

УДК 378.016:004.8

Н.І. Кунах,

доктор технічних наук,
професор, Київський коледж зв'язку,
м. Київ, Україна,

О.М. Ткаленко,

кандидат технічних наук, доцент,
Державний університет телекомунікацій,
м. Київ, Україна,

Л.О. Харлай,

викладач-методист вищої категорії,
голова комісії інформаційних
мереж зв'язку, Київський коледж зв'язку,
м. Київ, Україна,

ORCID ID 0000-0002-7633-933X

ОЦІНКА ЕФЕКТИВНОСТІ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ

Метою роботи є використання штучних нейронних мереж для покращення якості зображень при збільшенні їх дозволу. Проведено аналіз сучасних алгоритмів інтерполяції, які включають ініціалізацію парних точок значеннями зменшеного зображення, метрик оцінки якості роботи алгоритмів; використання штучних нейронних мереж для збільшення дозволу зображень; методів підвищення ефективності штучних нейронних мереж. Порівняння результатів за допомогою об'єктивних метрик показало, що генеративно-змагальна модель справляється з обробкою зображень ефективніше, ніж згортальна нейронна мережа.

Ключові слова: штучні нейронні мережі, інтерполяція, дозвіл зображень, якість зображень.

Целью работы является использование искусственных нейронных сетей для улучшения качества изображений при увеличении их разрешения. Проведен анализ современных алгоритмов интерполяции, которые включают инициализацию парных точек значениями уменьшенного изображения, метрик оценки качества работы алгоритмов; использование искусственных нейронных сетей для увеличения разрешения изображений; методов повышения эффективности искусственных нейронных сетей. Сравнение результатов с помощью объективных метрик показало, что генеративно-соревновательная модель справляется с обработкой изображений эффективнее, чем свертывающая нейронная сеть.

Ключевые слова: искусственные нейронные сети, интерполяция, разрешение изображений, качество изображений.

Вступ. Завдання якісного збільшення дозволу зображень є одним із найактуальніших питань цифрової обробки зображень. Для вирішення цього завдання використовуються різні методи, що належать до класу Single image super-resolution (супердозвіл на базі одного зображення). Але найчастіше застосування методів

інтерполяції, фільтрів і різних алгоритмів обробки не дає бажаного результату. Штучна нейронна мережа (ШНМ) – це математична модель, що описує систему з'єднаних між собою штучних нейронів, а також реалізації цієї моделі. Є багато метрик, які дозволяють отримати об'єктивну оцінку якості збільшеного зображення. Але, крім цього, важливо також враховувати такі фактори, як час роботи алгоритму, необхідні обчислювальні потужності, пам'ять, навантаження процесора і т.д. Оскільки моделі з'являються постійно, а сама сфера дослідження машинного навчання і ШНМ є сферою, яка динамічно розвивається, можна стверджувати, що ця задача є актуальною.

Основна частина. ШНМ становлять собою системи, які перетворюють інформацію подібно процесам, що відбуваються в мозку людини. Однією з основних характеристик ШНМ є її модель. Охарактеризувати цю модель можна за видами нейронів, структурою моделі і способом навчання. Для задачі масштабування зображень використовуються різні моделі ШНМ, але найкраще себе зарекомендували такі: згортальна нейронна мережа (Convolutional neural network, CNN) та генеративно-змагальна мережа (Generative adversarial network, GAN).

В основі згортальних нейронних мереж лежать такі ідеї: кожний нейрон шару отримує вхідний сигнал від локального рецептивного поля в попередньому шарі; кожний прихований шар нейронної мережі складається з багатьох карт ознак (feature maps). Однією з найбільш відомих моделей згортальних мереж, що застосовуються для вирішення завдання збільшення зображень, є модель нейронної мережі SRCNN (Super-Resolution Convolutional Neural Network). [2] На вхід SRCNN подається зображення, збільшене за допомогою методів інтерполяції, яке потім проходить процес покращення якості. Подібний метод застосовується в розробленому компанією Google методі RAISR. [3] Генеративно-змагальні мережі – це алгоритм машинного навчання без вчителя, побудований на комбінації з двох нейронних мереж, як правило, згортальних або прямого розповсюдження. Схема моделі змагальної мережі представлена на рис.1.

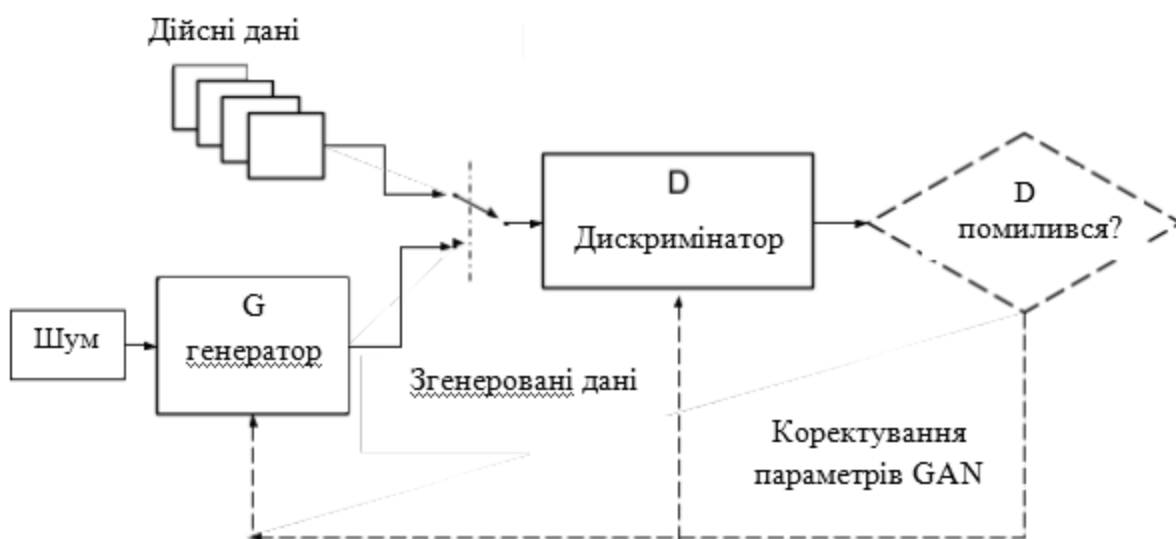


Рис. 1. Модель генеративно-змагальної ШНМ

На цей момент однією з останніх моделей, які розроблені для вирішення завдання збільшення дозволу зображення за допомогою ШНМ, є SRGAN (Super-Resolution Generative Adversarial Network). [4] Найбільш розповсюдженими на практиці є бікубічна інтерполяція та фільтр Ланцоша. Також варто відзначити, що ці алгоритми дають оптимальне співвідношення якості результату і часу роботи. Тому було прийнято рішення використовувати ці два алгоритми у програмній реалізації.

Аналізуючи моделі ШНМ, які використовуються для вирішення задачі збільшення зображень, відзначимо той факт, що найчастіше використовуються згортальні нейронні мережі, і однією з найпоширеніших моделей є SRCNN. На сьогодні генеративно-змагальні мережі, які з'явилися недавно, становлять великий інтерес і є дуже перспективними. Зокрема, модель SRGAN, створена для вирішення цього завдання, дозволяє отримати якісні результати.

Порівняння буде проводитися для таких алгоритмів і моделей: бікубічна інтерполяція; фільтр Ланцоша; модель ШНМ SRCNN; модель ШНМ SRGAN.

Для моделі SRGAN вхідними даними служать зображення. У результаті навчання необхідно отримати генеруючу функцію G . Генератор навчається як згортальна нейронна мережа прямого розповсюдження з параметрами θ_G :

$$Q_G = \{W_{1:L}; b_{1:L}\} \quad (1)$$

де $W_{1:L}$ і $b_{1:L}$ — це, відповідно, ваги і зміщення нейронної мережі глибиною в L шарів. Цей параметр обчислюється шляхом оптимізації складової функції втрат l_{SR} . Для навчання нейронної мережі знаходимо значення:

$$Q_G = \min \frac{1}{n} l_{SR}(G(I_n^{LR}), I_n^{HR}) \quad (2)$$

Для коректного порівняння якості роботи алгоритмів та нейронних мереж необхідно використовувати об'єктивні метрики. Вхідними даними для кожного методу буде служити зменшене зображення. Отримане в результаті роботи методів зображення можна порівняти з оригіналом на предмет появи розбіжностей. Найбільш розповсюдженими методами порівняння двох зображень є: метод середньо-квадратичної помилки моделі (Mean Squared Error, MSE); пікове відношення сигналу до шуму (Peak Signal-to-Noise Ratio, PSNR); індекс структурної схожості (Structural Similarity, SSIM). Найпростішим методом є MSE. В його основі лежить попиксельне порівняння двох зображень. Загальна формула:

$$MSE(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2, \quad (3)$$

де x_i — i -й піксель зображення x , y_i — i -й піксель зображення y , n — загальна кількість пікселів.

При порівнянні двох зображень значення MSE, рівне нулю, свідчить про те, що зображення повністю ідентичні. Чим вище значення цього показника, тим

більше відмінностей між зображеннями. Метрика розрахована на вікна розміром 8x8 пікселів. SSIM визначається локально, тому, щоб отримати глобальну оцінку подібності двох зображень, використовується усереднення SSIM по всіх вікнах. Результат, який виходить за допомогою методу SSIM, лежить у діапазоні від -1 до 1. Значення, рівне 1, досягається тільки при повній ідентичності зображень. Чим ближче значення до -1, тим більше відмінностей між зображеннями. Загальна формула має такий вигляд:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (4)$$

де μ_x – середнє зображення x , μ_y – середнє зображення y , σ_x^2 – дисперсія зображення x , σ_y^2 – дисперсія зображення y , σ_{xy} – коваріація x та y , C_1 та C_2 – поправочні коефіцієнти:

$$c_1 = (k_1, l)^2, c_2 = (k_2, l)^2, \quad (5)$$

де l – динамічний діапазон пікселів (зазвичай $2^{(\text{бітів на піксель})} - 1$), $k_1 = 0,01$, $k_2 = 0,03$ – константи. SSIM симетричний: $SSIM(x, y) = SSIM(y, x)$ [5]. Пікове співвідношення сигналу до шуму (PSNR) показує співвідношення між максимумом можливого значення сигналу і потужністю шуму, що спотворює значення сигналу. Оскільки багато сигналів мають широкий динамічний діапазон, PSNR прийнято зазвичай вимірювати в логарифмічній шкалі в децибелах. PSNR може приймати значення від нуля до ∞ . Чим більше значення PSNR, тим більшу схожість мають зображення, які порівнюються:

$$PSNR(x, y) = 10 \lg \frac{L^2}{MSE(x, y)}, \quad (6)$$

де x, y – зображення, які необхідно порівняти, $MSE(x, y)$ – середньоквадратична помилка двох зображень, L – максимальне значення яскравості. Зазвичай дорівнює 255 при використанні цілих чисел від 0 до 255, або 1, якщо використовуються речові значення від 0 до 1 [5]. Для вирішення різних завдань машинного навчання і обробки зображень використовуються різні мови програмування. Також існує велика кількість різних бібліотек, в яких реалізовані методи та алгоритми, які найбільш часто використовуються. Мовою програмування обираємо Python 3, використовуємо версію мови 3.6.2. Для програмної реалізації використовуємо основні бібліотеки: NumPy; OpenCV; TensorFlow. Вибір бази зображень пояснюється такими причинами. У цій базі поширюються не зображення, а посилання на них. Це дає можливість попередньої обробки зображень на етапі їх збереження. Наприклад, можна обмежити кількість завантажуваних зображень потрібною кількістю або, наприклад, при збереженні зменшувати розміри зображень.

Таким чином, можливо заощадити місце на жорсткому диску, що може бути істотною перевагою при роботі з великою кількістю зображень. Є можливість вказати тематику зображень. У цій базі представлені такі категорії, як фотографії людей, тварин, медичні зображення, фотографії природи і т.д. Щоб змінити тематику навчальної вибірки, досить замінити файл з посиланнями. Кількість повних проходжень тренувальних наборів для обох нейронних мереж (циклів, “epoch”, epoch) – 200. Після навчання моделі SRGAN проводимо повторне навчання, але в цей раз після обробки кожного тренувального набору підключається модуль тестування і результат обробки зберігається у вигляді зображення. Це зроблено для того, щоб мати можливість побачити динаміку поліпшення якості зображень, що генеруються, по мірі навчання нейронної мережі. Приклад такої динаміки наведений на рис. 2. На ньому цифра під зображенням позначає номер циклу навчання.

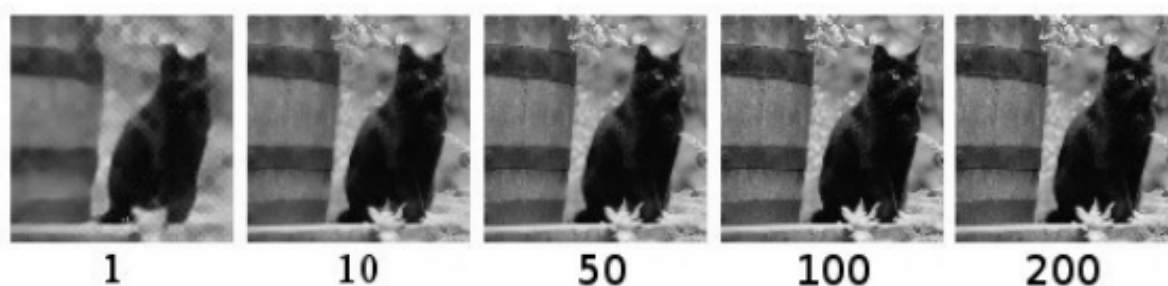


Рис. 2. Динаміка навчання ШНМ SRGAN

Як можна помітити з рис. 2, візуально складно побачити різницю між 100-м і 200-м циклом, але якщо подивитися на будь-яку з вибраних метрик якості, то на продовженні всього процесу навчання зберігається позитивна динаміка. Показники представлені в табл. 1.

Таблиця 1

Зміна значення метрик якості в процесі навчання

№ циклу	PSNR	SSIM	MSE
1	12,00	0,13	1577,13
10	22,01	0,73	291,49
50	27,01	0,87	125,32
100	27,37	0,89	118,07
200	27,47	0,90	116,02

Динаміка зміни показників на прикладі PSNR представлена на рис. 3. Як видно з рис. 3, швидкість навчання починає значно сповільнюватися вже після 10-го циклу, але тривала динаміка покращення показника зберігається аж до 200-го циклу. Після того, як був закінчений процес навчання і тестування нейронних мереж, були згенеровані результати роботи алгоритмів інтерполяції.

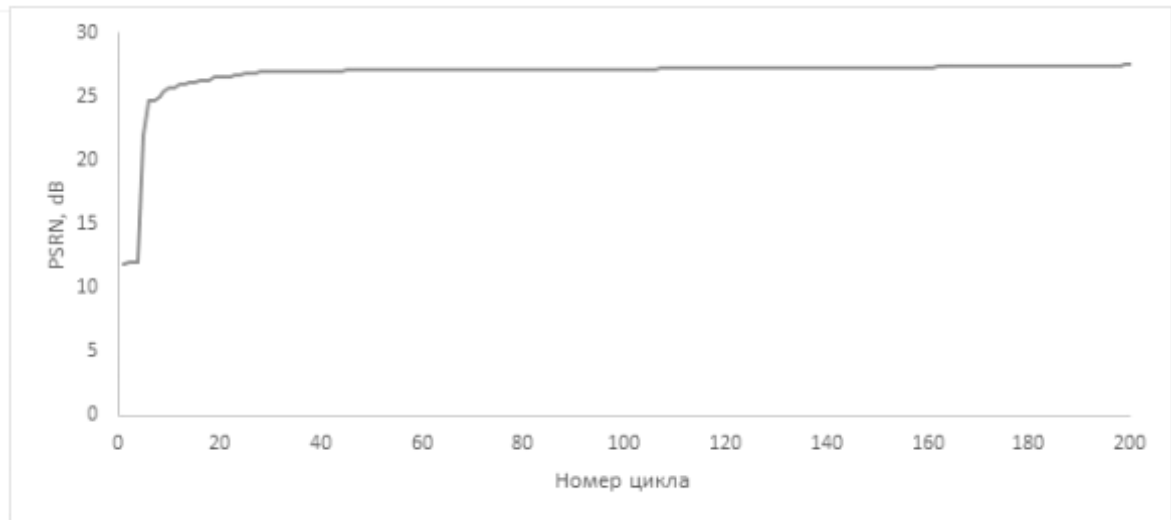


Рис. 3. Графік зміни показника PSNR у процесі навчання

Для моделі SRCNN час навчання склав 53 хвилини 26 секунд, тобто середній час на одну повну обробку тестових зображень – 16 секунд. У моделі SRGAN навчання зайняло значно більше часу – всього близько 8 години 43 хвилини. Відповідно, середній час на обробку всіх тестових наборів дорівнює 2 хвилинам 37 секундам. При проведенні тестування моделей час обробки тестових зображень – 27 секунд для моделі SRCNN і 33 секунди для SRGAN. Алгоритм бікубічної інтерполяції і фільтр Ланцоша на обробку тестових зображень зайняли 0,002005 і 0,561503 секунд відповідно. Файл із збереженим станом нейронної мережі SRCNN становить 441 мегабайт, SRGAN – 490 мегабайт. Приклад результату, який отримано внаслідок роботи алгоритмів та ШНМ, наведений на рис. 4.



Рис. 4. Результати роботи алгоритмів збільшення дозволу

Метрики для цього прикладу приведені у табл. 2.

Таблиця 2

Порівняння результатів на прикладі одного зображення

Метрика	Бікубічна інтерполяція	Фільтр Ланцоша	SRCNN	SRGAN
MSE	774,61	786,71	182,08	115,09
PSNR	19,24	19,17	25,53	27,52
SSIM	0,41	0,40	0,80	0,91

Середні показники результатів обробки всіх зображень представлені у табл. 3. Як видно з табл. 3, результат роботи моделі SRGAN має найкращі показники. Обидві моделі ШНМ значно перевершують алгоритми інтерполяції, проте між самими моделями відмінність в показниках незначна.

Таблиця 3

Порівняння середніх показників за всіма зображеннями

Метрика	Бікубічна інтерполяція	Фільтр Ланцоша	SRCNN	SRGAN
MSE	588,61	603,65	200,80	130,25
PSNR	21,26	21,15	26,94	27,63
SSIM	0,57	0,56	0,89	0,92

Якщо ж порівнювати моделі нейронних мереж, то з результатів видно, що генеративно-змагальні мережі перевершують згортувальні при порівнянні поза об'єктивними метриками.

Висновки. Аналіз сучасних підходів до вирішення завдання збільшення дозволу зображень показав, що, незважаючи на факт існування відносно швидких і якісних алгоритмів, постійно з'являються нові методи. У більшості досліджень використовуються різні реалізації моделей згортальних нейронних мереж. Але генеративно-змагальні нейронні мережі, які з'явилися відносно недавно, мають значну перспективу. Були досліджені алгоритми бікубічної інтерполяції, фільтра Ланцоша, моделі ШНМ SRCNN і SRGAN.

На вибір алгоритмів інтерполяції впливає той факт, що вони є досить ефективними за співвідношенням часу роботи і якості отриманого результату. Також на вибір впливає те, що ці алгоритми широко розповсюджені і використовуються в багатьох додатках. Вибір моделі SRCNN обґрунтований тим, що ця модель ШНМ є однією з найбільш розповсюджених і перевірених для вирішення задачі збільшення зображень. Модель SRGAN обрана через якісні результати роботи. Ця модель є новою і на сьогодні не має такої великої кількості програмних реалізацій, як модель SRCNN. При порівнянні результатів за допомогою об'єктивних метрик можна відзначити, що генеративно-змагальна модель справляється з обробкою зображень ефективніше, ніж згортальна нейронна мережа. Але для навчання такої моделі потрібно набагато більше часу. Використання навчальної моделі генеративно-змагальної мережі для обробки тестових зображень теж займає трохи більше часу, ніж згорткової. Але результат, який дозволяють отримати обидві моделі, значно перевершує алгоритми інтерполяції.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Горбачевская Е.Н. Классификация нейронных сетей. Вестник Волжского университета им. В.Н. Татищева. 2012. № 2. С. 128–134.
2. Kim J., Kwon Lee J., Mu Lee K. Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1646–1654 (2016).
3. ReymanMilanfar. Enhance. RAISR Sharp Images with Machine Learning. URL: [https://ai/googleblog.com](https://ai.googleblog.com) (дата звернення: 20.11.2018).
4. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. И. Д. Рудинского. М.: Горячая линия – Телеком, 2013. 384 с.

5. Wang Z., Bovik A.C., Sheikh H.R. and Simoncelli E.P. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, vol. 13, no. 4, p. 600–612, Apr. 2004.

REFERENCES

1. Gorbachevskaia Ye.N. (2012). Klassifikatsiia neyronnykh setey. "Classification of neural networks". Bulletin of the Volga University. V.N. Tatishcheva. No. 2. P. 128–134 [in Russian].
2. Kim J., Kwon Lee, J., Mu Lee (2016). K.: Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. P. 1646–1654 [in English].
3. PeymanMilanfar. Enhance. RAISR Sharp Images with Machine Learning. URL: [https://ai/googleblog.com](https://ai.googleblog.com) (date of application: 20.11.2018) [in English].
4. Rutkovskaya D., Pilinsky M., Rutkovsky L. (2013). Neyronnye seti, geneticheskiye algoritmy nechetkiye sistemy. "Neural networks, genetic algorithms and fuzzy systems": Trans. from polish I.D. Rudinsky. M.: Hotline-Telecom. P. 384 [in Russian].
5. Wang Z., Bovik A.C., Sheikh H.R. and Simoncelli E.P. (2004). Image quality assessment: From error visibility to structural similarity // IEEE Transactions on Image Processing. Vol. 13. No. 4. P. 600–612. Apr. [in English].

UDC 378.016:004.8

N.I. Kunakh,

Doctor of Technical Sciences, Professor, Kyiv College of Communication, Kyiv, Ukraine,

O.M. Tkalenko,

Candidate of Technical Sciences, Docent, State University of Telecommunications, Kyiv, Ukraine,

L.O. Kharlai,

Competitor, the teacher-methodologist of the highest category, Head of the Department, Kyiv College of Communication, Kyiv, Ukraine, ORCID ID: 0000-0002-7633-933X

EVALUATION OF THE EFFECTIVENESS OF ARTIFICIAL NEURAL NETWORKS

The purpose of the article is the use of artificial neural networks to improve the quality of images when increasing their permission. To achieve the goal, various types of most promising artificial neural networks were considered – mathematical models that describe the system of artificial neurons interconnected among them, as well as the implementation of these models. The analysis of modern interpolation algorithms, which include the initialization of paired points by the values of the reduced image, the metrics for assessing the quality of the algorithms; using artificial neural means to increase image resolution; methods of increasing the efficiency of artificial neural networks.

A software implementation has been created using several interpolation algorithms and models of artificial neural networks. An analysis of modern approaches to solving the problem of increasing the resolution of images has shown that despite the existence of relatively fast and high-quality algorithms, new methods are constantly emerging. In recent years, artificial neural networks are increasingly used to solve this problem. In most studies, different implementations of the models of the curtain neural networks are used.

The algorithms of bicubic interpolation, Lancosha filter, models of artificial neural networks SRCNN and SRGAN were studied in this work. The choice of interpolation algorithms is influenced by the fact that they are quite effective in terms of the ratio of the time of work and the quality of the result. Also, the choice is influenced by the fact that these algorithms are widespread and used in many applications.

When comparing results with objective metrics, it can be noted that generative-competitive model copes with image processing more efficiently than the creeping neural network. But adapting such a model requires much more time. Using the generic-competitive network model to handle test images also takes a little more time than the rollout. But the result, which allows both models, far outperforms interpolation algorithms.

Keywords: artificial neural networks, interpolation, image resolution, image quality.

Отримано 30.11.2018

Рецензент Хорошко В.О., д.т.н., проф.