[1]N. V. Shapoval,
[2]D. A. Lobanov

DETECTING HANDWRITTEN LINE FOR TIKZ CODE GENERATING

[1,2]National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine

E-mails: [1]shovgun@gmail.com, [2]mvch15.1.denya.lobanov1@gmail.com   ORCID 0000-0002-8509-6886

Abstract—*This work is devoted to the recognition of straight lines in geometric drawings created by hand with the help of neural networks. The LaTeX language is usually used for the design of articles, which takes more time, especially when writing formulas or constructing geometric drawings, unlike writing by hand. Automating the construction of drawings will make it possible to speed up the process of writing articles. The paper considers the recognition of straight lines as the most popular element of geometric drawings. The Hough transform for straight lines detection and its disadvantages are considered. The use of convolutional neural networks for this task is proposed as the best tool for working with images. To train the model, a dataset was created with handwritten lines and lines constructed in a graphics editor. The results of the neural network are given.*

Index Terms—Hough transform; convolutional neural networks; line detection.

I. INTRODUCTION

LaTeX is one of the most popular markup languages for document design and is especially popular among mathematicians, physicists, and other scientists for publishing scientific articles that often contain mathematical equations. Writing such equations in LaTeX takes more time than writing the same equations by hand on paper. Many studies have focused on the generation of LaTeX code based on a handwritten formula [1] – [3]. The same applies to drawing graphs, diagrams, etc., especially in the field of mathematics. There is Inkscape software that allows you to create high quality vector graphics. However, Inkscape has its drawbacks, such as the limited ability to quickly make changes to a drawing and the time it takes to master the program.

LaTeX allows users to connect various packages for high-quality document design, including TikZ. TikZ is a separate language for creating vector graphics. It allows LaTeX users to draw simple geometric shapes, graph functions, create diagrams, and other graphical elements. Although there are explanatory materials for learning TikZ, it has its drawbacks. First of all, the commands in TikZ can be difficult to understand, which takes time to learn. In addition, the construction of graphs and other graphic elements in TikZ can take a lot of time, including the time to compile a LaTeX document.

Modern approaches to recognizing handwritten mathematical formulas usually use convolutional neural networks, which show excellent results [4], [5]. Convolutional neural networks have the advantages of hierarchical feature learning and invariance to shift and scaling, making them effective for image analysis. At the same time, there are almost no scientific works devoted to the recognition of geometric handwritten drawings. This study is focused on the detection of straight lines in geometric drawings.

II. PROBLEM STATEMENT

A. Hough transform for line detection

The Hough transform is an algorithm used in image processing to detect lines and other geometric shapes in an image. It is based on a mathematical theory that allows you to convert the pixel coordinates of the image into a parametric form of straight lines that can be used to detect them. The Hough transform is widely used in computer vision and image processing to detect lines, circles, ellipses, and other geometric shapes.

In the Hough transform, the straight line is described by the equation $x\sin\varphi + y\cos\varphi = \rho$, where ρ is the distance from the origin to the nearest point on a straight line, and φ is the angle between the *x*-axis and the line connecting the origin to this nearest point. Lines are sought in the Hough space: plane (ρ, θ). For each pixel (x_i, y_i) we make a reflection in the Hough space. For a given point (pixel), the set of all straight lines passing through it corresponds to a sinusoidal curve on the (r, θ) plane unique to that point. If you map all the image points onto the Hough space, it will create many sinusoidal curves. If two points lie on the same straight line, their corresponding sine curves will intersect each other at a certain pair (ρ, θ). Thus, the Hough

transform algorithm detects lines by finding pairs (ρ, θ) whose number of intersections exceeds a certain threshold [6].

Convolutional neural networks are used in combination with the Hough transform to improve the quality of line recognition in images [7]. Thus, the authors of [8] use convolutional neural networks when mapping to Hough space.

B. Disadvantages of Hough transform

The Hough transform requires high computational costs and has unstable performance, since it is necessary to select a threshold for a specific task. When using the Hough transform to find a straight line in pictures, the problems shown in Figs 1 and 2 arise.
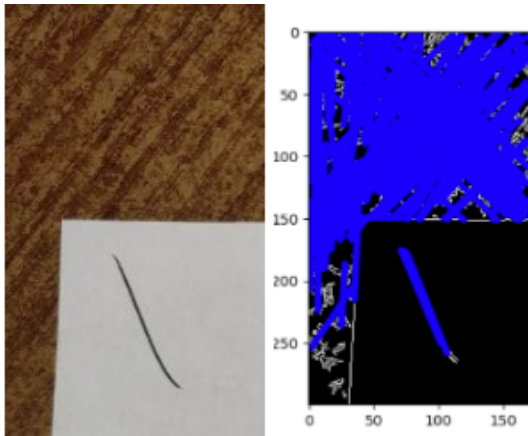
Fig. 1. The noises present in the image lead to the detection of extra lines

First and foremost, noise in the image can affect the line detection process, causing distortion or loss of points lying on the lines. This can lead to incorrect detection or low accuracy of line recognition.
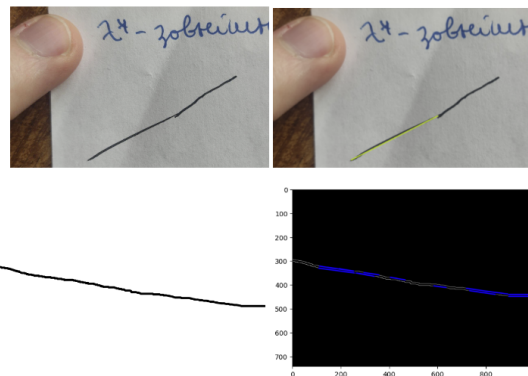
Fig. 2. Hough transform cannot detect whole line

Also, before using the Hough transform, it is necessary to use the methods of edge detection on the image for better detection of straight lines, and this

may lead to the loss of the sought straight lines. Lines drawn by hand may not be exactly straight, which makes it difficult to clearly define a straight line.

III. PROBLEM SOLUTION

A. Data collection and preprocessing

To solve the problem, a dataset of images was created, containing lines drawn by hand and in a graphic editor. A total of 300 images were collected, which were scaled to 128x128 and converted to grayscale. To expand the data sample, augmentation was performed: rotation and mirroring. A total of 90 options for rotations, from 0° to 89°, and 4 options for reflections: none, horizontal, vertical, both. A total of 360 options for applying augmentation to one image. Thus, the sample increased to 108,000 images. Marked data is necessary for training. So, for the initial 300 images, the coordinates of the ends of the straight lines were determined. Using rotational augmentation can cause the loss of straight line ends, as shown in Fig. 3.

The new edges of the line are calculated as follows: for the equation of a straight line $\dfrac{x - x_1}{x_2 - x_1} = \dfrac{y - y_1}{y_2 - y_1}$ the value is set $y = 0$ or $y = 128$ and new coordinates of x, are calculated, as in Fig. 4. Similarly for x. Then they obtained coordinates are normalized.
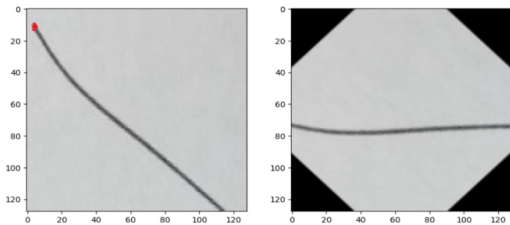
Fig. 3. Example of image after 45° rotation where the red edge is out of vision
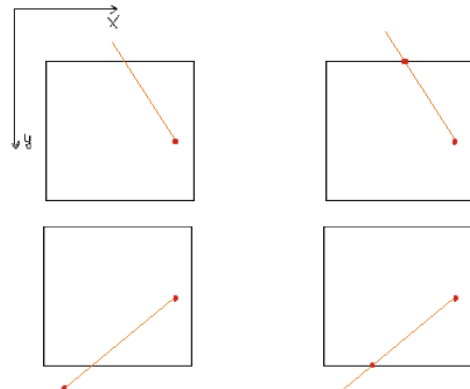
Fig. 4. New coordinates after rotation

If both edges of the line are outside the drawing area, then we update first x, then y, or vice versa,

until all edges are inside. If the updated coordinates of both edges coincide, it means that the line is outside the field of view of the image (Fig. 5). Hence this scenario must not be considered in the dataset.
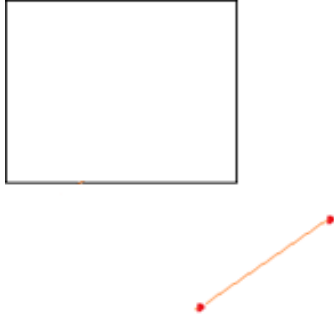

Fig. 5. A line which is out of vision

B. Experimental results

In order to solve the problem of line recognition in the drawing, a convolutional neural network was proposed. It consists of three convolutional layers with the number of filters 32, 64, 128 respectively, three fully connected layers. Adam optimizer and MAE loss function was used. The convolution layers have a filter size of 3 by 3, padding='same' to keep the image the same size. Max pooling layers are used, which reduce the image by half (Fig. 6).
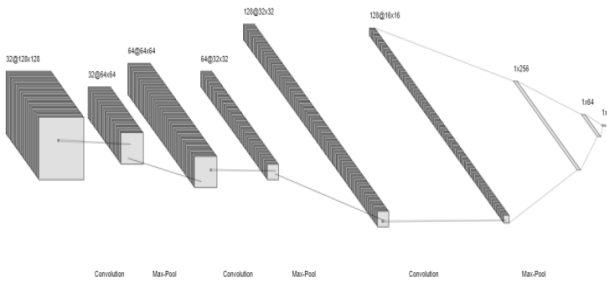

Fig. 6. Proposed architecture of convolutional neural network

After finding the coordinates of the edges of the line, the TikZ code is generated, which is used to draw the line. The results of the work are shown in Fig. 7.
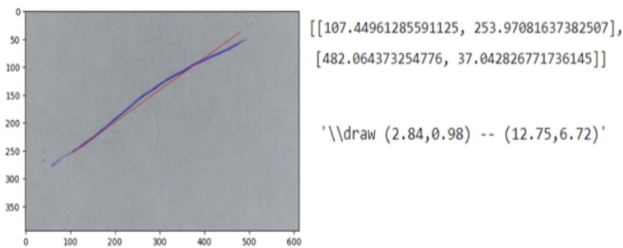

Fig. 7. Found the coordinates of the line and generated the line using the TikZ code

In 85% of cases, the network successfully recognized the lines and the corresponding TikZ code was generated. Nevertheless, for images with vertical lines, the network gave an unsatisfactory result.

If there are several straight lines in the picture, it is suggested to divide the image into several subimages on which there will be only one straight line (Fig. 8). After that, the obtained subimages will be fed to the convolutional neural network, and the received coordinates of the segments will be combined into one image of intersecting lines (Fig. 9).
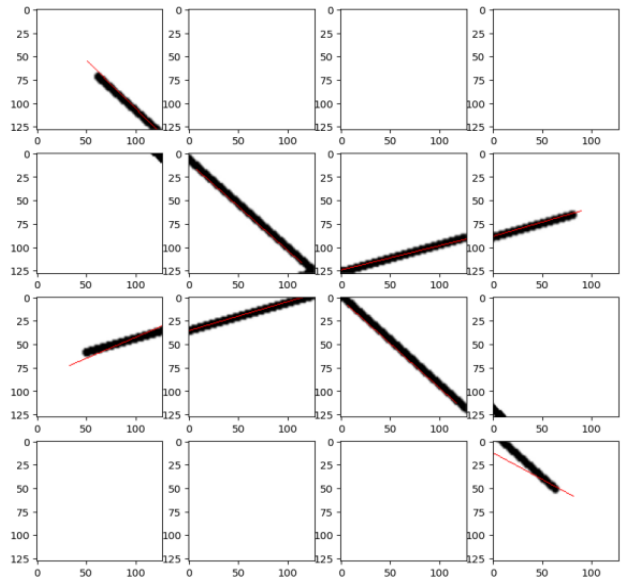

Fig. 8. Image partition, each of these subimages are sent to neural network to detect a line
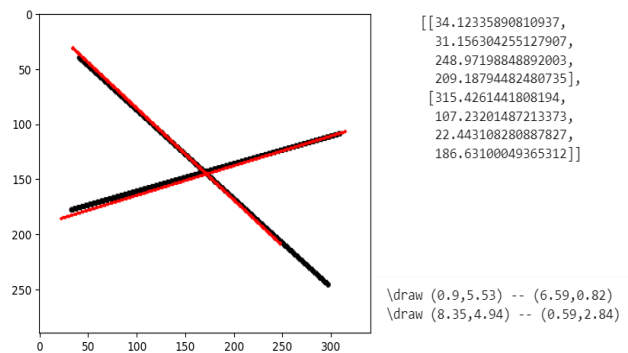

Fig. 9. Found the coordinates of more than 1 line

This approach allows you to detect intersecting lines in more than half of the cases.

IV. CONCLUSIONS

Recognition of handwritten geometric drawings is an interesting task that has received little attention. Creating an application that will be able to convert recognized geometric shapes into TikZ code will reduce the time it takes to design LaTeX documents. The use of convolutional neural networks avoids the

problems associated with the use of the Hough transform in line detection. A dataset for training a convolutional neural network was created, which consisted of drawn straight lines. To improve the performance of the network, it is necessary to expand the dataset with pictures that contains a different number of lines that can intersect. The solution of automatic recognition of handwritten geometric drawings is promising, and the definition of a straight line is the first step to solve this problem.

REFERENCES

[1] WANG Zelun and LIU Jyh-Charn, "Translating math formula images to LaTeX sequences using deep neural networks with sequence-level training," *International Journal on Document Analysis and Recognition (IJDAR)*, 2021, 24.1–2: 63–75. https://doi.org/10.1007/s10032-020-00360-2

[2] SPRINGSTEIN Matthias, MÜLLER-BUDACK Eric, and EWERTH Ralph, "Unsupervised training data generation of handwritten formulas using generative adversarial networks with self-attention," In: *Proceedings of the 2021 Workshop on Multi-Modal Pre-Training for Multimedia Understanding*, 2021, pp. 46–54. https://doi.org/10.1145/3463945.3469059

[3] BIAN Xiaohang, et al., "Handwritten mathematical expression recognition via attention aggregation based bi-directional mutual learning," In: *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022, pp. 113–121. https://doi.org/10.1609/aaai.v36i1.19885

[4] SIDDIQUE Fathma, SAKIB Shadman, and SIDDIQUE Md Abu Bakr, "Recognition of handwritten digit using convolutional neural network in python with tensorflow and comparison of performance for various hidden layers," In: *2019 5th international conference on advances in electrical engineering (ICAEE)*, IEEE, 2019, pp. 541–546. https://doi.org/10.1109/ICAEE48663.2019.8975496

[5] ANG Zelun and LIU Jyh-Charn, "Translating math formula images to LaTeX sequences using deep neural networks with sequence-level training,". *International Journal on Document Analysis and Recognition (IJDAR)*, 2021, 24.1–2: 63–75. https://doi.org/10.1007/s10032-020-00360-2

[6] HEROUT Adam, et al., "Review of Hough transform for line detection," *Real-Time Detection of Lines and Grids: By PClines and Other Approaches*, 2013, 3–16. https://doi.org/10.1007/978-1-4471-4414-4_2

[7] ADEM Kemal, "Impact of activation functions and number of layers on detection of exudates using circular Hough transform and convolutional neural networks," *Expert Systems with Applications*, 2022, 203: 117583. https://doi.org/10.1016/j.eswa.2022.117583

[8] ZHAO Kai, et al., "Deep hough transform for semantic line detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 44.9: 4793–4806.

Shapoval Nataliia. ORCID 0000-0002-8509-6886. Candidate of Sciences (Engineering).
National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (2010).
Research interests: computer vision, fuzzy neural network, deep neural network.
Publications: 6.
E-mail: shovgun@gmail.com

Lobanov Denys. Bachelor's degree student.
Department of Artificial Intelligence, Institute of Applied Systems Analysis, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine.
Research area: computer vision, deep neural network.
E-mail: mvch15.1.denya.lobanov1@gmail.com

Н. В. Шаповал, Д. А. Лобанов. Детекція прямих для генерації TikZ коду
Статтю присвячено розпізнаванню прямих ліній на намальованих від руки геометричних малюнках за допомогою нейронних мереж. Мова LaTeX зазвичай використовується для оформлення статей, що займає більше часу, особливо при написанні формул або побудові геометричних малюнків, на відміну від виконання тієї ж роботи від руки. Автоматизація побудови малюнків дасть змогу прискорити процес написання статей. У статті розглядається детекція прямих ліній, як найпопулярніших елементів геометричних малюнків. Розглянуто перетворення Гафа для виявлення прямих ліній та його недоліки. Використання згорткових нейронних мереж для цього завдання пропонується так як вони є найкращим інструментом для роботи із зображеннями. Для навчання моделі було створено набір даних із рукописними лініями та лініями, побудованими в графічному редакторі. Наведено результати роботи нейронної мережі.
Ключові слова: детекція прямої; згорткові нейронні мережі; перетворення Гафа.

Шаповал Наталія Віталіївна. ORCID 0000-0002-8509-6886. Кандидат технічних наук.

Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського», Київ, Україна.

Освіта: Національний технічний університет України "Київський політехнічний інститут", Київ, Україна, (2010).

Напрям наукової діяльності: комп'ютерний зір, нечіткі нейронні мережі, глибокі нейронні мережі.

Кількість публікацій: 6.

E-mail: shovgun@gmail.com

Лобанов Денис Артемович. Студент бакалаврату.

Кафедра штучного інтелекту, Інститут прикладного системного аналізу, Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського», Київ, Україна.

Напрям наукової діяльності: комп'ютерний зір, глибокі нейронні мережі.

E-mail: mvch15.1.denya.lobanov1@gmail.com