

УДК 621.391

В.Г. Дзюба

ЗАСТОСУВАННЯ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ ПОШУКУ ОБ'ЄКТІВ НА ЗОБРАЖЕННІ

Вступ

Одним із способів обробки інформації є розпізнавання. Під розпізнаванням розуміють класифікацію групи об'єктів, явищ, станів, що характеризуються спільними ознаками чи властивостями. Розпізнавання образів є однією з фундаментальних задач, розв'язання якої передбачає класифікацію об'єктів на зображеннях. Для розпізнавання образів створено системи технічного зору, за допомогою яких здійснюють пошук об'єкта на зображенні, а потім його ідентифікують [1, 2].

Методи пошуку переважно орієнтовані на розв'язанні конкретної задачі для певного класу об'єктів [3]. Так, у статті [4] запропоновано використання методу перетворення Хафа для знаходження плоских об'єктів простої геометричної форми (квадрат, коло, еліпс), які легко формалізуються для подання їх у вигляді, прийнятному для комп'ютерної обробки даних. Для більш складних об'єктів з точки зору формалізації, наприклад автомобільних номерних знаків, для їх пошуку був успішно використаний метод морфологічної обробки даних [5, 6]. Для пошуку очей на зображенні обличчя застосовувався метод лінійного дискримінантного аналізу [7]. У праці [8] запропоновано нейромережний підхід до вирішення проблеми пошуку, проте він є недостатньо точним і надійним. Створення інструменту, що є універсальним і здійснює пошук різних об'єктів на зображенні, є актуальною задачею.

У даній статті запропоновано варіант для розв'язання задачі пошуку різних класів об'єктів на зображенні за допомогою нейронних мереж, який дає можливість знаходити різні класи об'єктів з високою точністю та надійністю.

Постановка задачі

Мета статті полягає в розробці нейромережного алгоритму з високими класифікаційними характеристиками, що дасть змогу ефективно здійснювати пошук різних об'єктів на зображенні, а саме пошук пасивних елементів (котушок індуктивності) на зображенні друкованої плати, пошук автомобільних номерних знаків, а також пошук очей на зображенні.

Для досягнення поставленої мети треба розв'язати задачі формування навчальних вибірок для різних класів графічних образів, навчання нейронної мережі на цих вибірках, тестування та оцінки якості розв'язання кожної задачі пошуку.

Архітектура нейромережного детектора для пошуку об'єктів

Детектором називається алгоритм чи програма, що реалізує пошук об'єкта. Як ознаки, що піддаються класифікації нейронною мережею, застосовуються значення інтенсивності пікселів зображення.

Архітектура нейронної мережі, яка використовується як класифікатор образів, належить до класу асоціативних машин [9]. Детально ця нейронна мережа була досліджена в [10], де розв'язувалася задача пошуку обличчя на зображенні. Коротко наведемо основні особливості її архітектури.

Елементарною структурною одиницею мережі є нейрон. Активність нейрона на виході визначається за такою формулою:

$$F(x, p, t) = \begin{cases} 1, & \text{якщо } px \leq pt, \\ 0, & \text{в решті випадків,} \end{cases} \quad (1)$$

де x – вхідний сигнал; p – полярність; t – поріг. Полярність може набувати значень $+1$ і -1 , що дає змогу змінювати знак нерівності. Схематично нейрон з пороговою функцією активації наведений на рис. 1.

Величина сигналу на виході шару нейронів визначається так:

$$H(\mathbf{X}, \mathbf{W}, \mathbf{P}, \mathbf{T}, T_g) = \begin{cases} 1, & \text{якщо } \sum_{i=1}^M w_i F(x_i, p_i, t_i) \leq T_g \sum_{i=1}^M w_i, \\ 0, & \text{в решті випадків,} \end{cases} \quad (2)$$

де M – кількість нейронів в шарі; $\mathbf{X} = \{x_1, x_2, \dots, x_M\}$ – вектор вхідних сигналів; $\mathbf{W} = \{w_1, w_2, \dots, w_M\}$ – вектор вагових коефіцієнтів; $\mathbf{P} = \{p_1, p_2, \dots, p_M\}$ – вектор полярностей; $\mathbf{T} = \{t_1, t_2, \dots, t_M\}$ – вектор порогів; T_g – значення порогу для шару.

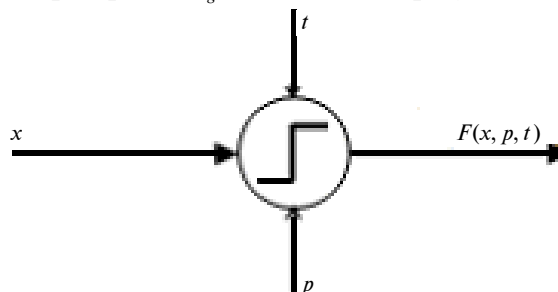


Рис. 1. Нейрон з пороговою функцією активації

На рис. 2 схематично зображено шар нейронів, функціонування якого описується формулами (1) і (2).

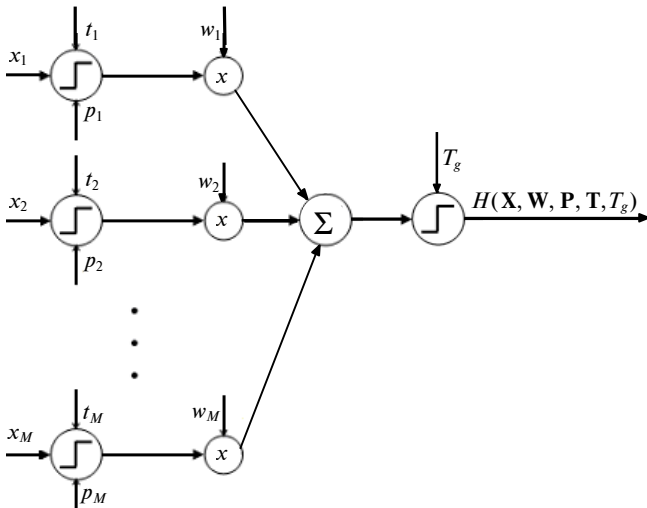


Рис. 2. Шар нейронів

Архітектура мережі може складатися з одного чи кількох шарів. Більш раціональним на практиці є застосування багатшарової мережі [9]. При цьому кожний шар послідовно, починаючи з першого, бере участь у класифікації ознак. Якщо на виході якогось із шарів з'являється нульовий сигнал, то класифікація завершується. Це означає, що ознаки належать до класу "фон". Ознаки можуть бути віднесені до класу "об'єкт" тільки в тому випадку, коли виходи всіх шарів мережі мають одиничні значення.

Навчання такої архітектури проводиться з використанням методу підсилення AdaBoost. Параметрами, що настроюються, є вектори \mathbf{W} , \mathbf{P} , \mathbf{T} та значення порогу T_g для шару нейронів [9].

Особливість структури нейронної мережі відображає закладену в ній стратегію розбиття складної задачі класифікації ознак на множину невеликих і простих задач, кожна з яких розв'язується окремим модулем (нейроном чи шаром нейронів). Це дає можливість проводити послідовне навчання кожного модуля, що зменшує кількість настроювальних параметрів мережі порівняно з класичними архітектурами типу багатшарового перцептрона. Помилка модульної нейронної мережі є мультиплікативною і визначається як добуток помилок кожного модуля. Теоретично доведено, що при навчанні загальна помилка прямує до нуля, якщо помилка кожного модуля не перевищує деякого наперед заданого значення [11]. З цього випливає, що потенційно якість класифікації нейронної мережі такої архітектури може бути достатньою для розв'язання задачі довільної складності.

Як інструмент для проведення дослідження було створено комплекс оригінальних програм для навчання та тестування нейронної мережі. Всі програми були написані на мові програмування C++ та скомпільовані в середовищі Microsoft Visual Studio 2005.

Далі подамо розв'язання трьох задач пошуку різних класів об'єктів. Слід відзначити, що перша задача пошуку котушок індуктивності є найпростішою, оскільки на зображенні котушки можна виділити прості графічні примітиви – квадрат і коло (див. далі рис. 3). Задача пошуку автомобільних номерів є складнішою. Хоча зображення номерного знака являє собою прямокутник із нанесеними на ньому символами, проте ця ознака не є домінуючою при розпізнаванні. З аналізу праць [5, 6] випливає, що морфологічний аналіз зображення підкреслює як ознаку високу контрастність області номерного знака. Третя задача, яка є найскладнішою, полягає в знаходженні очей на зображенні облич та характеризується складністю формалізації образу ока. Хоча запропоновані задачі відрізняються одна від одної за рівнем складності, однак покажемо, що кожна з них може ефективно розв'язуватися за допомогою нейронної мережі із розглянутою вище архітектурою.

Пошук котушок індуктивності на зображенні друкованої плати

Автоматичний пошук компонентів на друкованій платі може використовуватися на виробництві для проведення верифікації і контролю за якістю продукції. Розглянемо приклад розв'язання цієї задачі.

Для того щоб реалізувати автоматичний пошук об'єктів на зображенні, треба навчити класифікатор відрізняти два класи зображень – "об'єкт" і "фон". Процес навчання детектора починається з формування навчальної вибірки. Бажано, щоб ця вибірка достатньо повно представляла два класи зображень, які є в даній задачі, – клас "котушка індуктивності" і клас "фон" (назвемо їх відповідно як позитивні і негативні зображення або приклади). Фрагмент подібної вибірки наведено на рис. 3. Позитивні приклади вибірки формувалися виділенням фрагментів із зображень, що містили образи котушок індуктивності (рис. 3, а). Негативні приклади вибірки були утворені із зображень, які не містили образи котушок індуктивності (рис. 3, б). Слід зазначити, що від репрезентативності навчальної вибірки залежить якість подальшої роботи класифікатора.

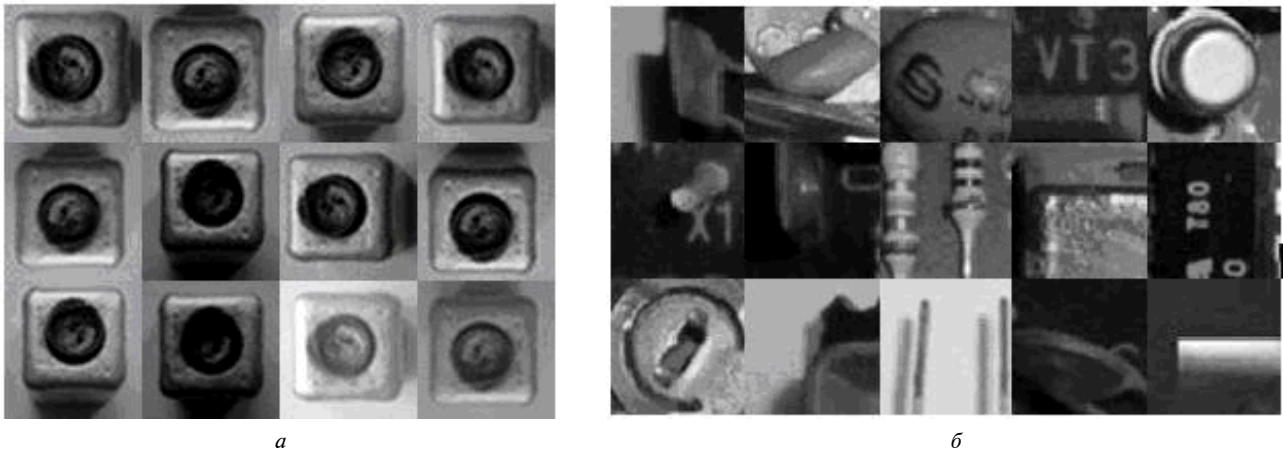


Рис.3. Фрагмент навчальної вибірки для задачі пошуку котушок індуктивності: *а* – позитивні приклади; *б* – негативні приклади

Були встановлені такі опції для навчання нейронної мережі:

- 1) загальна кількість прикладів: позитивних – 300, негативних – 987399;
- 2) кількість прикладів, що беруть участь у навчанні кожного шару нейронної мережі: позитивних – 300, негативних – 300;
- 3) цільове значення помилки навчання для кожного шару – 0,3;
- 4) розмір зображень прикладів – 45×45 пікселів.

В результаті навчання була оптимізована структура нейронної мережі, яка містила дев'ять шарів, загальна кількість нейронів у мережі дорівнювала 37. Час навчання нейронної мережі

на комп'ютері PENTIUM-IV 2,66 ГГц становив 24,4 хв.

Для оцінки якості розпізнавання було сформовано тестову вибірку з 23 зображень друкованих плат, на яких знаходилося 57 котушок індуктивності. В результаті тесту детектором було знайдено 55 із 57 котушок, що у відсотковому відношенні дорівнює 96,4%. Помилкових захватів виявлено не було. Під помилковим захватом розуміють фрагмент зображення, який помилково сприйнятий детектором за об'єкт. Результат роботи детектора на одному із зображень друкованої плати наведено на рис. 4, де знайдені місця розташування котушок індуктивності помічені градієнтними квадратами.

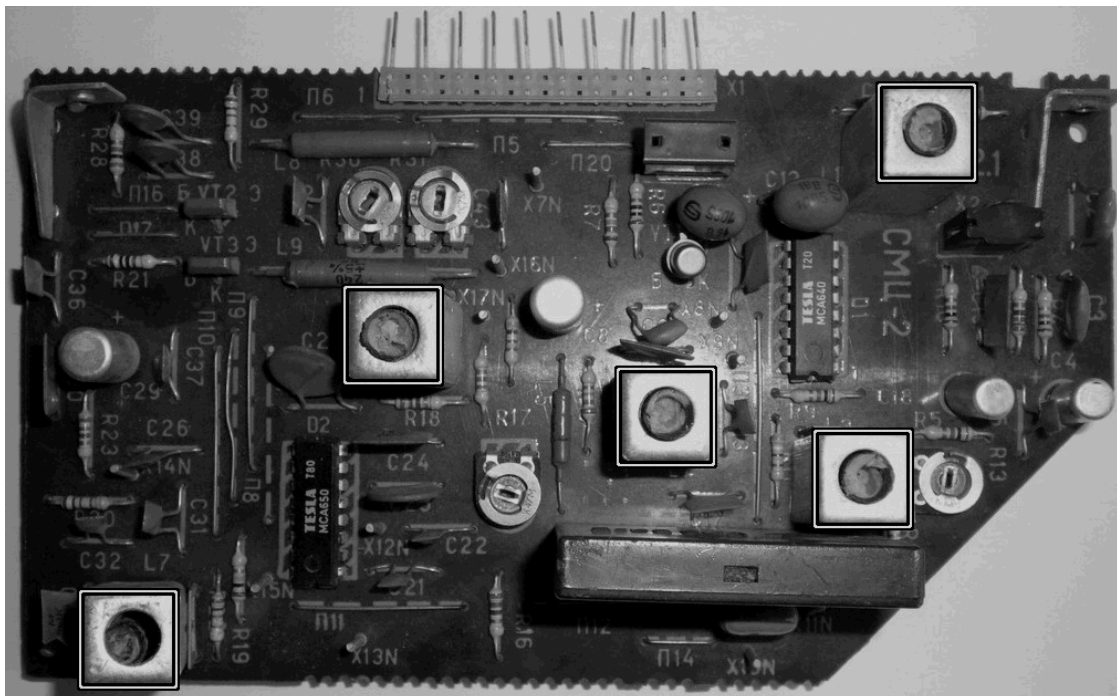


Рис. 4. Результат пошуку котушок індуктивності на зображенні

Аналізуючи тестові дані, можна зробити висновок, що нейронна мережа виявила високі властивості розпізнавання. Але слід зауважити, що ця оцінка є суб'єктивною, оскільки розмір тестової вибірки невеликий. Проте можна зробити висновок, що потенціал нейронної мережі достатньо високий для якісного розв'язання задачі пошуку котушок індуктивності на зображенні друкованої плати.

Пошук автомобільних номерних знаків на зображенні

Задача пошуку автомобільних номерних знаків є дуже актуальною. Це підтверджується великою кількістю як публікацій, так і компаній, в яких дана проблематика вирішується [12]. Пошук номера є першим етапом роботи системи автоматичного розпізнавання номерних знаків. Наведемо приклад розв'язання цієї задачі із застосуванням нейронної мережі.

Як і для попередньої задачі, першим кроком в її розв'язанні є створення навчальної вибірки. Позитивними прикладами в даному випадку є зображення, що містять номерні знаки, а негативними – будь-які фонові зображення (рис. 5).



a



b

Рис. 5. Фрагмент навчальної вибірки для задачі пошуку автомобільних номерних знаків: *a* – позитивні приклади; *b* – негативні приклади

Були встановлені опції навчання нейронної мережі:

- 1) загальна кількість прикладів: позитивних – 500, негативних – 5238746;
- 2) кількість прикладів, що беруть участь у навчанні кожного шару нейронної мережі: позитивних – 500, негативних – 500;
- 3) цільове значення помилки навчання для кожного шару – 0,3;
- 4) розмір зображень прикладів – 64×16 пікселів.

Навчена нейронна мережа містила 15 шарів, загальна кількість нейронів у мережі становила 285. Процес навчання зайняв 741,3 хв на комп'ютері PENTIUM-IV 2,66 ГГц. Тестування проводилось на власне створеній тестовій базі, що складалася із 108 зображень. На кожному зображенні був один номерний знак. Було знайдено 104 номерних знаків із 108, що становило 96,2%. Кількість фрагментів зображень, помилково прийнятих системою за номерний знак, дорівнювала 3. Приклад роботи детектора на двох зображеннях тестової вибірки наведено на рис. 6, на якому відображені місцезнаходження номерних знаків, а також помилковий захват (фрагмент забору).



Рис. 6. Результат пошуку автомобільних номерних знаків

Аналізуючи дані тестування, можна стверджувати, що нейронна мережа проявила високі властивості до розпізнавання автомобільних номерних знаків. Це дозволяє говорити про факт якісного розв'язання задачі пошуку номерів на зображенні.

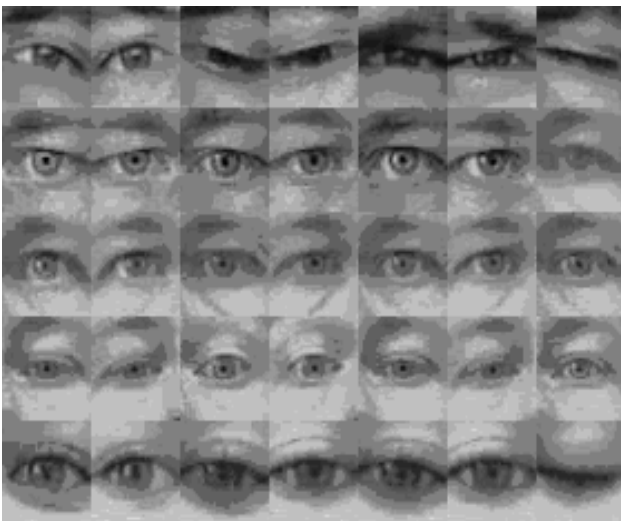
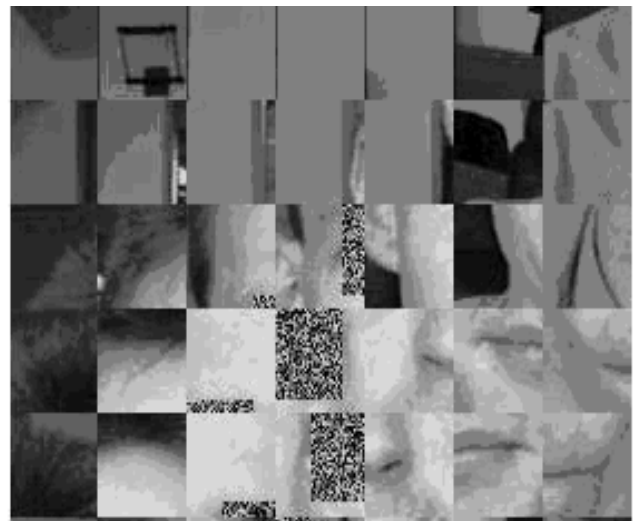
Пошук очей на зображенні

Проблема автоматичної локалізації очей полягає в знаходженні координат зіниць очей без сторонньої допомоги людини. Вирішення цієї проблеми може знайти застосування в задачах ідентифікації осіб. Координати зіниць очей зазвичай використовують як опорні точки при нормалізації зображень облич [13].

За аналогією з попередніми завданнями, позитивними прикладами в даному випадку є зображення, що містять очі, а негативними прикладами – фонові зображення (рис. 7).

Були встановлені такі опції навчання для цієї задачі:

- 1) загальна кількість прикладів: позитивних – 1000, негативних – 9328746;
- 2) кількість прикладів, що беруть участь у навчанні кожного шару нейронної мережі: позитивних – 1000, негативних – 1000;
- 3) цільове значення помилки навчання для кожного шару – 0,3;
- 4) розмір зображень прикладів – 8×8 пікселів.

*a**б*

На базі навченої нейронної мережі було створено детектор для здійснення автоматичного пошуку очей. Мережа після навчання містила 27 шарів, спільна кількість нейронів у мережі дорівнювала 1419. Час навчання нейронної мережі на комп'ютері PENTIUM-IV 2,66 ГГц тривав 1465,4 хв.

Тестування детектора проводилося на спеціалізованій базі даних “BioID Face Database”, яка складена компанією HumanScan, що спеціалізується на створенні технологій біометричної ідентифікації [14]. База складається з 1521 півтонового зображення з роздільною здатністю 384 на 286 пікселів. Кожне зображення містить фронтальний вид особи. Додатково база містить файли з координатами очей (еталонні файли). Саме ці дані були використані для проведення автоматичного тестування. При розпізнаванні кожного зображення з бази отримувалися від детектора координати місцезнаходження зіниць очей. Ці координати порівнювалися з координатами, що зберігалися в еталонних файлах. Програмою відстежувалися випадки збігу результатів та випадки виявлення розбіжностей. В результаті тесту було правильно знайдено 3009 із 3042 очей, що становило 98,9%. Кількість помилково знайдених очей дорівнювала 18. Приклад роботи нейронної мережі в складі детектора очей наведено на рис. 8. Місцезнаходження очей на рисунку помічено хрестиками.

Рис. 7. Фрагмент навчальної вибірки для задачі пошуку очей: *a* – позитивні приклади; *б* – негативні приклади

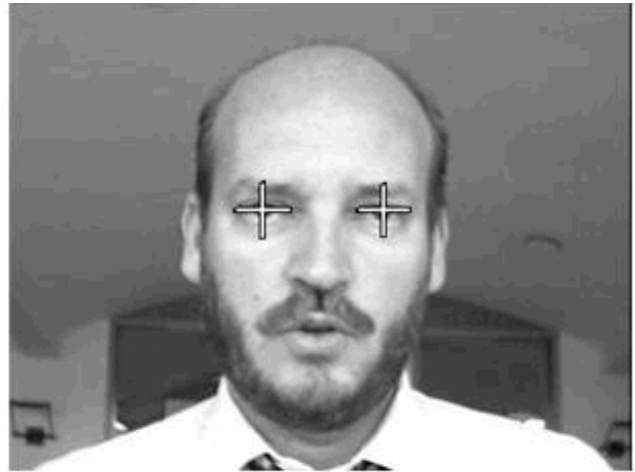


Рис. 8. Результат пошуку очей на зображенні

Аналізуючи результати тестування, можна з об'єктивною впевненістю сказати, що нейромережний детектор здатен з великою точністю розв'язувати задачу пошуку очей на зображенні.

Висновки

Запропонований варіант розв'язання задачі пошуку різних класів об'єктів на зображенні з використанням нейронної мережі відзначається високою ефективністю та якістю. Це підтверджується експериментальними результатами тестування, згідно з якими відсоток правильно знайдених об'єктів становив 96,4% при пошуку ко-

тушок індуктивності на зображенні друкованої плати, 96,2% – при пошуку номерів на зображенні та 98,9% – при пошуку очей на зображенні. Ці результати підкреслюють високі класифікаційні властивості нейронної мережі і підтверджують той факт, що нейронна мережа справді є тим універсальним інструментом, який забезпечує точний та ефективний пошук різних класів об'єктів на зображенні.

Подальші дослідження будуть направлені на розв'язання за допомогою нейронної мережі іншої актуальної задачі розпізнавання образів – задачі ідентифікації об'єктів.

В.Г. Дзюба

ПРИМЕНЕНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ ПОИСКА ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ

Представлен вариант решения задачи поиска различных классов объектов на изображении с помощью нейронных сетей. Экспериментально доказано, что нейронная сеть с высокой точностью и надежностью способна осуществлять поиск катушек индуктивности, автомобильных номерных знаков и глаз на изображении.

V.G. Dzyuba

USE OF NEURON NETWORK FOR OBJECTS DETECTION IN IMAGE

The paper considers the variant for problem solving of different object classes by neuron networks. Specifically, it was experimentally proven that a neuron network with high exactness and reliability can search for coils in image of circuit board, number plates, and eyes in image.

1. *Шапиро Л., Стокман Дж.* Компьютерное зрение / Пер. с англ. – М.: БИНОМ, 2006. – 752 с.
2. *Визильтер Ю.В., Желтов С.Ю., Князь В.А. и др.* Обработка и анализ цифровых изображений с примерами на LabVIEW и IMAQ Vision. – М.: ДМК Пресс, 2007. – 464 с.
3. *Методы компьютерной обработки изображений* / Под ред. В.А. Сойфера. – М.: Физматлит, 2001. – 784 с.
4. *Olson C.F.* Locating geometric primitives by pruning the parameter space // Pattern Recognition. – 2001. – **34**, N 6. – P. 1247–1256.
5. *Shih-Hao Yu, Jun-Wei Hsieh, Yung-Sheng Chen.* Morphology-based License Plate Detection from Complex Scenes // IEEE Proc. of Intern. Conf. on Pattern Recognition. – 2002. – **3**. – P. 176–179.
6. *Jun-Wei Hsieh, Shih-Hao Yu, Yung-Sheng Chen.* Morphology-based license plate detection in images of differently illuminated and oriented cars // J. of Electronic Imaging. – 2002. – **11**, N 4. – P. 507–516.
7. *Peng Wang, Qiang Ji.* Learning discriminant features for multi-view face and eye detection // Computer Vision and Pattern Recognition. – 2005. – **1**. – P. 373–379.
8. *Torralba A., Murphy K.P., Freeman W.T.* Sharing Visual Features for Multiclass and Multiview Object Detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2007. – **29**. – P. 854–869.
9. *Хайкин С.* Нейронные сети: Полный курс / Пер. с англ. – 2-е изд. – М.: Вильямс, 2006. – 1104 с.
10. *Якименко Ю.И., Дзюба В.Г.* Алгоритм автоматического захвата лиц // Тем. вып. “Проблемы электроники”. – К.: Электроника и связь, 2005. – Ч. 2. – С. 88–92.
11. *Freund Y., Schapire R.E.* Experiments with a new boosting algorithm // Machine Learning: Proc. of the 13th Intern. Conf. – 1996. – P. 148–156.
12. *Коваленко Е.Н., Сытник А.В.* Методы выделения номерного знака на изображении // V Междунар. конф. “Интеллектуальный анализ информации ИАИ-2005”. – 2005. – С. 167–177.
13. *Wechsler H.* Reliable Face Recognition Methods System Design, Implementation and Evaluation. – London: Springer Verlag, 2007. – 329 p.
14. *Jesorsky O., Kirchberg K., Frischholz R.* Robust Face Detection Using the Hausdorff Distance // Audio and Video based Person Authentication. – 2001. – P. 90–95.

Рекомендована Радою Науково-дослідного інституту прикладної електроніки НТУУ “КПІ”

Надійшла до редакції
24 квітня 2008 року