

DOI:

УДК 004.9

**О.М. Міхайлуца**, к.т.н., доцент, elenamikhaylutsa7@gmail.com

**А.В. Пожусь**, к.ф.-м.н., доцент, scorpio6828@gmail.com

**В.О. Небеснюк**, магістрант

Запорізька державна інженерна академія, м. Запоріжжя

## МОДЕЛЮВАННЯ СИСТЕМИ РОЗПІЗНАВАННЯ МАТЕМАТИЧНИХ ВИРАЗІВ У ГРАФІЧНОМУ ФОРМАТІ ПРЕДСТАВЛЕННЯ ДАНИХ

*В роботі розглядається задача розпізнавання рукописних математичних виразів. Визначено ряд способів, завдяки яким здійснюється виділення ознак математичних формул та розпізнавання математичної моделі на зображенні. Описується метод навчання програми розпізнавання, заснований на аналізі формул, та його технічна реалізація. Приділено увагу взаємовигідній людино-машинній взаємодії.*

**Ключові слова:** нейронні мережі, розпізнавання виразів, сегментація зображень, кластеризація.

*The task of recognition of the handwritten mathematical expressions is in-process examined. The row of methods due to that extraction of features of mathematical formulas and recognition of mathematical model comes true dark-and-light is certain. The method of studies of the program of recognition, based on the analysis of formulas, is described. The attention is paid to the mutually beneficial of human-machine interaction.*

**Keywords:** neural networks, spelling recognition, image segmentation, clusterization.

### Постановка проблеми

Незважаючи на те, що друкований текст, конкуруючи з рукописним, намагається витіснити останній, це навряд чи станеться повністю (принаймні в найближчі десятиліття). Ні ведення шкільних зошитів, ні письмова фіксація лекційного матеріалу, нехай навіть фрагментарно, ні ведення шкільних журналів і медичних карт не можуть повністю зжити себе в рукописній формі. Отже, існує безліч рукописних матеріалів, які потребують переведення в електронний варіант. Окрім того, в останній час набули популярності різноманітні тач-скрін прилади, такі як планшети, фаблети, смартфони та комп'ютери. Значно зручніше написати текст від руки, ніж друкувати на клавіатурі, яка може бути маленькою та незручною, особливо коли йде мова про математичні вирази.

Розпізнавання формул дає широкі можливості, такі як пошук наукових документів за математичними формулами і підформулами, формування бази знань формул, перенесення формул в інші документи, вилучення семантичної інформації з формули, інтегрування з системами комп'ютерної алгебри. В даний час не існує повномасштабного проекту аналізу математичних формул, безкоштовного для використання і з відкритим вихідним кодом. Розвинені проекти на зразок Infty [7] є пропрієтарними, як і інструменти оптичного розпізнавання, використовувани подібними системами. Використання контенту документу покликане знизити залежність від витратної і неточної процедури оптичного розпізнавання тексту.

Розпізнати формулу в документі означає поставити границю області, що містить всі елементи формули, розпізнати символи, що входять в дану область, згрупувати символи в слові одиниці, зазначені геометричним положенням, — токени, ідентифікувати зв'язок між токенами, відновити ієрархічну структуру формули. Отримана структура може бути трансльована в будь-який формат запису формули.

Математичні вирази являють собою двовимірну ієрархічну структуру даних. Формули можуть містити символи латинського алфавіту, грецького алфавіту, спеціальні символи і графічні об'єкти. Відзначимо, що зображення формул не формалізоване строго, в представленні формул відсутня канонічна форма запису. Точне визначення математичного виразу вимагає завдання формального списку математичних символів в базі знань формул.

Означені проблеми, безумовно, вимагають свого вирішення. З поступовою автоматизацією багатьох процесів в різних сферах більш пріоритетним є навчання комп'ютера розпізнаванню формул, а не підготовка спеціально навчених людей. Отже, розроблення інформаційної технології для автоматизації процесу розпізнавання математичних виразів у графічному форматі представлення даних є актуальним завданням.

#### **Аналіз останніх досліджень та публікацій**

Огляд основних методик аналізу та розпізнавання зроблено в роботах Н.С. Ісупова та А.В. Кучуганова, де авторами запропоновано новий підхід у розпізнаванні злитих рукописних текстів, заснований на поданні векторизованого тексту у вигляді нечітких навантажених графів [1]. Детальний розгляд задач розпізнавання рукописних математичних формул подано А.Ю. Сапаровим [2], [3]. Розроблено та досліджено засоби аналізу зображень зі складноструктурованими об'єктами і автоматичного розпізнавання рукописних математичних формул зі сканованих зображень, в яких одночасно можуть перебувати як рукописні формули, так і звичайний друкований текст. Описується метод навчання програми розпізнавання, заснований на аналізі формул, записаних різними почерками. Основна увага приділяється взаємовигідній людиномашинній взаємодії, де програма і людина отримують можливість взаємного навчання.

Серед останніх досліджень в галузі розпізнавання математичних формул виділяють роботи, в яких аналіз здійснюють з метою виявлення подібностей [4]. Для знаходження неповних дублікатів математичних формул запропоновано гібридний підхід, що передбачає використання шаблонів, створених відповідно до особливостей графічних редакторів, та спеціальних конверторів формул з різних форматів до канонічного формату.

Особливе місце займають задачі, в яких зроблена спроба врахувати формат документу. Розглядаються питання розпізнавання математичних формул в електронних документах наукових статей PDF формату, їх видалення та збереження в структурній формі в деякій «базі знань» математичних формул, а також нюанси відповідної програмної реалізації [5]. Серед властивостей, які допомагають ідентифікувати і розпізнавати формули, виділяють властивості математичної нотації, а саме геометричні властивості, властивості зображення, контекст, які дозволяють використовувати евристичні правила аналізу. До розпізнавання формул адаптуються статистичні методи розпізнавання: за навчальними екземплярами формул, гістограмою певних параметрів, методам геометричної декомпозиції формули, а також їх комбінації один з одним. Для розпізнавання відносин між токенами використовуються різноманітні розширення граматики.

Проаналізувавши ринок програмного забезпечення можна зробити висновок, що на даний час існують наступні програми, які дозволяють розпізнати математичні формули: Math input panel, InftyReader, GOCR, Wolfram. Math input panel є системою динамічного розпізнавання рукописних формул, тому не може бути застосована до обробки сканованих текстів. InftyReader і GOCR призначені для розпізнавання друкованих текстів з растрових зображень. Оскільки необхідна якість не досягається при скануванні зображень, тому точність розпізнавання невисока. Жодна з існуючих систем не розпізнає скановані рукописні формули.

Проект Wolfram надає можливість переведення природно-мовного питання в формат, зрозумілий для комп'ютерів. Він не повертає перелік посилань, заснований на результатах запиту, а обчислює відповідь, ґрунтуючись на власній базі знань, яка містить дані з математики, фізики, астрономії, хімії, біології, медицини, історії, географії, політики, музики, кінематографії, а також інформацію про відомих людей та інтернет-сайти. Він здатний переводити дані між різними одиницями виміру, системами числення, підбирати загальну формулу послідовності, знаходити можливі замкнуті форми для наближених дробових чисел, обчислювати суми, межі, інтеграли, розв'язувати рівняння і системи рівнянь, проводити операції з матрицями, визначати властивості чисел і геометричних фігур. На жаль, розрахунок на підставі власної бази є досить вразливий до помилок даних.

PhotoMath вирішує математичні задачі, використовуючи камеру вашого смартфона. Додаток вміє розв'язувати тільки не складні дроби, десяткові числа, лінійні рівняння і стандартну арифметику. Крім отриманого результату, PhotoMath також розпише весь процес вирішення рівняння по кроках. PhotoMath вміє розпізнавати лише друкований текст, рівняння "від руки" йому даються вкрай важко (звичайно, якщо вони не написані друкарським шрифтом).

### Формулювання мети дослідження

Дослідження систем з визначення тексту показало, що традиційно в основі лежать два підходи до розпізнавання: онлайн та офлайн. Найкращі результати з розпізнавання в обох підходах показують такі застосунки, що використовують нейронні мережі. Отже, метою дослідження є, по-перше, підвищення ефективності розпізнавання математичних виразів за рахунок класифікації та кластеризації зображень, а по-друге, вдосконалення існуючих методів розпізнавання математичних моделей, в алгоритм яких додано можливості багатоступеневої ідентифікації зображення різними методами оцінки математичного виразу. Розширення функціональних можливостей системи робить її більш доступною для використання іншими розробниками.

### Виклад основного матеріалу

#### 1. Розпізнавання структури математичних виразів

В он-лайн розпізнаванні рукописного математичного виразу зазвичай виділяють наступні стадії: збір і попередня обробка даних, розділення зображення на символи (сегментація виразу), розпізнавання окремих символів, розпізнавання структури математичного виразу.

Оскільки навчальна множина для конкретного класу складається з безлічі символів, а кожний символ зображується деякою кількістю розчерків у певному порядку, то метою попередньої обробки даних є перетворення «сирих» даних розчерків у формат, який дозволяє створити модель символу з метою його подальшої класифікації. Для спрощення моделі необхідно видалити ту інформацію, яка не потрібна при класифікації. Початкова безліч розчерків перетворюється в вектор певної і жорстко фіксованою розмірності.

Наступним кроком кожна точка розчерку зсувається так, що початкова точка є верхнім лівим кутом і обмежує рамки символу. Всі точки масштабуються зі збереженням відносин ширини до висоти так, що символ розташовується всередині квадрату. Набір точок кожного розчерку змінюється так, щоб точки розташовувалися рівновіддалено (методом лінійної інтерполяції). Кількість точок фіксована і дорівнює  $36/N$ , де  $N$  — кількість розчерків у символі. На завершення, координати точок збираються в єдиний вектор з 72 елементів, де перші 36 елементів представляють координати по  $x$ , а останні — координати по  $y$ . Отриманий таким чином набір векторів однієї розмірності, де кожному класу символів відповідає певна кількість векторів, можна використовувати, наприклад, для навчання нейронної мережі, або для розпізнавання символу за алгоритмом  $k$  — найближчих сусідів.

Класифікатор розпізнає символи окремо. Оскільки, за необхідністю розпізнавати поточе введення, він не може заздалегідь знати, які розчерки становлять символ, і скільки символів всього введено, можливості класифікатора необхідно доповнити визначенням кількості і розташуванням символів виразу, використовуючи сегментацію виразу [6]. Для визначення символів, що складаються з декількох розчерків, необхідно послідовно розглядати групи розчерків заданих розмірів. В ході структурного аналізу здійснюється перейменування символів, які неможливо розпізнати як різні (наприклад, знаки мінуса і дробни), позначаючи їх як один символ (горизонтальна риса). Алгоритм сегментації поліпшено за рахунок визначення допустимих символів для кожної кількості розчерків.

На етапі розпізнавання структури математичного виразу співвідношення математичних операторів визначаються в залежності від позиції та відносного розміру символу в виразі, використовуючи просторові області «зверху зліва», «зверху», «верхній індекс», «нижній індекс», «знизу», «знизу зліва» і «підвираз». Виходячи з простих атрибутів визначається поріг верхнього та нижнього індексів, а також центральна точка, яка визначає розташування символу в будь-якій з областей. З метою уникнення неоднозначностей при структурному аналізі у кожному класі символів атрибути визначаються по-різному. Окремий символ має посилання на базові лінії, кожна з яких являє собою список з горизонтального розташування символів у виразі, та які задовольняють різним просторовим відносинам. Функція визначення домінування символів побудована на правилі, що кожний наступний крок виконується, якщо на попередньому не вдалося визначити, який з символів домінує над яким. Читаючи математичний вираз за допомогою рекурсивної процедури спочатку визначається крайній лівий домінуючий символ в списку, потім наступний за ним домінуючий символ, і так далі, доки не будуть знайдені всі символи до-

мінуючої базової лінії, на яку не посилається жоден з них. На останньому кроці рекурсивним знаходженням домінуючих базових ліній у виразі, описаному впорядкованим списком символів, створюється структура даних, а саме дерево базових ліній.

### **2. Реалізація розпізнавання символу**

Задача сегментації зображення може бути зведена до кластеризації множини точок, що відповідають окремим пікселям. Алгоритми сегментації, що ґрунтуються на штучних нейронних мережах, відрізняються один від одного типом нейронної мережі та алгоритмом навчання. Слід зазначити, що більшість алгоритмів носить спеціалізований характер і орієнтована на зображення певного типу [7]. Це обмеження пов'язане з проблемою формування навчальної множини. Для ефективного навчання будь якої нейронної мережі необхідна досить велика кількість зображень, схожих з відомими результатами. Основним завданням стає не просто навчання, а знаходження нейронної мережі, яка найкращим (в певному сенсі) чином вирішує поставлену прикладну задачу.

Припустивши, що сітка складається з  $28 \times 28 = 784$  пікселів, вважаємо що є 784 нейрона, які містять різні числа від 0 до 1: чим ближче піксель до білого кольору, тим ближче відповідне число до одиниці (ці числа називають активації нейронів.) Перший шар нейромережі утворюють описані 784 нейрони, а останній шар містить 10 нейронів, кожен з яких відповідає одній з десяти цифр. У цих числах активація є також числом від нуля до одиниці, що відображає наскільки система впевнена, що вхідне зображення містить відповідну цифру. Також є пара середніх шарів, що називаються прихованими.

Принцип роботи нейромережі полягає в тому, що активація в одному шарі визначає активацію в наступному. Якщо передати навченій нейронній мережі на перший шар значення активації згідно яскравості кожного пікселя картини, ланцюжок активацій від одного шару нейромережі до наступного призведе до переважної активації одного з нейронів останнього шару, відповідного розпізнаній цифрі — вибору нейронної мережі. У процесі розпізнавання цифр зводяться воедино різні компоненти і кожен нейрон з другого шару співвідноситься з одним з цих компонентів, а перехід від другого прихованого шару до вихідного відповідає знанням про те, який набір компонентів якій цифрі відповідає. Отже, визначення граней і шаблонів корисно не тільки в задачі розпізнавання цифр, але й в задачах розпізнавання образів.

Оскільки бібліотеки машинного навчання оптимізовані під векторні обчислення, то будується математична модель нейронної мережі. Для її математичного опису всі активації першого шару об'єднуються в вектор-стовпець, а всі ваги — в матрицю, кожен рядок якої описує з'єднання між нейронами одного шару з конкретним нейроном наступного. Результатом множення матриці на вектор є вектор, що відповідає зваженим сумах активацій першого шару. До уваги береться той факт, що нейронам відповідають не просто числа-активації, а функції активації, які приймають значення з усіх нейронів попереднього шару і обчислюють вихідні значення в інтервалі від 0 до 1.

Найчастіше у якості активаційної функції використовують так званий сигмоїд, однією з переваг якого є диференційованість на усій осі абсцис та досить проста похідна. Незважаючи на те, що функція-сигмоїд є прикладом, який наслідує біологічні нейрони, запропоновано використання функції, що відповідає біологічній аналогії того, чи перебувають нейрони в активному стані. Цю функцію було протестовано і на глибоких багатосарових мережах. У загальному вигляді алгоритм полягає в тому, щоб показати нейромережі безліч тренувальних даних, що представляють пари зображень, складених з записаних від руки цифр та їх абстрактних математичних уявлень. В результаті навчання нейромережа правильним чином розрізняє числа з рідше не представлених тестових даних.

### **3. Аналіз можливостей запропонованого застосування**

Розроблена комп'ютерна система автоматичного розпізнавання рукописних формул. Алгоритм аналізу математичних моделей, до якого додано можливості багатоступеневої ідентифікації зображення різними методами оцінки математичного виразу, дозволяє перевіряти результат розпізнавання з точки зору відповідності мови математичних виразів.

Проект потребує зберігання даних про натреновану модель. Для кожного з прихованих шарів запропонованої моделі створюється окремий бінарний файл, в якому зосереджені усі ваги

зав'язків нейронів. Програмний продукт використовує зображення в якості вхідних даних, а завдяки застосованим бібліотекам дає змогу використовувати зображення з будь-яким розширенням. Реалізована можливість отримання результатів в LaTeX форматі.

#### **Висновки та перспективи подальших досліджень**

Після дослідження предметної області та детального аналізу поставленої проблеми встановлено актуальність розробки системи розпізнавання математичних виразів у цифровому форматі представлення даних. Шляхом аналізу нейронних мереж та проведення теоретичних і практичних досліджень сформовано набір методів для дослідження, проведено їх порівняння та аналіз, а також визначено стик технологій інструментів розробки, визначено їх відносну продуктивність, ступінь інтеграції та зручність використання. Розроблено алгоритм розпізнавання роздільно-рукописного математичного тексту та його експорту в електронний документ формату LaTeX.

Отримані результати можуть бути, зокрема, використані для розширення функціональних можливостей наявних систем автоматичного перекладу рукописного тексту в цифровий формат для інформаційних сайтів, а також для розробки більш клієнтоорієнтованих систем допомоги студентам та школярам.

У подальшому розвитку планується розробити систему адаптації програми до почерку конкретної людини. Це дозволить створювати зменшені бази еталонів, що, в свою чергу, прискорить процес розпізнавання.

#### **Список використаної літератури**

1. Исупов Н. С. Распознавание слитных рукописных текстов с использованием аппарата нечеткой логики / Н. С. Исупов, А. В. Кучуганов // Вестник Ижевского государственного технического университета имени М. Т. Калашникова. – 2012. – № 3. – С. 125–128.
2. Сапаров А. Ю. Применение регулярных выражений в распознавании математических текстов / А. Ю. Сапаров, А. П. Бельтюков // Вестник Удмуртского университета. Математика. Механика. Компьютерные науки. – 2012. – Вып. 2. – С. 63–73.
3. Сапаров А. Ю. Разработка пользовательского интерфейса для управления процессом распознавания рукописных математических формул / А. Ю. Сапаров, И. Ю. Ширококова // Вестник Удмуртского университета. Математика. Механика. Компьютерные науки. – 2016. – № 1. – С. 141–152.
4. Лізунов П. П. Гібридний підхід до аналізу та розпізнавання математичних формул з метою виявлення в них подібностей / П. П. Лізунов, А. О. Білощицький, Л. Е. Чала, С. В. Білощицька, О. Ю. Кучанський, С. Г. Удовенко // Управління розвитком складних систем. – 2016. – № 27. – С. 145–155.
5. Lin X., Gao L., Tang Z. Mathematical formula identification in PDF documents in Proc. of ICDAR '11. Washington, DC, USA: IEEE Computer Society, 2011, pp. 1419–1423.
6. Chandhok S. A novel approach to image segmentation using artificial neural networks and K-means clustering // Intern. J. of Engineering Research and Applications (IJERA). 2012. Vol. 2. Iss. 3. Pp. 274–279.
7. Круглов В.В. Искусственные нейронные сети. Теория и практика / В. В. Борисов, В.В. Круглов – М.: Горячая линия–Телеком, 2001. – 328 с.
8. Белим С. В. Алгоритм сегментации изображения с помощью искусственной нейронной сети без использования других изображений / С. В. Белим, С. Б. Ларионов // Радиооптика. МГТУ им. Н.Э. Баумана. Электрон. журн. – 2017. – № 03. – С. 43–53.

**SIMULATION OF THE RECOGNITION SYSTEM OF MATHEMATICAL EXPRESSIONS IN A GRAPHICAL FORMAT OF DATA PRESENTATION****Mikhailutsa O.M., Pozhuyev A.V., Nebesniuk V.O.****Abstract**

With gradual automation of many processes in various areas the priority is training a computer to recognition of a handwritten material, instead of preparation of specially trained people. The authors direct an attention on the problem of recognition of mathematical expressions, presented in graphical format.

This is especially useful by search of scientific documents on mathematical formulas, formation of the knowledge base of formulas, carry of formulas to other documents, withdrawal of the semantic information from the formula, integration with computer algebra systems.

Mathematical expressions represent bidimensional hierarchical data structure, which precise definition requires the task of the formal list of mathematical symbols in the knowledge base of formulas. The research of text definition systems has shown that traditionally in a basis two approaches to recognition lay: online and offline. The best recognition results in both approaches show such applications that use neural networks. So, the purpose of the research is, firstly, to increase the efficiency of recognition of mathematical expressions due to the classification and clustering of images, and secondly, to improve the existing methods of recognition of mathematical models, in which algorithm possibilities multistage identifications of the image by various methods of an assessment of mathematical expression are added.

By analyzing neural networks and conducting theoretical and practical research, a set of research methods was formed, compared and analyzed, and a stack of technologies and development tools was determined, their relative performance, degree of integration, and usability were determined. The algorithm for recognition of the separate handwritten mathematical text and its export to the electronic document of the LaTeX format has been developed.

The received results can be used, in particular, to expand the functionality of existing systems for machine translation of the handwritten text into digital format for information sites, as well as to develop more client-oriented systems of assistance to students and schoolchildren. In the further progress it is planned to develop a system of adaptation of the program to the handwriting of the specific person.

**References**

- [1] Isupov N. S., Kuchuganov A. V. *Raspoznavaniye slitnykh rukopisnykh tekstov s ispol'zovaniyem apparata nechetkoy logiki* [Recognition of coherent handwritten texts using the apparatus of fuzzy logic]. The Bulletin of Izhevsk State Technical University named after MT Kalashnikov. – 2012. – № 3. – pp. 125–128.
- [2] Saparov A. Yu., Beltyukov A. P. *Primeneniye regul'yarnykh vyrazheniy v raspoznavanii matematicheskikh tekstov* [The use of regular expressions in the recognition of mathematical texts]. The Bulletin of Udmurt University. Maths. Mechanics. Computer science. – 2012. – Vol. 2. – pp. 63–73.
- [3] Saparov A. Yu., Shirobokova I. Yu. *Razrabotka pol'zovatel'skogo interfeysa dlya upravleniya protsessom raspoznavaniya rukopisnykh matematicheskikh formul* [Development of the user interface for managing the process of recognizing handwritten mathematical formulas]. The Bulletin of Udmurt University. Maths. Mechanics. Computer science. – 2016. – № 1. – pp. 141–152.
- [4] Lizunov P.P., Biloshchitsky A.A., Chala L.E., Beloshchitskaya S.V., Kuchanskiy O.Yu., Udovenko S.G. *Hibrydnyy pidkhid do analizu ta rozpoznavannya matematychnykh formul z metoyu vyavleniya v nykh podibnostey* [A hybrid approach to the analysis and recognition of mathematical formulas in order to identify similarities in them]. Management of complex systems development. – 2016. – No. 27. – pp. 145–155.
- [5] Lin X., Gao L., Tang Z. Mathematical formula identification in PDF documents in Proc. of ICDAR '11. Washington, DC, USA: IEEE Computer Society, 2011, pp. 1419–1423.

- [6] Chandhok C. A novel approach to image segmentation using artificial neural networks and K-means clustering // Intern. J. of Engineering Research and Applications (IJERA). 2012. Vol. 2. Iss. 3. pp. 274–279.
- [7] Kruglov V.V., Borisov V.V. *Iskusstvennyye neyronnyye seti. Teoriya i praktika* [Artificial neural networks. Theory and Practice]. – M.: Hotline – Telecom, 2001. – 328 p.
- [8] Belim S.V., Larionov S.B. *Algoritm segmentatsii izobrazheniya s pomoshch'yu iskusstvennoy neyronnoy seti bez ispol'zovaniya drugikh izobrazheniy* [Algorithm of image segmentation using an artificial neural network without using other images]. Radio-optics. MGTU them. N.E. Bauman. Electron. journals – 2017. – № 03. – pp. 43–53.