



МЕТОДИ РОЗПІЗНАВАННЯ ОБЛИЧЧЯ У СИСТЕМАХ ВІДЕОСПОСТЕРЕЖЕННЯ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ

В. Мрак ^[ORCID: 0009-0002-6066-5592]

Національний університет “Львівська політехніка”, вул. С. Бандери, 12, Львів, 79013, Україна

Відповідальний за рукопис: Мрак Василь (e-mail: vasyi.b.mrak.@lpnu.ua).

(Подано 1 Серпня 2023)

Стаття присвячена дослідженню методів розпізнавання облич і спрямована на визначення найбільш підходящого для системи безпеки на основі розпізнавання облич з камер відеоспостереження. Проаналізовано часові витрати цих методів та їх стійкість до спотворень геометричного масштабу та поворотів у різних площинах. Для експериментів згенеровано власні набори даних. Особливу увагу приділено компромісу між швидкістю та точністю розглянутих методів для їх використання як першого етапу системи безпеки на основі розпізнавання обличчя у відеопотоці. Проведені дослідження показали, що найефективнішими виявилися методи RetinaFace-MobileNet0.25, FaceBoxes, SCRFD500MF, CenterFace; RetinaFaceResNet125, DSFD, RetinaFaceMobileNet0.25 які стійкі до повороту обличчя. Крім того, при виборі найбільш оптимального методу розпізнавання обличчя для застосування в системі безпеки враховувалася наявність інформативних параметрів обличчя, а також той факт, що методи розпізнавання, які використовуються на наступному етапі, мають свої обмеження щодо стійкості до афінних перетворень.

Ключові слова: розпізнавання обличчя; MTCNN; FaceBoxes; DSFD; RetinaFace; CenterFace; SCRFD; система безпеки; відеоспостереження.

УДК: 621.126

1. Методи виявлення обличчя та їх аналіз

Дослідження методів виявлення облич спрямоване на вибір найбільш оптимального алгоритму та архітектури нейронної мережі для системи безпеки підприємства на основі розпізнавання осіб з камер відеоспостереження. Відповідно до вимог системи безпеки особливу увагу було приділено компромісу між швидкістю та надійністю аналізованих методів для їх використання як першого етапу системи безпеки. Після дослідження джерел для аналізу були обрані наступні навчені моделі нейронних мереж (табл. 1):

- Багатозадачні каскадні згорткові мережі (MTCNN) [1];
- FaceBox [2];
- Подвійний детектор обличчя (DSFD) [3];
- RetinaFace [4];
- CenterFace [5];
- Одноступінчастий каскадний детектор залишкового обличчя (SCRFD) [6].

У табл. 1 подано короткий опис досліджуваних моделей. RetinaFace розглядався у двох варіантах з використанням Resnet50 і Mobilenet0.25. Для мережі SCRFD обрано модель SCRFD-500MF, яка потребує 500 мегафлопсів для вхідного зображення з роздільною здатністю VGA (640×480).

Кожна з нейронних мереж має певні характеристики, які наведено у роботах [1]–[6]. MTCNN використовує каскад із трьох згорткових нейронних мереж для виявлення обличчя на зображеннях. Він особливо добре виявляє обличчя різних розмірів та орієнтації [1]. Головною особливістю моделі FaceBoxes є здатність розпізнавати обличчя дуже швидко. Ця здатність моделі робить її ідеальною для додатків, що потребують високої швидкості обробки, таких як системи безпеки [2]. DSFD – це модель, яка використовує два етапи для розпізнавання обличчя. Вона має дві гілки згорткових нейронних мереж, кожна з яких ідентифікує області з різним масштабом і виявляє обличчя у відповідних областях. Такий підхід дає цій моделі змогу виявляти обличчя різних розмірів та під довільним кутом на зображеннях [3]. Головна характеристика моделі RetinaFace ResNet125 полягає у тому, що вона використовує дві згорткові нейронні мережі: RetinaFace для виявлення обличчя і ResNet125 для виділення ознак. Завдяки цьому вона дає змогу виявляти обличчя дуже точно та визначає детальнішу інформацію про обличчя на зображеннях. RetinaFace MobileNet0.25 – це модель, подібна до RetinaFaceResNet125, але замість ResNet125 вона використовує простішу мережу MobileNet0.25, тому її доцільніше використовувати на мобільних пристроях [4]. CenterFace – це модель, яка використовує центральну точку обличчя для визначення його положення на зображенні. Вона здатна виявляти маленькі та повернуті обличчя на зображеннях [5]. SCRFD – це модель, яка застосовує кілька згорткових шарів для виявлення обличчя на зображеннях. Характеризується високою швидкістю обробки та може виявляти обличчя різних розмірів на зображеннях [6].

Таблиця 1

Коротка інформація про моделі виявлення

Назва методу	Рік створення	Наявність орієнтирів	СТ у наборі перевірки для WIDER FACE (Hard)
MTCNN	2016	+	0,598[1]
FaceBoxes	2018	–	0,395 [1]
DSFD	2018	–	0,904 [1]
RetinaFaceResNet125	2019	+	0,918 [1,5]
RetinaFaceMobileNet0.25	2019	+	0,78 [5]
CenterFace	2020	+	0,875 [1]
SCRFD-500MF	2021	–	0,685 [2, 3]

Середня точність (СТ) часто використовується для порівняння методів виявлення. СТ – це площа під кривою точності запам'ятовування, яку можна розрахувати як наближення кривою точності запам'ятовування із прямокутниками [7]

У цьому дослідженні СТ ми не розраховували, її взято із публікацій і репозиторіїв, де моделі є у відкритому доступі: [5] для MTCNN, FaceBoxes, DSFD, CenterFace; [4], [5] для RetinaFace; [6], [8] для SCRFD-500MF. Для всіх моделей СТ розрахована за набором даних перевірки WIDER FACE (Hard) [9].

Усі ці мережі можуть працювати з обличчями різного розміру та змінними кутами. Дослідження, які б описували діапазон значень масштабного коефіцієнта та кутів повороту обличчя, в межах яких точність методів заявлена, наведено в [17].

Крім того, швидкість моделей, поданих в джерелах, була виміряна на різних пристроях і для різних розмірів зображення, що унеможливило порівняння часових витрат цих моделей. Наприклад, у [1] швидкість методу MTCNN становила 16 FPS (2,60 ГГц CPU) і 99 FPS (Nvidia Titan Black GPU); швидкість FaceBox 20 FPS (Intel Xeon E5-2660v3@2.60GHz) [12]; модель RetinaFace

ResNet125 становила 13 FPS для зображень VGA (640×480 пікселів), RetinaFaceMobileNet0.25 – 40 FPS (зображення 4K (4096×2160 пікселів) на графічному процесорі NVIDIA Tesla P40), 20 FPS (HD-зображення (1920×1080 пікселів)) на процесорі Intel i7-6700K), 16 FPS (зображення VGA на ARM-RK3399) [4]. У роботі [5] наведено порівняння швидкості роботи моделей DSFD та CenterFace на NVIDIA GTX2080TI. Отримано такі результати: 78,08 мс та 5,51 мс для зображень із роздільною здатністю VGA відповідно. Також показано, що модель CenterFace демонструє швидкість 30 FPS на CPU I7-6700@2.6 із найбільшою точністю. Документи [6] визначають, що SCRFD0.5GF перевершує RetinaFaceMobileNet0.25 на 45,57 % у часі. Швидкість методу DSFD дуже низька, його не можна використовувати без GPU.

Наведена інформація дає підстави зробити висновок, що модель SCRFD-0.5GF можна назвати найшвидшою з розглянутих моделей, а DSFD, очевидно, найповільнішою. Проте питання швидкості потребує подальшого дослідження. Необхідно вимірювати швидкість для моделей MTCNN, FaceBoxes, DSFD, RetinaFaceResNet125, RetinaFaceMobileNet0.25, CenterFace і SCRFD-500MF на одному пристрої для зображень однакового розміру.

2. Мета та зміст дослідів

Оскільки мета цього дослідження – вибір найефективнішого методу розпізнавання, який буде використовуватися на першому етапі системи розпізнавання осіб, то його результат істотно вплине на кінцевий результат розпізнавання, тому ми дослідили такі параметри методів виявлення:

- стійкість до поворотів у різних площинах (кут зміни навколо вертикальної та горизонтальної осей від –90 до 90 градусів);
- стійкість до спотворень геометричного масштабу (зміна розміру зображень обличчя від 20×20 до 310×310);
- витрати часу (для визначення обличчя у двох розмірах 640×480 (VGA) пікселів та 1280×720 (HD) пікселів).

В експериментах використано нейронні моделі із загальнодоступних репозиторіїв: MTCNN [10], FaceBoxes [11], DSFD [12], RetinaFace [12], CenterFace [13], SCRFD-500MF [14]. Значення точності ідентифікації облич, отримані цими моделями, використано для оцінки якості виявлення. Експерименти виконано на наборах даних, створених за допомогою редакторів Generated [15] і Character Creator [16]. Часові затрати виміряно на відеокарті NVIDIA GeForce GTX 780 на шині PciExpress 2.0.

3. Дослідження залежності точності розпізнавання від повороту обличчя

Для порівняння методів виявлення обличчя за різних кутів повороту голови людини створено власний набір даних, у якому людські обличчя згенеровано за допомогою Generated Photos [15]. На основі цих зображень створені 3D-моделі. Приклади зображень зі згенерованого набору даних показано на рис. 1.



Рис. 1. Приклади зображень, створених за допомогою редактора *Generated Photos*

Цю 3D-модель відрендерено в графічному редакторі Character Creator [16]) з кроком обертання на один градус по двох осях окремо зліва направо від -90 градусів до $+90$ градусів (обертання навколо осі X) і кут підйому знизу вгору від -90 градусів до $+90$ градусів (обертання навколо осі Y). Це означає, що для кожної 3D-моделі створено 181 зображення. На рис. 2 подано приклади зображень із синтезованого набору даних отриманих 3D-моделей.

На рис. 3 і 4 наведено значення середньої точності за обертання навколо осі Y та X відповідно.



Рис. 2. Приклади зображень зі штучного набору даних, створених за допомогою редактора Character Creator: поворот ліворуч навколо осі Y на -50 градусів (рух вгору – вниз); вправо – поворот навколо осі X на -20 градусів (рух вліво – вправо)

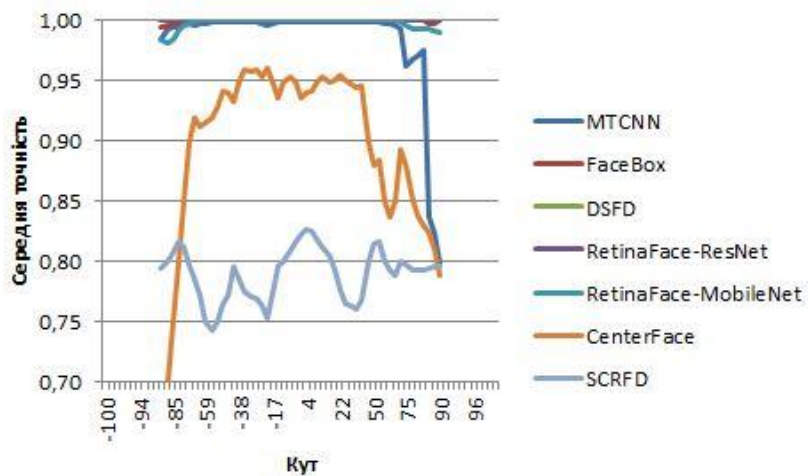


Рис. 3. Залежність точності ідентифікації обличчя від кута повороту навколо осі Y (рух вліво – вправо)

Таблиця 2

Експериментальні результати щодо точності виявлення в умовах просторового повороту

Метод	Діапазон кутів повороту навколо осі Y (рух вліво – вправо), середня точність $\geq 0,9$	Діапазон кутів повороту навколо осі X (рух вгору – вниз), середня точність $\geq 0,9$
MTCNN	$[-89;83]$ (range=172)	$[-31;47]$ (range=78)
FaceBoxes	$[-90;90]$ (range=180)	$[-62;40]$ (range=102)
DSFD	$[-90;90]$ (range=180)	$[-69;50]$ (range=119)
RetinaFaceResNet125	$[-90;90]$ (range=180)	$[-65;62]$ (range=127)
RetinaFaceMobileNet0.25	$[-90;90]$ (range=180)	$[-60;55]$ (range=115)
CenterFace	$[-71;46]$ (range=117)	$[-42;37]$ (range=79)
SCRFD0.5GF	\emptyset	\emptyset

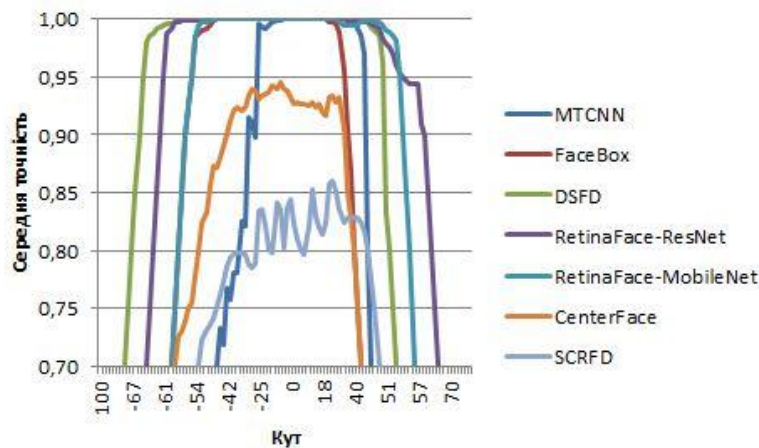


Рис. 4. Залежність точності ідентифікації обличчя від кута повороту навколо осі X (рух вгору – вниз)

У випадку поворотів навколо осі Y (рух ліворуч – праворуч) середнє значення точності перевищує 0,9 і максимально можливий діапазон змін кута $[-90; 90]$ у FaceBoxes, DSFD, моделі RetinaFaceResNet125, RetinaFaceMobileNet0.25. MTCNN виявляє із середнім значенням точності понад 0,9 у діапазоні $[-89; 83]$. Наступна – модель CenterFace з діапазоном, значно меншим за $[-71; 46]$ і середньою точністю не вище ніж 0,97. Найнестабільнішою виявилася модель SCRFD-500MF, середня точність якої змінювалася синусоїдально на 0,7 значення і не перевищувала значень 0,83. Крім того, FaceBoxes, DSFD і RetinaFaceResNet125 мають середнє значення точності, близьке до 0,99 у всьому діапазоні. У моделей RetinaFaceMobileNet0.25 і MTCNN середня точність близько 0,99 у діапазоні $[-75; 75]$ і $[-77; 70]$ відповідно.

Що стосується обертання навколо осі X (рух вгору – вниз), то моделі FaceBoxes, DSFD, RetinaFaceResNet125 і RetinaFaceMobileNet0.25 також показали кращі результати. Однак навколо осі X діапазон значень кута повороту був набагато меншим порівняно з віссю Y. В експериментах із середньою вірогідністю вище за 0,9 найкращими моделями були RetinaFaceResNet125, DSFD, RetinaFaceMobileNet0.25 з діапазонами $[-65; 62]$, $[-69; 50]$, $[-60; 55]$ відповідно, наступні – FaceBox з діапазоном $[-62; 40]$. Моделі MTCNN і CenterFace мають подібний діапазон $[-31; 47]$ і $[-42; 37]$ відповідно, але середня точність для CenterFace не перевищує 0,95. Середнє довірче значення SCRFD500MF змінюється нестабільно, графік синусоїдальний і не перевищує 0,87. Також можна зауважити, що всі розглянуті моделі, крім CenterFace, SCRFD-500MF, досягають середніх значень точності, близьких до 0,99 (DSFD, RetinaFaceResNet125, RetinaFaceMobileNet0,25 в діапазоні $[-50; 47]$, FaceBoxes – $[-48; 25]$, MTCNN – $[-20; 40]$).

Загалом, за точністю виявлення у разі повороту обличчя розглянуті методи можна ранжувати та: RetinaFaceResNet125, DSFD > RetinaFaceMobileNet0.25 > FaceBoxes > MTCNN >> CenterFace >> SCRFD-500MF, де “>” означає, що метод ліворуч надійніший за той, що праворуч, “>>” означає, що відмінність між надійністю методів істотна.

4. Залежність точності виявлення від розміру обличчя

Сформовано додатковий штучний набір даних для вивчення точності методів розпізнавання обличчя залежно від розміру обличчя. Спочатку було створено 10 облич за допомогою редактора Generated Photos. На наступному кроці для кожного обличчя були згенеровані тестові зображення, щоб подати їх у різних розмірах. Розмір обличчя змінювали з кроком 10 пікселів. Найменший розмір – 20×20 пікселів, найбільший – 310×310 пікселів.

Для моделей виявлення, які використовують немаксимальне придушення (NMS) для об'єднання результатів, для NMS було встановлено поріг 0,3. Зображення передавали в методи без додаткової нормалізації або зміни розміру.

На рис. 5 і 6 подано кілька прикладів результатів виявлення зображень зі штучного набору даних. Виявлені обличчя оточені прямокутником, колір якого залежить від значення точності, яке повертає модель виявлення, і значення точності відображається поруч із ним у вигляді цифр.

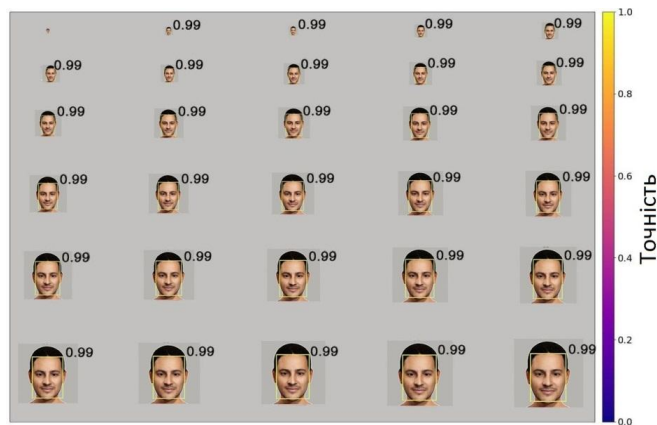


Рис. 5. Залежність точності для MTCNN від розміру зображення, який змінюється від 20×20 до 310×310 (px)

Середні значення точності розраховали для кожного розміру обличчя на основі значень точності для всіх експериментів. На рис. 7 наведено залежність середньої точності від зміни розміру обличчя.



Рис. 6. Залежність точності для CenterFace від розміру зображення, який варіюється від 20×20 до 310×310 (px)

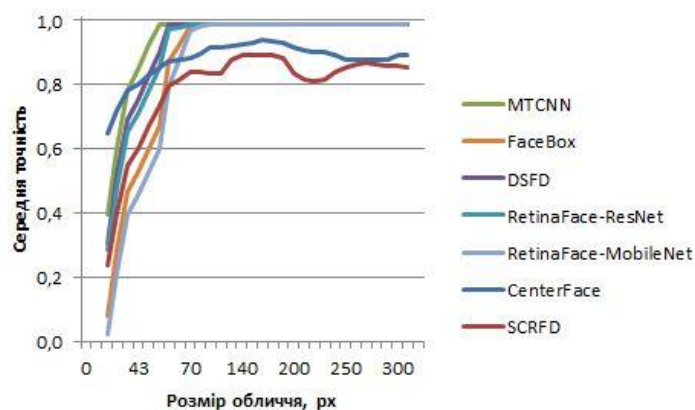


Рис. 7. Залежність середніх значень точності від зміни розміру зображення від 20×20 до 310×310 (px)

Можна зробити висновок, що моделі MTCNN, FaceBoxes, DSFD, RetinaFaceRes-Net125, RetinaFaceMobileNet0.25 дуже стабільно визначають обличчя розміром від 75×75 до 310×310 (px) (середня точність $\geq 0,99$). Крім того, модель MTCNN дає середню точність $\geq 0,99$, починаючи з розміру обличчя 50×50 (px), а в деяких випадках вона виявляла обличчя 30×30 (px) з точністю 0,99. Це означає, що MTCNN виявляє найменший розмір обличчя з високою точністю, часто навіть 30×30 (px), як показано на рис. 6, мають середню точність $\geq 0,99$. Для DSFD, RetinaFace-ResNet125, розмір першого обличчя має бути принаймні 60×60 (px), щоб отримати середню точність $\geq 0,99$. Моделі FaceBoxes, RetinaFace-MobileNet0.25 мають середню точність $\geq 0,99$ для розмірів зображення, починаючи з 70×70 (px). Середня точність моделі CenterFace не перевищує 0,95 у всьому діапазоні, але демонструє найвищу середню точність 0,62 для дуже маленьких зображень 20×20 (px). Для порівняння: моделі MTCNN, RetinaFaceResNet125, FaceBoxes, DSFD для зображень такого розміру дають набагато нижчу середню точність. Середня точність CenterFace вища за 0,90, лише починаючи з 80×80 (px). Однак цей метод зміг знайти найменший розмір 20×20 (px) і характеризується меншим розкидом ймовірностей зі зміною масштабу. Модель SCRFD0.5GF виявляє зображення, починаючи з 30×30 (px), але середня точність не перевищує 0,9, досягаючи 0,9 лише в невеликому діапазоні розмірів від 140×140 до 170×170 (px). Крім того, під час експериментів модель показала нестабільну роботу.

У табл. 3 наведено результати ідентифікації у разі зміни мінімального розміру обличчя із середньою точністю більше ніж 0,9 та 0,99.

Таблиця 3

Результати досліджень методів виявлення у разі зміни розмірів обличчя

Метод	Мінімальний розмір обличчя (px), середня точність $\geq 0,9$	Мінімальний розмір обличчя (px), середня точність $\geq 0,99$
MTCNN	45×45	50×50
FaceBoxes	60×60	70×70
DSFD	50×50	60×60
RetinaFaceResNet125	53×53	60×60
RetinaFaceMobileNet0.25	63×63	70×70
CenterFace	80×80	∅
SCRFD0.5GF	140×140	∅

Підсумовуючи, можна сказати, що моделі можна ранжувати за найбільшим діапазоном зміни розміру обличчя із середньою точністю 0,99 або вище, в такій послідовності: MTCNN>DSFD, RetinaFace-ResNet> FaceBoxes, RetinaFace-MobileNet0.25>> CenterFace> SCRFD-500MF.

5. Порівняння витрат часу на ідентифікацію

Методи порівнювали стосовно затрат часу на двох зображеннях з різною роздільною здатністю: VGA (640×480 пікселів) і HD (1280×720 пікселів). Це пов'язано з тим, що ці розміри є звичайними для камер, і, як показано в попередніх тестах, виявлення можливе на обличчях розміром лише 20×20 (px). Тому необхідно досліджувати продуктивність мереж за різних умов, оскільки можна зменшити або збільшити роздільну здатність кадрів і, отже, змінити відповідно витрати часу. Всі вимірювання виконували з попереднім "розігріванням" мережі, пропускаючи перше виявлення, на відеокарті NVIDIA GeForce GTX 780.

Результати вимірювання швидкості розпізнавання обличчя для 100 зображень наведено на діаграмах, де по осі X відкладено значення часу (мс), а по осі Y – нормалізована кількість кадрів з певним часом опрацювання кадру (рис. 8, 9).

На рис. 9 наведено графік, що відображає основну статистику вимірювань часу, таких як медіани, перший 0,25 % і третій 0,75 % квартилі та викиди, для всіх методів із зображеннями 640×480 (px). Статистичні моделі для зображень 1280×720 подібні. Однак метод DSFD дає ще більший розкид, а метод RetinaFace-ResNet125 – 26-кратне збільшення розсіювання (рис. 10).

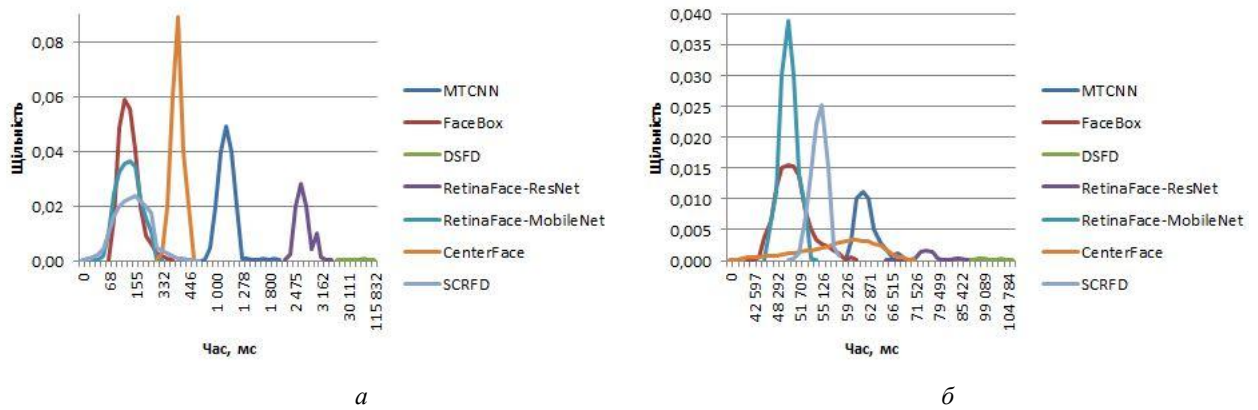


Рис. 8. Щільність розподілу часу оброблення кадру для зображень а) VGA; б) HD

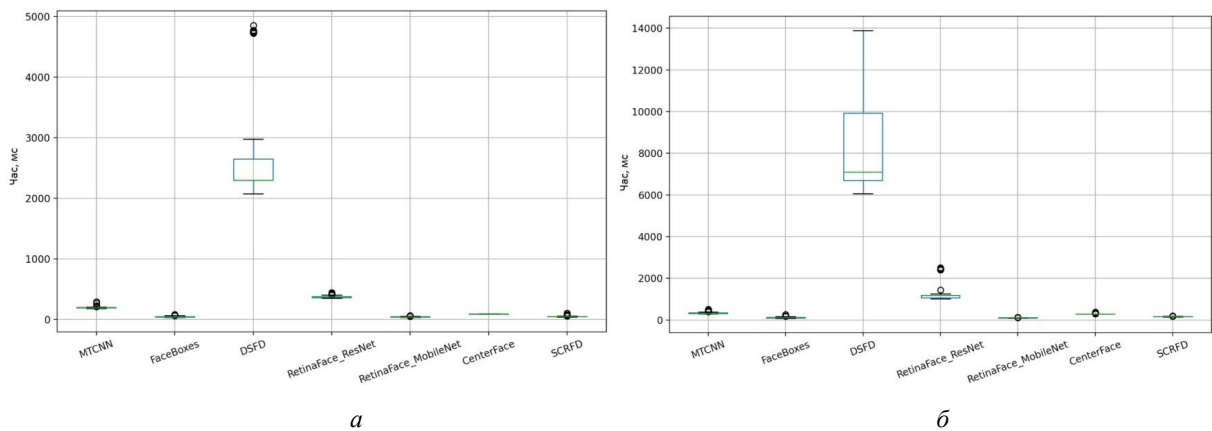


Рис. 9. Статистика часу оброблення кадру для зображень у форматі: а – VGA; б – HD

Таблиця 4

Значення середнього часу, медіани та стандартне відхилення часу виявлення обличчя

Метод	Середній час обробки кадру, мс 640×480	Медіана часу обробки кадру, мс 640×480	Стандартне відхилення часу обробки кадру, мс 640×480	Середній час обробки кадру, мс 1280×720	Медіана часу обробки кадру, мс 1280×720	Стандартне відхилення часу обробки кадру, мс 1280×720
MTCNN	193	191	12	342	315	56
FaceBoxes	42	40	8	112	105	31
DSFD	2863	2292	1005	8754	7090	2802
RetinaFace-ResNet	375	370	17	1373	1175	493
RetinaFace-MobileNet	40	38	4	102	101	4
CenterFace	86	86	1	276	272	15
SCRFD-500MF	43	42	7	155	153	7

На основі рис. 7–10 і даних табл. 4 можна зробити висновок, що найшвидші моделі – RetinaFaceMobileNet0.25, FaceBoxes і SCRFD-500MF, оскільки у них подібний середній час обробки кадрів, який становить 42 мс, 44 мс і 46 мс для зображень 640×480 пікселів і 103 мс, 115 мс і 157 мс для зображень 1280×720 пікселів відповідно. Крім того, зміна середнього часу оброблення кадру для цих моделей дуже мала від експерименту до експерименту. Далі йде модель CenterFace, із часом виявлення 89 мс і 274 мс для зображень 640×480 пікселів і 1280×720 пікселів відповідно. Цей

метод у 2,1 разу швидший за RetinaFace-MobileNet0.25 для 640×480 і в 2,7 разу швидший для 1280×720. Виявлення моделі MTCNN у 2,2 разу швидше, ніж CenterFace для зображень 640×480, але лише в 1,2 разу швидше для зображень 1280×720. З RetinaFace-ResNet125 розпізнавання обличчя займає ще більше часу. Ця мережа вже працює в дев'ять разів довше, ніж RetinaFaceMobileNet0.25 для 640×480, і в 11,5 разу довше для 1280×720. Метод DSFD істотно відрізняється від інших, він визначає к 68 разів повільніше, ніж RetinaFace-MobileNet0.25 для 640×480 пікселів і в 85 разів повільніше для 1280×720 пікселів. Крім того, DSFD притаманний великий розкид порівняно з іншими методами, який становить 1007 мс (640×480). Наприклад, для методу CenterFace розкид становить 1 мс, RetinaFaceResNet125 – 19 мс.

На підставі виконаних експериментів розглянуті методи можна ранжувати за часовими витратами так: RetinaFace-MobileNet0.25, FaceBoxes, SCRFD-500MF > CenterFace > MTCNN > RetinaFaceResNet125 >>> DSFD.

Висновки

Практичне значення отриманих результатів полягає у дослідженні моделей MTCNN, FaceBoxes, DSFD, RetinaFaceResNet125, RetinaFaceMobileNet0.25, CenterFace, SCRFD-500MF для подальшого використання оптимальної для системи безпеки на основі розпізнавання обличчя з відеопотоків камер спостереження.

Здійснено експерименти щодо стійкості моделей до повороту обличчя у різних площинах, зміни розміру обличчя, а також оцінено витрати часу.

Найкращими моделями за обертанням стали RetinaFaceResNet125, DSFD, RetinaFaceMobileNet0.25, які точно (точність $\geq 0,9$) визначають обличчя із поворотами в діапазоні $[-45; 45]$, що є достатньою вимогою для використання у системі безпеки. Моделі MTCNN, FaceBoxes, CenterFace і SCRFD-500MF також працюють із поворотом обличчя в різних площинах, але у них нижча точність і менший діапазон кута обличчя.

В експериментах з різними розмірами обличчя найкращі результати показали моделі MTCNN, DSFD, RetinaFace-ResNet, FaceBoxes, RetinaFace-MobileNet0.25, які виявляють зображення, починаючи з розміру 75×75 px, з точністю $\geq 0,9$. Моделі CenterFace і SCRFD0.5GF мають значно нижчу точність, але в деяких випадках виявляють дуже маленькі обличчя.

Вимірювання часу виявлення показало, що найшвидшими є RetinaFace-MobileNet0.25 і FaceBox, які витрачають менше ніж 47 мс на опрацювання одного зображень VGA і 100 мс для розміру HG. На наступному місці SCRFD-500MF, CenterFace. Для моделей MTCNN і RetinaFaceResNet125 час виявлення перевищує 100 мс навіть для зображень VGA. Модель DSFD не можна застосовувати в реальному часі навіть для зображень VGA.

Список використаних джерел

- [1] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10), 1499–1503. DOI: 10.1109/lsp.2016.2603342.
- [2] Zhang, S., Zhu, X., Lei, Z., Shi, H., Wang, X., & Li, S. Z. (2017). FaceBoxes: A CPU real-time face detector with high accuracy. 2017 *IEEE International Joint Conference on Biometrics (IJCB)*. DOI: 10.1109/bias.2017.8272675.
- [3] Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137–1149. <https://doi.org/10.1109/tpami.2016.2577031>.
- [4] Deng, J., Guo, J., Verweras, E., Kotsia, I., & Zafeiriou, S. (2020). Retinaface: Single-shot multilevel face localisation in the wild. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. <https://doi.org/10.1109/cvpr42600.2020.00525>
- [5] Xu, Y., Yan, W., Yang, G., Luo, J., Li, T., & He, J. (2020). Centerface: Joint face detection and alignment using face as point. *Scientific Programming*, 2020, 1–8. <https://doi.org/10.1155/2020/7845384>

- [6] Guo, J., Deng, J., Lattas, A., & Zafeiriou, S. (2021, May 10). *Sample and computation redistribution for efficient face detection*. *arXiv.org*. Retrieved July 15, 2023, from <https://arxiv.org/abs/2105.04714>
- [7] Zhang, E., & Zhang, Y. (2009). *Average precision*. *Encyclopedia of Database Systems*, 192–193. https://doi.org/10.1007/978-0-387-39940-9_482
- [8] Deepinsight. (n. d.). *Insightface/model_zoo at master • deepinsight/insightface*. GitHub. Retrieved July 15, 2023, from https://github.com/deepinsight/insightface/tree/master/model_zoo
- [9] Yang, S., Luo, P., Loy, C. C., & Tang, X. (2016). *Wider face: A face detection benchmark*. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). <https://doi.org/10.1109/cvpr.2016.596>
- [10] Timesler (n. d.). *Timesler/facenet-pytorch: Pretrained Pytorch Face Detection (MTCNN) and facial recognition (InceptionResnet) models*. GitHub. Retrieved July 15, 2023, from <https://github.com/timesler/facenet-pytorch>
- [11] Zisianw (n. d.). *Zisianw/faceboxes.pytorch: A pytorch implementation of faceboxes*. GitHub. Retrieved July 15, 2023, from <https://github.com/zisianw/FaceBoxes.PyTorch>
- [12] Hukkelas (n. d.). *Hukkelas/DSFD-pytorch-inference: A high-performance pytorch implementation of face detection models, including RetinaFace and DSFD*. GitHub. Retrieved July 15, 2023, from <https://github.com/hukkelas/DSFD-Pytorch-Inference>
- [13] Star-Clouds (n. d.). *Star-Clouds/Centerface: Face detection*. GitHub. Retrieved July 15, 2023, from <https://github.com/Star-Clouds/CenterFace>
- [14] Deepinsight (n. d.). *Insightface/python-package at master • deepinsight/insightface*. GitHub. Retrieved July 15, 2023, from <https://github.com/deepinsight/insightface/tree/master/pythonpackage>
- [15] Unique, worry-free model photos. *Generated Photos* (n. d.). Retrieved July 15, 2023, from <https://generated.photos/>
- [16] *Character Creator (CC) is a full character creation solution for designers to easily generate, import and customize stylized or realistic character*. Retrieved July 10, 2023, from <https://www.reallusion.com/character-creator/>
- [17] Olena Yakovleva, Andrii Kovtunencko, Valentyn Liubchenko, Vadym Honcharenko and Oleg Kobylin. *Face Detection for Video Surveillance-based Security System COLINS-2023: 7th International Conference on Computational Linguistics and Intelligent Systems, April 20–21, 2023, Kharkiv, Ukraine, 69–86*.

FACE RECOGNITION METHODS IN VIDEO SURVEILLANCE SYSTEMS USING MACHINE LEARNING

Vasyl Mrak

Lviv Polytechnic National University, 12, S. Bandery str., Lviv, 79013, Ukraine

The article is dedicated to the investigation of face identification methods and aims to determine the most suitable one for a security system based on facial recognition from surveillance cameras. The time costs of these methods and their robustness against geometric scale distortions and rotations in various planes have been analyzed. Custom datasets have been generated for experimentation purposes. Special attention has been given to striking a balance between the speed and accuracy of the examined methods for their utilization as the initial stage of a security system based on facial recognition in a video stream. The conducted research has revealed that the most effective methods are RetinaFace-MobileNet0.25, FaceBoxes, SCRFD500MF, and CenterFace; RetinaFaceResNet125, DSFD, and RetinaFaceMobile0.25 which are resilient to facial rotations. Furthermore, when selecting the most optimal facial recognition method for application within a security system, the presence of informative facial parameters was taken into account, as well as the fact that the recognition methods used in the subsequent stage have their limitations concerning resilience to affine transformations.

Key words: *face detection; embeddings; MTCNN; FaceBoxes; DSFD; RetinaFace; CenterFace; SCRFD; security system; video surveillance.*