

АНАЛІЗ ЕФЕКТИВНОСТІ ЗАСТОСУВАННЯ БАГАТОШАРОВИХ НЕЙРОННИХ СХЕМ ДЛЯ РОЗПІЗНАВАННЯ ОБРАЗІВ

У статті проводиться порівняння нейросхемних та нейромережевих алгоритмів на теоретичному та прикладному рівні з метою визначення оптимального методу розпізнавання символічної інформації на графічних зображеннях.

In article is conducted comparison neuroschem and neural network algorithms on theoretical and practical levels for the reason determinations of the optimal method of the recognition to symbol information on graphic pictures.

Впродовж багатьох років розвитку кібернетики велика увага приділялася вивченню здатності людини швидко і безпомилково класифікувати і виділяти окремі образи навколишнього світу, ідентифікувати об'єкти незалежно від їх положення в просторі, освітлення і інших умов. Перспектива реалізації комп'ютерного аналога такої системи дуже прикладна, адже вона б дозволила замінити людину в багатьох областях її діяльності, де робота є небезпечною чи монотонною [1, 2].

У цілому проблема розпізнавання образів складається із двох частин: навчання й розпізнавання. Навчання здійснюється шляхом показу окремих об'єктів із зазначенням їхньої приналежності тому або іншому образу. В результаті навчання система, що розпізнає, повинна придбати здатність реагувати однаковими реакціями на всі об'єкти одного образу й різними – на всі об'єкти різних образів. Дуже важливо, що процес навчання повинен завершитися тільки шляхом показів кінцевого числа об'єктів без яких-небудь інших підказок. Об'єктами навчання можуть бути або картинки, або інші візуальні зображення (букви), або різні явища зовнішнього світу, наприклад звуки, стани організму при медичному діагнозі, стан технічного об'єкта в системах керування й ін.. При цьому в процесі навчання вказуються тільки самі об'єкти і їхня приналежність образу. За навчанням слідує процес розпізнавання нових об'єктів, що характеризує дії вже навченої системи [3]. Автоматизація цих процедур і становить проблему навчання розпізнаванню образів. У тому випадку, коли людина сама складає або визначає, а потім нав'язує машині правило класифікації, проблема розпізнавання вирішується частково, тому що основну й головну частину проблеми (навчання) людина бере на себе [4].

У загальному випадку, розпізнавання тексту складається з наступних етапів:

- дообробка;
- сегментація;
- розпізнавання.

Сьогодні відомо три підходи до розпізнавання символів – шаблонний, структурний і признаковий.

Шаблонні методи перетворюють зображення окремого символу в растрове, порівнюють його зі всіма шаблонами, наявними в базі і вибирають шаблон з найменшою кількістю крапок, відмінних від вхідного зображення. Шаблонні методи досить стійкі до дефектів зображення і мають високу швидкість обробки вхідних даних, але надійно розпізнають тільки ті шрифти, шаблони яких їм „відомі”. І якщо розпізнаний шрифт хоч трохи відрізняється від еталонного, шаблонні методи можуть робити помилки навіть при обробці дуже якісних зображень.

У структурних методах об'єкт описується як граф, вузлами якого є елементи вхідного об'єкту, а дугами – просторові відношення між ними. Методи, що реалізують подібний підхід, зазвичай працюють з векторними зображеннями. Структурними елементами є складові лінії символу. Так, для букви „р” – це вертикальний відрізок і дуга. До недоліків структурних методів слід віднести їх високу чутливість до дефектів зображення, які порушують відношення складових елементів. Також векторизація сама по собі може додати додаткові дефекти. Крім того, для цих методів, на відміну від шаблонних і признакових, досі не створені ефективні автоматизовані процедури навчання. Тому структурні описи найчастіше доводиться створювати вручну.

У признакових методах усереднене зображення кожного символу представляється як об'єкт в n-мірному просторі ознак. Тут вибирається алфавіт ознак, значення яких обчислюються при розпізнаванні вхідного зображення. Отриманий n-мірний вектор порівнюється з еталонними, і зображення відноситься до найбільш підходящого з них.

Також існує ряд методів, побудованих як синтез трьох викладених підходів. Нижче розглядаються найпопулярніші, добре вивчені і часто застосовні на практиці різні методи розпізнавання символів.

Нейромережі. Медичні дослідження людського мозку спонукали учених вести розробки моделей, що розпізнають, спираючись на структуру зорової кори. К.Фукушима в 1975 році розробив когнітрон – гіпотетичну модель біологічної системи сприйняття і розпізнавання, інваріантну до поворотів, переміщень, змін масштабів образів. Подальшим розвитком нейромережі когнітрон стала архітектура стандартного неокогнітрона, яка, не дивлячись на подальші розробки, як і раніше залишається основоположною [5].

Неокогнітрон показав відмінні результати по розпізнаванню образів, підданих зсувам в позиції, зашумленню, спотворенню форми. Проте, неокогнітрон не справлявся із завданням розпізнавання образів, повернених на деякий кут, а процес розпізнавання займав досить тривалий час.

Протягом багатьох років неокогнітрон як технологія удосконалився завдяки праці учених. Були введені дві моделі неокогнітронів, стійких до повороту вхідного образу:

- R-неокогнітрон, з динамічним числом нейронів;
- Td-r-неокогнітрон, з постійним числом нейронів.

R-неокогнітрон втрачає свою ефективність при великому числі класів, до яких можуть належати образи, оскільки час розпізнавання зростає експоненціально залежно від числа нейронів.

Td-r-неокогнітрон виконує т.з. ментальний поворот образу, визначаючи кут і порівнюючи з еталоном. Даний принцип має багато спільного з поведінкою людини: коли людина дивиться на картину, повернену на 90 градусів, вона визначає кут і прагне нахилити свою голову, щоб зрозуміти, що намальоване на картині.

Неокогнітрон є багатошаровою нейронною мережею (Рис 1). У ній розрізняють два види шарів – S і C (Simple – прості і Compatible – складні). Дані шари парами утворюють один рівень в системі. Кожен шар (S і C) містить декілька площин нейронів. У них використовуються два типи нейронів: збудливі і гальмуючі. Перші прагнуть збільшити вихідну реакцію шару, другі навпаки – орієнтовані на зменшення вихідного сигналу. Будь-який збудливий нейрон в площині шару типу S і C отримує набір вхідних сигналів з нейронів попереднього шару (для U_{s1} входом служить шар U_0), які утворюють його область зв'язку. Цю ж область зв'язку має і гальмуючий нейрон, вихід якого також поступає на вхід збудливого нейрона. Таким чином, результуючий вихід генерує тільки збудливий нейрон. На величину вихідного сигналу також впливає „сила збудження” сусідніх нейронів на поточному шарі, які утворюють область конкуренції даного нейрона. Кількість областей конкуренції рівна числу нейронів на площині і тому вони частково перекриваються. Сенс такої організації зв'язків полягає у визначенні найбільш сильно збудженого нейрона („переможця”), який міг би повною мірою охарактеризувати всю область конкуренції. Переможець в області конкуренції завжди один.

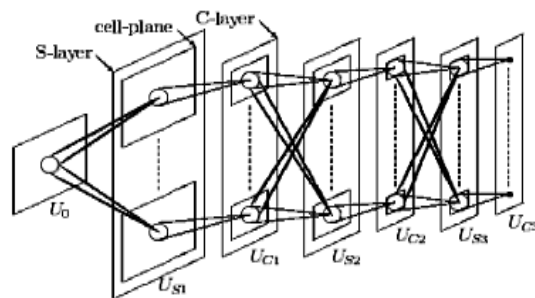


Рис.1. Багатошарова структура неокогнітрона

Області зв'язку призначені для виділення характерних частин вхідного образу. Вони теж частково перекриваються, що забезпечує у разі відмови в роботі одного з нейронів заміну його іншим нейроном з цієї області конкуренції.

Кожній площині нейронів в шарі S відповідає своя площина в шарі C. Збудливі і гальмуючі нейрони шару C отримують сигнали з певних областей зв'язку попереднього шару S. Семантика S і C шарів полягає в наступному: S виділяє особливості (фрагменти) вхідного образу, визначає серед них, за допомогою областей конкуренції, найбільш значущі і передає їх парному шару C, який у свою чергу додає цим особливостям меншу позиційну залежність. Дана тенденція простежується від рівня до рівня. При цьому розмір площин нейронів поступово зменшується і на шарі C останнього рівня кожна площина складається з одного нейрона відповідного одному класу вхідних образів.

Існує декілька версій реалізації неокогнітрона, що мають незначні відмінності:

- Моделі, засновані на методі навчання „без вчителя”;
- Моделі, засновані на методі навчання „з вчителем”.

Моделі, засновані на методі навчання з вчителем володіють вищою швидкістю навчання і стабільністю. Проте їх недолік в тому, що передбачається використання додаткової інформації про приналежність всього вхідного образу цілком і його окремих частин до того або іншого класу.

Моделі, які навчаються без вчителя, є більш адаптивними і оптимальними. Вони базуються на уявленнях про зорову систему.

Нейросхеми. Нейронні мікросхеми – алгоритмічно інший підхід до розпізнавання образів, ніж нейронні мережі. З програмно-апаратної точки зору мозок і нервова система складаються з великого набору нейронів – основної структурної і функціональної одиниці нервової системи. Робота нервової системи полягає у взаємодії нейронів між собою. Нейрон складається з тіла і відростків, що відходять від нього, – коротких дендритів і довгого аксона, які виконують функції транспортних магістралей, тобто щось схоже на електричні провідники на друкованій платі комп'ютера. Прийнято вважати, що по аксону передається витікаюча від нейрона інформація, а по дендритам – вхідна. Образно кажучи, через свій аксон нейрон „розмовляє” з іншими нейронами, а „слухає” через свої дендрити. Тому деяку ділянку нервової системи можна представити у вигляді друкованої плати, де ролі мікросхем виконують тіла нейронів, а електричними провідниками є аксони і дендрити, підключені один до одного в якій-небудь послідовності.

На програмному рівні збудження і гальмування штучних нейронів може бути єдиною властивістю нейрона, якою визначається характер його функціонування. Якщо властивість приймає значення збудження, то нейрон включається й починає працювати в схемі нервового вузла або центру. Якщо ж приймає значення повного гальмування, то нейрон відключається. Схема нервового вузла або центру продовжує працювати, але через цей нейрон сигнали взагалі не проходять. Аналогом є випадок, коли на друкованій платі комп'ютера перегоріла або тимчасово відключилася певна мікросхема. Проте якщо властивість нейрона приймає значення гальмування до слабого реагування, то значення вхідних нервових сигналів пригнічуються на величину, еквівалентну ступеню гальмування.

Основу програмної моделі штучної нервової системи нейросхемної моделі складає невеликий набір програмних нейронних мікропроцесорів, які самі є програмними об'єктами, що мають свої властивості, методи і події [6]. Якщо на один з дендритів мікропроцесора приходить сигнал з даними, то дана подія викликає внутрішню подію, яка запускає виконання певної внутрішньої базової операції, внаслідок чого в аксоні мікросхеми з'являються вихідні дані. Аксон підключений до дендритів інших мікропроцесорів, а значить на їх дендритах з'являються сигнали з даними, які поступили з аксона. Ці сигнали викликають внутрішні події усередині початкових мікросхем. Результати їх роботи через аксони подаються на дендрити наступних мікропроцесорів і так далі. У такій віртуальній схемі безперервно виникають події, що є наслідками минулих подій, що примушує схему жити своїм життям. Подібний підхід до побудови штучного інтелекту є об'єктно-орієнтованим і фрактальним, що дає можливість швидко створювати складні нейросхемні структури й динамічно корегувати їх архітектуру. Значною перевагою нейронних схем є їх властивість паралельної обробки однорідної інформації різними сегментами.

З нейромережами нейросхеми мають низьку схожість, яка обмежується лише тим, що в нейромережах нейронні зв'язки (які називають „вагами”) зовні нагадують з'єднання дендритів з аксонами в нейроланцюгах. Тому нейронні мікросхеми – це не те, що зазвичай називають нейромережами. Відповідно, не коректне також твердження, що нейрочіпи – це нейромережі в чистому вигляді, адже вони мають віддалену спорідненість. Стосовно питання, чи інший це алгоритмічно підхід, то на сучасному етапі це не має великого значення: припускається, що нейромережі – один підхід, нейроланцюги – інший підхід, програмні алгоритми – третій підхід.

Аналіз ефективності застосування нейромереж та нейросхем для вирішення прикладних задач. При застосуванні мережі неокогнітрон та нейросхем на практиці [7], а саме використання їх в задачі розпізнавання локалізованих автомобільних номерів (Рис. 2), були визначені наступні особливості.

Спостерігається низька швидкодія мережі неокогнітрон у порівнянні з іншими технологіями штучного інтелекту, що зумовлено великою кількістю нейронів в прошарках та площинах, хоч нейрони і не зв'язувались один з одним. При спробі зменшити кількість прошарків значно знижувалась ефективність мережі по розпізнаванню образів, а при спробі збільшити кількість прошарків для збільшення ефективності, час навчання та роботи мережі збільшувався експоненційно з кожним новим шаром, тому було знайдено компроміс між швидкістю та ефективністю, а саме 4 шари площин з нейронами. Якщо зважити на те, що ця мережа найточніше на сьогоднішній день моделює роботу зорової кори мозку, й на швидкий ріст продуктивності сучасних комп'ютерів, то можна зробити висновок, що мережа неокогнітрон є достатньо ефективною в задачах розпізнавання образів, хоч і не є оптимальною для розглядуваної області застосування – розпізнаванню локалізованих автомобільних номерів. Адже в цьому типі задач їх специфікою одне є значна увага, що приділяється критерію швидкодії.

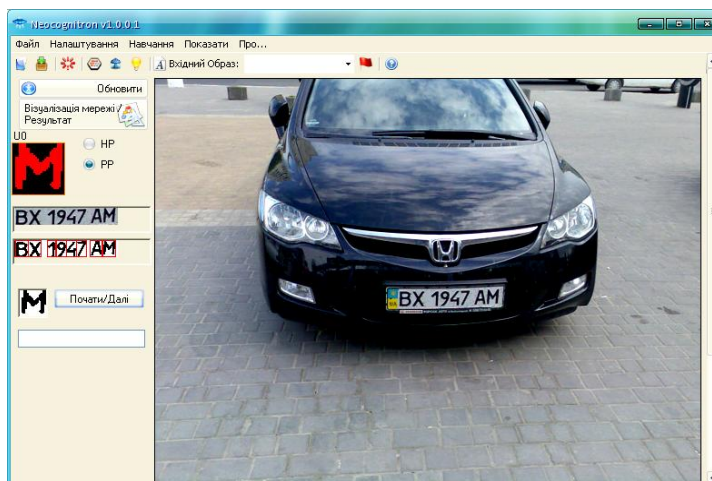


Рис. 2. Застосування мережі неокогнітрон та нейросхем для розпізнавання локалізованих автомобільних номерів

Застосовуючи нейросхемне вирішення в задачі розпізнавання образів автомобільних номерних знаків, відмічено значно вищу швидкодію (у середньому 15-кратна різниця у швидкості роботи по розпізнаванню) в порівнянні з неокогнітронним аналогом. Також відзначено більшу гнучкість нейролацюгових вирішень у порівнянні з нейромережевими. Проте варто зауважити, що при роботі з нейросхемою необхідна більш потужна дообробка символів та їх стандартизація, ніж при роботі з нейромережею, оскільки в самій конструкції мережі закладено механізми вилучення особливостей того чи іншого образу та запам'ятовування саме їх, на відміну від нейроланцюга. Проте, як зазначалося вище, швидкодія нейросхем значно вища за нейромережу, що в даній області застосування є одним з головних критеріїв вибору, тому нейросхемний аналог краще підходить для цієї конкретної задачі. При цьому процес стандартизації та дообробки символів може сповільнювати роботу всієї системи взагалом, тому на прикладному рівні вирішенню цього питання слід приділяти особливу увагу.

Література

1. Hubel D.h., Wiesel T.n. Receptive fields and functional architecture in two nonstriate visual area (18 and 19) of the cat. J. Neurophysiol. 28, 1965, pp.229-289.
2. Перетворення знань. Переклад з японського / під ред. С.Осуги, Ю.Саеки. – М.: Мир, 1990. – 304с.
3. Сотник С.Л, Конспект лекцій по курсу „Основы проектирования систем искусственного интеллекта”, Москва, 1998.
4. Горбань А.Н., Сенашова М.Ю. Погрешности в нейронных сетях // Вычислительный центр СО РАН в г.Красноярске. Красноярск, 1997. – 38с.
5. Fukushima K. Neocognitron: a self-organising neural network for mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics 36, 1980, pp. 193-202.
6. Мазурець О.В., Ковальчук С.С. Застосування багатопшарових нейронних схем для вирішення складних технологічних задач // Збірник наукових праць за матеріалами другої всеукраїнської науково-технічної конференції «Актуальні проблеми комп'ютерних технологій 2008» – Хмельницький – ХНУ, 2008. – Том 2 – С.27–35.
7. Романюк О.П. Аналіз ефективності застосування багаторівневих нейронних мереж типу неокогнітрон для розпізнавання образів // Збірник наукових праць за матеріалами другої всеукраїнської науково-технічної конференції «Актуальні проблеми комп'ютерних технологій 2008» – Хмельницький – ХНУ, 2008. – Том 2 – С.122–125.