

ВДОСКОНАЛЕНИЙ НЕЙРОМЕРЕЖЕВИЙ ПІДХІД ДЛЯ ШВИДКОГО ВИЯВЛЕННЯ ОБЛИЧ

© Палій І., 2007

Описано підхід до виявлення облич за допомогою сегментації за кольором шкіри, Хаар-подібних рис та нейронних мереж. На етапах виявлення облич-кандидатів знаходяться області вхідного зображення, які можуть містити обличчя. Потім каскад нейронних мереж з різними локальними рецепторними полями класифікує ці області з використанням покращеної стратегії пошуку облич за масштабом та розміщенням.

The paper describes the face detection approach using skin color segmentation, Haar-like features and neural networks. The face candidate selection stages label promising input image areas that may contain faces. Then, the cascade of the neural networks with different local receptive fields performs the final classification of these image windows using improved face search strategy across scale and position.

Вступ. Виявлення людських облич (ВО) є дуже важливою галуззю досліджень, яка стрімко розвивається і має багато сфер застосування: розпізнавання облич, відеоконференції, індексація баз даних зображень, відеоспостереження тощо. ВО є складною задачею завдяки таким факторам, як масштаб, розташування, орієнтація, ступінь повороту та умови освітлення облич. За минулі роки було запропоновано ряд підходів до ВО, що ґрунтуються на знаннях, на інваріантних рисах, на порівнянні з шаблоном та на зовнішньому вигляді [1].

Останні методи ВО, що ґрунтуються на зовнішньому вигляді, показують відмінні результати на еталонних тестових наборах із зображеннями різноманітних облич на складному фоні. Зокрема, Сан та Поджіо розробили розподілений підхід до ВО, який був першим точним методом на основі зовнішнього вигляду [2]. Тренувальні набори отримані шляхом створення віртуальних прикладів та алгоритму bootstrapping. Всі зображення облич та необлич нормалізується та групуються у кластери. Класифікація здійснюється багаторівневим перцептроном на основі відстаней від вхідного зображення до кластерів.

Перший нейромережевий підхід, який показав добрі результати на великому і складному тестовому наборі, запропоновано Роулі та ін. [3]. Він став де-факто стандартом для порівняння з іншими методами виявлення фронтальних облич. Авторами використано нейронну мережу зі з'єднаннями за принципом сітківки ока, а попередня обробка вхідного зображення відбувається подібно до системи [2]. Зображення сканується вікном 20x20 пікселів у всіх можливих позиціях та масштабах. Щоб скоротити кількість хибних виявлень, авторами поєднано ансамбль нейронних мереж із системою прийняття рішення на основі евристик.

Віола та Джонс запропонували новий надзвичайно швидкий алгоритм ВО, який використовує AdaBoost для відбору важливих Хаар-подібних рис і каскад "слабких" класифікаторів [4].

Сучасні методи [3, 4] мають низку недоліків. Наприклад, система ВО на основі [4] пропускає частково перекриті або сильно затінені обличчя і показує гірші результати, ніж в [3], тоді як підхід, запропонований в [3], занадто повільний для роботи з відеопотоком. Для подолання цих недоліків запропоновано поєднати вищенаведені підходи, а саме: використати декілька Хаар-подібних рис [4] для вибору облич-кандидатів і вдосконалений нейромережевий метод ВО для остаточної класифікації. Також пропонується застосувати сегментацію за кольором шкіри для прискорення процесу ВО (у випадку відеопотоку – додаткову сегментацію за рухом [5, 6]). Після етапів

попередньої обробки, остаточне виявлення виконується каскадом нейронних мереж із різними локальними рецепторними полями за допомогою вдосконаленої стратегії пошуку облич за масштабом та розташуванням.

Вибір облич-кандидатів. Вибір облич кандидатів на основі сегментації за кольором шкіри. Людська шкіра має характерний колір, який легко розпізнається людьми, тому використання інформації про колір шкіри (КШ) може значно полегшити процеси виявлення, локалізації та відслідковування облич [7]. Колір дає можливість швидко обробити вхідне зображення, а також він володіє стійкістю до геометричних варіацій обличчя. У цій роботі за основу взято піксельну сегментацію [7] як метод моделювання КШ, на основі явного визначено границі кластера КШ, оскільки він є простим у використанні, швидким і достатньо точним.

Існує ряд кольорових просторів, які з успіхом можна застосувати для сегментації: RGB, nRGB, HSV, TSL, HSI, YIQ, YCbCr та ін. На основі експериментальних досліджень встановлено, що найкращу сегментацію забезпечує поєднання RGB та TSL кольорових просторів (рис. 1).

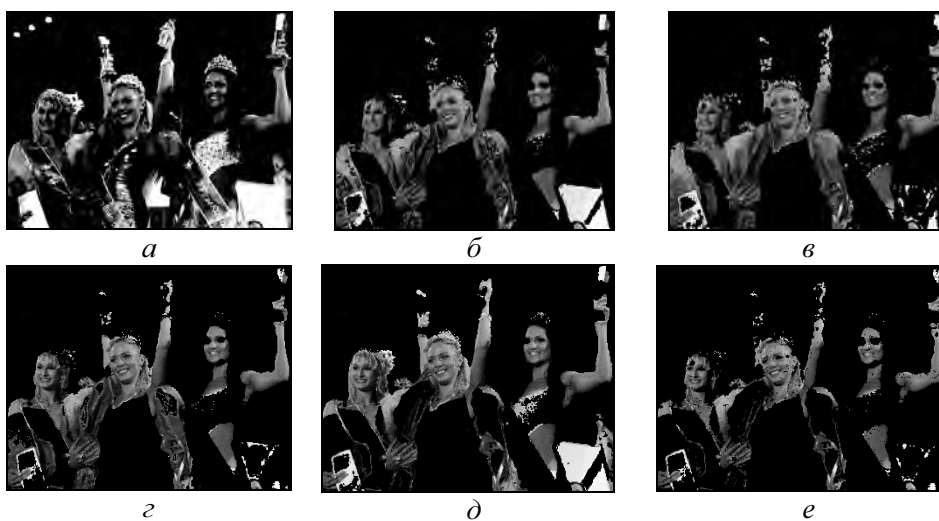


Рис. 1. Результати сегментації вхідного зображення (а) з використанням кольорових просторів RGB (б), TSL (в), YCbCr (г), YIQ (д) та результат їхнього поєднання (е)

Додаткове застосування інших просторів (YCbCr, YIQ) дає змогу відкинути ще деяку незначну частину фону, але при цьому сповільнюється робота блоку сегментації. Для RGB використано наступне явне визначення границь кластера КШ (для кожного з каналів R, G, B) [8]:

$$\begin{aligned}
 & \% \text{ Модель кольору шкіри при денному освітленні} \\
 & R > 95 \text{ and } G > 40 \text{ and } B > 20 \text{ and} \\
 & \max \{R, G, B\} - \min \{R, G, B\} > 15 \text{ and} \\
 & |R - G| > 15 \\
 & \text{and } R > G \text{ and } R > B \\
 & \text{OR} \\
 & \% \text{ Модель кольору шкіри при яскравому освітленні} \\
 & R > 220 \text{ and } G > 210 \text{ and } B > 170 \text{ and} \\
 & |R - G| \leq 15 \text{ and} \\
 & R > B \text{ and } G > B
 \end{aligned} \tag{1}$$

Перед сегментацією доцільно провести балансування кольору, а після неї застосувати до сегментованого зображення морфологічні операції (відкриття, закриття, фільтрація) для покращання якості (рис. 2).



a



б

Рис. 2. Сегментоване зображення до (а) та після (б) морфологічних операцій

У результаті проведення сегментації за КШ значно звужиться область пошуку облич (в 5–20 разів на тестовому наборі), що пришвидшить процес ВО.

Вибір облич кандидатів за допомогою Хаар-подібних рис. Для зменшення області пошуку облич використано деякі Хаар-подібні риси, запропоновані в [4] (рис. 3). Розмір та положення кожної риси вибрано експериментально так, щоб помилка не перевищувала 1 % на навчальному наборі зображень.

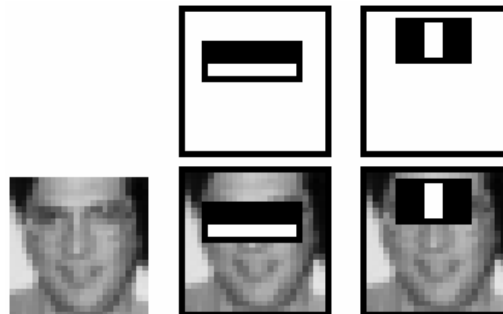


Рис. 3. Перші дві Хаар-подібні риси [4]

Порівняно з [3] використання цих рис різко зменшує кількість вікон, які необхідно надалі аналізувати за допомогою нейронних мереж.

Архітектура та навчання нейромережевого класифікатора. Метод, описаний у [3], використовує нейронні мережі зі з'єднаннями за принципом сітківки ока (retinally connected neural networks), які використовують різні локальні рецепторні поля (рис. 4).

На перший погляд, це складна нейронна мережа (НМ), яка може бути розділена на три простіші (і швидші) мережі з подібними рецепторними полями [6]. Ці НМ можна об'єднати у каскад для пришвидшення процесу ВО [4]. Кожен рівень каскаду може бути представлений однією і більше НМ (рис. 5). У цій роботі використано ансамблі з двох НМ з двома або трьома копіями прихованих нейронів.

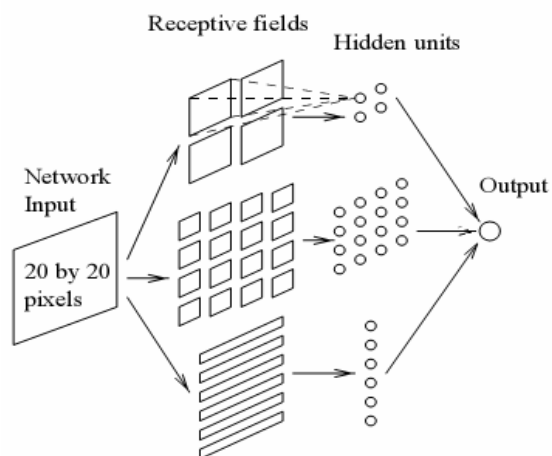


Рис. 4. Архітектура нейронної мережі зі з'єднаннями за принципом сітківки ока

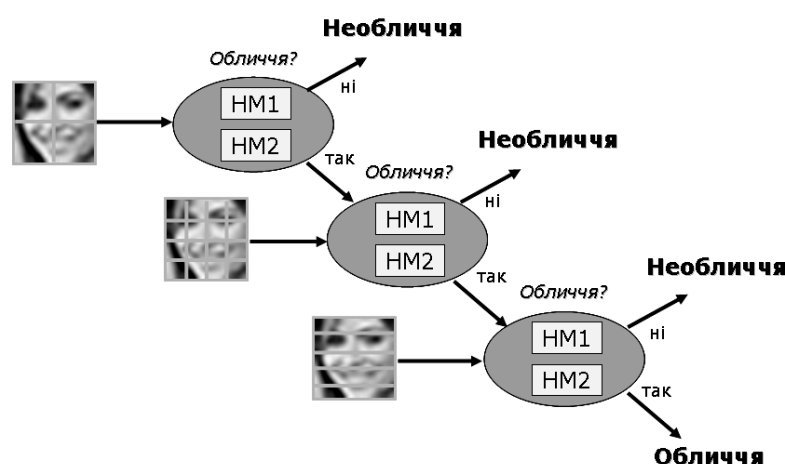


Рис. 5. Процес виявлення облич з використанням нейромережевого каскаду

Кожен наступний рівень каскаду є точнішим, але разом з тим і повільнішим, ніж попередній. Швидкодія рівня залежить від структури НМ, яку він використовує, тобто від кількості параметрів мережі (синапси та пороги). У табл. 1 показано складність НМ каскаду для однієї копії прихованих нейронів у кожній мережі.

Таблиця 1

Структурна складність нейронних мереж каскаду

№ рівня	Розмір рецепторного поля, пікселів	Кількість синапсів	Кількість порогів	Загальна кількість параметрів НМ
1	10x10	404	5	409
2	5x5	416	17	433
3	24x5	606	7	613

Каскадна структура нейромережевого класифікатора передбачає, що кожний рівень відкидає певну кількість необлич (фону) і обробляє зображення, “прийняті” попереднім. Загалом три рівні відкидають приблизно 97–99 %, 94–97 % та 60–85 % вікон вхідного зображення відповідно. Усі ансамблі рівнів використовують логічне правило прийняття рішень “АБО”, тоді як рівні з’єднані за допомогою правила “І”. Така структура класифікатора значно зменшує час для обробки фону, приділяючи більше уваги областям, що схожі на обличчя. У табл. 2 подано результати ВО на зображенні розміром 320x240 пікселів (196057 вікон), виконані ансамблем НМ [3] та

нейромережевим каскадом (програмне забезпечення для експериментів розроблено у середовищі Visual C++ з використанням бібліотеки OpenCV [9]).

Таблиця 2

Порівняльний аналіз швидкодії нейромережевих класифікаторів для виявлення облич

Класифікатор	Час обробки, сек.	Кількість симуляцій НМ
Ансамбль (2 НМ зі з'єднаннями за принципом сітківки ока)	108	392114
Каскад (3x2 НМ)	38	394819
Каскад (3x2 НМ) з використанням двох Хаар-подібних рис	9	84733

Зображення для тренувального набору взято із мережі Інтернет з подальшим їхнім масштабуванням до розміру 20x20 пікселів. Також набір розширено за допомогою створення віртуальних зображень [2, 3] шляхом випадкового відображення, повороту, зміни масштабу, гістограми пікселів, зсуву та згладжування кожного з оригінальних навчальних зображень, – це дає змогу збільшити репрезентативність навчальних облич.

Для навчання НМ використано алгоритм активного навчання з деякими модифікаціями [3, 6, 10] поєднано із процедурою bootstrapping. Він дає змогу працювати з відносно невеликим тренувальним набором, а також дає можливість НМ самій відбирати приклади необлич. Активне навчання триває відведену кількість ітерацій або при досягненні заданої мінімальної помилки на оцінковому наборі.

Вдосконалена стратегія пошуку облич. Класична стратегія пошуку передбачає поступове зменшення вхідного зображення та його сканування фіксованим вікном із постійним кроком (який переважно дорівнює одиниці) [2, 3]. У роботі використано вдосконалену стратегію пошуку [10, 11], яка дає змогу обробляти значно меншу кількість вікон, шляхом використання оберненої піраміди зображень за масштабом, адаптивного кроку проходження зображенням та процедури “прийняття” облич.

Піраміда зображень за масштабом будується у зворотному порядку, тому спершу відбувається пошук великих облич. “Прийняття” облич передбачає видалення області зображення, яка визнається класифікатором як обличчя при певній пороговій кількості виявлень за розташуванням та масштабом (рис. 6). Надалі ця область більше не аналізується щодо наявності облич.



Рис. 6. “Прийняття” облич

Для уникнення обробки раніше прийнятих вікон використано адаптивний крок проходження по зображенню, який збільшується при потраплянні у прийнятну область. Вдосконалена стратегія є досить ефективною за наявності у вхідному зображенні великих облич, які швидко “приймаються” і вилучаються із подальшого розгляду.

Висновки та подальша робота. У роботі запропоновано ряд алгоритмів попередньої обробки зображення та вдосконалений нейромережевий метод виявлення облич. Пошук облич-кандидатів здійснюється за допомогою сегментації за кольором шкіри та Хаар-подібних рис. Вдосконалений метод виявлення облич використовує каскад нейронних мереж, швидкодія якого майже втричі (а при використанні декількох Хаар-подібних рис – в 12 разів) вища, ніж у відомого

нейромережевого підходу. Вдосконалена стратегія пошуку облич пришвидшує процес виявлення особливо за наявності великих облич на вхідному зображенні.

Подальші дослідження пов'язані із розпаралеленням алгоритму навчання нейронних мереж, а також інтеграцією Хаар-подібних рис у процес тренування.

1. *Ming Hsuan Yang: Recent Advances in Face Detection, IEEE ICPR 2004 Tutorial, Cambridge, United Kingdom (2004).* 2. *K. K. Sung and T. Poggio: Example-based learning for view-based human face detection, IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 20, No. 1 (1998) 39-51.* 3. *H. Rowley, S. Baluja, and T. Kanade: Neural network-based face detection. In IEEE Patt. Anal. Mach. Intell., Volume 20 (1998) 22–38.* 4. *P. Viola, M. Jones: Robust Real-Time Face Detection, International Journal of Computer Vision 57(2) (2004) 137–154.* 5. *A. Sachenko, I. Paliy, Y. Kurylyak, V. Kapura, R. Sadykhov, D. Lamovsky. Face Detection Algorithms for Video-Surveillance Systems // Proceedings of the Ninth International Scientific Conference «Pattern Recognition and Information Processing» (PRIP 2007), May 22–24, 2007, Minsk, Belarus, pp. 141–145.* 6. *A. Sachenko, Y. Kurylyak, V. Kapura, I. Paliy, R. Sadykhov, D. Lamovsky. Combined Approach to Face Detection for Video-Surveillance // Proceedings of the IEEE Fourth International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2007), September 6-8, 2007, Dortmund, Germany, pp. 594–598.* 7. *V. Vezhnevets, V. Sazonov, A. Andreeva: A Survey on Pixel-Based Skin Color Detection Techniques, Graphics and Media Laboratory, Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, Russia (2003).* 8. *Peter Peer, Jure Kovac, Franc Solina: Human Skin Colour Clustering for Face Detection, EUROCON 2003 - International Conference on Computer as a Tool, Eds. B. Zajc, Volume 2. Ljubljana, Slovenia (2003) 144–148.* 9. *Open Computer Vision Library, <http://sourceforge.net/projects/opencv/>.* 10. *Палій І.О., Куриляк Ю.О., Саченко А.О. Вдосконалений нейромережевий метод виявлення облич // Вісник Хмельницького національного університету. – 2007. – №2. Т.1. – С. 145–148.* 11. *Yuriy Kurylyak, Ihor Paliy, Anatoly Sachenko, Kurosh Madani, Amine Chohra. Improved Neural Network-based Face Detection Method using Color Images // Proceedings of the Third International Workshop on Artificial Neural Networks and Intelligent Information Processing (ANIIP 2007), May 10, 2007, Angers, France, pp. 107–114.*