

зображень [3, 2392-2399], [4, 1067–1074], [5, 390-396]. З недавніми зростанням потужності обробки комп'ютерної графіки MLP замінили згорткові нейронні мережі (CNN), особливо стосовно завдань обробки зображень.

Згорткова нейромережа (ЗНМ) має спеціальну архітектуру, яка дозволяє їй максимально ефективно аналізувати образи. Сама ідея ЗНМ базується на чередуванні згорткових та субдискретизуючих шарів (pooling), а її структура є одноправленою. ЗНМ отримала свою назву від операції згортки, яка передбачає, що кожен фрагмент зображення буде помножено на ядро згортки поелементно, при цьому отриманий результат повинен бути підсумований та записаний у схожу позицію вихідного зображення. Така архітектура забезпечує інваріантність розпізнавання відносно зсуву об'єкта, поступово збільшуючи «вікно», на яке «дивиться» згортка, виявляючи все більші й більші структури і шаблони у зображенні.

Розглянемо моделі процесів спотворення та відновлення зображень. Як показано на рис. 1, процес спотворення зображення полягає в тому, що деякий оператор спотворення H діє на вихідне зображення $f(x, y)$, так, що в результаті додавання адаптивного шуму отримуємо спотворене зображення $g(x, y)$.

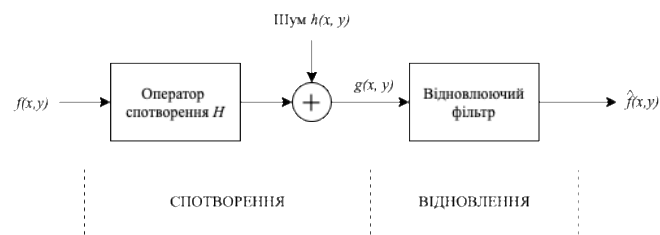


Рис. 1 – Модель процесу спотворення/відновлення зображень

Процес відновлення полягає в побудові деякого наближення вихідного зображення по спотвореному зображенню $g(x, y)$, за деякою інформацією щодо оператора спотворення H і адитивного шуму $\eta(x, y)$. При цьому наше наближення має бути якомога ближче до вихідного зображення. Спотворене зображення може бути представлено в просторовій області у вигляді (1).

$$g(x, y) = h(x, y) * f(x, y) + \eta(x, y), \quad (1)$$

де $h(x, y)$ – функція, що представляє спотворює оператор в просторовій області, а символом «*» позначається згортка.

Згортка в просторовій області еквівалентна множенню в частотній області, тому вище наведена рівність (1) може бути представлена в частотній області у такому вигляді:

$$G(u, v) = H(u, v)F(u, v) + N(u, v), \quad (2)$$

де зазначені функції – Фур'є-образи відповідних їм функцій в (1).

Так як спотворення зображення це результат згортки, то для його відновлення необхідно знайти такий фільтр, застосування якого призводить б до зворотного процесу. Тому для позначення лінійного процесу відновлення часто використовується термін реконструкція (деконволюція) зображень. Аналогічно, фільтри для відновлення часто називаються реконструюючими фільтрами.

Нами передбачається розробка алгоритму видалення шумів на цифрових зображеннях за допомогою ЗНМ та здійснення його програмної реалізації засобами Python з використанням фреймворка для глибокого навчання Keras. Даний фреймворк є надбудовою над TensorFlow – відкритою платформою для машинного навчання, що має гнучку екосистему інструментів, бібліотек та ресурсів, та дозволяє дослідникам просувати найсучасніші технології машинного навчання, а розробники можуть легко розгортати додатки.

Список літератури

1. J. Najeer Ahamed, V. Rajamani Design of hybrid filter for denoising images using fuzzy network and edge detecting // American Journal of Scientific Research, Issue 3, 2009, pp 5-14.
2. Методи удалення шумов на изображениях на основе применения искусственных нейронных сетей [Електронний ресурс] URL: <https://www.lib.tpu.ru/fulltext/c/2010/C01/V2/170.pdf>
3. H. C. Burger, C. J. Schuler, and S. Harmeling, "Image denoising: Can plain neural networks compete with BM3D?" in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012, pp. 2392–2399.
4. C. J. Schuler, H. C. Burger, S. Harmeling, and B. Scholkopf, "A machine learning approach for non-blind image deconvolution," in Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. IEEE, 2013, pp. 1067–1074.
5. R. G. Pires, D. F. S. Santos, L. A. M. Pereira, G. B. De Souza, A. L. M. Levada, and J. P. Papa, "A robust restricted boltzmann machine for binary image denoising," in 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). Ieee, 2017, pp. 390–396.

Дмитрюк В., Липовецький О.

Черкаський національний університет ім. Б. Хмельницького

РОЗРОБКА СИСТЕМИ ДЕТЕКЦІЇ РУХУ

Метою роботи є аналіз та вибір технології і створення програмного забезпечення віддаленого відеоспостереження із виявленням руху для системи на основі смартфонів, а також проведення експериментального дослідження створеної системи.

Смартфон з операційною системою Android можна перетворити на IP-камеру для відеоспостереження за об'єктами за допомогою Wi-Fi-з'єднання. IP-камеру можна придбати в магазині за \$ 20-50, але, перетворивши свій пристрій Android на IP-камеру, користувач зможе завантажити її безкоштовно.

IP-камера на базі смартфона може використовуватися як камера спостереження. Для перетворення смартфона на IP-камеру, використовується мобільний пристрій Android, програмне забезпечення та доступ до Wi-Fi. IP-камера працює, коли смартфон, на якому встановлена програма, підключений до мережі Wi-Fi або до мобільного оператора. Після підключення IP-камери та початку роботи, користувач зможе отримати до неї доступ через Telegram.

При реалізації проекту створення відеоспостереження за допомогою IP-камери зі звичайного смартфона виникає питання про цілеспрямованість самої ідеї реалізації. З одного боку, є безкоштовний телефон, який шкода викидати, оскільки, з іншого боку, на цей пристрій існують обмеження, які можуть спричинити негативний сценарій використання.

На рис. 1 наведено, запропонований нами алгоритм роботи програми. Клас Reader відповідає за зчитування кадрів із відеопотоку. Після прочитання кадру клас Detection відповідає за визначення та відстеження об'єктів. Обробка зображень, фільтрація, створення фонові моделі, визначення об'єктів виконуються за допомогою методу detectObjects, який містить масив об'єктів. Кожен об'єкт відеоспостереження є екземпляром класу Object. Клас Object містить поля, що зберігають інформацію про об'єкт: скільки кадрів пройшло з моменту першого визначення, кількість кадрів, в яких програма фіксує переміщення об'єкта. Для кожного об'єкта створюється фільтр Калмана, який є екземпляром класу SingleKalmanFilter.

Максимальний розмір зображення – 1920 пікселі в довжину і 1080 в ширину. Послідовність містить об'єкт, який йде по криволінійній траєкторії. Протягом усієї послідовності програма надає можливість бачити зміни в становищі об'єкта, що контролюється. Отже, незважаючи на невеликі розміри об'єкта та спотворення реальної камери, результат цього прикладу повністю виправдовує очікування.

У результаті роботи була розпочата розробка програми для визначення та утримання об'єктів під час відеоспостереження (рис.2). Проаналізовано методи та технології ідентифікації об'єктів. Програмні тести проводились на трьох прикладах, які показали переваги та недоліки розробленого алгоритму.

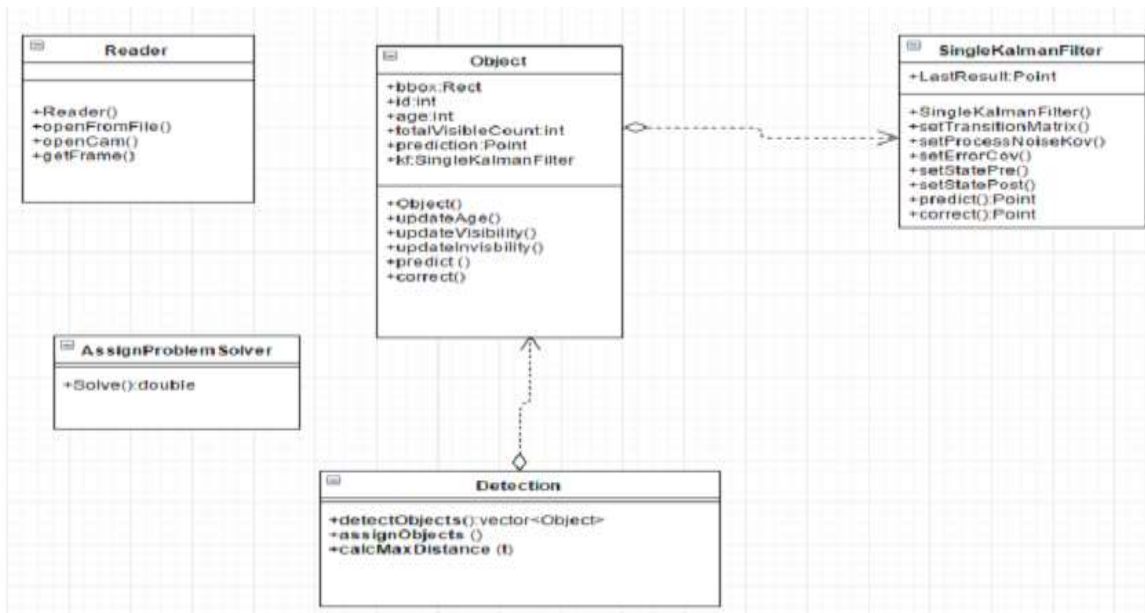


Рис. 1 Алгоритм роботи програми



Рис.2 Скріншоти програми

Науковий керівник: д.пед.н., доц. Луценко Г.В.

Юрко О. В.

Черкаський національний університет ім. Б. Хмельницького

ЗАБЕЗПЕЧЕННЯ ГНУЧКОСТІ ОСВІТНІХ ПРОГРАМ В УМОВАХ ДИСТАНЦІЙНОГО НАВЧАННЯ

Реалії сьогодення свідчать, що всі сфери людської діяльності переходять в онлайн і це не є виключенням для освітнього процесу. Сфера освіти зазнала суттєвих змін

внаслідок стрімкого розвитку інформаційно-комунікаційних технологій. На зміну традиційним, усталеним формам організації освітнього процесу (стаціонарна, заочна) прийшла дистанційна.

Для успішного старту дистанційного навчання першочерговим є формування нормативно-правового поля для забезпечення дистанційної освіти. Іншим важливим аспектом є розробка адекватних освітніх програм для використання в електронній формі з можливістю інтерактивної взаємодії між студентом і викладачем. Саме інформаційно-комунікаційні технології є інструментом забезпечення такої взаємодії. Сучасні студенти активно використовують мобільні девайси для здобуття освіти. (Мойсеєнко, 2015) Можливість доступу до навчальних курсів у зручний для студента час є запорукою успішного опанування значної частини інформації. Здобувач освіти, перебуваючи у будь-якому місці, може проходити теоретичну частину освітньої програми, закріплюючи її підсумковими тестами.

Дистанційне навчання розглядається як комплекс сучасних інформаційних технологій, що доповнюють принципи відкритого навчання. Така побудова процесу освіти за допомогою інформаційно-комунікаційних технологій надає студентам можливість доступу до всієї теоретичної бази для успішного проходження навчальної програми.

На сьогодні науковці поділяють принципи для дистанційної системи навчання на три групи: (Шупта, 2011)

–*Загальні принципи.* Включають принципи науковості, систематичності, послідовності, свідомості та доступності знань;

–*Принципи дистанційного навчання.* Утворюють діяльність, з формуванням середовища «м'яких» і «твердих» форм пізнавальної діяльності студентів, створення інтелектуальних продуктів з індивідуальним підходом;

–*Специфічні принципи дистанційного навчання.* Ці принципи складається з комунікації, інтерактивності, що обумовлюється інтенсивним використанням інформаційних технологій.

Гнучкість освітнього процесу – це можливість здобувачами освіти самостійно вибрати місце, час, темп проходження освітньої програми, який розглядається як дієвий спосіб формування якостей самореалізації спеціаліста. (Луценко, 2020)

Звертаючись до наведених вище принципів дистанційного навчання, зазначимо, що остання група вдало поєднує в собі педагогічну доцільність засобів новітніх інформаційно-комунікаційних технологій, відкритості та гнучкості навчання. У першу чергу гнучкість надання освітніх програм полягає в правильно відведеному часі на виконання тих чи інших завдань. Таким чином, поетапне опрацювання теоретичного матеріалу, який підкріплюється проміжним тестуванням, відповідає вимогам оцінювання здобувачів освіти.

Основна роль викладача в системі забезпечення гнучкості освітнього процесу – організація і надання можливих для дисципліни, теоретичних та практичних завдань і своєчасна підтримка (надання feedback) при перевірці даних робіт.

Гнучкість освітніх програм стала більш адаптивною в умовах дистанційної освіти. При правильній організації самостійної роботи студентів маємо результат з великою кількістю опанованої інформації при послідовно складеному навчальному плані.

Список використаних джерел

1. Луценко, Г. В. (2020). Гнучкість освітнього процесу як складова професійної підготовки майбутніх фахівців в умовах становлення індустрії 4.0, [с. 16-17]. Черкаси.
2. Мойсеєнко, В. І. (2015). Застосування сучасних інформаційних технологій для організації самостійної роботи. у В. А. Бриксіні (Ред.), [с. 102]. Київ: Національна бібліотека України імені В.І. Вернадського.
3. Шупта, О. В. (2011). Дидактичні принципи дистанційного навчання. [с. 4-6]. Хмельницький: Інституційний репозитарій Хмельницького національного університету.

Науковий керівник: д.п.н., доц. Г.В.Луценко

КОМП'ЮТЕРНІ НАУКИ

PRE-PROCESSING METHODS EFFECTIVENESS FOR HANDWRITING RECOGNITION BY PHOTO

Abstract. This article presents results of effectiveness comparison of image pre-processing methods for handwriting recognition from an image. It describes recognition accuracy increase/decrease depending on pre-processing methods with the artificial neural network consisting of 5 convolutional layers that use max-pooling between layers and 2 recurrent layers with 256 neurons at the hidden level. The resulting methods combination improves the final accuracy by 7.5% comparing to the initial accuracy.

Keywords: Artificial Neural Network, Convolutional Neural Network, Handwriting Recognition, Image Pre-Processing.

Introduction. Even today's world, there is still a lot of information that exists only in handwriting, although, given the development of information technology, it is much more convenient to analyse, store and transmit information in electronic form. And for this, it is necessary to have reliable mechanisms for converting handwritten text on paper into an electronic version.

The research aims to create an intelligent image pre-processing system for handwriting recognition by analysing image processing methods related to text recognition, their improvement and combination. The task is to analyse the existing methods of image pre-processing for handwriting recognition by photo and to develop intelligent technology by improving and combining existing methods.

When performing the work the following tasks needs to be performed:

- Review and comparative analysis of existing methods of pre-processing images with text.
- Identify the methods that are best for pre-processing images with handwritten text, highlight methods to improve the visibility of the text, divide the photo into individual words and recognition.
- Identify solutions of existing problems in existing methods.
- Design and develop a system for analysis and verification of the implementation of the found methods of image pre-processing for handwriting recognition.
- Conduct experiments using the developed system with selected methods and their combinations, analyse the results of experiments and choose the most optimal mechanisms.

Materials and Methods. Methods such as comparative analysis and experiment were used in the study. Comparative analysis was used to study the feasibility of using certain methods of image pre-processing for handwriting recognition, and the experiment was used to test the effectiveness of different implementations of intelligent pre-processing technology using the analyzed methods and their combinations.

Usually the task of handwriting recognition is not limited to the recognition of one word, which is already perfectly cut from the text, aligned, free of digital noise, etc. The most difficult task is to divide the document into segments that can be recognized by the neural network.

The tasks of preparing a manuscript image for recognition by a neural network can be divided into several groups: general image processing, text alignment, word selection (which in turn, depending on the method, can be divided into line selection and word selection from these lines).

The general processing of the image includes such operations as binarization, reduction of the image to shades of gray, noise reduction, blurring and sharpening of the image, normalization [1], etc.

Results. *Designing the experimental system.*

Experiments are needed to identify the effectiveness of the methods described above, to identify the best of them, and to improve their effectiveness. But given that experiments require the training of a neural network of the same configuration with different sets of pre-processing methods, it is first necessary to develop a system that will allow these experiments to be performed.

Therefore, a neural network model was developed to test the effectiveness of image pre-treatment methods for handwriting neural network training. [2]

It consists of 5 layers of CNM, 2 layers of LNM, CTC level loss, and decoding.

- The input image is gray and has a size of 128x32 [3]
- 5 layers of ZNM display the input image to a sequence of objects of size 32x256
- 2 layers of RNM with 256 units distribute information through the sequence and display the sequence to a matrix of 32x80. Each matrix element represents an estimate for one of the 80 symbols at one of the 32-time steps

• The CTC layer calculates the value of the losses specified by the matrix and the true text (during training) or decodes the matrix to the final text with the decoding of the best path. [6]

Implementation depends only on the numpy, cv2, and tensorflow libraries. The experiments will be carried out taking into account the following points:

- The batch size is set to 50
- One epoch - 500 games, so one epoch will include 25,000 images
- IAM Handwriting Database [5] with 118 thousand images will be used for neural network training
- For training and validation, the dataset is divided into 95% for training and 5% for validation. Data for validation remain unchanged, 25 thousand images are selected from 95% randomly
- 2 parameters will be evaluated: the average error in word recognition and the error in the characters in the recognized words
- If the recognition accuracy does not improve after 10 epochs, the training ends
- Accuracy will be considered the average error of character recognition.

Base model

First you need to determine the starting point from which you can start. To do this, we will train the neural network, the pre-processing of the image for which will include only the conversion of the image into shades of gray.

In order to make sure that the results are not random and based on regularity, training will be conducted several times to determine how the accuracy of accuracy results when learning from scratch with the same data set and the same method of pre-processing.

The learning results of the neural network are presented in table 1 and in Fig. 1.

Table 2.1

	Base Model - 1	Base Model - 2
Word error	28,80%	28,94%
Symbol error	11,96%	11,95%

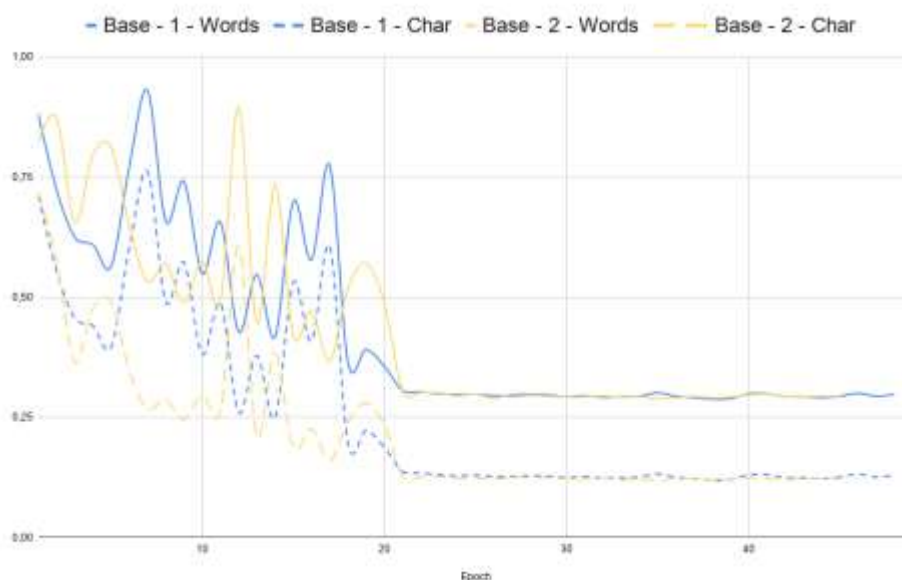


Fig.1. Diagram of the average error of the neural network depending on the era for the base model

As can be seen from the results, the accuracy of the neural network hardly changes when re-learning with random initial weights in the model, but with the same set of data for training and the same method of pre-processing.

The first approximately 20-25 epochs of learning are unstable with constant fluctuations in accuracy, but after 25 epochs only slightly improves accuracy. It is worth noting this feature of the basic model and compare the methods used in the following experiments.

The difference in results between the two re-training of NN is only 0.01%, which can be considered as an error in future experiments. Also, any change greater than 0.1% in character recognition accuracy can be considered as an effect of the pre-processing method.

For the model for comparison the second run of training will be used as accuracy it turned out a little higher.

Final model

After experiments, it was found that binarization not only does not improve the result in the case of handwriting recognition, but also worsens the results of handwriting recognition from the image.

At normalization it is possible to see obvious improvement, both on stability of results, and on the maximum accuracy. Although the absolute value of accuracy has not changed much, but the stability of the result has significantly improved. This improvement will reduce training time when using very large amounts of data, as it allows you to stop learning before the maximum accuracy has been reached (and determining such a maximum for a large sample is also a difficult task).

When using augmentation, the absolute accuracy result improved markedly relative to the base model. The change in the relatively normalized model is not so noticeable, but it is also there. But the other side of improving accuracy was the much lower stability of the results.

Taking into account the previous results, only normalization and swelling of the images were chosen for the combination. Normalization significantly improves the stability of learning and at the same time increases the accuracy of recognition. When the images are inflated, an increase in accuracy can be achieved, but the stability of the results deteriorates even compared to the base model.

The results in comparison with the base model are presented in Table 2 and in Fig. 2.

Table 2

	Base Model	Normalization + Augmentation
Word error	28,94%	26,92%
Symbol error	11,95%	10,88%

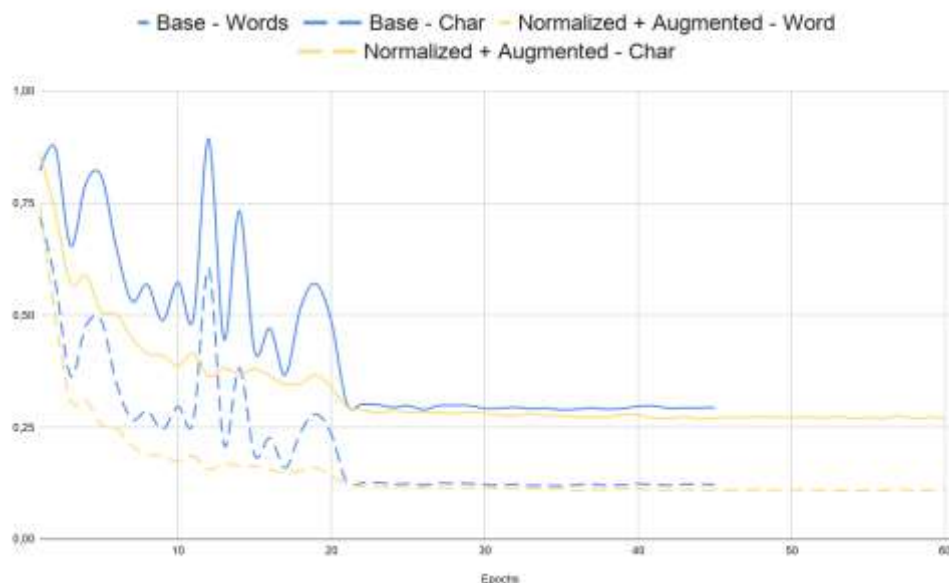


Fig.2. Diagram of the average error for the base image and using inflation

The combination of the methods described above is expected to give results, both in a marked increase in accuracy and in improving the stability of learning relative to simply inflated images.

The final result showed an improvement of 7.5% in word recognition accuracy and 9.83% in character recognition accuracy relative to the base model.

Summary & Conclusions.

The article carried out a theoretical study of image pre-processing methods for handwriting recognition and formed a set of methods, the effectiveness of which was analysed at the stage of experiments.

An experiment was conducted to analyse the effectiveness of various methods of image pre-processing to teach a neural network for handwriting recognition. As a result of the experiment, the result relative to the basic model trained in grayscale images without additional processing was able to improve the result by 7.5% in word recognition accuracy and 9.83% in character recognition accuracy, which can be considered a significant result.

As a result, the following prerequisites are formed for handwriting recognition systems:

- For neural network training, binarization leads to data loss (which does not preclude the use of binarization in the early stages of processing)
- Normalization and bloating can significantly increase the accuracy of recognition.

Discussion. This approach has further prospects for improving accuracy. This can be achieved by sharpening images, whitening the background, increasing saturation, deslanting, etc. This will allow the algorithm to be used to digitize library archives automatically. Also, using this algorithm, you can optimize processes in schools by automating handwriting verification.

References

1. Rafael C. González, Richard Eugene Woods (2007). Digital Image Processing. Prentice Hall. p. 85. ISBN 978-0-13-168728-8.
2. Yun Wang, Juncheng Li, Florian Metz. Comparing the Max and Noisy-Or Pooling Functions in Multiple Instance Learning for Weakly Supervised Sequence Learning Tasks. April 3, 2018. pp. 5 URL: <https://arxiv.org/pdf/1804.01146.pdf>.
3. Blumenstein, Michael, Brijesh Verma, and Hasan Basli. "A novel feature extraction technique for the recognition of segmented handwritten characters." In Document Analysis and Recognition, 2003. Proceedings. Seventh International Conference on, pp. 137-141. IEEE, 2003
4. Eikvil, Line. "Optical character recognition." citeseer. ist. psu. Edu / 142042. Html (1993).
5. IAM Handwriting Database," IAM Handwriting Database - Computer Vision and Artificial Intelligence. [Online]. Available: <http://www.fki.inf.unibe.ch/databases/iam-handwriting-database>. [Accessed: 18-Sep-2019].
6. GuifangLina, WeiShen. Research on convolutional neural network based on improved Relu piecewise activation function. Volume 131, 2018, Pages 977-984 URL: <https://www.sciencedirect.com/science/article/pii/S1877050918306197>.
7. Casey, Richard G., and Eric Lecolinet. "A survey of methods and strategies in character segmentation." Pattern Analysis and Machine Intelligence, IEEE Transactions on 18, no. 7 (1996): 690-706.

Bushyn I. M.

The Bohdan Khmelnytsky National University of Cherkasy

Анотація. У цій статті представлені результати порівняння ефективності методів попередньої обробки зображень для розпізнавання рукописного вводу із зображення. Він описує збільшення / зменшення точності розпізнавання залежно від методів попередньої обробки за допомогою штучної нейронної мережі, що складається з 5 згорткових шарів, які використовують макс-пул між шарами та 2 повторюваних шари з 256 нейронами на прихованому рівні. Отримана комбінація методів покращує кінцеву точність на 7,5% порівняно з початковою точністю.

Ключові слова: розпізнавання рукописного тексту, штучна нейронна мережа, згорткова нейронна мережа, попередня обробка зображень.

ОГЛЯД НАУКОВИХ ДОСЛІДЖЕНЬ ФОРМУВАННЯ НАВЧАЛЬНОЇ ВИБІРКИ ДЛЯ СЕГМЕНТАЦІЇ ОРГАНІЗАТОРІВ ТЕНДЕРІВ НА ОСНОВІ МАШИННОГО НАВЧАННЯ

Тендери є важливим інструментом у сучасному механізмі ринкового господарства: сприяють розвитку внутрішньої торгівлі, яка, у свою чергу, є засобом забезпечення більш високих темпів зростання національної економіки. Учаснику тендерів, важливо знати, про організатора ґрунтовну інформацію, скільки організатор проводив аукціонів, скільки з них успішних та на яку суму. Тому для цього потрібна автоматизована система для вибору організаторів тендерів на основі машинного навчання, що також дасть можливість автоматизувати процес участі у тендері. Для цього потрібно сформуванати навчальну вибірку, яка зможе навчити систему розпізнавати організаторів тендерів, без додаткових обрахунків.

Багато авторів проводять дослідження державних тендерних закупок [19] в сферах: охорони здоров'я [14]; відновлюваних джерел енергії [15]; харчовій промисловості [16]; будівельній галузі [17]; гірничої промисловості [20] та по окремих країнах [18, 21-23].

У дослідженнях [1, 2, 5] запропоновано адаптивний метод формування класифікованої навчальної вибірки на основі використання порогової оцінки коефіцієнта кореляції міжканалів комбінованих перешкод. У дослідженнях [3, 4] запропоновано алгоритм розробки навчального набору, який найкраще описує об'єкти розпізнавання.

Дослідження [6, 8, 10, 12] засноване на моделі RFM (Recency, Frequency and Monetary) та використовує принципи сегментації набору даних за допомогою алгоритму K-Means. Дослідження [7, 9, 11] має на меті розробити новий методологічний підхід, інтегруючи "Recency, Frequency and Monetary" із рідкісним K -означає алгоритм кластеризації Віттена та Тібшірані. Запропонований підхід придатний для обробки великих, великих розмірів та розріджених споживчих даних. Дослідження [13] вдосконалює дослідження щодо розробки стратегій цифрового маркетингу на основі рекомендацій, надаючи компаніям прогнозу модель. Ця модель використовує науку даних, включаючи методи машинного навчання та великі дані, для персоналізації фінансових стимулів для користувачів на основі якості нових клієнтів, яких вони посилають на веб-сайт з поверненням готівки.

Дослідження сегментації клієнтів, є надзвичайно популярною темою, що і підтверджено аналізом вище. Проте жоден з них не розглядає сегментацію організаторів тендерних аукціонів, адже це дасть учасникам аукціонів проаналізувати: нові ринки збуту своєї продукції; прозорість та чесність системи електронних торгів – виграє кращий; вибір найбільш привабливого організатора, що дасть можливість слідкувати за найвигіднішими пропозиціями. Також на основі отриманих даних, можна розробити систему для вибору організаторів тендерів на основі машинного навчання, що дасть можливість автоматизувати процес участі у тендері.

Список використаних джерел

1. Piza, D. M.; Bugrova, T. I.; Lavrentiev, V. N.; Semenov, D. S. (2018). Method of Forming Classified Training Sample in Case of Spacial Signal Processing under Influence of Combined Interference. *Radioelectronics and Communications Systems*, Springer, 61(7), 325–331. doi:10.3103/S0735272718070051
2. Piza, D. M.; Moroz, G. V. (2018). Methods of Forming Classified Training Sample for Adaptation of Weight Coefficient of Automatic Interference Compensator. *Radioelectronics and Communications Systems*, Springer, 61(1), 32–37. doi:10.3103/S0735272718010041
3. Pastukhov, Aleksey A.; Prokofiev, Aleksander A. (2017). Clustering algorithms application to forming a representative sample in the training of a multilayer perceptron. *St. Petersburg Polytechnical University Journal: Physics and Mathematics*, Elsevier, 3(2), 127–134. doi:10.1016/j.sppjm.2017.05.004
4. Kamilov, M.; Hudayberdiev, M.; Khamroev, A. (2019). Algorithm for the Development of a Training Set that Best Describes the Objects of Recognition. *Procedia Computer Science*, Elsevier, 150, 116–122. doi:10.1016/j.procs.2019.02.024

5. Arunachalam, Deepak; Kumar, Niraj (2018). Benefit-based consumer segmentation and performance evaluation of clustering approaches: An evidence of data-driven decision-making. *Expert Systems with Applications*, Elsevier, 111, 11-34. doi:10.1016/j.eswa.2018.03.007
6. Anitha, P.; Patil, Malini M. (2019). RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University - Computer and Information Sciences*, Elsevier, 1319-1578. doi:10.1016/j.jksuci.2019.12.011
7. Zhou, Jian; Zhai, Linli; Pantelous, Athanasios A. (2020). Market segmentation using high-dimensional sparse consumers data. *Expert Systems with Applications*, Elsevier, 145, 113136–doi:10.1016/j.eswa.2019.113136
8. Piskunova, O., & Klochko, R. (2020). Classification of e-commerce customers based on Data Science techniques. *MPSESM-W 2020. CEUR-WS. Vol-2649*. 6-20.
9. Huang, Yong; Zhang, Mingzhen; He, Yue (2020). [IEEE 2020 5th International Conference on Computational Intelligence and Applications (ICCIA) - Beijing, China (2020.6.19-2020.6.21)] 2020 5th International Conference on Computational Intelligence and Applications (ICCIA) - Research on improved RFM customer segmentation model based on K-Means algorithm. 24–27. doi:10.1109/ICCIA49625.2020.00012
10. Wu, J., Shi, L., Lin, W. P., Tsai, S. B., Li, Y., Yang, L., & Xu, G. (2020). An Empirical Study on Customer Segmentation by Purchase Behaviors Using a RFM Model and K-Means Algorithm. *Mathematical Problems in Engineering*, Volume 2020, Article ID 8884227. <https://doi.org/10.1155/2020/8884227>
11. Pu, X., Song, C., & Huang, J. (2020, September). Research on Optimization of Customer Value Segmentation Based on Improved K-Means Clustering Algorithm. In *2020 IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE) IEEE*. 538-542.
12. Zare, Hamed; Emadi, Sima (2020). Determination of Customer Satisfaction using Improved K-means algorithm. *Soft Computing*, Springer, 24, 16947–16965. doi:10.1007/s00500-020-04988-4
13. Ballestar, M.T., Grau-Carles, P. & Sainz, J. (2019) Predicting customer quality in e-commerce social networks: a machine learning approach. *Rev Manag Sci*, Springer, 13, 589–603. <https://doi.org/10.1007/s11846-018-0316-x>
14. Qendri, V., Bogaards, J.A. & Berkhof, J. (2019). Pricing of HPV vaccines in European tender-based settings. *Eur J Health Econ*, Springer, 20, 271–280 <https://doi.org/10.1007/s10198-018-0996-9>
15. Bayer, Benjamin; Schäuble, Dominik; Ferrari, Michele (2018). International experiences with tender procedures for renewable energy – A comparison of current developments in Brazil, France, Italy and South Africa. *Renewable and Sustainable Energy Reviews*, Elsevier, 95, 305–327. doi:10.1016/j.rser.2018.06.066
16. H. Yang *et al.*, (2019), Tender Tea Shoots Recognition and Positioning for Picking Robot Using Improved YOLO-V3 Model," in *IEEE Access*, vol. 7, pp. 180998-181011. doi: 10.1109/ACCESS.2019.2958614.
17. Agyekum, K., Adinyira, E. and Amudjie, J. (2021), "Ethical misconducts within the invitation to tender and tender evaluation and award stages of construction contracts in Ghana", *Journal of Engineering, Design and Technology*, Vol. ahead-of-print No. ahead-of-print. <https://doi.org/10.1108/JEDT-07-2020-0274>
18. Haddadi, Tarik El; Mourabit, Taoufik; Haddadi, Anass El (2021). Sustainable public procurement in morocco: An investigative survey regarding tender preparation. *Sustainable Production and Consumption*, Elsevier, 26, 33–43. doi:10.1016/j.spc.2020.09.002
19. Holma, Anne-Maria; Vesalainen, Jukka; Söderman, Anne; Sammalmaa, Johanna (2019). Service specification in pre-tender phase of public procurement - A triadic model of meaningful involvement. *Journal of Purchasing and Supply Management*, Elsevier, 100580. doi:10.1016/j.pursup.2019.100580
20. Love, P.E.D.; Zhou, J.; Matthews, J.; Sing, M.C.P.; Edwards, D.J. (2017). System information modelling in practice: Analysis of tender documentation quality in a mining mega-project. *Automation in Construction*, Elsevier, 84, 176–183. doi:10.1016/j.autcon.2017.08.034
21. Cheng, Wenjuan; Appolloni, Andrea; D'Amato, Alessio; Zhu, Qinghua (2017). Green Public Procurement, Missing Concepts and Future Trends – A Critical Review. *Journal of Cleaner Production*, Elsevier, Volume 176, 1 March 2018, Pages 770-784 doi:10.1016/j.jclepro.2017.12.027
22. Uyarra, Elvira; Zabala-Iturriagoitia, Jon Mikel; Flanagan, Kieron; Magro, Edurne (2020). Public procurement, innovation and industrial policy: Rationales, roles, capabilities and implementation. *Research Policy*, Elsevier, 49(1), 103844. doi:10.1016/j.respol.2019.103844
23. Sönnichsen, Sönnich Dahl; Clement, Jesper (2019). Review of green and sustainable public procurement: Towards circular public procurement. *Journal of Cleaner Production*, 118901. doi:10.1016/j.jclepro.2019.118901

Мисюра Ю. О.

Черкаський національний університет ім. Б. Хмельницького

ПРОГРАМА ВІДНОВЛЕННЯ ТЕКСТОВИХ ПОВІДОМЛЕНЬ З БАГАТОКРАТНИМ ВИПРАВЛЕННЯМ ПОМИЛОК

Сучасні інформаційні системи ставлять досить високі вимоги до надійності і безпеки передачі інформації. Досить часто з різних причин переданий файл (наприклад, текстове

повідомлення) може виявитись пошкодженим. Тому необхідно якимось чином підвищити надійність передачі інформації.

Існують різні методи забезпечення цілісності повідомлень. Найчастіше вони полягають у додаванні деякої надлишкової інформації у повідомлення перед відправленням. Прикладами таких кодів є лінійні блокові коди Хеммінга [1, 134] та коди Ріда-Маллера [1, 144]. Але вони працюють з бітами, яких у великому повідомленні дуже багато. Це може спричинити великі часові затримки при кодуванні і декодуванні великих текстових повідомлень.

Тому ми будемо працювати з групами символів. Їх зручно представляти у вигляді одновимірних матриць. Для того, щоб закодувати групу символів, помножимо матрицю, що її представляє, на деяку іншу, яку назвемо хеш-матрицею. Ми маємо такі вимоги до хеш-матриці – вона мусить мати кількість рядків, рівну довжині групи символів, а також бути квадратною і невиродженою (її визначник не повинен дорівнювати нулю) для того, щоб ми могли отримати обернену до неї матрицю [2].

Розглянемо алгоритм кодування повідомлення. Для того, щоб закодувати повідомлення, ми розбиваємо його на групи символів, які записуємо у одновимірні матриці, а потім множимо кожен з них на хеш-матрицю. В результаті матимемо файл з такою структурою:

$$d_1 d_2 \dots d_n C_1 C_2 \dots C_n,$$

де d_i – i -та група символів повідомлення, C_i – i -та матриця, що є результатом описаної вище операції кодування.

Для декодування і виправлення помилок у отриманому повідомленні ми знову виконуємо операцію кодування – множення d_i на хеш-матрицю, яку ми позначимо як M . Якщо отримана матриця дорівнює C_i , то i -тий блок було передано без помилок.

Якщо ж отримана матриця не дорівнює C_i , то для виправлення помилки скористаємося властивостями оберненої матриці [2]:

$$M \cdot M^{-1} = E,$$

де M^{-1} – обернена до M матриця, E – одинична матриця.

Отже, отримуємо правильний блок повідомлення за формулою:

$$d_i \cdot M \cdot M^{-1} = C_i \cdot M^{-1},$$
$$d_i = C_i \cdot M^{-1}.$$

Для програмної реалізації модулів кодування і декодування ми використали бібліотеку Math.NET Numerics [3], яка містить необхідні нам операції з матрицями. Складністю було те, що ця бібліотека може записати у файл лише одну матрицю за раз, тому ми мусили скласти результати кодування повідомлення в одну матрицю за допомогою операції конкатенації.

Отже, ми проаналізували і дослідили один із способів кодування тестових повідомлень із подальшою можливістю виправлення помилок. Отримані нами результати дослідження можуть допомогти у розробці прикладних систем передачі даних, комп'ютерної безпеки та захисту документів різних типів.

Список використаної літератури

1. У. Питерсон, Э. Уэлдон. Коды, исправляющие ошибки. Москва : Издательство «Мир», 1976. 595 с.
2. Обернена матриця [Електронний ресурс]. – Режим доступу: <http://ua.onlinesechool.com/math/library/matrix/inverse/>
3. Math.NET Numerics [Електронний ресурс]. – Режим доступу: <https://numerics.mathdotnet.com/>

Науковий керівник: старший викладач, к.т.н. Ярмілко А. В.