FONT VISUAL CLASSIFICATION SYSTEM
USING NEURAL NETWORKS

UDC 004.921

MERKULOVA Kateryna
Candidate of Science (technic), Associated professor, Head of the Department of Computer Technology of Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine, **E-mail:** k.merkulova@donnu.edu.ua, ORCID ID: 0000-0001-6347-5191.

ZAITSEVA Elina
Candidate of Science (technic), Associated professor of Department of Computer Technology of Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine, **E-mail:** zaytseva.elina@gmail.com, ORCID ID 0000-0002-2135-8146.

KOVALCHUK Artem
master`s program student of specialty 122 Computer Science of Vasyl' Stus Donetsk National University, Vinnytsia, Ukraine, **E-mail:** hardkovalchuk@gmail.com.

Abstract. Context and objective. The purpose of the paper is research and development of font visual classification system, that will allow users to work with fonts more effectively. That includes: select the required font using preview, apply various filters and font management tools. With the help of classification system users will be able to search fonts by slope, basic style and weight according to their actual representation avoiding the problem of incorrect sub-family styles names.

Research methods. In this paper the basic approaches of fonts classification with relevant advantages and disadvantages were described. After analyzing the approaches, it was decided to apply neural network that takes the images with font symbols as input. Neural network selects patterns (filters, features maps), to be considered in classification process based on the teacher's correct answers. Possible use of alternative fonts classification techniques was suggested, and the main related problems were described.

Results. Computer fonts classification system driven by convolutional neural networks, that allows to classify fonts by slope, basic style and weight according to their graphical representation, was developed. Percentage of fonts that were correctly classified by the system is: in determining slope — 96%, in determining basic style — 92%, in determining weight — 91%. The developed system can be applied to solving fonts classification tasks as an assistance tool for determination of digital documents structure, and as font selection system for design creation. Based on results, it may be concluded that neural networks were successfully applied to the fonts classification problem solving.

The scientific novelty and practical significance. In the work were investigated the main approaches to the classification of fonts, analyzed their advantages and disadvantages. The effectiveness of mulilayer perceptronsand convolutional neural networks were tested. Experimentally revealed the most optimal parameters of the models of neural networks.

A visual font classification system that works on convolutional neural networks was developed and integrated into the font organizer. Control tests were carried out confirming the high accuracy of classification of the developed neural network models.

Keywords: *machine learning, convolutional neural network, font classification, image classification.*

Introduction. The computer font is a file that contains description of a set of alphabetical, digital, reserved and pseudo-graphic characters and serves for displaying these characters (particularly, the text) by the program or the operating system. Every font has its weight (100-900), slope and one of the basic typefaces: Serif, Sans-Serif, Script, Decorative. There is an enormous number of fonts, however there is not a single standard of their classification. Fonts with a common subfamily name often considerably differ one from another by weight as much as by slope. In some cases, when development of all types of weight is not needed, the authors of the font families make so called "shifts", when actual weight of the font corresponds to another weight. Also, after developing the single font in the sub-family, author names it Regular regardless of the font type.

The lack of the standard causes the problem of font search and selection. For example, when searching the font by style Regular, the results may contain fonts more similar to Black or even Thin Italic.

Problem definition. The purpose of the paper is research and development of font visual classification system, that will allow users to work with fonts more effectively. That includes: select the required font using preview, apply various filters and font management tools. With the help of classification system users will be able to search fonts by slope, basic style and weight according to their actual representation avoiding the problem of incorrect sub-family styles names. Developed system must correctly classify >= 90% of fonts.

Problem solution. Classification problem consists of 3 sub-tasks: recognition of "basic" slope, classification by weight, classification by style.

To determine the font's weight it is possible to apply a simple approach: one can calculate the number of black pixels in the image and assign font by this sum to the closest class (by the average value).

However, during the fonts analysis the fonts with weight that depends on image edges rather than number of black pixels were found. Those are outlined fonts with edges being hollow inside. Given that to solve this issue Edge Detection methods could be used, the attempt was made to apply them. The result of symbol image processing using convolutional matrix is shown in Fig. 1.
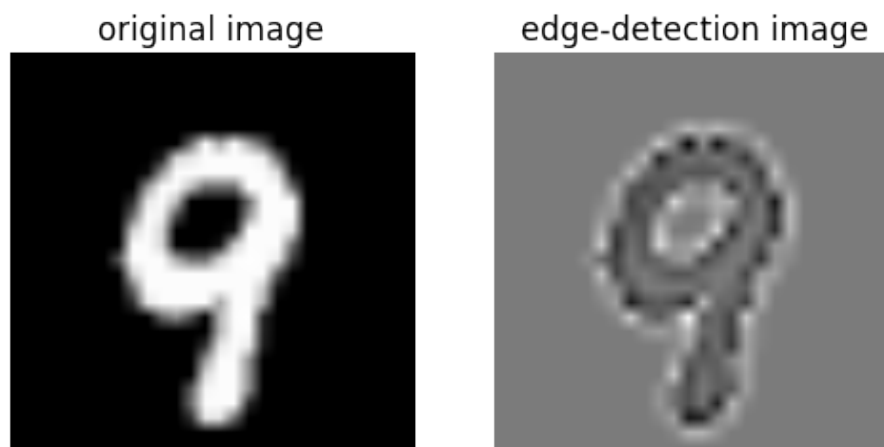

Fig. 1 – The result of symbol image processing using Edge Detection convolutional matrix

The study shows that used methods are not sufficient since they don't consider the symbol's width. Also, some fonts have combined edges that are difficult to interpret, therefore these fonts can be classified only based on patterns of human perception of symbols.

Serif and Sans-Serif fonts in most cases have a sharp slope line, however the study of Script and Decorative fonts shows collisions – one symbol can consist of straight and slope lines at the same time, making the classification more complicated even for a human.

The problem of classification by style is the most difficult of the sub-tasks listed above because of the vast number of various fonts that cannot be easily described by basic distinct features which allow to associate the font with one certain style (fig. 2).

Fig. 2 – Example of variability of "a" symbol

Considering the complexity of the problems and inability to describe the features of various font classes without assistance, it was decided to apply methods of machine learning, namely neural networks. Neural networks are the most frequently used algorithms for solving the image classification problems, as they demonstrate the best results of patterns recognition and classification.

After analyzing the structure features of different font styles, specifically how much their design differs, it was decided to use Convolutional Neural Network (CNN). This type of network allows to recognize fonts features regardless of their position and deformation. Network itself selects patterns that are important for classification (filters, feature maps), considering teacher's correct responses [2].

Because of wide symbol variation by size and design, the use of typical Multi Layer Perceptron (MLP) network would lead to difficulty in high accuracy classification, as it would require construction of a big number of neurons, and therefore would require power efficiency.

To solve the slope classification problem, i.e. if font belongs to Italic type, firstly it was decided to apply MLP network, that properly coped with the task, however during further data increasing by more complex fonts of Script style, network started to reduce its accuracy. Therefore, it was decided to use CNN network as well (fig. 3).
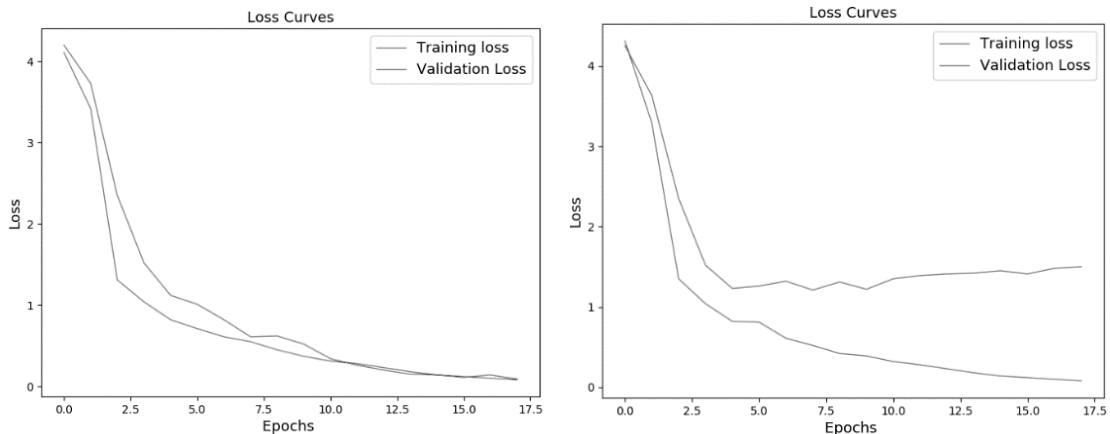

Fig. 3 – Value of Loss function during training of MLP network before and after addition of Script type fonts

Network build and training were performed using Python library called Keras. Inspired by "VGG NET" architecture, one of the most accurate networks at this time [4], it was decided to take an approach of sequential application

of layer combinations: Conv > Conv > MaxPooling. This approach doesn't require lots of filters and layers to solve the problems of given complexity, so it is well-suited for execution on the client's side. Number and size of kernels were selected on an individual basis for every task with priority to processing speed with condition that final accuracy of classification should be at least 90%. During the experiment it was established that for slope classification for Conv layers 8 kernels with a size 5x5 work best, and for classification by weight and style – 16 kernels with a size 3x3. Popular technique Dropout was used to avoid fast retraining of neural network, with network's element inactivation based on certain probability.

ReLU activation function was selected for all layers except for the last one. It can be described by concise formula: $f(x) = max(0, x)$. There is a research that indicates that this function can increase the convergence speed of stochastic gradient descent by factor 6 compared to sigmoid function and hyperbolic tangent. It is thought that this feature based on linear nature and absence of the function saturation [3].

For the last layer softmax function is used, therefore confidence of belonging to one of the classes depends on confidence in other classes. It can be defined by the formula:

$$y_i = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}} ,$$ (1)

where z_i – input value of i-th neuron
y_i – output value of i-th neuron

As error determination function was selected categorical_crossentropy, that is used for Multi Class classification, where the only one class can belong to the font. This function's expression is:

$$H(p,q) = -\sum_x p(x) log q(x)$$ (2)

where p and q – values of expected and received result.

Adadelta (or adaptive learning rate method) algorithm was used as an algorithm of weight update. The method considers gradient values history and weight update history. It presents way over fast convergence than regular SGD during the training on simplified data sample. The example of neural network model is given in Fig. 4.
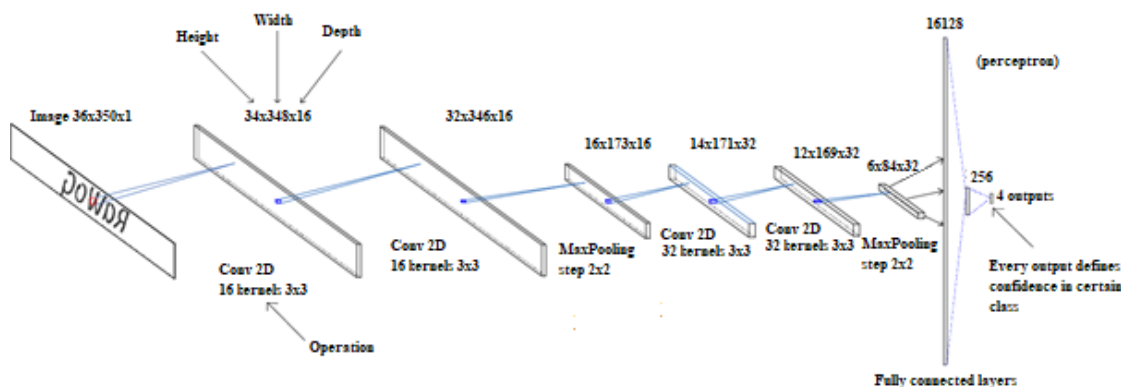

Fig. 4 – Neural network model of style classification

For font classification by slope "IW" symbols were used, and "RoWGa" were used for classification by weight and style. Corresponding image sizes are 90x36 and 350x36. These symbols were selected analytically, as they appropriately depict the most typical features of corresponding classes.

The number of fonts in one class can be substantially lower comparing to other classes. Network trains badly and outputs incorrect metrics due to data imbalance, and small data set leads to early retraining. To solve this issue the application that performs operations over images with their content being intact was developed.

Operations over images

Operation	Allowability of applying
Rotate	Allowable for all sub-tasks, except for slope determination
Scale	Allowable for all sub-tasks, except for weight determination
Horizontal flip	Allowable for all sub-tasks
Vertical flip	Allowable for all sub-tasks
Vertical and Horizontal flip at the same time	Allowable for all sub-tasks
Shift, Wrap	Non-allowable for all sub-tasks

The basic operations over images were studied and their capabilities for the data augmentation were determined.

Operations like Rotate and Scale are allowable for certain data samples, however their usage assumes enlarge-ment of image palette to prevent the loss of fonts features. Enlargement of image palette will cause the increasing of required power supply, that is not appropriate for this type of tasks when computer resources are quite limited.

Number of fonts that were collected and classified is as follows: 826 fonts were classified by style, 664 – by weight, 1128 – by slope. The result of image generation and apply-ing operations for image number augmentation and data balancing is: 12754 images were classified by style (using 3 types of mirroring and letter reordering); 12074 images were classified by weight (also with use of 3 types of mir-roring and letter reordering); 4628 images were classified by slope (using letter reordering and size modification).

The example of neural network training is given in Fig. 5. Figure shows that network has correct training, with validation loss value is close to training loss value (i.e. there is no retraining effect). Also, from plot the stochastic gradi-ent descent appliance can be seen, with validation loss curve that seems to roll down, hence it may be concluded that configuration is advantageous. Values of Loss and Accuracy functions for every data samples are given in table 2.
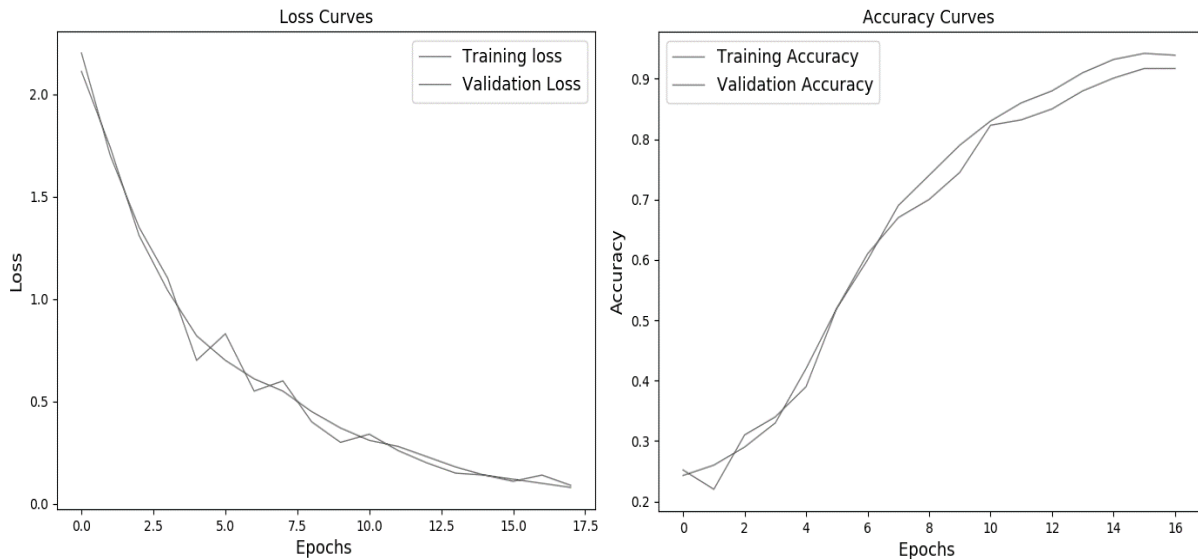

Fig. 5 – Plot of Loss and Accuracy functions during the training of classification by style

Visualization method called Gradient-weighted Class Activation Mapping was used to analyze and understand on what basis neural networks establish the association of the font with one or the other class [6]. It allows to represent the areas that attract neural network attention and the attention intensity with the heat map.

Amount of data in samples and the result of network training

Training type	Training images	Test images	loss_train	acc_train	loss_test	acc_test
By style	12274	480	0.09	0.924	0.12	0.909
By weight	11594	480	0.1	0.937	0.13	0.91
By slope	4168	460	0.08	0.946	0.1	0.938

Fig. 6 – Heat map of class activation that is responsible for Decorative style recognition

Fig. 7 – Heat map of class activation that is responsible for Serif style recognition

Fig. 6 shows the neural network focuses attention on unusual FasterOne font lines, that indicate the font belongs to Decorative style. Activation in serif location areas that is typical for Serif class is shown in fig. 7.

This method has allowed to examine activation for each class. Examples of arguable points, where neural network wasn't confident if the font belongs to one or the other style, are presented below.

Analyzing of IndieFlower Script font from Guided Backpropagation

Style	Confidence	Neural network activation
Script	0.54	
Decorative	0.46	
Serif or Sans-serif	0.0	

Table 4

Analyzing WireOne Sans-serif without Guided Backpropagation

Style	Confidence	Neural network activation
Serif	0.02	
Sans-serif	0.44	
Decorative	0.22	
Script	0.33	

Table 5

Analyzing Overpass Thin Italic font without Guided Backpropagation

Style	Confidence	Neural network activation
Italic	0.85	
Non-italic	0.15	

To check the accuracy of classification the Accuracy-Checker module that assesses the percentage of correctly classified fonts was developed. 100 fonts were collected for each classification by style, weight and slope. Module classifies the font using neural network and then compares its classification with the correct answer, then counts and displays the percentage of the right answers. The result of module execution is given in table 3.5. Comparison of network training metric with conducted accuracy check is shown in Fig. 8.

Table 3.5

The result of classification accuracy check

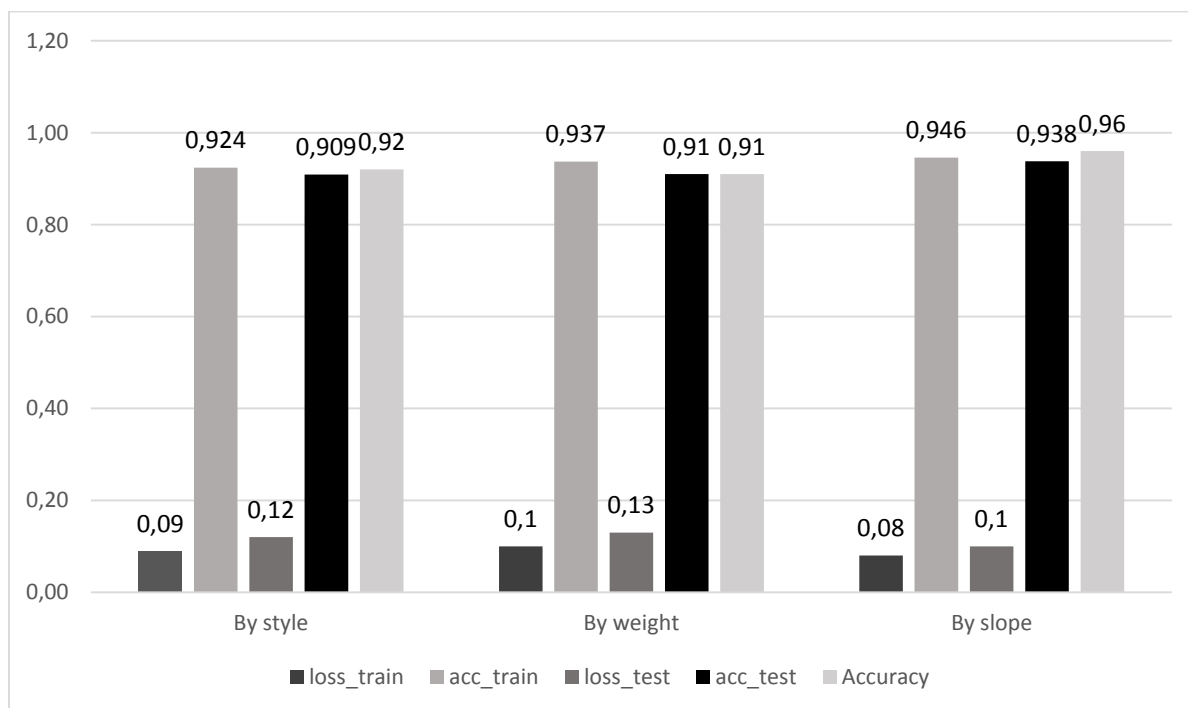| Classification	Test fonts	Classes	Accuracy
By slope	100	2	96%
By weight	100	9	91%
By style	100	4	92%

Fig. 8 – Comparison of networks training metric with their classification accuracy

Main results and conclusions. The main problem of fonts classification consists in fuzziness of features that allow to classify the font, and their strong variation. In this paper the basic approaches of fonts classification with relevant advantages and disadvantages were described. After analyzing the approaches, it was decided to apply neural network that takes the images with font symbols as input. Neural network selects patterns (filters, features maps), to be considered in classification process based on the teacher's correct answers. Possible use of alternative fonts classification techniques was suggested, and the main related problems were described.

Computer fonts classification system driven by convolutional neural networks, that allows to classify fonts by slope, basic style and weight according to their graphical representation, was developed. Percentage of fonts that were correctly classified by the system is: in determining slope – 96%, in determining basic style – 92%, in determining weight – 91%.The developed system can be applied to solving fonts classification tasks as an assistance tool for determination of digital documents structure, and as font selection system for design creation.

Based on results, it may be concluded that neural networks were successfully applied to the fonts classification problem solving.

REFERENCES:

1. Matusugu, Masakazu; Katsuhiko Mori; Yusuke Mitari; Yuji Kaneda. (2013) Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Networks*, 16 (5), 555–559.
2. Cameron Chapman. Understanding the nuances of typeface classification. Retrieved from https://www.toptal.com/designers/typography/typeface-classification.
3. Alex Krizhevsky. ImageNet Classification with Deep Convolutional Neural Networks. Retrieved from http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf
4. Tao, Dapeng & Lin, Xu & Jin, Lianwen & Li, Xuelong. (2015). Principal Component 2-D Long Short-Term Memory for Font Recognition on Single Chinese Characters. *IEEE transactions on cybernetics*, 46. 10, 1109.

5. Zhu, Yong & Tan, Tieniu & Wang, Yunhong. (2001). Font Recognition Based on Global Texture Analysis. Pattern Analysis and Machine Intelligence, *IEEE Transactions on,* 23, 1192-1200.
6. Xavier Glorot, Yoshua Bengio. (2010) Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, *PMLR*, 9, 249-256.
7. Devin Soni. Dealing with Imbalanced Classes in Machine. Retrieved from https://towardsdatascience.com/dealing-with-imbalanced-classes-in-machine-learning-d43d6fa19d2.
8. Königsberg Z.R. (2011) A mixed lyapunov-max-plus algebra approach to the stability problem for discrete event dynamical systems modeled with timed petri nets. *Neural Parallel & Scientific Computations*, 19. 1-2, 35-50.
9. Aumasson, J.-P., Çalık, Ç., Meier, W., Özen, O., Phan, R. C.-W., Varıcı, K. (2009). Improved Cryptanalysis of Skein. *Tokyo, Japan: International Conference on the Theory and Application of Cryptology and Information Security ASIACRYPT,* 542-559.
10. Sebastian Ruder. An overview of gradient descent optimization algorithms. Retrieved from http://ruder.io/optimizing-gradient-descent/index.html.

МЕРКУЛОВА Катерина Володимирівна

кандидат технічних наук, завідувач кафедри комп'ютерних технологій, Донецький національний університет імені Василя Стуса,
м. Вінниця, E-mail: k.merkulova@donnu.edu.ua, ORCID ID: 0000-0001-6347-5191.

ЗАЙЦЕВА Еліна Євгенівна

кандидат технічних наук, доцент кафедри комп'ютерних технологій, Донецькій національний університет імені Василя Стуса,
м. Вінниця, E-mail: zaytseva.elina@gmail.com, ORCID ID 0000-0002-2135-8146.

КОВАЛЬЧУК Артем Юрійович

студент СО «Магістр» спеціальності 122 Комп'ютерні науки, Донецький національний університет імені Василя Стуса,
м. Вінниця, E-mail: hardkovalchuk@gmail.com

СИСТЕМА КЛАСИФІКАЦІЇ ШРИФТІВ ЗА ДОПОМОГОЮ НЕЙРОННИХ МЕРЕЖ

Анотація. Актуальність та мета статті. Метою даної статті є дослідження та розробка системи візуальної класифікації шрифтів, що дасть змогу користувачам ефективніше працювати з шрифтами, а саме: обирати необхідні шрифти використовуючи попередній перегляд, застосовувати різноманітні фільтри та засоби організації шрифтів. Завдяки системі класифікації користувачі зможуть шукати шрифти за нахилом, базовим стилем і вагою, відповідно до їх дійсного відображення, цим самим уникаючи проблему некоректно вказаних стилів.

Методи дослідження. В даній статті розглянуто основні підходи класифікації шрифтів, їх переваги та недоліки. Зваживши розглянуті підходи, було вирішено застосувати нейронну мережу, на вхід якої надходять зображення з символами шрифтів. Нейронна мережа підбирає патерни (фільтри, карти ознак), на які слід звертати увагу при класифікації, зважаючи на правильні відповіді вчителя. Запропоновано можливе використання альтернативних методів класифікації шрифтів з описом їх проблемних місць.

Результати. Створено систему класифікації комп'ютерних шрифтів на основі згорткових нейронних мереж, що дозволяє класифікувати шрифти за нахилом, базовим стилем і вагою, відповідно до їх графічного відображення. Відсоток коректно класифікованих шрифтів системи: в визначенні нахилу – 96%; в визначенні базового стилю – 92%, в визначенні ваги – 91%. Розроблена система може застосовуватись для вирішення задач класифікації шрифтів як допоміжний інструмент визначення структури оцифрованих документів, а також у якості системи підбору шрифтів для створення дизайну. Виходячи з результату, можна судити про успішне застосування нейронних мереж для вирішення задач класифікації шрифтів.

Наукова новизна та практична значимість. В роботі були досліженні основні підходи до класифікації шрифтів, проаналізовані їх недоліки та переваги. Була перевірена ефективність застосування багатошарових перцептронів та згорткових нейронних мереж. Експериментальним шляхом виявлені найбільш оптимальні параметри моделей нейронних мереж. Була розроблена та інтегрована в органайзер система візуальної класифікації шрифтів, що

працює на згорткових нейронних мережах. Проведені контрольні тести, що підвердили високу точність класифікації розроблених моделей нейронних мереж.

Ключові слова: машинне навчання, згорткова нейронна мережа, класифікація шрифтів, класифікація зображень.

МЕРКУЛОВА Катерина Владимировна

кандидат технических наук, заведующая кафедрой компьютерных технологий,
Донецкий национальный университет имени Василия Стуса, Винница, E-mail: k.merkulova@donnu.edu.ua, ORCID ID: 0000-0001-6347-5191.

ЗАЙЦЕВА Элина Евгеньевна

кандидат технических наук, доцент кафедры компьютерных технологий, Донецкий национальный университет имени Василия Стуса,
Винница, E-mail: zaytseva.elina@gmail.com, ORCID ID 0000-0002-2135-8146.

КОВАЛЬЧУК Артем Юрійович

студент СО «Магистр» специальности 122 Компьютерные науки, Донецкий национальный университет имени Василия Стуса,
Винница, E-mail: hardkovalchuk@gmail.com

СИСТЕМА КЛАССИФИКАЦИИ ШРИФТОВ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

Аннотация. Актуальность и цель статьи. Целью данной статьи является исследование и разработка системы визуальной классификации шрифтов, что позволит пользователям эффективно работать со шрифтами, а именно: выбирать необходимые шрифты, используя предварительный просмотр, применять различные фильтры и средства организации шрифтов. Благодаря системе классификации пользователи смогут искать шрифты по наклону, базовым стилем и весом, в соответствии с их действительным отображением, тем самым избегая проблему некорректно указанных стилей.

Методы исследования. В данной статье рассмотрены основные подходы классификации шрифтов, их преимущества и недостатки. Взвесив рассмотреные подходы, было решено применить нейронную сеть, на вход которой поступают изображения с символами шрифтов. Нейронная сеть подбирает паттерны (фильтры, карты признаков), на которые следует обращать внимание при классификации, учитывая правильные ответы учителя. Предложено возможно использование альтернативных методов классификации шрифтов с описанием их проблемных мест.

Результаты. Созданная система классификации компьютерных шрифтов на основе сверточных нейронных сетей, позволяет классифицировать шрифты по наклону, базовым стилям и весам, в соответствии с их графическим отображением. Процент корректно классифицированных шрифтов системы: в определении наклона - 96%; в определении базового стиля – 92%, в определении веса – 91%. Разработанная система может применяться для решения задач классификации шрифтов как вспомогательный инструмент определения структуры оцифрованных документов, а также в качестве системы подбора шрифтов для создания дизайна. Исходя из результата, можно судить об успешном применении нейронных сетей для решения задач классификации шрифтов.

Научная новизна и практическая значимость. В работе были исследованы основные подходы к классификации шрифтов, проанализированы их недостатки и преимущества. Была проверена эффективность применения многослойных перцептронов и сверточных нейронных сетей. Экспериментальным путем выявлены наиболее оптимальные параметры моделей нейронных сетей. Была разработана и интегрирована в органайзер система визуальной классификации шрифтов, работающий на нейронных сетях. Проведенные контрольные тесты подтвердили высокую точность классификации разработанных моделей нейронных сетей.

Ключевые слова: машинное обучение, сверточная нейронная сеть, классификация шрифтов, классификация изображений.

ЛІТЕРАТУРА:

1. Matusugu, Masakazu; Katsuhiko Mori; Yusuke Mitari; Yuji Kaneda. Subject independent facial expression recognition with robust face detection using a convolutional neural network. Neural Networks, 2013, №16 (5), p. 555–559.

2. Cameron Chapman. Understanding the nuances of typeface classification. URL: https://www.toptal.com/designers/typography/typeface-classification.

3. Alex Krizhevsky. ImageNet Classification with Deep Convolutional Neural Networks. URL:http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf .

4. Tao, Dapeng & Lin, Xu & Jin, Lianwen & Li, Xuelong. Principal Component 2-D Long Short-Term Memory for Font Recognition on Single Chinese Characters. IEEE transactions on cybernetics, 2015, Vol 46. 10, p.110-118.

5. Zhu, Yong & Tan, Tieniu & Wang, Yunhong. Font Recognition Based on Global Texture Analysis. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2010, Vol. 23, P. 1192-1200.

6. Xavier Glorot, Yoshua Bengio. Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, PMLR, 2010, p. 249-256.

7. Devin Soni. Dealing with Imbalanced Classes in Machine. URL: https://towardsdatascience.com/dealing-with-imbalanced-classes-in-machine-learning-d43d6fa19d2.

8. Königsberg Z.R. A mixed lyapunov-max-plus algebra approach to the stability problem for discrete event dynamical systems modeled with timed petri nets / Z.R. Königsberg // Neural Parallel & Scientific Computations, 2011, vol. 19, no. 1-2., p. 35-50.

9. Aumasson, J.-P., Çalık, Ç., Meier, W., Özen, O., Phan, R. C.-W., Varıcı, K. Improved Cryptanalysis of Skein. Tokyo, Japan: International Conference on the Theory and Application of Cryptology and Information Security ASIACRYPT ,2009, p. 542-559.

10. Sebastian Ruder. An overview of gradient descent optimization algorithms. URL http://ruder.io/optimizing-gradient-descent/index.html.