

УДК 004.032.26

Т. О. Савчук, Є. О. Ярема

Вінницький національний технічний університет
вул. Хмельницьке шосе, 95, 21021 Вінниця, Україна

Використання нейронних мереж для розпізнавання символів

Проведено дослідження методів розпізнавання символів за допомогою нейронних мереж. Створено інтелектуальний модуль штучних нейронних мереж, що функціонує за принципом перцептрона, та має можливість розпізнавати рукописні символи із зашумленістю до 40 %.

Ключові слова: розпізнавання символів, штучна нейронна мережа, адаптивна резонансна теорія, перцептрон, синапс.

Вступ

На сьогоднішній день існує багато підходів (алгоритмів) щодо вирішення проблеми розпізнавання символів, але більшість із них або вузько спрямовані на певну сферу розпізнавання (вони показують високі результати розпізнавання, але не є універсальними), або якість розпізнавання дуже низька, й алгоритм працює повільно. Тому для вирішення задач розпізнавання символів часто використовуються високоінтелектуальні системи на основі штучних нейронних мереж.

Однак штучні нейронні мережі не є інструментом для вирішення задач будь-якого типу. Вони не підходять для виконання таких задач як нарахування заробітної платні, але їм віддається перевага у великому класі задач розпізнавання символів, з якими погано або взагалі не справляються звичайні комп'ютери.

Аналіз методів розв'язання задач розпізнавання символів

На сьогодні розроблено багато систем, що демонструють можливості штучних нейронних мереж: мережі здатні подавати текст фонетично, розпізнавати рукописні букви, стискувати зображення. Більшість потужних мереж, що здатні розпізнавати символи та звуки, використовують як основу принцип зворотного поширення [1], що є систематичним методом для навчання багатосарових мереж і, тим самим, переборює обмеження, зазначені Мінським [1]. Нейронні мережі, які працюють за вищевказаним принципом прийшли на зміну системам, що склались з одного шару штучних нейронів і використовувались для широкого класу зав-

дань, у тому числі для організації штучного зору. Однак такі системи мають недоліки. Насамперед немає гарантії, що мережа може бути навчена за скінченний час.

Кожен із розроблених алгоритмів навчання мереж має свої специфічні переваги, але спільним недоліком є обмеження у своїх можливостях навчатися й згадувати.

При навчанні мережі кожна вхідна (або вихідна) множина сигналів розглядається як вектор. Навчання здійснюється шляхом послідовного пред'явлення вхідних векторів з одночасним налаштуванням ваг відповідно до певної процедури. У процесі навчання ваги мережі поступово стають такими, щоб кожний вхідний вектор відповідно виробляв вихідний. Навчальні алгоритми можуть бути класифіковані як алгоритми навчання без учителя й з учителем [1, 2].

У першому випадку, при пред'явленні вхідних символів мережа самоорганізується за допомогою настроювання своїх ваг у відповідності з визначеним алгоритмом. Унаслідок відсутності вказівки наперед визначеного виходу, у процесі навчання результати непередбачені з погляду визначення збудливих символів для конкретних нейронів. При цьому, мережа організується у формі, що відбиває істотні характеристики навчального набору. Наприклад, вхідні символи можуть бути класифіковані відповідно до ступеня їхньої подібності так, що символи одного класу активізують той самий вихідний нейрон.

Зазначенні методи навчання мережі передбачають, що відомі лише вхідні вектори, а на їхній основі мережа вчиться давати найкращі значення вихідної функції.

У другому випадку існує вчитель, що надає вхідні символи мережі, порівнює результуючі виходи з необхідними, а потім налаштовує ваги мережі таким чином, щоб зменшити розходження.

Методи навчання мережі з учителем передбачають, що існують пари «вхід-вихід», тобто відомі значення вхідних векторів, і значення вихідних векторів, що їм відповідають.

Таким чином, нейронна мережа, яка навчається з учителем, є більш надійною, оскільки при певному вхідному сигналі на виході формується відповідне вихідне значення.

Навчаючи мережу розпізнавати нові символи, дуже часто знищуються або змінюються результати попереднього навчання. Якщо існує тільки фіксований набір навчальних векторів, вони можуть пред'являтися при навчанні циклічно. У мережах зі зворотним поширенням, навчальні вектори подаються на вхід мережі послідовно, поки мережа не «навчиться» всьому вхідному набору. Проте цілком навчена мережа повинна запам'ятати новий навчальний вектор, і він може змінити ваги настільки, що буде потрібно повне перенавчання мережі.

Адаптивна резонансна теорія (АРТ) [1] є одним із рішень цієї проблеми. Мережі й алгоритми АРТ зберігають пластичність, необхідну для вивчення нових образів, у той же час, запобігаючи зміні раніше запам'ятованих образів, і тому вони можуть бути ефективними при розпізнаванні символів.

Мережа АРТ являє собою векторний класифікатор. Вхідний вектор класифікується в залежності від того, на який із множини раніше запам'ятованих образів він схожий. Своє класифікаційне рішення мережа АРТ виражає у формі збудження одного з нейронів розпізнавального шару. Якщо вхідний вектор не відповідає

жодному із запам'ятованих образів, створюється нова категорія за допомогою запам'ятовування образу, ідентичного новому вхідному вектору. Якщо визначено, що вхідний вектор схожий на один із раніше запам'ятованих векторів із погляду визначеного критерію подібності, запам'ятований вектор буде змінюватися (навчатися) під впливом нового вхідного вектора, щоб стати більш схожим на цей вхідний вектор. Запам'ятований символ не буде змінюватися, якщо поточний вхідний вектор не виявиться досить схожим на нього.

Таким чином, вирішується проблема «стабільність–пластичність» роботи мережі при виконанні дій із розпізнавання символів. Новий образ може створювати додаткові класифікаційні категорії, однак новий вхідний образ не може змусити змінитися існуючу пам'ять.

Мережа АРТ навчається без учителя за допомогою зміни ваг таким чином, що пред'явлення мережі вхідного вектора змушує мережу активізувати нейрони в шарі розпізнавання, пов'язані з подібним запам'ятованим вектором. Крім цього, навчання проводиться у формі, що не руйнує запам'ятовані раніше образи, запобігаючи тимчасовій нестабільності, яка регулюється вибором критерію подібності. Новий вхідний образ (який мережа «не бачила» раніше) не буде відповідати запам'ятованим образам із погляду параметра подібності, тим самим формуючи новий образ, що запам'ятовується. Вхідний образ, у достатньому ступені відповідний одному із запам'ятованих образів, не буде формувати нового екземпляра, він просто буде модифікувати той, на який він схожий. Таким чином, при відповідному виборі критерію подібності запобігається запам'ятовування раніше вивчених образів і тимчасова нестабільність.

Для навчання нейронної мережі АРТ пропонується метод, за яким особа, що керує мережею, бере особисту участь у навчанні мережі, вона сама задає еталонні зображення всіх символів, а також перекручені зображення еталонів (зашумлені копії).

Раніше цей метод застосовувався до імовірнісних нейронних мереж, але використання цього методу для мереж багатошарових нейронів значно підвищить його продуктивність із розпізнавання символів.

Розробка системи розпізнавання символів

В якості головного компонента системи виступає видозмінений (багатошаровий) перцептрон (рис. 1). Кількість нейронів перцептрона визначається кількістю вивчених ним образів (m), кількість синапсів (n) визначається розширенням рисунків, які розпізнаються.

Розглянемо алгоритм зворотного поширення для перцептрона. Символ подається рисунком, розміри якого $79,38 \times 52,92$ мм, тип — бітовий, розширення *.bmp. Мережею рисунок розбивається на 600 частин (по горизонталі на 20, по вертикалі на 30). Отже, кількість синапсів для одного нейрона буде незмінним числом і визначатися як:

$$n = 20 * 30 = 600 .$$

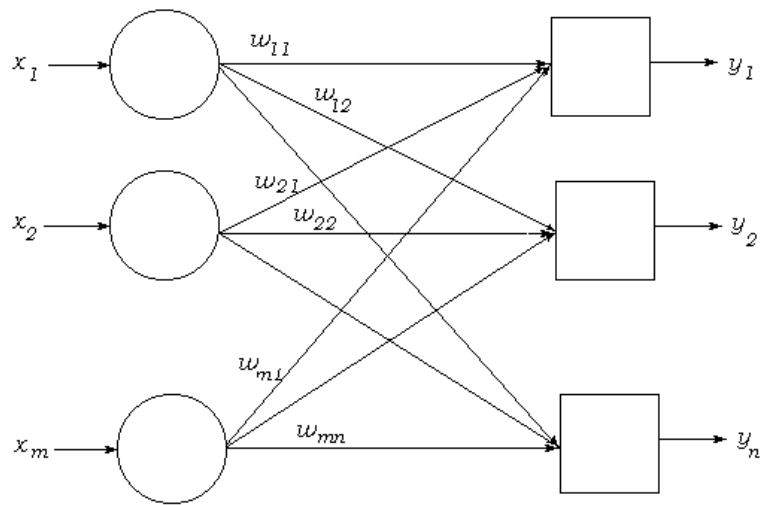


Рис. 1. Структура видозміненого перцептрона

Навчання перцептрона з урахуванням запропонованого підходу можна подати наступним чином [1].

1. Ініціалізуються всі ваги мережі в малі ненульові величини.

2. На вхід мережі подається вхідний навчальний вектор X вхідних символів і обчислюється сигнал NET вихідної функції від кожного нейрона, використовуючи формулу:

$$NET_j = \sum_i x_i w_{ij} ,$$

де x_i — i -й елемент вхідного вектора X ; w_{ij} — вага входу i нейрона j ; NET_j — рівень збудження післясинаптичного нейрона.

3. Обчислюється значення граничної функції активації для сигналу NET від кожного нейрона в такий спосіб:

$$OUT_j = 1, \text{ якщо } NET_j \text{ більше ніж поріг } \theta_j ;$$

$$OUT_j = 0 \text{ — у протилежному випадку,}$$

де θ_j являє собою поріг, що відповідає нейрону j (у найпростішому випадку всі нейрони мають той самий поріг).

4. Обчислюється помилка для кожного нейрона:

$$error_j = target_j - OUT_j ,$$

де $error_j$ — помилка для j -го нейрона; $target_j$ — необхідний вихід j -го нейрона; OUT_j — отриманий вихід j -го нейрона.

5. Кожна вага модифікується в такий спосіб:

$$w_{ij}(t+1) = w_{ij}(t) + \alpha x_i \text{error}_j,$$

де $w_{ij}(t+1)$ — вага входу i нейрона j в момент часу $(t+1)$; $w_{ij}(t)$ — вага входу i нейрона j в момент часу t ; α — нормуючий коефіцієнт навчання, початкове значення якого 0,1, і який зменшується в процесі навчання.

Повторюються кроки з другого до п'ятого, доки помилка не стане досить малою.

Робота інтелектуального модуля з розпізнавання символів полягає в підрахуванні виходів перцептрона й, якщо два або більше нейронів на виході отримали максимальне значення, то переглядається який нейрон має більшу суму на виході його суматора, і вносяться зміни до результату.

Активаційна функція нейронів має вигляд лінійного порога :

$$Y = S \text{ при } S < T ; \\ Y = 0 \text{ при } S \geq T ,$$

де Y — значення активаційної функції; S — аргумент активаційної функції; T — величина порога.

Подальше навчання перцептрона здійснюється під час роботи інтелектуального модуля штучних нейронних мереж із розпізнавання символів. Ваги синапсів корегуються відповідно Дельта-правилу:

$$\delta = (D - Y),$$

де δ — рівна різниці між необхідним або цільовим виходом D і реальним виходом Y .

Дельта-правило модифікує ваги відповідно до необхідного й дійсного значень виходу кожної полярності як для безперервних, так і для бінарних входів і виходів.

Алгоритм навчання зберігається якщо δ збільшується на величину кожного входу x_i , і цей добуток додається до відповідної ваги. З метою організації управління середньою величиною зміни ваг уведемо коефіцієнт «швидкості навчання» η , що збільшується на δx_i при навчанні:

$$\Delta_i = \eta \delta x_i , \\ w_i(n+1) = w_i(n) + \Delta_i ,$$

де Δ_i — корекція, пов'язана з i -м входом x_i ; η — коефіцієнт швидкості навчання; $w_i(n+1)$ — i -е значення ваги після корекції; $w_i(n)$ — i -е значення ваги до корекції.

Дельта-правило модифікує ваги відповідно до i -го необхідного дійсного значення виходу. Таким чином, використання зворотного поширення для перцептрона, дозволяє досягти ефективного розпізнавання навіть зашумлених символів.

Розробка алгоритму функціонування інтелектуального модуля розпізнавання символів

Розпізнавання символів повинно проводитись за таким алгоритмом:

- 1) завантаження параметрів по замовчуванню під час запуску програми;
- 2) створення еталонних зображень (створивши еталони, користувач може перенавчити мережу так, щоб створений ним образ символу був еталоном);
- 3) завантаження чи створення зображення символу для розпізнавання;
- 4) розпізнавання символу (зашумленість якого може бути до 40 %) системою;
- 5) виведення результату розпізнавання.

Зазначений алгоритм може бути реалізований у середовищі Borland Delphi 7, на мові Object Pascal, яка дозволяє використовувати бібліотеки, що реалізують нейронні мережі. Схему алгоритму функціонування модуля розпізнавання символів з урахуванням обраної мови програмування наведено на рис. 2.

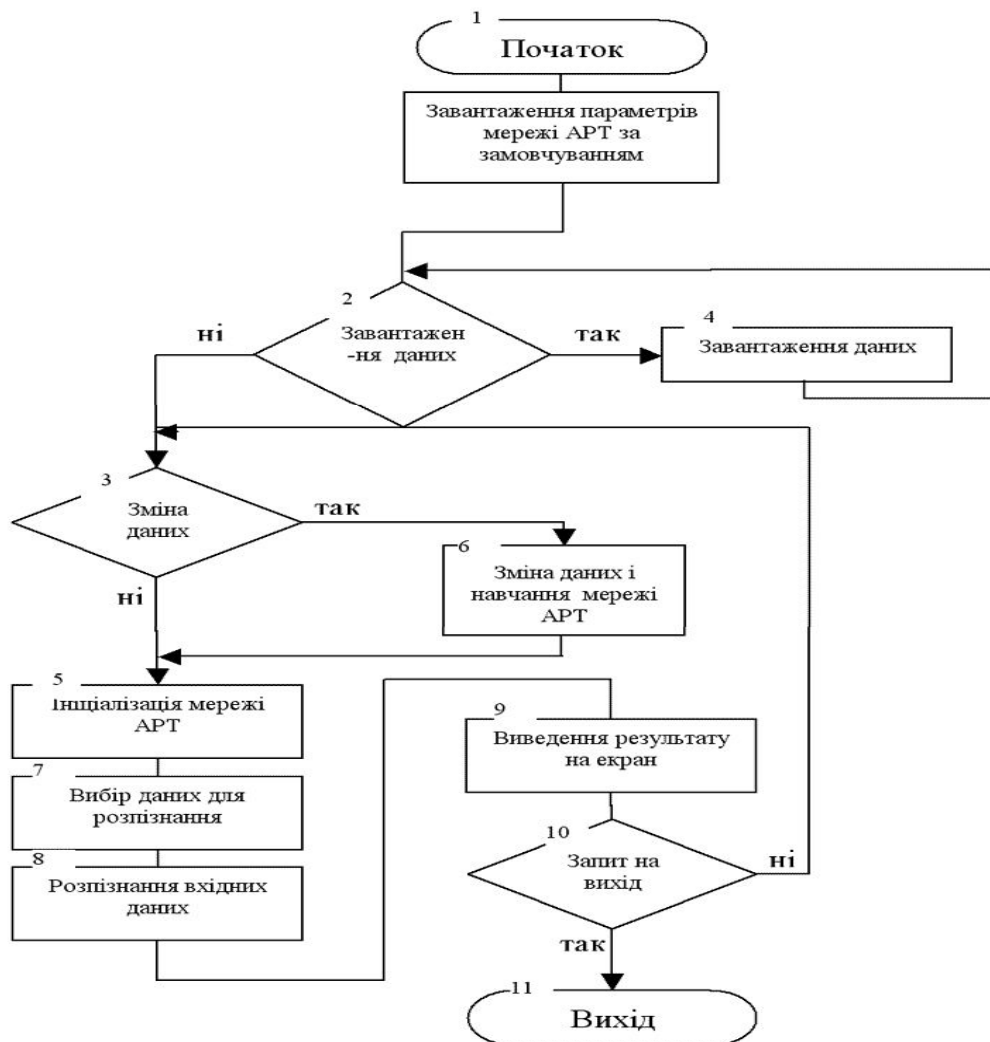


Рис. 2. Схема алгоритму нейронної мережі АРТ для розпізнавання

Результат реалізації приведеного алгоритму в середовищі програмування Borland Delphi 7 надав можливість реалізувати інтелектуальний модуль штучних нейронних мереж із розпізнавання символів.

Результати роботи інтелектуального модуля можна відобразити на рисунку, на якому зображені еталони образів і зашумлені копії, при яких програмний комплекс може правильно визначити образ (рис. 3). Зображення поділяються за відсотком зашумленості.

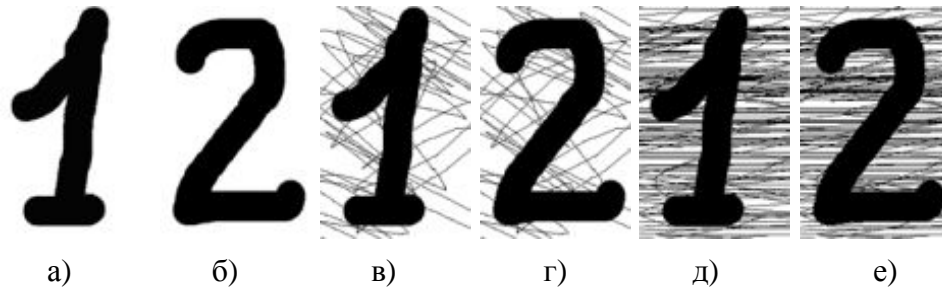


Рис. 3. Символи для роботи інтелектуального модуля штучних нейронних мереж із розпізнавання символів: а), б) — еталонні символи; в), г) — символи із зашумленістю до 30 %; д), е) —символи із зашумленістю до 40 %

У результаті реалізації інтелектуального модуля, що функціонує у відповідності із запропонованим алгоритмом, розпізнаються символи, які належать до будь-якого алфавіту, цифри й знаки в залежності від об'ємів пам'яті інтелектуальної системи, що містить відповідні еталони із зашумленістю до 40 %.

1. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика: Пер. на русский язык / Ю.А.Зуев, В.А.Точенов, 1992. — 184 с.

2. Месюра В.І., Ваховська Л.М. Основи проектування систем штучного інтелекту: Навчальний посібник. — Вінниця: ВДТУ, 2000. — 96 с. — укр. мовою.

Надійшла до редакції 24.11.2005