилось навчання. Перша частина складається з 250000 негативних зображень. Для пришвидшення тестування таку вибірку спершу було збережено у бінарний файл і зображення були попередньо масштабовані до розміру 20×20 пік-

селів у відтинках сірого. На цьому файлі послідовно запускаються відповідні класифікатори і відслідковується кількість зображень фону, що розпізнані як обличчя.

Позитивні вибірки для кожного типу детектора складаються із зображень, що містять обличчя з відповідними кутами повороту. Всі зображення позитивних вибірок також оброблюються аналогічно до негативних зображень, після чого тестуються на побудованих класифікаторах. Кожна із позитивних вибірок містить по 10000 зображень, що не брали участі у навчанні.

У нашому дослідженні відслідковувалися залежності часу навчання та якості натренованих детекторів від кількості етапів навчання, розміру та композиції вибірки. Для порівняння під час кожного експерименту всі параметри, окрім досліджуваного, фіксувалися.

Спочатку було досліджено вплив розміру навчальної вибірки на криві якості навчання. Отримані результати подані на графіках (рис. 6–9), у яких по осі абсцис відкладено розмір вибірки у тисячах зображень, а по осі ординат – значення критерію якості. Кількість етапів навчання позначається певним типом лінії.

Точність детекції обличчя можна інтерпретувати як імовірність того, що знайдене обличчя є насправді обличчям, а не фоном. З графіків на рис. 6 видно, що точність детекції обличчя підвищується зі збільшенням розміру вибірки та зменшенням етапів навчання каскаду. Так, наприклад, для досягнення точності в 95 % при 14 етапах навчання необхідна вибірка розміром у 15 тисяч зображень.

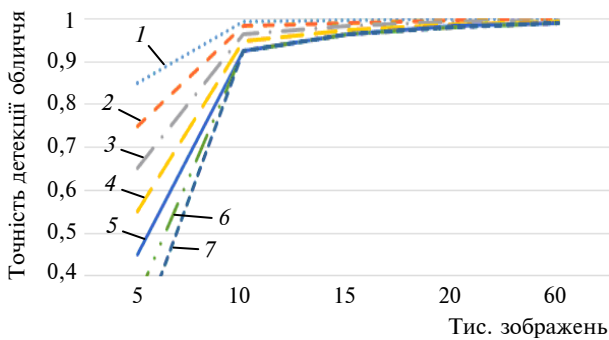


Рис. 6. Графік точності детекції обличчя залежно від розміру вибірки. Кількість етапів навчання: 1 – 6; 2 – 8; 3 – 10; 4 – 12; 5 – 14; 6 – 16; 7 – 20

Імовірність правильної класифікації показує, наскільки точно детектор вміє правильно розпізнавати об'єкти: чим більше хибних спрацювань, тим менша ймовірність (рис. 7).

З рис. 7 видно, що при шести етапах навчання для досягнення ймовірності правильної класифікації обличчя більш ніж 85 % необхідно мати не менш ніж 60 тисяч позитивних зображень у вибірці. Системи продуктового рівня потребують значно вищої ймовірності правильної класифікації, наприклад, вона повинна становити не менше 99,9999 %. Звідси можна зробити висновок, що кількість рівнів каскаду та, відповідно, етапів навчання для цієї бази облич повинна вибиратися не меншою 14.

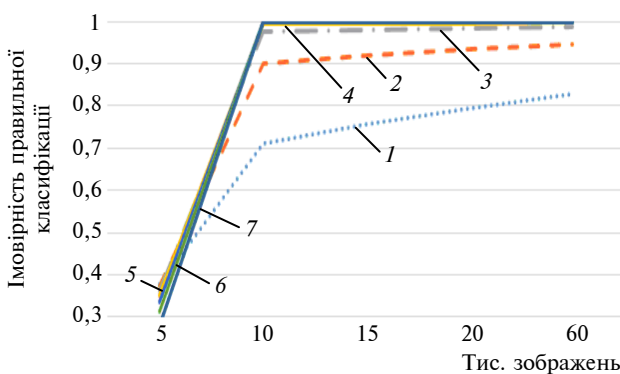


Рис. 7. Ймовірність правильної класифікації залежно від розміру вибірки. Кількість етапів навчання: 1 – 6; 2 – 8; 3 – 10; 4 – 12; 5 – 14; 6 – 16; 7 – 20

Імовірність розпізнавання обличчя показує відсоток правильно розпізнаних облич серед усіх облич у вибірці (рис. 8).

Наприклад, із графіка на рис. 8 видно, що для досягнення детектором 95 %-ної ймовірності розпізнавання обличчя при наявній вибірці зобра-

жень розміром у 10 тисяч необхідно не менше восьми етапів навчання класифікатора. При наявній вибірці у 60 тисяч зображень 95 %-на ймовірність розпізнавання досягається вже на шостому етапі навчання.

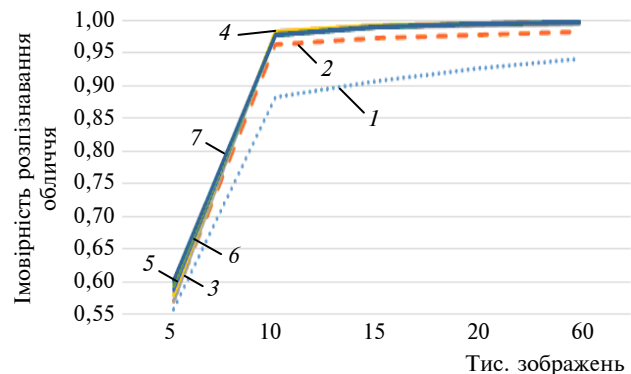


Рис. 8. Ймовірність розпізнавання обличчя зображення залежно від розміру вибірки. Кількість етапів навчання: 1 – 6; 2 – 8; 3 – 10; 4 – 12; 5 – 14; 6 – 16; 7 – 20

Можна помітити, що при збільшенні розміру вибірки якість класифікаторів суттєво зростає, наприклад, особливо помітний стрибок ймовірності при збільшенні розміру вибірки з 5 до 10 тисяч. Отже, використання вибірок із кількістю зображень близько 5 тисяч не забезпечує необхідної якості.

Проаналізуємо час, витрачений на навчання класифікаторів залежно від розміру вибірки (рис. 9).

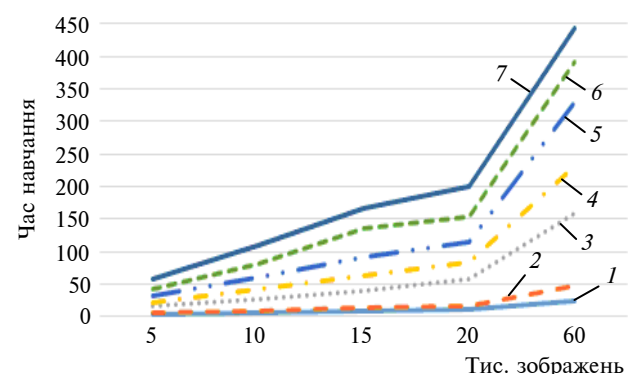


Рис. 9. Час навчання класифікаторів залежно від розміру вибірки. Кількість етапів навчання: 1 – 6; 2 – 8; 3 – 10; 4 – 12; 5 – 14; 6 – 16; 7 – 20

Як видно з рис. 9, час навчання класифікаторів росте експоненціально до розміру вибірки, тому зрозуміло, що чим більший розмір вибірки,

тим вища якість навчання, але й тим більшим є час, витрачений на навчання класифікатора.

Спираючись на сформовані вимоги, автори побудували низку класифікаторів, серед яких найбільш якісним є класифікатор, який містить вибірку в кількостях 60 тисяч позитивних зображень та негативних зображень у розмірі 60 % (від позитивних) протягом 20 етапів навчання.

Порівняння фронтального та бокового класифікаторів

Для виявлення залежностей якості детектора обличчя від варіативності навчальної вибірки було навчено декілька груп по 4 класифікатори кожного типу за однакової величини вибірки в 10000 по-

зитивних зображень і 8000 негативних. Групи відрізняються кількістю етапів навчання. Кількість етапів навчання при побудові класифікаторів змінювались у діапазоні (6-8-10-12-14-16-20).

На рис. 10–12 наведені приклади детекції класифікаторами 2–4.

Для кожного з навчених класифікаторів результуючі криві якості наведені на графіках, зображених на рис. 13–15, де на осях абсцис позначені кількості етапів навчання, а по осях ординат – значення відповідного показника.

З рис. 13 видно, що зі збільшенням кількості етапів навчання точність детекції обличчя зменшується. Це відбувається через те, що з кожним наступним етапом збільшується кількість ознак Хаара, які враховуються для класифікатора. Тому

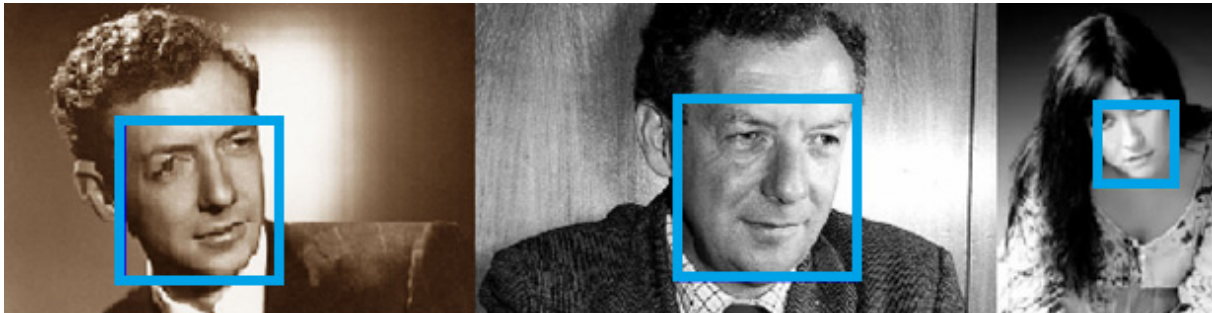


Рис. 10. Приклади детекції класифікатором 2



Рис. 11. Приклади детекції класифікатором 3

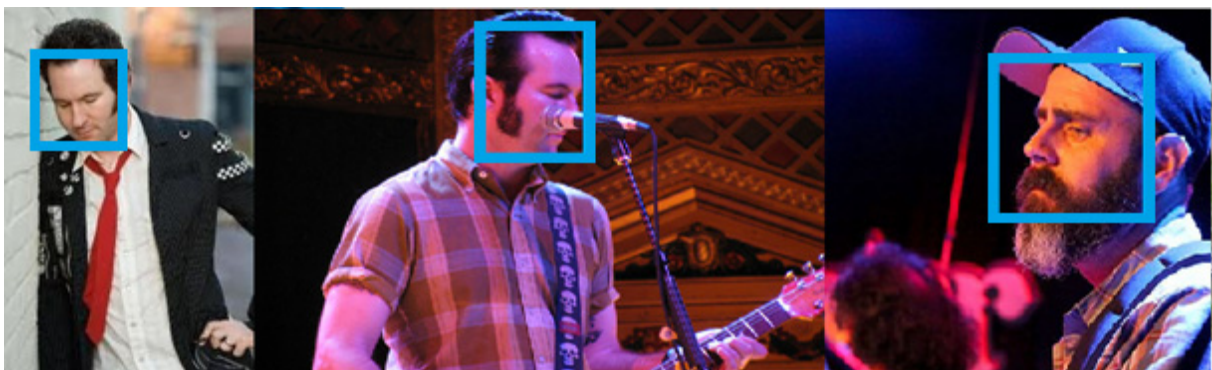


Рис. 12. Приклади детекції класифікатором 4

збільшується число відбракованих зображень, що насправді є обличчями.

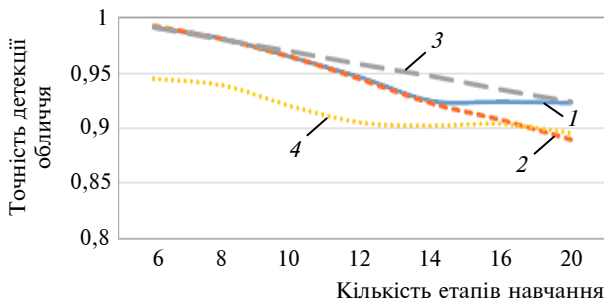


Рис. 13. Графіки точності детекції обличчя різними типами класифікаторів: 1 – класифікатор 1; 2 – класифікатор 2; 3 – класифікатор 3; 4 – класифікатор 4

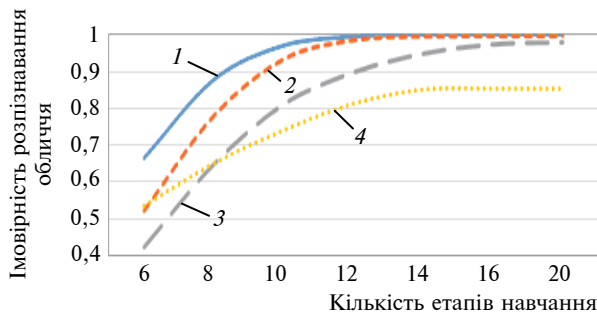


Рис. 14. Ймовірність розпізнавання обличчя: 1 – класифікатор 1; 2 – класифікатор 2; 3 – класифікатор 3; 4 – класифікатор 4

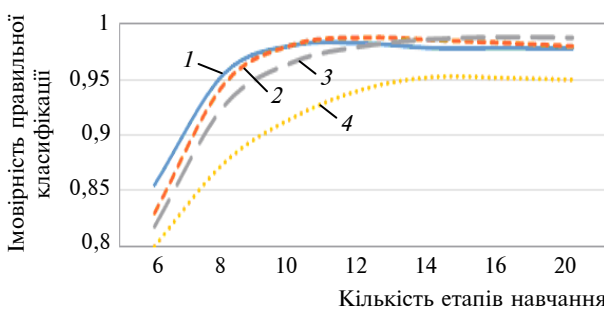


Рис. 15. Ймовірність правильної класифікації зображення: 1 – класифікатор 1; 2 – класифікатор 2; 3 – класифікатор 3; 4 – класифікатор 4

Як видно з графіка на рис. 14, найкращі результати показує фронтальний детектор (№ 1), а найгірші – боковий (№ 4).

Зі збільшенням кількості етапів навчання ймовірність детекції обличчя та ймовірність правильної класифікації зростають і прямують до 1 для фронтального класифікатора та до 0,95 для бокового.

Аналізуючи рис. 13–15, можна помітити, що найвищі показники ймовірності детекції обличчя та ймовірності правильної класифікації має класифікатор 1 – фронтальний, а найнижчі – боковий. Таку розбіжність можна пояснити тим, що якість класифікатора залежить від якості вибірки, тобто чим об'єкт детекції більш стандартний, тим якість його класифікатора буде вищою. У нашому випадку очевидно, що фронтальне обличчя при невеликих поворотах змінюється значно менше, ніж бокове.

Також проведемо аналіз ROC-кривих для класифікаторів за різного розміру вибірки (рис. 16).

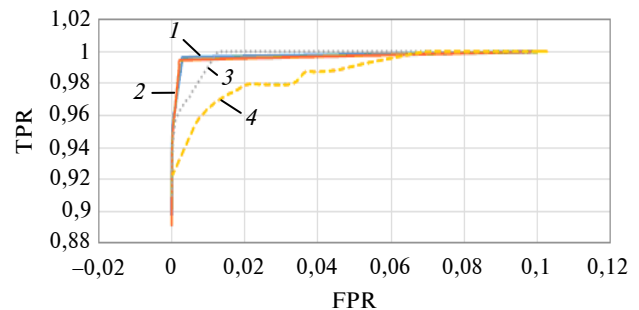


Рис. 16. Графіки ROC-кривих: 1 – класифікатор 1; 2 – класифікатор 2; 3 – класифікатор 3; 4 – класифікатор 4

Можемо помітити, що якість класифікаторів 1 і 2 є вищою порівняно з іншими класифікаторами, якість класифікатора 4 – найнижча. В абсолютних значеннях якість класифікаторів така:

$$AUC_1 = 0,999741959, \quad AUC_2 = 0,9996639,$$

$$AUC_3 = 0,999663526, \quad AUC_4 = 0,996025978.$$

Це означає, що побудовані класифікатори мають високу якість і можуть використовуватися для побудови реальних систем.

Висновки

У статті представлено побудовані детектори обличчя у вигляді каскадів на ознаках Хаара, що навчені на підвибірках зображень із різними діапазонами кутів нахилу голови.

Для дослідження впливу композиції тренувальної вибірки на результуючу якість класифікаторів у роботі виділено чотири типи детекторів обличчя: два фронтальних, а також два бокових, які побудовані на основі методу Віоли–Джонса.

Показано, що фронтальні класифікатори обличчя мають майже однакові параметри якості

та криві навчання, які є кращими за параметри бокових детекторів за тих самих умов навчання. Навчальна вибірка для одного з фронтальних класифікаторів охоплювала набагато ширший діапазон кутів нахилу голови, але можна зазначити, що фронтальні обличчя відрізняються не настільки сильно, як обличчя в профіль, що і зумовлює швидше навчання та вищу якість отриманих фронтальних детекторів.

Встановлено, що ймовірність правильної класифікації обличчя прямує до 100 % за великої кількості рівнів каскаду (більше 10) вже при розмірі вибірки в 20 тисяч обличчя, а при шести рівнях ймовірність становить лише 95 % навіть при розмірі навчальної вибірки в 60 тисяч зображень. У той самий час точність детекції знижується при збільшенні кількості рівнів каскаду.

З порівняння ROC-кривих видно, що боковий класифікатор має найгірші показники –

$AUC = 0,9960$, фронтальний є найкращим – $AUC = 0,9997$. Відсоткове розходження між параметрами AUC фронтального та профільного детекторів становить 0,37 %.

Таким чином, побудовано ряд класифікаторів, серед яких найбільш якісним є класифікатор, побудований на вибірці у 60 тисяч позитивних зображень і кількості негативних зображень у розмірі 60 % від позитивних протягом 20 етапів навчання.

Подальші дослідження у цьому напрямі можуть дати змогу формувати навчальну вибірку для детекторів обличчя із наперед заданими характеристиками якості. Автори планують також розширити тематику досліджень у цьому напрямі, включивши задачі детекції елементів обличчя та їх порівняння між собою. Це дасть змогу розробити системи для більш якісного розпізнавання виразів обличчя та емоцій людини.

Список літератури

1. Modi M., Macwan F. Face Detection Approaches: A Survey [Online]. – Available: https://www.ijirset.com/upload/2014/april/36_Face.pdf
2. Viola P., Jones M.J. Rapid object detection using a boosted cascade of simple features // Proc. IEEE Conf. Comput. Vision Pattern Recog. – 2001. – P. 511–518.
3. Viola P., Jones M.J. Robust real-time face detection // Int. J. Comp. Vision. – 2004. – 57, № 2. – P. 137–154.
4. 60 Facial Recognition Databases [Online]. – Available: <https://www.kairos.com/blog/60-facial-recognition-databases>
5. Bansal A., Nanduri A., Castillo C. UMD Faces: An Annotated Face Data set for Training Deep Networks [Online]. – 2016. – Available: <http://umdfaces.io>
6. Everingham M. The VOC 2005 Database: Test set 2 [Online]. – 2005. – Available: <http://host.robots.ox.ac.uk/pascal/VOC/databases.html#TUD>

References

- [1] M. Modi and F. Macwan. (2014). *Face Detection Approaches: A Survey*. [Online]. Available: https://www.ijirset.com/upload/2014/april/36_Face.pdf
- [2] P. Viola and M.J. Jones, “Rapid object detection using a boosted cascade of simple features”, in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recog.*, 2001, pp. 511–518. doi: 10.1109/CVPR.2001.990517
- [3] P. Viola and M.J. Jones, “Robust real-time face detection”, *Int. J. Comp. Vision.*, vol. 57, no. 2, pp. 137–154, 2004. doi: 10.1023/B:VISI.0000013087.49260.fb
- [4] C. Calistra. (2015, May 7). *60 Facial Recognition Databases* [Online]. Available: <https://www.kairos.com/blog/60-facial-recognition-databases>
- [5] A. Bansal. *UMDFaces: An Annotated Face Data Set for Training Deep Networks* [Online]. Available: <http://umdfaces.io>
- [6] M. Everingham. (2005). *The VOC 2005 Database: Test Set 2* [Online]. Available: <http://host.robots.ox.ac.uk/pascal/VOC/databases.html#TUD>

С.С. Ніколаєв, Ю.О. Тимошенко, К.Ю. Матвіїв

ЗАЛЕЖНІСТЬ ЯКОСТІ ДЕТЕКТОРА ОБЛИЧЧЯ НА ОЗНАКАХ ХААРА ВІД ВАРІАТИВНОСТІ НАВЧАЛЬНОЇ ВИБІРКИ

Проблематика. При навчанні узагальнених детекторів обличчя на основі каскадів з ознаками Хаара виникає проблема довгого навчання результатуючих каскадів та їх низької якості. Тому на практиці прийнято тренувати окремо фронтальні та профільні детектори обличчя, що ускладнює системи розпізнавання.

Мета дослідження. Необхідно порівняти вплив композиції навчальної вибірки на якість натренованих детекторів за різних кутів нахилу обличчя.

Методика реалізації. Пропонується натренувати ряд детекторів облич на підвбірках, що покривають різні діапазони кутів нахилу обличчя. При цьому всі інші параметри навчання будуть фіксовані. У результаті порівнюються час навчання та якість отриманих каскадів.

Результати дослідження. Оцінено якість і час навчання класифікаторів облич залежно від композиції вибірки, а також здійснено порівняння якості фронтальних і бокових класифікаторів за однакової величини тренувальної вибірки. Показано, що метрика *AUC* має розходження у 0,003 між фронтальним та профільним детекторами.

Висновки. Експериментально показано, що чим більш варіативний об'єкт (обличчя повернуте профілем відносно фронтального положення), тим довше і гірше навчається каскад Хаара за однакового розміру навчальних вибірок. Використовуючи запропонований підхід, можна керувати якістю фінального класифікатора підбором відповідної композиції тренувальної вибірки.

Ключові слова: детектор облич; ознаки Хаара; тренувальна вибірка; бустинг; композиція навчальної вибірки.

С.С. Николаев, Ю.А. Тимошенко, Е.Ю. Матвиив

ЗАВИСИМОСТЬ КАЧЕСТВА ДЕТЕКТОРА ЛИЦ НА ПРИЗНАКАХ ХААРА ОТ ВАРИАТИВНОСТИ ОБУЧАЮЩЕЙ ВЫБОРКИ

Проблематика. При обучении обобщенных детекторов лиц на каскадах с признаками Хаара возникает проблема длительного обучения результирующих каскадов и их низкого качества. Поэтому на практике принято тренировать отдельно фронтальные и профильные детекторы лиц, что усложняет системы распознавания.

Цель исследования. Необходимо сравнить влияние композиции обучающей выборки на качество натренированных детекторов при различных углах наклона лица.

Методика реализации. Предлагается натренировать ряд детекторов лиц на подвыборках, которые покрывают различные диапазоны углов наклона лица. При этом все остальные параметры обучения будут фиксированы. В результате сравниваются время обучения и качество полученных каскадов.

Результаты исследования. Оценены качество и время обучения классификаторов лиц в зависимости от композиции выборки, а также проведено сравнение качества фронтальных и боковых классификаторов при одинаковой величине тренировочной выборки. Показано, что метрика *AUC* имеет расхождение в 0,003 между фронтальным и профильным детекторами.

Выводы. Экспериментально показано, что чем более варіативний об'єкт (лицо в профіль відносно фронтального положення), тем довше і гірше навчається каскад Хаара при однаковому розмірі навчаючих вибірок. Існуючи запропонований підхід, можна керувати якістю фінального класифікатора шляхом підбору відповідної композиції навчаючої вибірки.

Ключевые слова: детектор лиц; признаки Хаара; тренировочная выборка; бустинг; композиция обучающей выборки.

Рекомендована Радою
Навчально-наукового комплексу "Інститут
прикладного системного аналізу"
КПІ ім. Ігоря Сікорського

Надійшла до редакції
21 вересня 2017 року