

## ПІДВИЩЕННЯ РОЗДІЛЬНОЇ ЗДАТНОСТІ ЗОБРАЖЕНЬ НА ОСНОВІ НЕЙРОННОЇ МЕРЕЖІ

Олексієвець О.В., студент

Демчишин А.А., к.т.н., доцент, НН ІАТЕ

Національний технічний університет України

«Київський політехнічний інститут ім. І. Сікорського»

(Україна, м. Київ)

***Анотація** – предметом дослідження є нейромережеві методи вирішення задачі підвищення роздільної здатності зображень. У статті розглянуто визначено особливості обробки зображення нейронними мережами та запропоновано можливі способи оптимізації навчання мережі надроздільної здатності. Методами дослідження виступають методи інтерполяції зображень, метод зворотного поширення помилки, метод нормалізації пакетів, метод зниження вагових коефіцієнтів мережі, метод випадання нейронів. Результати дослідження порівняно з результатами роботи методу бікубічної інтерполяції для збільшення зображень та продемонстровано, що запропонована модифікована мережа надроздільної здатності дає кращий результат, ніж результат, отриманий після обробки зображення бікубічною інтерполяцією. Розглянута в статті модифікація мережі надроздільної здатності дозволяє покращити якість результату роботи мережі та пришвидшує швидкість навчання мережі.*

***Ключові слова** – підвищення роздільної здатності, методи інтерполяції, нейронна мережа, мережі глибокого навчання, SRCNN, нормалізація пакетів, випадання нейронів, зменшення вагових коефіцієнтів.*

**Постановка проблеми.** Зображення є одним із найбільш поширених та важливих типів даних в сучасному світі. Вони використовуються в багатьох галузях, таких як медицина, промисловість, розваги, наука та ін. Одним з найважливіших параметрів зображення є його роздільна здатність, яка визначається кількістю пікселів на одиницю довжини. Підвищення роздільної здатності зображень може значно покращити їх якість та деталізацію, що є важливим для багатьох застосувань. Нейронні мережі в наш час є доволі потужним інструментом для роботи з зображеннями та обробки інших типів даних. Існує багато архітектур нейронних мереж, які можуть вирішити цю проблему, проте вибір структури мережі та методів її оптимізації, що позитивно вплинули б на навчання мережі та дозволили уникнути проблеми перенавчання, залишається викликом.

**Аналіз останніх досліджень.** Сьогодні для досягнення більшої роздільної здатності зображень найчастіше застосовують методи інтерполяції, які включають в себе такі відомі методи як: інтерполяція методом найближчого сусіда, білінійна інтерполяція, бікубічна інтерполяція. Серед перелічених методів найкращий результат дає обробка зображення методом бікубічної інтерполяції, проте цей метод поступається іншим у необхідному часі для виконання обробки [1]. Крім методів інтерполяції сьогодні існує досить багато нейронних мереж глибокого навчання, кожна з яких має свої особливості, які по різному впливають на кінцевий результат: від рекурсії, до багат шаровості [2]. В рамках дослідження розглядається мережа надроздільної здатності (SRCNN), яка складається з 3 шарів, які схожі за будовою, проте виконують різні функції [3]: перший шар відповідає за виокремлення з зображення окремих ділянок пікселів та представлення кожної ділянки, як окремого багатовимірного вектора, другий шар відповідає за нелінійне відображення отриманих у першому шарі векторів у інші багатовимірні вектори, третій шар відповідає за з'єднання отриманих на попередніх шарах ділянок зображення для створення кінцевого зображення.

**Формулювання цілей (Постановка завдання).** Оптимізувати процес навчання нейронної мережі надроздільної здатності, користуючись методами нормалізації і регуляризації нейронних мереж.

**Основна частина.** Навчання нейронної мережі – це процес, який може займати доволі багато часу, особливо при роботі з зображеннями. Тому цей процес потребує оптимізації та регуляризації для запобігання перенавчанню мережі. В дослідження запропоновано використання наступних методів для оптимізації мережі: методу випадання нейронів, методу нормалізації пакетів та методу зменшення вагових коефіцієнтів.

Нормалізація пакетів є одним із методів регуляризації нейронних мереж, який забезпечує стабільність та швидкість навчання. Він зменшує взаємозв'язок між шарами нейронної мережі та забезпечує сталу нормалізацію ваг та зміщень.

Складається цей метод з нормалізації векторів активації прихованих шарів нейронної мережі за допомогою використання середнього значення та дисперсії поточного пакету. Нормалізація використовується перед або після використання нелінійної функції.

Під час навчання мережі трансформація сигналу за допомогою нормалізації пакетів відбувається наступним чином:

1. Спочатку визначається середнє значення  $\mu$  (1) та дисперсія значень  $\sigma^2$  (2) активації у пакеті:

$$\mu = \frac{1}{n} \sum_i Z^{(i)} \quad (1)$$

$$\sigma^2 = \frac{1}{n} \sum_i (Z^{(i)} - \mu)^2 \quad (2)$$

2. Далі нормалізується вектор активації  $Z^{(i)}$  (3). Тобто кожне вихідне значення нейрону буде відповідати нормальному розподілу у пакеті:

$$Z_{norm}^{(i)} = \frac{Z^{(i)} - \mu}{\sqrt{\sigma^2 - \epsilon}} \quad (3)$$

3. Обраховується вихідне значення шару  $\check{Z}$  (4) шляхом використання лінійного перетворення з параметрами  $\gamma$  і  $\beta$ , які також навчаються (рис. 1).

$$\check{Z} = \gamma \cdot Z_{norm}^{(i)} + \beta \quad (4)$$

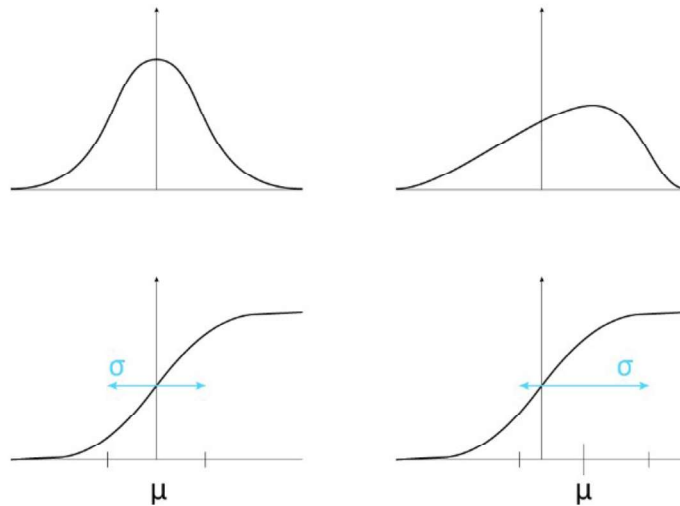


Рис. 1. Переваги використання параметрів  $\gamma$  і  $\beta$ . Модифікація розподілу (графіки у верхній частині рисунку) дозволяє використовувати різні режими для нелінійних функцій (графіки у нижній частині рисунку)

Такий підхід дозволяє мережі вибрати оптимальний розподіл для кожного прихованого шару шляхом регулювання цих параметрів,  $\gamma$  відповідає за регулювання стандартного відхилення,  $\beta$  дозволяє регулювати зсув, зміщуючи криву вправо чи вліво.

Нормалізація пакетів забезпечує нормалізацію вихідних значень по всій групі зображень і зменшує кореляцію між різними каналами зображення. Це зменшує вплив варіацій між зображеннями та допомагає моделі швидше зберігати та відтворювати деталі зображення.

Таким чином, застосування нормалізації пакетів допомагає збільшити роздільну здатність зображень у нейронних мережах, знижує вплив шуму та варіацій між зображеннями, і забезпечує більш швидке та стабільне навчання.

Випадання нейронів – це метод регулювання мережі, при якій певна кількість одиниць (нейронів) випадково ігнорується (“випадає”). Таке випадання відбувається на кожному оновленні вагових коефіцієнтів мережі, що дозволяє на кожній ітерації отримувати шар з різними видами з’єднання з попередніми шарами.

Під “випаданням” мають на увазі, що ці одиниці не беруться до уваги при проходженні даних через них вперед або назад (рис. 2).

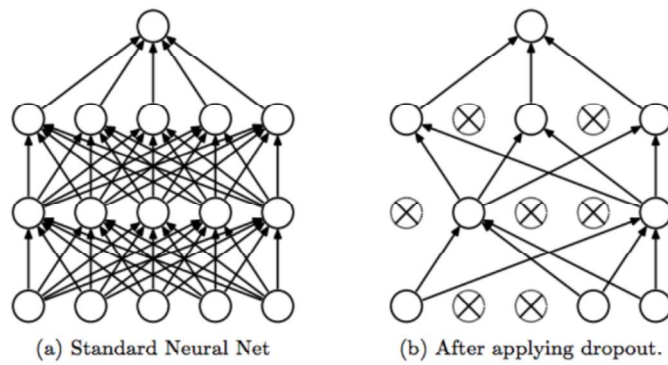


Рис. 2. Приклад використання техніки випадання нейронів на одній з ітерацій навчання мережі

На кожному оновленні мережі окремі вузли або випадають з мережі з ймовірністю  $1 - p$ , або зберігаються з ймовірністю  $p$  – залишається зменшена мережа.

Основна мета використання цієї техніки – запобігання перенавчанню мережі. При включенні усіх нейронів розвивається залежність нейронів одного шару від нейронів іншого, що обмежує індивідуальну потужність кожного нейрона та призводить до перенавчання мережі.

Для імплементації нормалізації пакетів до мережі на етапі навчання на кожній ітерації навчання для кожного прихованого шару ігнорується (обнуляється) випадкова частина  $p$  вузлів. На фазі тестування мережі використовуються всі активації, але при цьому зменшуються на коефіцієнт  $p$ .

Зменшення ваги це техніка регуляризації шляхом додавання певного штрафу до функції втрат (5).

$$L = \frac{1}{N} * \sum (y - f(x))^2 + \lambda * \sum w^2, \quad (5)$$

де  $N$  - кількість прикладів у тренувальному датасеті,  $y$  - цільові значення,  $f(x)$  - прогнозовані значення,  $w$  - вага моделі.

Використання цього методу:

1. дозволяє уникнути перенавчання мережі;
2. зменшує вплив великих значень ваг на модель;
3. забезпечує більш точні прогнози на нових даних;
4. дозволяє зберегти значення вагових коефіцієнтів малими та уникнути вибуху градієнта: оскільки штраф додається до функції втрат, то на кожній ітерації мережа намагатиметься оптимізувати вагові коефіцієнти невідривно від самої функції втрат, це дозволить зберігати якомога менші значення вагових коефіцієнтів, і таким чином уникнути стрімкого зростання градієнту.

Для оцінки результатів роботи модифікованої нейронної мережі використовується метрика пікового співвідношення сигналу до шуму (PSNR) (6).

$$PSNR = -10 \log_{10} \frac{MSE}{S^2}, \quad (6)$$

де  $S$  – максимальне значення, яке приймає піксель зображення.

Результати обробки зображення модифікованою нейронною мережею порівняно з результатами обробки методом бікубічної інтерполяції (рис.3).

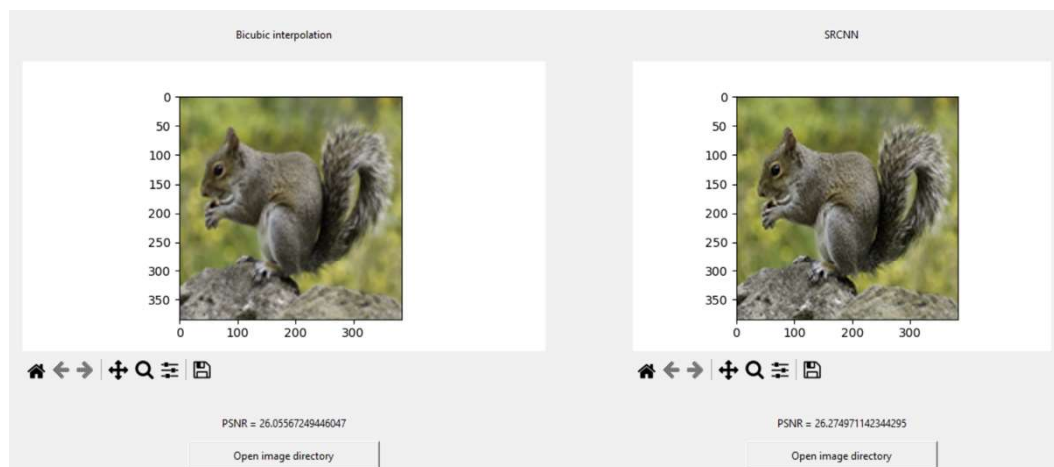


Рис. 3. Результат виконання обробки зображення модифікованою нейронною мережею надроздільної здатності

**Висновки.** Дослідження показало, що імплементація методів оптимізації навчання до нейронної мережі з простою структурою і малою кількістю шарів дає кращий результат в середньому на 0.2-0.3 dB для різних зображень, розглянутий підхід і методи можуть в подальшому використовуватися у більш глибоких нейронних мережах для вирішення інших задач пов'язаних із зображеннями, наприклад, розпізнавання об'єктів.

### *Бібліографічний список*

1. Dianyuan H. Comparison of commonly used image interpolation methods. Atlantis Press | Atlantis Press Open Access Publisher Scientific Technical Medical Proceedings Journals Books. URL: <https://www.atlantispress.com/article/4822.pdf>.
2. Anwar S., Khan S., Barnes N. A deep journey into super-resolution: a survey. URL: <https://arxiv.org/pdf/1904.07523.pdf>.
3. Dong C., Loy C. C. Image super-resolution using deep convolutional networks. URL: <https://arxiv.org/pdf/1501.00092.pdf>.