

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ СТИСНЕННЯ ЗОБРАЖЕНЬ НА ОСНОВІ НЕЙРОННОЇ МЕРЕЖІ ПРЯМОГО ПОШИРЕННЯ

Вінницький національний технічний університет

Анотація

Розглянуто актуальність задачі стиснення зображень на основі нейронної мережі прямого поширення для інтелектуального модуля стиснення зображень. Наведено покращений алгоритм роботи і навчання нейронної мережі.

Ключові слова: стиснення зображень, компресія, алгоритм стиснення, нейронна мережа.

Abstract

We consider the relevance of image compression algorithms based on neural network direct distribution for intelligent image compression module. An improved algorithm and neural network training depicted.

Keywords: compression scheme, vector quantization, discrete cosine transform, neural network.

Вступ

З розвитком цифрової обробки даних [1] зросли вимоги щодо розміру зображень, швидкості стискання та якості їх передачі по каналах зв'язку, але всі існуючі на сьогодні алгоритми не у змозі задовольнити таку вимогу. Саме тому виникає потреба пошуку нових методів компресії, які виконують більш інтелектуальний аналіз даних та є кращими по тих чи інших характеристиках стиснення. За таких умов перспективним виглядає нейромережевий підхід [2,3] до стиснення даних.

Отже виникає потреба провести дослідження штучних нейронних мереж як засобу для стиснення зображень, в ході якого буде запропоновано та реалізовано алгоритм стиснення зображень.

Методи стиснення зображень на основі нейронних мереж

На відміну від традиційних методів стиснення [4,5] - математичного обчислення і видалення надмірності - нейронна мережа при вирішенні задачі стиснення виходить з міркувань нестачі ресурсів. Топологія мережі і її алгоритм навчання такі, що дані великої розмірності потрібно передати з входу нейронної мережі на її виході через канал порівняно невеликих розмірів. Для реалізації стиснення такого роду може використовуватися багат шаровий перцептрон наступної архітектури: кількість нейронів у вхідному і вихідному шарі однаково і дорівнює розмірності стиснених даних; між цими шарами розташовуються один або більше проміжних шарів меншого розміру. Число проміжних шарів визначає ступінь складності перетворення даних. Наприклад, мережа з трьома проміжними шарами може краще виконувати стиснення на навчальних даних, але може дати найгірший результат в реальних ситуаціях. Це пов'язано з тим, що у вихідних даних може випадково утворитися якась залежність, яка не має ніякого відношення до реальності.

Вихідні дані для мережі складаються таким чином, щоб на виходах був завжди той же набір сигналів, що і на вході. В процесі роботи алгоритм зворотного поширення помилки мінімізує помилку. Це означає, що ваги зв'язків від вхідного шару нейронів і, приблизно, до серединного шару будуть працювати на компресію сигналу, а решта - на його декомпресію. При практичному використанні отриману мережу розбивають на дві [6]. Висновок першої мережі передають по каналу зв'язку і подають на вхід другої, яка здійснює декомпресію.

Іншим способом вирішення задачі стиснення є застосування автоасоціативної пам'яті, такої як мережа Хопфілда, так як вона має здатність відновлювати сигнал по його пошкодженому образу [7].

Розробка алгоритму стиснення зображень

На основі аналізу типів нейронних мереж, методів та засобів стиснення зображень пропонується використати для вирішення цієї задачі мережу прямого поширення з двома шарами нейронів (Рисунок 1) в комбінації з алгоритмом Левенберга-Марквардта.

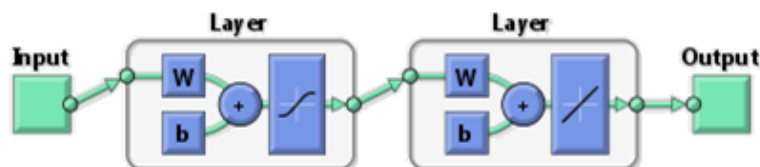


Рисунок 1 – Архітектура нейронної мережі прямого поширення з двома шарами нейронів.

Стиснення зображення з використанням нейронної мережі прямого поширення складається з наступних кроків:

- Зображення F ділиться на $R \times C$ блоків пікселів. Кожен блок потім сканується для формування вхідного вектора $x(n)$ розміру $p = r \times c$;
- Передбачається, що прихований шар мережевого рівня складається з L нейронів кожен з P синапами і він характеризується відповідно до обраної вагової матриці W_h .
- Всі N блоків вхідного зображення передаються через прихований шар, щоб отримати приховані сигнали h_n , які представляють собою закодований набір вхідних блоків зображення, $x(n)$, якщо $L < P$, то таке кодування забезпечує стиснення зображення.
- Передбачається, що вихідний шар складається з $m = p = r \times c$ нейронів, кожен з яких має L синапсів. Нехай W_y буде, відповідним чином обрана, матриця виведення ваги. Весь N прихований вектор h_n представляє кодоване зображення H , яке передається через вихідний шар, щоб отримати вихідний сигнал u_n . Вихідні сигнали повторно зібрані в $p = r \times c$ блоків зображення, щоб отримати відтворене зображення, F_r .

Існують дві матриці помилок, які використовуються для порівняння різних методів стиснення зображення. Це матриця середньоквадратичної похибки (MSE) та максимального співвідношення сигналу до шуму (PSNR). MSE - це сукупність квадратичної помилки стисненого та вихідного зображення, тоді як PSNR є мірою пікової помилки [4,6].

$$MSE = I / MN \sum_{y=1}^m \sum_{x=1}^n [I(x, y) - I'(x, y)]^2$$

Якість кодування зображень, як правило, оцінюється по відношенню сигналу до шуму (PSNR) і визначається як [2]:

$$PSNR = 20 \lg \frac{\max_i P_i}{\sqrt{MSE}}$$

Навчання проводиться для представницького класу зображень з використанням алгоритму Левенберга-Марквардта.

Після того, як вагові матриці були відповідним чином обрані, будь-яке зображення може бути швидко закодоване з використанням матриці W_h , а потім декодоване з використанням матриці W_y .

Висновки

Найбільшою перевагою нейронних мереж в задачі стиснення зображень є висока швидкість обробки, яка забезпечується за рахунок паралельної реалізації [8].

Методи стиснення, які засновані суто на штучних нейронних мережах використовуються лише для певних типів даних та мають ряд недоліків та обмежень. Штучні нейронні мережі – це потужний інструмент, але сам по собі не дає ефективних результатів при стисненні зображень. Тому їх варто застосовувати в комбінації з іншими методами, або як проміжний етап в алгоритмі стиснення.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Савчук Т. О. Інформаційна технологія масштабування хмарного застосунку з неперіодичними піками навантаження / Т. О. Савчук, С. І. Петришин // Вісник Національного університету «Львівська політехніка». – 2014. – № 783 : Інформаційні системи та мережі. – С. 187–193. – Бібліографія: 5 назв.
2. Руденко О.В. Штучні нейронні мережі: Навчальний посібник / О.В.Руденко, Є.В.Бодяньський. - Харків : ТОВ «Компанія СМІТ», 2006. — 404 с. - ISBN 966-8630-73-Х.
3. В.Ф.Бардаченко, О.К.Колесницький, С.А.Василецький. Перспективи застосування імпульсних нейронних мереж з таймерним представленням інформації для розпізнавання динамічних образів// УСiМ.-2003-№6.- С. 73-82.
4. Сэломон, Дж. Сжатие данных, изображений и звука [Текст] / Дж. Сэломон : пер. с англ. – Москва : Техносфера, 2004. – 366 с.
5. Майданюк В. П., Кожемяко К. В., Арсенюк І. Р. Нейроподібні методи ущільнення зображень. // Оптико-електронні інформаційно-енергетичні технології - 2009 - №1(17). - С. 37 - 41.
6. Руденко О.Г. Сжатие изображений на основе нейронной сети ART / О.Г. Руденко, М.С. Сныткин // Кибернетика и системный анализ. – 2008. – № 6. – С. 10-16.
7. Олефір А.О. Застосування штучних нейронних мереж типу перцептрон та мереж Хопфілда для стиснення даних [Текст] / А.О. Олефір, М.В. Превір. – Прикладна математика та комп'ютеринг 2015. Збірник тез доповідей,. – Київ : Просвіта, 2015. – с. 211-216.
8. Колесницький О. К. Аналітичний огляд апаратних реалізацій спайкових нейронних мереж / О. К. Колесницький // Математичні машини і системи. – 2015. – №1, С.3-19. ISSN 1028-9763.

Міщанчук Андрій Павлович — студент групи 2КН–16м, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, м. Вінниця, e-mail: adriy.mishchanchuk@gmail.com;

Науковий керівник: **Колесницький Олег Костянтинович**— к. т. н., доцент, доцент кафедри комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця.

Andriy P. Mishchanchuk — Department of Information Technology and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: adriy.mishchanchuk@gmail.com;

Supervisor: **Kolesnytskyj K. Oleg** — Cand. Sc. (Eng), Assistant Professor, Assistant Professor of the Chair of Computer Science, Vinnytsia National Technical University, Vinnytsia.