

УДК 004.932.75`1

ЗАСТОСУВАННЯ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ З ПРОСТОЮ АРХІТЕКТУРОЮ ДЛЯ РОЗПІЗНАВАННЯ РУКОПИСНИХ ЦИФР

Лукович В.В.

*Міжнародний науково-навчальний центр інформаційних технологій та систем
НАН і МОН України*

Вступ

Відомо, що багат шарові перцептрони, які навчаються за методом зворотного розповсюдження похибки (back propagation), можуть реалізувати складні, багатовимірні та нелінійні відображення. Це дозволяє застосовувати такі мережі для розв'язування практичних задач розпізнавання. Але пряме використання неструктурованих багат шарових перцептронів для розпізнавання зображень не є оптимальним. У такому разі ігноруються суттєві властивості зображень: наявність локальних ознак та помітна кореляція значень сусідніх елементів зображення. Для розпізнавання зображень більш доцільно використовувати нейронну мережу, певним чином спеціалізовану для роботи саме з зображеннями.

Мета

Метою даної роботи було дослідження простої архітектури нейронної мережі в задачі розпізнавання зображень рукописних цифр. Робота виконувалася з використанням нейромережевого підходу, який розвивається у відділі нейросіткових технологій на основі ідей М.М. Амосова [1], [2] та його наукової школи[3].

Згорткові нейронні мережі

Одним з варіантів спеціалізованої нейронної мережі для розпізнавання зображень є архітектура згорткової нейронної мережі [4]. Основні принципи побудови цієї архітектури полягають в наступному:

- 1) Мережа в цілому є багат шаровим перцептроном з прямими зв'язками.
- 2) Нейрони у кількох перших шарах мають 2-вимірні локальні рецептивні поля. Ці нейрони формують 2-вимірні карти ознак способом, подібним до математичної операції згортки (convolution). Над картами ознак виконується процедура зменшення просторової роздільної здатності.
- 3) Входом першого шару нейронів-детекторів ознак є вхідне зображення. Входом кожного з наступних шарів детекторів ознак є карти ознак з зменшеною

просторовою роздільною здатністю, сформовані попереднім шаром детекторів ознак.

Архітектура згорткової нейронної мережі інтегрує виділення елементарних ознак зображення, формування більш складних ознак на вищих рівнях обробки та розпізнавання. Навчання всієї мережі виконується єдиним алгоритмом зворотного розповсюдження похибки (back propagation).

Дослідження виконувалися з згортковою нейронною мережею, яка мала 20 нейронів у першому шарі, 80 нейронів у третьому та 10 нейронів у четвертому шарі. Другий шар забезпечував зменшення просторової роздільної здатності першої карти ознак, нейрони другого шару мали постійні параметри. Загальна кількість параметрів для навчання мережі становила 61390. Для навчання мережі був використаний метод back propagation. В якості функції помилки мережі використано крос-ентропію. Перерахунок параметрів нейронів мережі здійснювався за методом стохастичного градієнту (on-line), тобто параметри нейронів модифікувалися після кожного пред'явлення зразка з навчальної вибірки. Черговість подачі зразків змінювалася для кожної ітерації навчання шляхом випадкової перестановки зразків у послідовності. Навчання мережі на поточному зразку здійснювалося не тільки у випадку неправильної його класифікації, але й тоді, коли абсолютне значення помилки на виході перевищувало 0.1. Для покращення результату навчання нейронної мережі було використане штучне збільшення обсягу навчальної вибірки шляхом деформації зразків, що належать до неї. Параметри деформації генерувалися випадково та незалежно для кожного пред'явлення зразка, таким чином навчальна вибірка віртуально стала необмеженою. Зображення деформувалися за методом, близьким до запропонованого в [9].

Аналогічно до роботи [7], для навчання нейронної мережі були використані перших 50000 зображень цифр з навчальної вибірки. Решта 10000 зображень цифр використовувалися як валідаційна вибірка. Навчання мережі виконувалося протягом 100 ітерацій. Після кожної ітерації виконувалося розпізнавання валідаційної вибірки. Для ітерації з мінімальною кількістю помилок на валідаційній вибірці стан мережі фіксувався та вважався результатом процесу навчання. Після навчання мережі виконувалося розпізнавання тестової вибірки. Кращий результат розпізнавання тестової вибірки з серії експериментів наведений в таблиці 1 (0.38 відсотка помилок або 38 помилок з 10000 зображень, що належать до тестової вибірки).

База даних

Для досліджень була використана база даних MNIST, яку склав Y. LeCun [5]. Вона має широке застосування для оцінювання якості алгоритмів розпізнавання та їх порівняння. База даних містить 70000 зображень рукописних

цифр від 0 до 9. 60000 зображень складають навчальну вибірку, контрольна вибірка має 10000 зображень цифр. В Інтернеті цю базу даних можна знайти за адресою <http://yann.lecun.com/exdb/mnist/>.

Таблиця 1. Якість розпізнавання контрольної вибірки з бази даних MNIST для різних методів побудови класифікатора.

Метод	Відсоток помилок
Linear classifier [5]	12
kNN Euclidean [5]	5.0
LeNet1 (CNN) [4]	1.7
LeNet4 (CNN) [5]	1.1
LeNet5 (CNN) [5]	0.8
My CNN w/RBF [8]	0.57
PCNC classifier [10]	0.44
SVM w/RBF kernels [11]	0.42
Simard's Simple CNN [7]	0.4
LeNet6 (CNN) [6]	0.38
My Simple CNN (ця робота)	0.38
Classifier combination [12]	0.35
Huge MLP [13]	0.35

Результати

Кращий результат розпізнавання тестової вибірки з серії експериментів наведений в табл. 1 (0.38 відсотка помилок або 38 помилок з 10000 зображень, що належать до тестової вибірки).

Висновки

В даній роботі запропоновано варіант архітектури простої згорткової нейронної мережі для розпізнавання зображень рукописних цифр. Експериментальна перевірка розпізнавання рукописних цифр на базі даних MNIST дала результат, близький до найкращих відомих на цей час. Це означає, що запропонована проста згорткова нейронна мережа спроможна ефективно розв'язувати задачу розпізнавання рукописних цифр.

Література

1. Амосов Н.М. Моделирование мышления и психики / Н.М. Амосов. – Киев: Наукова думка, 1965. – 304 с.
Amosov N. M. Modelling of thinking and the mind / N. M. Amosov. – New York: Spartan Books, 1967.

2. Амосов Н.М. Алгоритмы разума / Н.М. Амосов. – Киев: Наукова думка, 1979. – 223 с.
Amosov, N.M. Algorithms of the Mind / N. M. Amosov. – Kiev: Naukova Dumka, 1979. 223 p.
3. Kussul E. M. Neural Networks and Micro Mechanics / E.M. Kussul, T.N. Baidyk, D.C. Wunsch. – New York: Springer, 2010. – 210 p.
4. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., and Jackel, L.D. Backpropagation applied to handwritten zip code recognition, *Neural Computation*, vol. 1, No 4, p. 541-551, 1989.
5. LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. Gradient-Based Learning Applied to Document Recognition / LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. // *Proceedings of the IEEE*, Vol. 86, No. 11, November 1998, p. 2278-2324.
6. Bottou, L., Chapelle O., DeCoste D., and Weston J. Large-Scale Kernel Machines/ Bottou, L., Chapelle O., DeCoste D., and Weston J. // MIT Press, Cambridge, USA, 2007. - 416 pp.
7. Simard, P.Y., Steinkraus, D., and Platt, J.C. Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis / Simard, P.Y., Steinkraus, D., and Platt, J.C. // *Proceedings of the Seventh International Conference on Document Analysis and Recognition*, August 03 - 06, 2003 Edinburgh, Scotland, p. 958-962.
8. Лукович В. Згорткова нейронна мережа для розпізнавання рукописних цифр / Лукович В. // Пр. 8-ї міжнар. конф. УкрОбраз'2006, Київ, 2006. – С.135-138.
Lukovych V. Convolutional Neural Network for Handwritten Digit Recognition / Lukovych V. // *Proceedings of the 8-th all-ukrainian international conference UkrOBRAZ'2006* Kiev, Aug. 28-31 2006, pp. 135-138.
9. Gosselin, B., Improved Hand-written Character Recognition thanks to a New Geometric Distortion Method / Gosselin, B. // *Proc. of the 6th Int. Conf. On Image Processing and its Applications*, 1997, Dublin, Ireland, vol. 1, p. 327-331.
10. Kussul, E., Baidyk, T., Wunsch, D. C., Makeyev, O., and Martin, A. Permutation Coding Technique for Image Recognition Systems / Kussul, E., Baidyk, T., Wunsch, D. C., Makeyev, O., and Martin, A. // *IEEE Transactions on Neural Networks*. - 2006. – vol. 17, No 6. - p.1566-1579.
11. Liu, C.-L., Nakashima, K., Sako, H., Fujisawa, H. Handwritten Digit recognition: benchmarking of state-of-the-art techniques / Liu, C.-L., Nakashima, K., Sako, H., Fujisawa, H. // *Pattern Recognition*, 36(10):p. 2271-2285, 2003.
12. Keysers, D. Comparison and Combination of State-of-the-art Techniques for Handwritten Character Recognition: Topping the MNIST Benchmark / Keysers, D. // *Technical Report*, IUPR Research Group, DFKI and Technical University of Kaiserslautern, May 2006.

13. Ciresan, D.C., Gambardella, L.M., Schmidhuber, J.Meier, U., Deep Big Simple Neural Nets Excel on Handwritten Digit Recognition, available at: <http://arxiv.org/pdf/1003.0358v1>.