

УДК 004.032.26; 004.855.5

ДОСЛІДЖЕННЯ РОЗПІЗНАВАННЯ РУКОПИСНИХ ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ ПОВНОЗВ'ЯЗНИХ І НЕПОВНОЗВ'ЯЗНИХ НЕЙРОМЕРЕЖ ТА ГЕНЕТИЧНИХ АЛГОРИТМІВ

А. Фаренюк¹, З. Любунь¹, О. Фаренюк²

¹ Львівський національний університет ім. І. Франка,
вул. Ген. Тарнавського, 107, 79017, Львів, Україна

² Інститут фізики конденсованих систем НАН України
ул. Свенцицького 1, 79001 Львів, Україна

Одним з найскладніших завдань автоматизованих систем є розпізнавання рукописних символів. Незважаючи на актуальність, воно ще дуже далеке від вирішення. Поки не запропоновано ефективних алгоритмів розпізнавання, та й не відомо, чи вони взагалі існують, тому природно залучити системи штучного інтелекту, використавши властиву їм адаптивність. Для розпізнавання ми використали як повнозв'язні, так і не повнозв'язні багатошарові нейромережі (багатошарові перцептрони). Для вибору оптимального режиму роботи мережі застосовано генетичні алгоритми. Навчальна вибірка складається з наборів цифр від «0» до «9», та знаків «+», «-» і «,», написаних різними людьми.

Ключові слова: нейромережа, перцептрон, генетичний алгоритм, адаптивні алгоритми, розпізнавання рукописних символів.

Концепція нейромереж виникла внаслідок дослідження процесів мислення та створення штучного інтелекту для моделювання роботи окремих нейронів та їхніх зв'язків [1]. Після десятиліть розвитку, серії підйомів і спадів зацікавленості та багатьох відкриттів нейромережі хоч і не дали нам повного розуміння процесів мислення, проте стали важливим інструментом розв'язування задач, що традиційно називають інтелектуальними. Однією із таких задач є розпізнавання образів [2].

Наше завдання – побудувати нейромережу, яка б уміла розпізнавати рукописні цифри від 0 до 9, а також знаки «+», «-» та «,». Приклад набору зображень наведено на Рис. 1. Для розпізнавання використано багатошаровий перцептрон, який навчали методом зворотного поширення похибки. Цей метод можна застосовувати для навчання одно- та багатошарових нейронних мереж, у яких немає прямих і зворотних зв'язків, що перестрибують через шари. Кожен нейрон попереднього шару зв'язаний тільки з нейронами наступного шару. Метод зворотного поширення похибки полягає у послідовному поширенні похибки від

вищих шарів до нижчих. Відповідно, корекцію ваг визначають за допомогою градієнтних методів. Для всіх нейронів використовували сигмоїдальну функцію активації.

Структурна схема повнозв'язної нейромережі показана на рис. 2

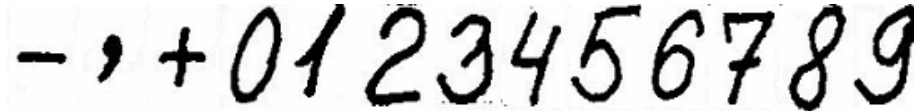


Рис 1. Приклад набору символі для навчання та розпізнавання

Корекцію ваг під час навчання нейромережі обчислюють за формулою

$$\Delta w^t = -\alpha \frac{\partial E}{\partial w} + M \Delta w^{t-1},$$

де t – номер кроку навчання; Δw^t – значення корекції ваг на кроці t ; $\frac{\partial E}{\partial w}$ –

градієнт похибки; α – швидкість навчання, M – момент навчання.

Нейромережі названого типу та процес їхнього навчання характеризують такими загальними глобальними параметрами:

швидкість навчання α – відображає, наскільки ми враховуємо запропоновані алгоритмом поправки;

момент навчання M – інертність алгоритму, міра врахування попередніх значень ваг кожного нейрона;

очікувана похибка ε – значення похибки, у разі досягнення якого навчання мережі припиняється;

коефіцієнт сигмоїди;

Кількість нейронів прихованого шару.

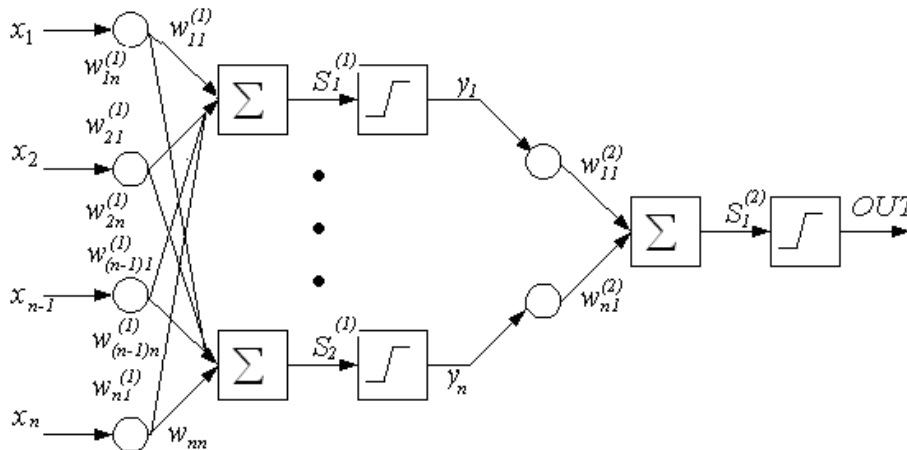


Рис 2. Структура повнозв'язного перцептрона.

Для оцінки роботи нейромережі використовують *функцію ефективності розпізнавання* (ефективність розпізнавання невідомих символів) – відношення кількості правильно розпізнаних символів до загальної кількості символів, заданих для розпізнавання.

Повнозв'язна нейромережа загалом добре розпізнавала знаки, на яких її навчали, однак погано працювала на символах, які не використовували для її навчання. Також вона дуже чутлива до геометричних спотворень зображень, хоч добре справляється з шумом. Крім того, її навчання потребує великих затрат машинного часу.

Унаслідок цього надалі використовували неповнозв'язну нейромережу, яку можна уявити як сукупність кількох повнозв'язних нейромереж. Причому кожен її структурний елемент відповідає за розпізнавання лише одного символа і повинен уміти виділяти його серед будь-яких інших. Така мережа працювала дещо ліпше, однак також не давала очікуваної ефективності розпізнавання.

Для поліпшення розпізнавання невідомих символів було вирішено підібрати оптимальні глобальні параметри нейромережі та процесу навчання, перераховані вище. Такий підбір, який є багатовимірною задачею оптимізації, вирішили виконати за допомогою генетичних алгоритмів (ГА) [4]. ГА є ефективним способом пошуку екстремумів багатовимірних функцій, особливо для випадків, коли достатньо знайти хороший результат і нема потреби шукати ідеальний (глобальний екстремум). Оптимізацію проводили за ефективністю розпізнавання невідомих символів – тобто мінімізували відносну частку неправильно розпізнаних символів. Однак цілком можливою є також оптимізація за часом навчання, часом розпізнавання, використанням машинних ресурсів, що є важливим для вбудованих пристроїв і т.д. Від нейромережі також домагаються доброго розпізнавання зразків, на яких вона навчалася – так обходили потребу оптимізації кількох функцій одночасно: функцій ефективності для відомих та невідомих зразків.

Доцільність використання такого підбору параметрів проілюструємо на прикладі залежності ефективності розпізнавання нейромережі від двох параметрів – *швидкості навчання* α та *моменту навчання* M за фіксованих інших. Навчальну вибірку ділили на дві рівні частини. Спочатку нейромережу вчили розпізнавати на першій частині, а тестували на другій. Потім навпаки. Розпізнавання вибірки, на якій відбувалося навчання в обох випадках, – практично 100 %. Як бачимо з рис.3, 4, рельєф *функції ефективності* дуже складний, значно залежить від навчальної вибірки і навряд чи піддається аналітичному дослідженню, не говорячи про загальноприйняте “вгадування” параметрів. Зазначимо, що табулювання зі щораз меншим кроком виявляє нові вузькі ущелини та піки.

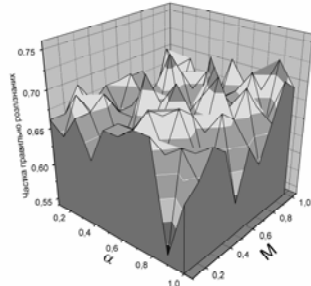


Рис.3

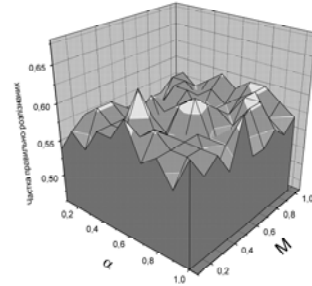


Рис. 4.

Функція ефективності розпізнавання
(Коефіцієнт. сигмоїди - 1, нейронів у прихованому шарі – 50, очікувана похибка – 0,2)

Крім того, для тієї ж навчальної вибірки, але різних значень очікуваної похибки залежність ефективності розпізнавання від швидкості та моменту навчання теж буде дуже відрізнятися, що проілюстровано на рис. 5, 6, очікувана похибка для яких становить 0,1 та 0,2, відповідно.

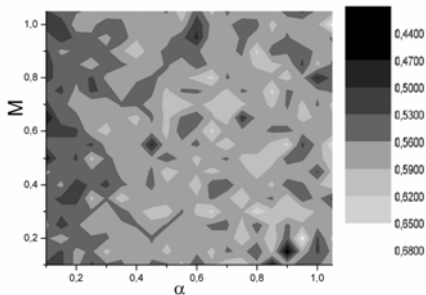


Рис. 5

очікувана похибка – 0.1

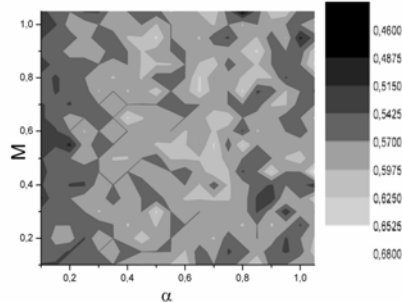


Рис. 6

очікувана похибка – 0.2

Функція ефективності розпізнавання
(Коефіцієнт сигмоїди - 1, нейронів у прихованому шарі – 50)

Також оптимальне розпізнавання різних символів потребує різних наборів параметрів. Це підтверджує доцільність використовувати неповнозв'язні перцептрони. Зокрема, для першої вибірки та того ж набору параметрів “+” – розпізнається добре, легко досягти 100 % розпізнавання, а от “,” – розпізнається погано, що видно з рис. 7, 8, відповідно.

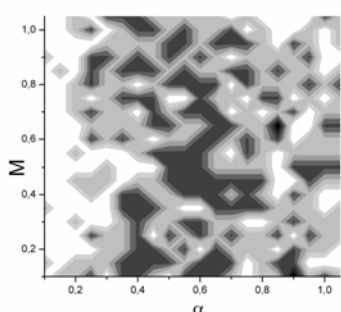


Рис.7

“+” – розпізнається добре

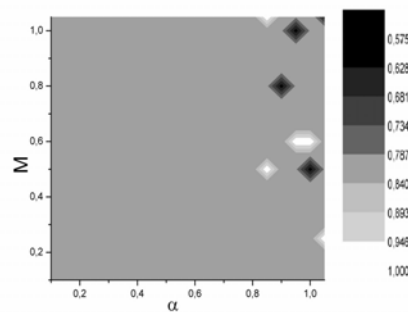


Рис.8

“;” – розпізнається погано.

Функція ефективності розпізнавання

(Коефіцієнт сигмоїди - 1, нейронів у прихованому шарі – 50)

Аналогічна ситуація з розпізнаванням конкретного символу. Ефективність розпізнавання без попередньої підгонки параметрів дуже чутлива до конкретних способів написання тих же символів різними людьми. Це відображено на рис. 9, 10, на прикладі цифри “9”.

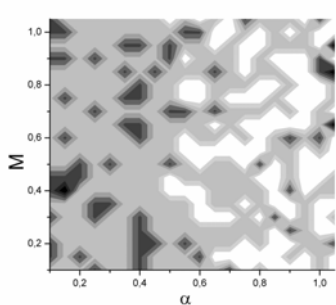


Рис.9

“9”, розпізнається добре.

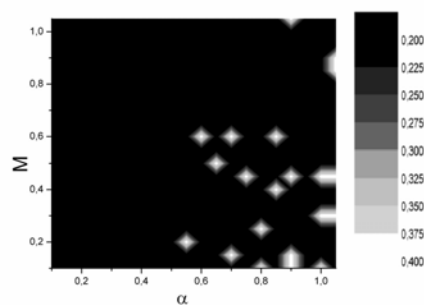


Рис.10

“9”, розпізнається погано.

Функція ефективності розпізнавання

(Коефіцієнт сигмоїди - 1, нейронів у прихованому шарі – 50)

Дослідження оптимальних значень глобальних параметрів відбувалося для повнозв'язних і неповнозв'язних нейромереж, при тому в останньому випадку для кожної окремої підмережі підбирали свій набір параметрів. Використовували простий генетичний алгоритм. Щоб отримати загальне уявлення про простір пошуку, виконували серію послідовних запусків. Оскільки нам потрібні оптимальні параметри для роботи в середовищі з великим впливом випадкових, малопередбачуваних факторів, зокрема неможливо сказати наперед, які варіанти написання будуть, то для роботи ГА використовували таку стохастичну процедуру.

1. Створюють популяцію різноманітних наборів параметрів.
2. Для кожного такого набору – особи в термінах ГА, навчальну вибірку випадковим способом розбивають на дві групи: група для навчання нейромережі та група для пробного розпізнавання навченою мережею. Мережа навчається, від неї домагаються хорошого (не обов'язково ідеального) розпізнавання зразків, на яких вона вчилася. Визначають ефективність розпізнавання нейромережею зразків, що не входили в навчальну вибірку. Процес повторюють декілька разів, визначають середню ефективність розпізнавання нейромережі для різних навчальних наборів.
3. Отриману на попередньому етапі ефективність розпізнавання нейромережею невідомих зразків використовують як пристосованість цієї особи. Виконують один крок простого ГА, генеруючи нову популяцію.
4. Переходять до кроку 2.

Розбиття навчальної вибірки на дві випадкові групи є важливим, оскільки в іншому випадку – фіксованого розбиття, ми отримаємо мережу, оптимізовану під конкретну задачу – вчитися на одній множині так, щоб розпізнавати іншу.

Завдяки використанню такого методу вдалося досягнути ефективності розпізнавання 86 ± 5 % для повнозв'язної мережі. Для неповнозв'язної мережі результати ще кращі 92 ± 4 %, при чому кожна з підмереж досягає ефективності 95 – 98 % розпізнавання “свого” символу. Для порівняння, підбираючи параметри вручну, вдається досягнути ефективності розпізнавання 55 – 60 %. Також не виключено існування ще ліпших наборів параметрів, які будуть знайдені в процесі подальшої роботи ГА. Однак просунутися суттєво далі в цьому напрямі навряд чи вдасться, і для подальшого поліпшення ефективності очевидно є необхідність попередньої обробки символів – підгонку під один масштаб, кут нахилу тощо, що спрощуватиме задачу для нейромережі.

У процесі роботи з нейромережами експериментально з'ясовано, що: важко досягти якісного розпізнавання таких символів, як “1”, “3”, “6”, “8”, “9”; легко досягти розпізнавання таких символів, як “4”, “7”, “2” та “5”.

Також продемонстровано наявність широкого простору для оптимізації за часом навчання та роботи нейромережі. Цей висновок ґрунтується на тому, що приблизно тієї ж ефективності розпізнавання можна досягнути при різних наборах параметрів та різних кількостях нейронів. Наприклад, для цифри “3” отримано ефективність розпізнавання (99 % / 93 %) за використання 22 нейронів і певного фіксованого набору параметрів мережі, проте для досягнення трішки меншої ефективності розпізнавання (98 % / 92 %) достатньо всього 11 нейронів та іншого набору параметрів.

1. *Уоссермен Ф.* Нейрокомпьютерная техника: Теория и практика. М.: Мир, 1992. 184 с.
2. *Egmont-Petersen, M., de Ridder, D., Handels, H.* Image processing with neural networks – a review. // *Pattern Recognition*. 2002. N35 (10): С. 2279–2301.
3. *Rojas R.* Neural networks - A Systematic Introduction. Berlin, New-York, Springer-Verlag, 1996. 502 p.
4. *Goldberg D.E.* Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, USA, 1989.

**INVESTIGATION OF RECOGNITION OF HAND-WRITTEN SYMBOLS,
USING NEURAL NETWORKS AND GENETIC ALGORITHMS****A. Ya. Farenyuk¹, Z.M. Lubun¹, O. Ya. Farenyuk²***¹Ivan Franko National University of L'viv
Tarnavsky Str. 107, UA-79017 Lviv, Ukraine**²Institute for Condensed Matter Physics, National Academy of Sciences of Ukraine
Svientsitskii Str. 1, UA-79011 Lviv, Ukraine*

Automated recognition of hand-written symbols is one of most difficult computing tasks. Although being very important, it is far from being solved. There are no known algorithms for it, and no one can be sure, that they exist at all. So the most practical way is to use some Artificial Intelligence systems and utilize their inherent adaptability. In this paper both fully connected and not fully connected multilayered neural networks (multilayered perceptrons) were used. Optimal network modes were selected by using Genetic Algorithms. Symbols for recognition was numbers from "0" to "9" and symbols "+", "-", ",", written by many different peoples.

Key words: neural net, perceptron, genetic algorithms, adaptive algorithms, handwritten text recognition.

**ИССЛЕДОВАНИЕ РАСПОЗНАВАНИЯ РУКОПИСНЫХ ИЗОБРАЖЕНИЙ
С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ И ГЕНЕТИЧЕСКИХ
АЛГОРИТМОВ****А.Я. Фаренюк¹, З.М. Любунь¹, О.Я. Фаренюк²***¹ Львовский национальный университет имени Ивана Франка
ул. Ген. Тарнавського, 107, 79017 Львов, Украина**² Институт физики конденсированного состояния НАН Украины,
ул. Свенцицкого 1, 79001 Львов, Украина*

Распознавание рукописных символов – одна из наиболее сложных задач для автоматизированных систем. Не смотря на большую актуальность, она все еще очень далека от решения. До сих пор не предложено эффективных алгоритмов ее решения, да и неизвестно, существуют ли они вообще, поэтому возникает желание использовать те или иные системы искусственного интеллекта, используя характерную для них адаптивность. Для распознавания использовано полносвязные и неполносвязные нейронные сети (многослойные перцептроны), для выбора оптимального режима работы сети – генетические алгоритмы. Учебная выборка состояла из набора цифр от «0» до «9» и символов «+», «-», «,», написанных разными людьми.

Ключевые слова: нейронная сеть, перцептрон, генетические алгоритмы, адаптивные алгоритмы, распознавание рукописных символов.

Стаття надійшла до редколегії 03.03.2010

Прийнята до друку 25.05.2010