DOI: 10.15276/aait.01.2021.6
UDC 004.93.1

Borys I. Tymchenko
ORCID: http://orcid.org/0000-0002-2678-7556; tymchenko.b.i@opu.ua
Odessa National Polytechnic University, 1, Shevchenko Ave., Odesa, Ukraine, 65044

NEURAL NETWORK METHODS FOR PLANAR IMAGE ANALYSIS IN AUTOMATED SCREENING SYSTEMS

ABSTRACT

Nowadays, means of preventive management in various spheres of human life are actively developing. The task of automated screening is to detect hidden problems at an early stage without human intervention, while the cost of responding to them is low. Visual inspection is often used to perform a screening task. Deep artificial neural networks are especially popular in image processing. One of the main problems when working with them is the need for a large amount of well-labeled data for training. In automated screening systems, available neural network approaches have limitations on the reliability of predictions due to the lack of accurately marked training data, as obtaining quality markup from professionals is very expensive, and sometimes not possible in principle. Therefore, there is a contradiction between increasing the requirements for the precision of predictions of neural network models without increasing the time spent on the one hand, and the need to reduce the cost of obtaining the markup of educational data. In this paper, we propose the parametric model of the segmentation dataset, which can be used to generate training data for model selection and benchmarking; and the multi-task learning method for training and inference of deep neural networks for semantic segmentation. Based on the proposed method, we develop a semi-supervised approach for segmentation of salient regions for classification task. The main advantage of the proposed method is that it uses semantically-similar general tasks, that have better labeling than original one, what allows users to reduce the cost of the labeling process. We propose to use classification task as a more general to the problem of semantic segmentation. As semantic segmentation aims to classify each pixel in the input image, classification aims to assign a class to all of the pixels in the input image. We evaluate our methods using the proposed dataset model, observing the Dice score improvement by seventeen percent. Additionally, we evaluate the robustness of the proposed method to different amount of the noise in labels and observe consistent improvement over baseline version.

Keywords: Image analysis; automated screening; multi-task machine learning; loss functions; label noise; noise models

***For citation*: Tymchenko B. I. Neural Network Methods for Planar Image Analysis in Automated Screening Systems. *Applied Aspects of Information Technology*. 2021; Vol.4 No.1: 71–79. DOI: 10.15276/aait.01.2021.6**

INTRODUCTION

Nowadays, means of preventive management in various spheres of human life are actively developing. Modern medicine and production are actively developing and becoming more complex. Modern climatology can predict floods and hurricanes. Early diagnosis of diseases, early detection of defects in products, and timely prediction of cloud movement can reduce costs, conserve natural resources and respond in a timely manner to changes in the environment.

The task of *automated screening* is to detect hidden problems at an early stage without human intervention, while the cost of responding to them is low [1]. Among the undesirable effects of screening is the possibility of misdiagnoses, creating a false sense of confidence in the absence of the problem. For these reasons, screening studies should have sufficient sensitivity and an acceptable level of specificity.

Software and hardware play one of the main roles in the task of automated screening. Thus, with the help of cameras, steel sheets are continuously examined for defects [2]. The use of computerized ophthalmoscopes can increase the volume of testing

for eye diseases in developing countries [3]. Special satellites transmit images of the earth's surface in real time, which helps scientists predict the weather [4].

Planar images are the images of objects that have a constant scale, and for which the depth and perspective distortion can be neglected [5].

Semantic segmentation is the task of assigning an object class to each pixel for the input image [6].

Thus, classification can indicate the presence of a problem, and semantic segmentation – its localization, which simplifies further human work. Classification and semantic segmentation of images are methods of image analysis, as well as an important component of automated screening systems and allows selecting objects of a certain type in the image. Deep artificial neural networks (ANNs) are especially popular in image processing tasks. One of the main problems when working with ANN is the need for a large amount of well-marked data for training [7].

LITERATURE OVERVIEW

Recent research in the field of automatic screening is connected with state-of-the-art deep learning approaches in image analysis, there are much fewer works with classical machine learning and handcrafted features. Classical approaches are the most used in the medical field, for example, Mustafa et al. [8] created an approach with manually extracted features (GrabCut for melanoma segmentation) and trained SVM with a radial basis kernel to discriminate cancerous lesions. Also, Nasiri et al. [9] tried to augment images with different algorithms and trained k-nearest neighbor models to solve the task.

The use of multi-task learning in classification and segmentation problems is not common in the literature. Typically, in works that offer a solution to the segmentation problem, the classification problem is solved at the stage of image pre-processing to weed out false-positive segmentation results and reduce system operating time as a whole through the use of light classification models [10-11].

The disadvantage of this approach is the possibility of spreading the error from the classifier, which in the screening tasks requires increasing the recall of the classifier, with the subsequent operation of the segmentation network.

Additionally, in tasks that have small objects in images or unbalanced datasets, this approach can reduce the accuracy of the system compared to its parts [12]

In contrast to sequential processing, the paper SegTHOR [13] simultaneously solves the problems of classification and segmentation, complementing each other. Thus, the authors propose a new neural network architecture based on the Unet [14] architecture, but has an additional classification output from the last layer of the segmentation network. Obtaining global classes from the segmentation output is achieved through the arithmetic mean of all pixels for each of the classes. The architecture of the SegTHOR neural network is shown in Fig. 1.

In the SegTHOR neural network, the classification output is used only during training as an additional task, and is not used when predicting segmentation masks.

In [15], the authors proposed to apply multi-task approach learning to the classification and segmentation of skin tumors. Unlike conventional architectures for image segmentation, the paper proposes segmentation of any lesion, with a parallel classification of melanoma and seborrheic keratosis as separate tasks. Authors show that such a neural network architecture slightly improves reliability in both classification and segmentation problems. In this architecture, the results of different tasks are used separately during forecasting.

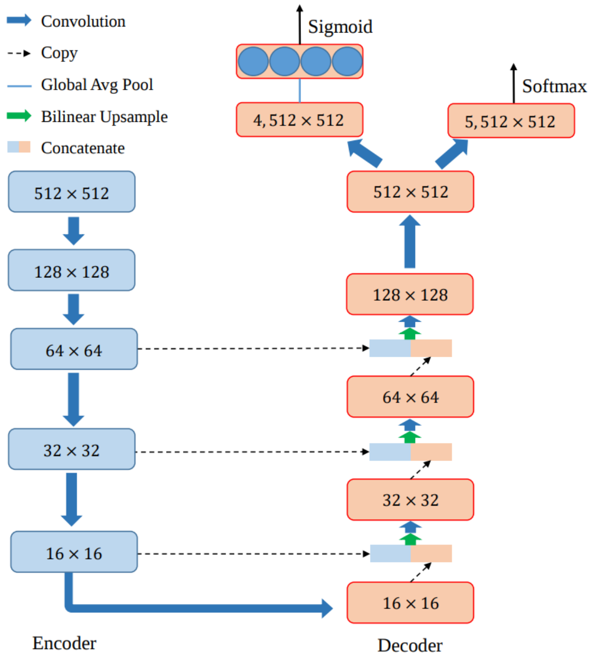

Fig. 1. **SegTHOR architecture [13]**
***Source:* [13]**

The authors of [16] proposed to simultaneously learn the problems of classification and segmentation for the localization of cancerous tumors in mammography. Like [15], this paper uses a neural network consisting of one encoder and two decoders – for classification and segmentation, respectively. The neural network architecture is shown in Fig. 2.

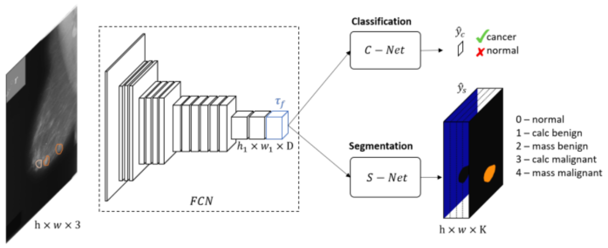

Fig. 2. **Multi-task network architecture from [16]**
***Source:* [16]**

PROBLEM STATEMENT

Our previous research [17-18] showed that the complexity of developing automated screening systems arises from the need to solve the following tasks:

1. Inconsistency and imbalance of training datasets for automated screening tasks occurs due to the relatively small number of anomalous examples in populations and the significant level of labeling errors. It creates a problem in teaching of deep neural networks by standard methods [19].

ISSN 2617-4316 (Print)
ISSN 2663-7723 (Online)

Therefore, when building models of classification and segmentation, it is necessary to solve the problem of the partially erroneous markup.

2. For automatic screening, it is important to reduce the number of false-positive diagnostic results, due to the relatively low prevalence of abnormal cases in populations and further actions associated with false-positive results [20]. Therefore, when building classification and segmentation systems, it is necessary to focus on reducing false-positive results.

3. Previous research has shown that automated screening systems based on multitasking in-depth learning methods can increase the reliability of classification and segmentation through joint learning on these tasks. However, previous studies [15-16] do not take into account the possible combination of the results of different tasks to increase reliability in forecasting.

In automated screening systems, available neural network approaches have limitations on the reliability of predictions due to the lack of accurately marked training data, as obtaining quality markup from professionals is very expensive, and sometimes not possible in principle [21]. Therefore, there is a **contradiction** between increasing the requirements for the precision of predictions of neural network models without increasing the time spent on the one hand, and the need to reduce the cost of obtaining the markup of educational data.

The precision of predictions is defined as a Dice-Sorensen measure [22], or an F1 measure on a validation set with accurate markup. The Dice measure can be used to compare the pixel-wise agreement between a predicted segmentation and its corresponding ground truth.

The formula is given by:

$$\mathrm{D}=\mathrm{F}_1=\frac{2\cdot(X\cap Y)}{(X)+(Y)},$$

where X is the predicted set of pixels and Y is the ground truth set of pixels.

The Dice measure is defined to be 1 when both X and Y are empty. The overall score is the mean of the Dice coefficients for every pair of images and labels in the dataset.

The **aim** of the work is to increase the precision of classification and semantic segmentation of planar images in automated screening systems through the development, improvement and development of methods of analysis of planar images based on artificial neural networks.

The **research object** is the process of intelligent image analysis of planar images.

The **research subject** is methods of semantic segmentation and classification of planar images, structure and models of deep convolution neural networks.

We propose to solve following **research tasks:**

1. Develop a parametric data set model with noisy markup for classification and semantic segmentation problems
2. Develop a neural network method for the analysis of planar images by learning on several semantically similar tasks simultaneously to increase the accuracy of classification and semantic segmentation of planar images
3. Develop a method of combining semantically similar tasks at the forecasting stage, in order to increase the reliability of classification and segmentation of planar images without increasing time.
4. Improve the method of segmentation of important for the classification of image features in the absence of semantic segmentation markup in the training data set

THE MODEL OF LABEL NOISE IN SEMANTIC SEGMENTATION

In order to study the influence of noise level in the data, and the corresponding impact of proposed solutions, we developed a model of noisy data for the segmentation problem, which corresponds to the estimates of noise models in various automated screening problems.

In [18] was found that the main problems with data labeling for screening tasks are.

1. Segmentation masks that capture neighboring pixels;
2. Segmentation masks that do not completely cover the object;
3. No segmentation masks for some objects;
4. Extra masks in places where there are no objects.

To model these labeling defects, we propose a method of controlled generation for both images and segmentation masks, which are then modified with randomly selected flaws. Controlled image generation is performed in two stages: background generation and object placement. To generate the background both regular natural images and synthetic textures, or a constant color fill can be used.

We use Imagenette [23] dataset for background generation and MNIST [24] of FashionMNIST [25] for object placement.

We use these datasets as objects due to three factors:

1. Ability to easily separate the object from the background.

2. Presence of similar elements in different classes (for example, numbers 1 and 7, or classes T-Shirt and Dress).

3. High accuracy of modern neural networks on these datasets, which allows us to focus on the impact of noise in the markup.

We introduce labeling errors artificially by random application of morphological erosion and dilation operations with a square kernel to the masks of segmentation of individual objects before adding them to the overall mask.

In addition to the data sets of the background image χ_b and the objects χ_f, the model has the following parameters:

- Image size S_i in pixels.
- The average size of the object S_o in pixels.
- Size limit of the object σ_o.
- The maximum number of objects in the image N.
- Probability of erosion and dilatation of the mask of each of the objects P_e and P_d.
- Shapes of erosion and dilatation kernels of masks of all objects S_e and S_d.

We generate train and validation datasets based on the model using the following algorithm, all images are generated independently.

1. Select a random background image: $x_b \sim \chi_b$.

2. Select the number of objects in the image: $n_o \sim U(1, N)$.

3. Initialization of the segmentation mask: $M = 0_{S_i,S_i}$.

According to the number of images n_o perform the following steps:

1. Select the image of the object $x_f \sim \chi_f$.

2. Select object dimensions: $s \sim U(S_o - \sigma_o, S_o + \sigma_o)$.

3. Resize the image of the object using bilinear interpolation:

$$\hat{x_f} = R_{bilinear}(x_f).$$

4. Select the coordinates of the object:

$$i_f \sim U(0, S_i - s) j_f \sim U(0, S_j - s).$$

5. Place the image of the object on the background image

$$x_b[i_f \dots i_f + s, j_f \dots j_f + s] =$$

$$= max(x_b[i_f \dots i_f + s, j_f \dots j_f + s], x_f).$$

Images generated with the model are shown in the Fig. 3.

Using the proposed model for generation, we generate the dataset to perform experiments on.

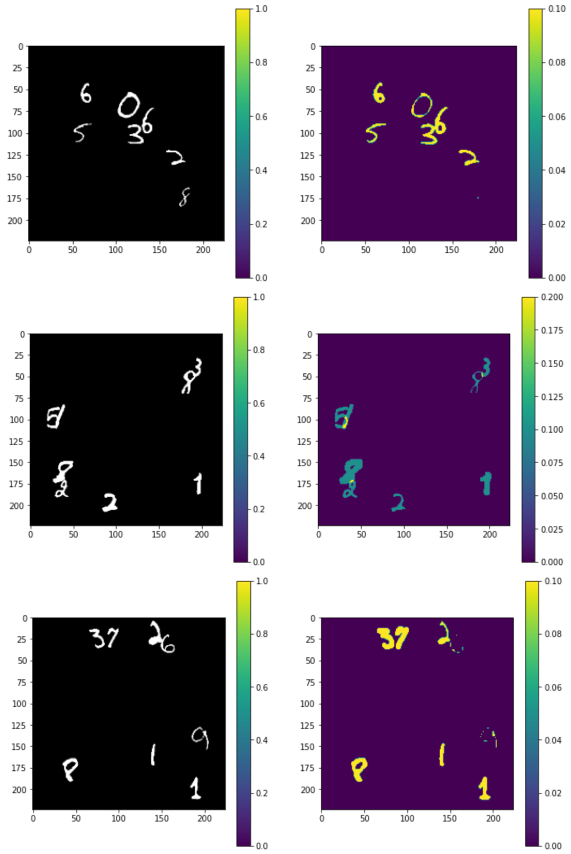

Fig. 3. **Examples of generated images**
***Source*: compiled by the author**

MULTI-TASK TRAINING

To improve neural networks precision on datasets with noisy labels, we propose a method of multi-task learning using more generalized auxiliary tasks derived from the original. The proposed method is based on the assumption that there is a semantically close more general problem, for which labels in training data are more accurate than for the original task.

Semantically-similar tasks are the tasks that operate on the same input space and share similar objectives [26]. In contrast to [26], which are based on the training on more detailed semantically similar tasks, in the proposed method we use more accurate data for more general problems. This allows improving the representation of internal representations of the neural network, which in turn improves initial task. Additionally, because the tasks are semantically similar, there is no gradient conflict, which is typical when learning semantically heterogeneous tasks [27].

We propose to use classification task as a more general to the problem of semantic segmentation. As semantic segmentation aims to classify each pixel in the input image, classification aims to assign a class to all of the pixels in the input image. In this context, classification can be reduced to the task of multiple-instance learning [28]: instead of marking each of the objects for all classes in the image, the image is represented as a bag with one or more objects and appropriate labeling, whether objects of specified classes in the image.

Since inaccurate labeling in segmentation problems is the presence of extra, or the absence of some marked pixels, the presence of such inaccurate markup allows you to create an accurate markup for classification: if the image $x \in \chi$ has at least one marked object of class C, the label of the corresponding class $y_c \in (0,1)$is set for a classification task.

We use shared trunk [29] architecture for multi-task learning. The architecture of a neural network consists of shared encoder, and two decoders for classification and segmentation tasks, respectively.

Given the input image $x \in \chi$, encoder $f_{encoder}$generates the latent feature vector v_f:

$$v_f = f_{encoder}(x).$$

This feature vector is then used by classification and segmentation decoders:

$$M = f_{seg}(v_f), \quad (1)$$
$$C = f_{cls}(v_f), \quad (2)$$

where: $f_{seg}$and $f_{cls}$are segmentation and classification decoders respectively, and M and C are segmentation mask logits and classification result respectively.

For RGB images and K classes, $x \in R^{3\times H\times W}$, $M \in R^{K\times H\times W}andC \in R^{K}$.

The general structure of the neural network is shown in Fig. 4.

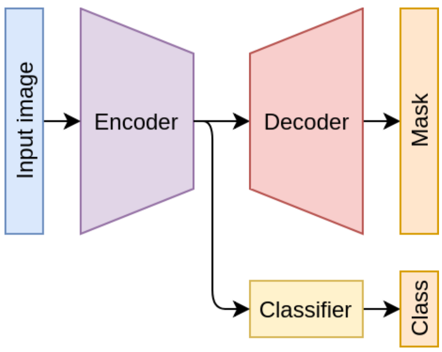

Fig. 4. **General shared trunk MTL architecture**
***Source*: compiled by the author**

Additionally, this architecture allows simle use of transfer learning from pretrained encoder architectures.

Label noise is disruptive for log-based loss functions, as they assign exponentially higher values to erroneous samples, which can be erroneous because of wrong label and correct prediction, which decreases performance of learning.

In order to decrease the influence of such erroneous labels on the learning process, we propose the modification to any loss function, which allows excluding these samples from training.

The modification consists of clipping the loss function with a special *min* function:

$$L] = min(L, \theta).$$

As for the default *min* function the gradient is defined only on the interval, we define it as 0 outside of this interval.

Thus, the gradient becomes:

$$\nabla_{min(L,\theta)} = \begin{cases} 1 & L \in (-\infty, \theta] \\ 0 & L \in (\theta, +\infty) \end{cases}.$$

Using this operation, erroneously-labeled samples are effectively excluded from training. Also, correctly-labeled samples can be occasionally excluded from training, if they are too hard for the neural network to learn in early training stages. In this case, we still have other training objective from the other task, which helps to incorporate knowledge about these samples into the network.

Plots of different clipped functions is shown in the Fig. 5.

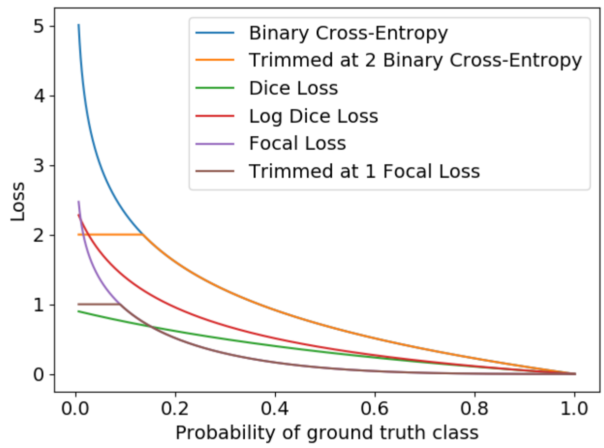

Fig. 5. **Plot of loss functions (clipped and basic)**
***Source*: compiled by the author**

In practice, we found that clipping is necessary only for very confident erroneous predictions, so suitable range of θ is 10 ... 100.

MULTI-TASK INFERENCE

With the use of the proposed multi-task training method, we propose the method for multi-task inference that reduces the number of false-positives in the semantic segmentation task. As the neural network already has classification and segmentation decoders, we propose to merge their predictions in a probabilistic fashion contrary to sequential merging, which uses classifier with a threshold to select, if the image has to be segmented.

Given the logits of classification C and segmentation M (equations 1, 2), we transform them to uncalibrated confidence scores using the elementwise sigmoid activation function:

$$M_p = \sigma(M),$$
$$C_p = \sigma(C).$$

These scores have same dimensions as masks and classes: $M_p \in R^{K\times H\times W}$ and $C_p \in R^{K\times 1\times 1}$.
To retrieve the refined segmentation mask, we re-weight mask channels with classification results by performing element-wise multiplication along the first dimension:

$$M_{ref} = M_p \circ C_p.$$

In order to get refined classification results, we propose to aggregate and normalize M_{ref} by channels:

$$C_{ref} = \frac{\sum M_{ref}}{\sum M_{ref}}.$$

Graphical representation of the network graph is shown in the Fig. 6.

As the classifier branch was trained with more precise data labeling, this refining decreases the influence of the small false-positive regions in the segmentation mask if they have low confidence.

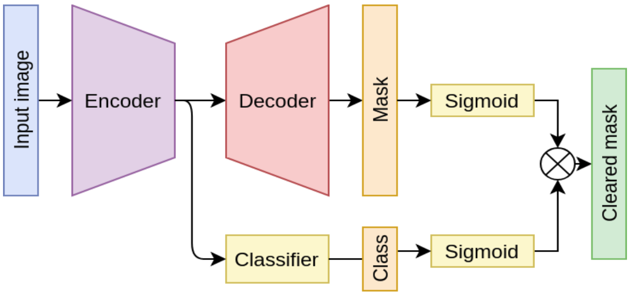

Fig. 6. **Inference neural network graph**
***Source*: compiled by the author**

UNSUPERVISED SALIENT REGIONS SEGMENTATION

Following the approach of multi-task inference, we propose the approach of semi-supervised learning, in which we use the classification task to train both classification and segmentation decoders.

Now, we use segmentation mask to re-weight classification logits along spatial dimensions:

$$M_{unsup} = \sigma(M) \circ C,$$
$$C_{unsup} = \frac{\sum_h^H \sum_w^W M_{unsup(hw)}}{\sum_h^H \sum_w^W \sigma(M_{hw}) + c}.$$

This re-weighting makes M_{unsup} to assign higher values to the important regions of the image in order to propagate them to the classification result. To improve numerical stability during training, we add small constant $c = 10^{-5}$.

Mask output M_{unsup} from CNN is continuous. As we train the mask segmentation branch in an unsupervised fashion, we cannot directly predict the range of the output. To alleviate the calibration of the model predictions, we binarize the mask using the Otsu threshold [30].

After the mask has been binarized, we apply a morphological opening to reduce the number of small false-positive regions.

EXPERIMENTAL RESULTS

We perform experiments using the proposed noisy dataset model with following parameters.

Table 1. **Parameters of the experimental model**

χ_b	Imagenette
χ_f	MNIST
S_i	224 pixels
S_o	32 pixels
σ_o	4 pixels
N	10 objects
P_e	0.5
P_d	0.5
S_e	5 pixels
S_d	5 pixels

***Source*: compiled by the author**

We generate training dataset of 10000 images with noise and testing dataset of 3000 images without label noise. Additionally, we use non-overlapping subsets of Imagenette and MNIST for training and testing datasets.

We use pretrained on Imagenet ResNet18 [31] as the encoder $f_{encoder}$, UNet decoder for segmentation (f_{seg}) and two-layer decoder for classification (f_{cls}). The structure of the decoder is shown in the Fig. 7.

For experiment, we compare our developed method to a baseline, for which a standard single-task learning is used. As a baseline, we use the same architecture, but without classification decoder and classification task.

Training performance is evaluated with Dice coefficient for segmentation. Plot of Dice coefficient on validation dataset is show in the Fig. 8. As seen from plot, our proposed method converges faster and to a better score.

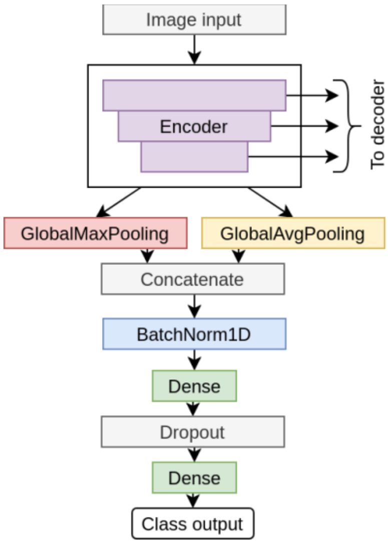

Fig. 7. **Experimental classification decoder**
***Source*: compiled by the author**

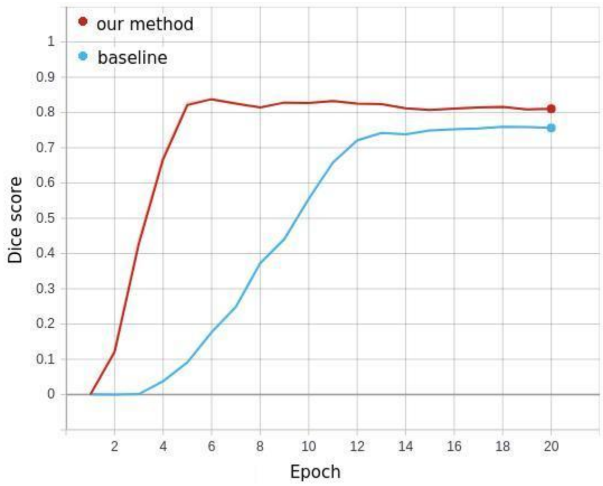

Fig. 8. **Plot of Dice score on the validation dataset**
***Source*: compiled by the author**

In order to check the robustness of our method to label noise, we vary P_e and P_d and re-train the same model. Results are shown in the Table 2.

Table 2. **Robustness to noise**

P_e	P_d	Dice score Baseline model	Dice score Proposed method
0.1	0.1	0.84	0.96
0.1	0.25	0.81	0.91
0.25	0.1	0.65	0.83
0.5	0.25	0.51	0.62
0.25	0.5	0.57	0.69
0.5	0.5	0.3	0.47
0.75	0.75	0.12	0.23

***Source*: compiled by the author**

Robustness test shows, that the proposed method is more robust to noise, than the baseline version. Both baseline and the proposed method are more sensitive to the higher probability of erosion, which corresponds to the absence of labeling for certain pixels.

Here, we demonstrate the results of unsupervised segmentation on the simplified dataset with the constant background (Fig. 9).

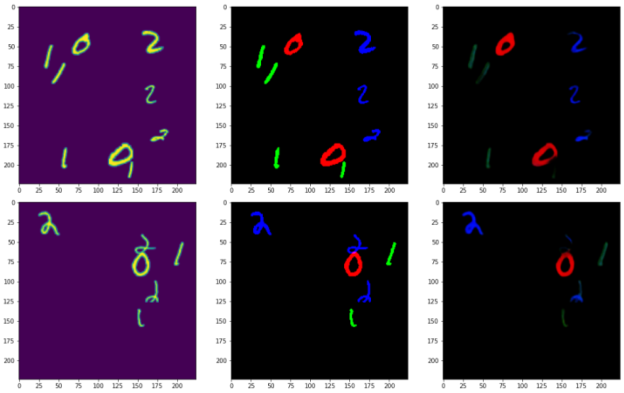

Fig. 9. **Examples of unsupervised segmentation (input, ground truth, output)**
***Source*: compiled by the author**

We used Pytorch [32] along with Catalyst [33] frameworks to implement out dataset model and neural networks.

Presented approach was used in several practical automated screening tasks, what was presented in our previous research. As real world tasks do not have noise-free labeling in their validation datasets, it's hard to get clear causation of improvement brought by our method over more successful fitting of label noise.

Namely, we tested it for melanoma detection [18], diabetic retinopathy detection [3], cloud patterns detection [17] and mixed protein patterns classification [20] with slight variations depending on the task. Averaged, presented method shows improvement of Dice/F1 scores over baseline of 17%.

CONCLUSIONS

In this work, we propose the model of dataset for semantic segmentation with noisy labels, as well as the multi-task learning approach to improve the precision of the semantic segmentation while training on noisy labels.

Presented approach uses semantically-similar tasks to both train deep neural network and infer predictions using several tasks at the same time. Based on this multi-task approach, we propose a semi-supervised approach to segment important regions in the image without labels for semantic segmentation.

We evaluate these approaches using a dataset, that is generated with the use of our model.

Experiments show faster convergence as well as improvement of Dice score by 17 %.

REFERENCES

1. Nørgaard, M. F. & Grauslund, J. "Automated Screening for Diabetic Retinopathy – A Systematic Review". *Ophthalmic Res.* 2018. 60(1): 9–17. DOI: 10.1159/000486284.

2. Jingwen, Fu, Xiaoyan, Zhu & Yingbin, Li. "Recognition Of Surface Defects On Steel Sheet Using Transfer Learning", eprint arXiv:1909.03258. USA. 2019.

3. Tymchenko, B., Marchenko, P. & Spodarets, D. "Deep Learning Approach to Diabetic Retinopathy Detection", eprint arXiv:2003.02261. USA. 2020.

4. Ohring, G., Lord, S. & Derber, J. "Applications of satellite remote sensing in numerical weather and climate prediction". *Advances in Space Research.* 2012; 30(11): 2433-2439. DOI: 10.1016/S0273-1177(02)80298-8.

5. Hartley, R. & Zisserman, A. "Multiple View Geometry in Computer Vision". Australian National University. University of Oxford. Canberra: Australia. 2011. p.112-114. DOI: 10.1017/CBO9780511811685.

6. Liu, X., Deng, Z. & Yang, Y. "Recent progress in semantic image segmentation", eprint arXiv:1809.10198. USA. 2018.

7. "CS231n: Convolutional Neural Networks for Visual Recognition". – Available from: http://cs231n.stanford.edu/. – [Accessed: Jan, 2021].

8. Mustafa, S. & Kimura, A., "A SVM-based diagnosis of melanoma using only useful image features". *2018 International Workshop on Advanced Image Technology (IWAIT).* Chiang Mai: Thailand. 2018. p.1–4. DOI: 10.1109/IWAIT.2018.8369646.

9. Nasiri, S., Jung, M., Helsper, J. & Fathi M. "Detect and Predict Melanoma Utilizing TCBR and Classification of Skin Lesions in a Learning Assistant System". *Bioinformatics and Biomedical Engineering, Lecture Notes in Computer Science. Springer International Publishing.* 2018; Vol. 10813: 531–542.

10. Yousefikamal, P. "Breast Tumor Classification and Segmentation using Convolutional Neural Networks", eprint arXiv:1905.04247. USA. 2019.

11. Alahyari, A., Hinneck, A., Tariverdi, R. & Pozo, D. "Segmentation and Defect Classification of the Power Line Insulators: A Deep Learning-based Approach", eprint arXiv:2009.10163. USA. 2020.

12. Biertimpel, D., Shkodrani, S., Baslamisli, A. & Baka, N. "Prior to Segment: Foreground Cues for Novel Objects in Partially Supervised Instance Segmentation", eprint arXiv:2011.11787. USA. 2020.

13. Tao He, Jixiang Guo, Jianyong Wang, Xiuyuan Xu & Zhang Yi. "Multi-Task Learning For The Segmentation Of Thoracic Organs At Risk In CT Images". *Machine Intelligence Laboratory*. Sichuan University. Sichuan: China. 2019.

14. Ronneberger, O., Fischer, P. & Brox, T. "U-Net: Convolutional Networks for Biomedical Image Segmentation", eprint arXiv:1505.04597. USA. 2015.

15. Yang, X., Zeng, Z., Yeo, S., Tan, C., Tey, H. & Su, Y. "A Novel Multi-task Deep Learning Model for Skin Lesion Segmentation and Classification", eprint arXiv:1703.01025. USA. 2017.

16. Le, T., Thome, N., Sylvain, B. & Bismuth, V. "Multitask Classification and Segmentation for Cancer Diagnosis in Mammography". *Medical Imaging with Deep Learning*, eprint arXiv: 2019,1909.05397

17. Tymchenko, B., Marchenko, P. & Spodarets, D. Segmentation of cloud organization patterns from satellite images using deep neural networks. *Herald of advanced information technology*. *Publ. Science i Technical*. Odesa: Ukraine. 2020; Vol. 3 No. 1: 352–361. DOI: 10.15276/hait 01.2020.2.

18. Borys Tymchenko, Eugene Khvedchenya, Philip Marchenko, Dmitry Spodarets. "Classification of skin lesions using multi-task deep neural networks". *Herald of advanced information technology*. *Publ. Science i Technical*. Odesa: Ukraine. 2020; Vol. 3 No. 3: 136–148. DOI: 10.15276/hait.03.2020.3.

19. Danquah, R. A. "Handling Imbalanced Data: A Case Study for Binary Class Problems", eprint arXiv:2010.04326. USA. 2020.

20. Tymchenko, B., Hramatik, A. & Tulchiy, H."Classifying mixed patterns of proteins in microscopic images with deep neural networks". *Herald of advanced information technology*. *Publ. Science i Technical*. Odesa: Ukraine. 2019; Vol. 2 No. 1: 29–36. DOI://10.15276/hait.02.2019.3.

21. Volkova, N. "Detector Quasi-Periodic Texture Segmentation Method for Dermatological Images Processing". *Herald of Advanced Information Technology. Publ. Science i Technical*. Odesa: Ukraine. 2019; Vol.2 No.4: 268–277. DOI:10.15276/hait.04.2019.3.

22. Bertels, J., Eelbode, T., Berman, M., Vandermeulen, D., Maes, F., Bisschops, R. & Blaschko, M. "Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory & Practice", eprint arXiv:1911.01685. USA. 2019.

23. "fastai/imagenette". – Available from: https://github.com/fastai/imagenette. – [Accessed: Jan, 2021].

24. "MNIST Database". – Available from: http://yann.lecun.com/exdb/mnist/. – [Accessed: Jan, 2021].

25. "zalandoresearch/fashion-mnist". – Available from: https://github.com/zalandoresearch/fashion-mnist. – [Accessed: Jan, 2021].

26. Tian, Y., Luo, P., Wang, X. & Tang, X. “Pedestrian Detection aided by Deep Learning Semantic Tasks”, eprint arXiv:1412.0069. USA. 2014.

27. Yu, T., Kumar, S., Gupta, A., Levine, S., Hausman, K. & Finn, C. “Gradient Surgery for Multi-Task Learning”, eprint arXiv:2001.06782. USA. 2020.

28. Carbonneau, M. A., Cheplygina, V., Granger, E. & Gagnon, G. “Multiple Instance Learning: A Survey of Problem Characteristics and Applications”, eprint arXiv:1612.03365. USA. 2016.

29. Caruana, R. “Multitask Learning”. *Machine Learning* 28. 1997. p. 41–75. DOI: 10.1023/A:1007379606734.

30. Otsu, N. “A Threshold Selection Method from Gray-Level Histograms”. *IEEE Transactions on Systems, Man, and Cybernetics.* Jan. 1979; Vol. 9 Issue 1: 62–66. DOI: 10.1109/TSMC.1979.4310076.

31. He, K., Zhang, X., Ren, S. & Sun, J. “Deep Residual Learning for Image Recognition”, eprint arXiv:1512.03385. USA. 2016.

32. “PyTorch”. – Available from: https://pytorch.org/. – [Accessed: Sep, 2020].

33. “Accelerated DL R&D”. – Available from: https://github.com/catalyst-team/catalyst/. – [Accessed: Sep, 2020].

Conflicts of Interest: The authors declare no conflict of interest

Received 14.01.2021
Received after revision 23.02.2021
Accepted 12.03.2021

DOI: 10.15276/aait.01.2021.6
УДК 004.93.1

НЕЙРОМЕРЕЖЕВІ МЕТОДИ АНАЛІЗУ ПЛАНАРНИХ ЗОБРАЖЕНЬ В СИСТЕМАХ АВТОМАТИЗОВАНОГО СКРИНІНГУ

Борис Ігорович Тимченко
ORCID: http://orcid.org/0000-0002-2678-7556; tymchenko.b.i@opu.ua
Одеський національний політехнічний університет, пр. Шевченка, 1, Одеса, 65044,Україна

АНОТАЦІЯ

У наш час активно розвиваються засоби превентивного управління в різних сферах людського життя. Завдання автоматизованого скринінгу полягає у виявленні прихованих проблем на ранній стадії без втручання людини, тоді як вартість реагування на них низька. Візуальний огляд часто використовується для виконання скринінгу. Глибокі штучні нейронні мережі особливо популярні при обробці зображень. Однією з головних проблем при роботі з ними є потреба у великій кількості добре розмічених даних для навчання. В автоматизованих системах скринінгу доступні нейромережеві підходи мають обмеження по точності прогнозів через відсутність точно розмічених навчальних даних, оскільки отримання якісної розмітки від професіоналів є дорогим, а іноді неможливо в принципі. Отже, виникає протиріччя між підвищенням вимог до точності прогнозів моделей нейронних мереж без збільшення витрачається часу, з одного боку, і необхідністю зниження витрат на отримання розмітки навчальних даних. У цій роботі ми пропонуємо параметричну модель набору даних сегментації, яка може бути використана для формування навчальних даних; а також метод многозадачного машинного навчання для навчання і прогнозування глибоких нейронних мереж в задачі семантичної сегментації. На основі запропонованого методу запропонований підхід полуавтоматіческго навчання завдання сегментації важливих для класифікації регіонів. Основною перевагою запропонованого методу є те, що він використовує семантично схожі загальні завдання, які мають кращу маркування, ніж вихідна, що дозволяє користувачам знизити вартість процесу маркування. В роботі запропоновано використовувати завдання класифікації як більш загальну до завдання семантичної сегментації. Оскільки семантична сегментація спрямована на класифікацію кожного пікселя вхідного зображення, класифікація спрямована на присвоєння класу всім пікселям вхідного зображення. Робота методу оцінена, використовуючи запропоновану модель набору даних, спостерігається поліпшення коефіцієнта Дайса-Соренсена на сімнадцять відсотків. Крім того, оцінена стійкість запропонованого методу до різній кількості шуму в розмітці, що показало поліпшені результати щодо методу однозадачних навчання.

Ключові слова: аналіз зображень; автоматичний скринінг; багатозадачне машинне навчання; функції втрат; зашумлена розмітка; моделі шуму

ABOUT THE AUTHOR

Borys Igorovych Tymchenko, PhD student of the Institute of Computer Systems, Odessa National Polytechnic University, 1, Shevchenko Ave., Odesa, Ukraine
ORCID: http://orcid.org/0000-0002-2678-7556; tymchenko.b.i@opu.ua
Research field**:** Computer Science

Борис Ігорович Тимченко, аспірант інституту Комп’ютерних систем, Одеський національний політехнічний ун-т, пр. Шевченка, 1, Одеса, 65044, Україна