Improving sign language processing via few-shot machine learning

Grigory F. Shovkoplias[1], Dmitriy A. Strokov[2], Daniil V. Kasantsev[3],
Aleksandra S. Vatian[4], Arip A. Asadulaev[5], Ivan V. Tomilov[6],
Anatoly A. Shalyto[7], Natalia F. Gusarova[8]✉

[1,2,3,4,5,6,7,8] ITMO University, Saint Petersburg, 197101, Russian Federation

[1] gfshovkoplias@itmo.ru, https://orcid.org/0000-0001-7777-6972
[2] dmitry.strokov@mail.ru, https://orcid.org/0000-0003-1924-0621
[3] kazantsev@itkvar.ru, https://orcid.org/0000-0001-7974-0922
[4] alexvatyan@gmail.com, https://orcid.org/0000-0002-5483-716X
[5] aripasadulaev@itmo.ru, https://orcid.org/0000-0002-2581-935X
[6] ivan-tomilov3@yandex.ru, https://orcid.org/0000-0003-1886-2867
[7] anatoly.shalyto@gmail.com, https://orcid.org/0000-0002-2723-2077
[8] natfed@list.ru✉, https://orcid.org/0000-0002-1361-6037

Abstract

Improving the efficiency of communication of deaf and hard of hearing people by processing sign language using artificial intelligence is an important task both socially and technologically. One of the ways to solve this problem is a fairly cheap and accessible marker method. The method is based on the registration of electromyographic (EMG) muscle signals using bracelets worn on the arm. To improve the quality of recognition of gestures recorded by the marker method, a modification of the marker method is proposed — duplication of EMG sensors in combination with a low-frame machine learning approach. We experimentally study the possibilities of improving the quality of processing of sign language by duplicating EMG sensors as well as by reducing the volume of the dataset required for training machine learning tools. In the latter case, we compare several technologies of the few-shot approach. Our experiments show that training with few-shot neural nets on 56k samples we can achieve better results than training on random forest with 160k samples. The use of a minimum number of sensors in combination with few-shot signal processing techniques provides the possibility of organizing quick and cost-effective interaction with people with hearing and speech disabilities.

УДК 004.85

Повышение эффективности обработки жестового языка посредством малокадрового машинного обучения

Григорий Филиппович Шовкопляс[1], Дмитрий Андреевич Строков[2],
Даниил Владимирович Казанцев[3], Александра Сергеевна Ватьян[4],
Арип Амирханович Асадулаев[5], Иван Вячеславович Томилов[6],
Анатолий Абрамович Шалыто[7], Наталия Федоровна Гусарова[8]✉

[1,2,3,4,5,6,7,8] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

[1] gfshovkoplias@itmo.ru, https://orcid.org/0000-0001-7777-6972
[2] dmitry.strokov@mail.ru, https://orcid.org/0000-0003-1924-0621

Научно-технический вестник информационных технологий, механики и оптики, 2022, том 22, № 3
Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2022, vol. 22, no 3

559

[3] kazantsev@itkvar.ru, https://orcid.org/0000-0001-7974-0922
[4] alexvatyan@gmail.com, https://orcid.org/0000-0002-5483-716X
[5] aripasadulaev@itmo.ru, https://orcid.org/0000-0002-2581-935X
[6] ivan-tomilov3@yandex.ru, https://orcid.org/0000-0003-1886-2867
[7] anatoly.shalyto@gmail.com, https://orcid.org/0000-0002-2723-2077
[8] natfed@list.ru✉, https://orcid.org/0000-0002-1361-6037

Аннотация

Предмет исследования. Повышение эффективности коммуникации глухих и слабослышащих людей путем обработки жестового языка средствами искусственного интеллекта – важная задача в социальном и в технологическом планах. Одним из направлений решения этой проблемы является применение достаточно дешевого и доступного маркерного метода. **Метод.** Метод основан на регистрации электромиографических сигналов мышц, регистрируемых с помощью браслета, надеваемого на руку. Для повышения качества распознавания жестов применяется модификация маркерного метода, сущность которого состоит в дублировании датчиков электромиографических сигналов в сочетании с малокадровым подходом машинного обучения. **Основные результаты.** Экспериментально изучены возможности повышения качества обработки жестового языка за счет дублирования электромиографических датчиков и уменьшения набора данных, необходимого для машинного обучения. Выполнено сравнение нескольких технологий малокадрового подхода. Показано, что при обучении малокадровых нейронных сетей на наборе данных объемом 56 000 образцов можно достичь лучших результатов, чем при обучении классификатора типа «случайный лес» на наборе данных объемом 160 000 образцов. **Практическая значимость.** Использование минимального количества датчиков электромиографических сигналов в сочетании с малокадровыми методами их обработки обеспечивает возможность организации быстрого и экономичного взаимодействия с людьми с нарушениями слуха и речи.

Ключевые слова

обработка жестового языка, малокадровое машинное обучение, браслеты, маркерные методы

Introduction

According to the World Health Organization, practically or completely deaf and dumb people make up more than 5 % of the world's population. Improving communication efficiency with people with hearing and speech disabilities is one of the most important goals of real-world applications of Intelligent Techniques (IT) and Artificial intelligence (AI). This problem becomes especially important in the context of medical care for patients with hearing and speech disabilities, when medical personnel needs to obtain the most prompt and accurate information about the status of a potentially infected or already ill person.

The main means of communication for the deaf and hearing-impaired people is Sign Language (SL), which, apparently, most medical personnel do not know. Therefore, there is a great need for the development and widespread introduction into medical practice of IT tools for SL processing, that is, computer SL interpretation systems that convert a sequence of gestures (sign speech) into text in a spoken language.

There are two ways to recognize gestures using technical means: markerless and marker [1–3]. Markerless methods use video cameras and infrared transmitters that allow tracking movements, with subsequent processing of the video stream. Marker methods use sensors recording the electromyographic (EMG) signals of muscles which correspond to the performance of certain gestures. These are either special gloves or devices on the wrists, such as bracelets. The marker method is fundamentally simpler than markerless one, and for reasons of user friendliness

and affordability, they have taken a leading position in the market in recent years [4–6]. A typical scenario for using the marker method for medical needs is as follows. The patient (a native speaker of SL) puts a specialized bracelet [7] on his forearm and spells the words in his usual way; each gesture is classified by means of AI. The resulting sequential set of tokens is converted into text using NLP and then displayed on the screen of the doctor's mobile phone or converted into audio speech.

However, in real practice, the execution of this scenario encounters a number of difficulties. First, the EMG signal corresponding to the same gesture depends significantly on the individual characteristics of the person's forearm, such as the presence of fatty tissue, sweat and hair, as well as on the position of the sensor on it [8]. Second, SL is context sensitive, i.e., the result of token recognition and their interpretation will depend on the nationality of the speaker, on the communication situation, etc. As a result, when processing SL by means of AI, there are uncertainties arising both at the stage of signal pickup and at the stage of converting a sequence of tokens into a coherent text. It significantly reduces the efficiency of SL processing or even makes it impossible, which is especially important for organizing full-fledged communication in the provision of emergency medical care.

Our contribution to the solution of this problem is twofold. We experimentally study the possibilities of improving the quality of processing of SL by duplicating EMG sensors, as well as by reducing the volume of the dataset required for training AI tools. In the latter case, we compare several technologies of the few-shot approach [9].

560

Научно-технический вестник информационных технологий, механики и оптики, 2022, том 22, № 3
Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2022, vol. 22, no 3

Background and Related Works

The problem of increasing robustness in the marker-based SL processing is constantly in the field of view of researchers. In recent years, various approaches have been proposed for this purpose, including ensemble neural network learning [10], multiple-agent voting [11], bilinear models [12], spectral decomposition and common modes analysis [13], back propagation neural networks [14], building of fuzzy matrix model [15], and support vector machines [8]. In general, the listed approaches rely heavily on the availability of a large amount of training data.

A number of studies focus on working with small input data. In these cases, it is suggested to use information fusion based on convolutional neural networks [16], transfer learning [17], and data augmentation schemes [18]. But, as mentioned in the review [19], the robustness of marker-based SL processing remains underexplored. In this regard, the few-shot learning approach as a variant of domain adaptation with the goal of inferring the required output based on just one or a few training examples should be highlighted. For example, the work [20] shows that the few-shot learning approach quickly generalizes after seeing very few examples from each class. Namely, for 5 classes sampled from whole set of labels and 5 examples extracted sampled from each of those 5 classes, the classification accuracy of 73–86 % was obtained depending on a specific problem statement.

There are two main approaches to fulfill few-shot learning — optimization-based and metric-based. The first approach [21, 22] uses gradient descent-based methods as the optimization process in the meta-learning, sharing knowledge across all tasks. Metric-based methods [23, 24] classify samples based on the distance between them; they require less computing resources compared to optimization-based. In general, though the few-shot methods are claimed for the classification of datasets containing a small number of examples in each class, their high accuracy is achieved only on large datasets such as American SL [25], and it drops sharply with a decrease in the size of the dataset (74.7 % for Flemish SL corpus [26] and even less, up to 48.6 % [27]).

The influence of the sensors position and configuration on the EMG signal is stated in a number of works [8, 28–30]. For example, placing sensors on the flexor carpi radialis, flexor carpi ulnaris, extensor digitorum communis and extensor carpi ulnaris significantly (several times) changes the amplitude of the EMG signal as well as its shape for the same gesture [30]. According to [28], the sensor position on the flexor carpi ulnaris muscle has a critical contribution to the authentication performance of the EMG signal. Obviously, in the conditions of real medical practice, it is not possible to position the sensors on the patient's arm, taking into account the location of specific muscles. In this regard, it is proposed to use a set of sensors covering the entire arm of the patient to a certain extent and to process the averaged signal. The work [31] even proposes a distributed signal pickup using a sensitive sleeve. Much more technological and economical is the simultaneous use of several widespread sensors such as [7]. However, the study of the influence of the number of sensors on the efficiency of processing SL by the marker method has not been presented in the literature.

Thus, the task of the article is to compare the effectiveness of SL processing methods depending on the amount of data available (few-shot method vs. traditional random forest method) and on the number of simultaneously used EMG sensors.

Materials and Methods

Traditional few-shot learning techniques fulfill classification within a subset of classes named episode (usually of 2–10 classes in each episode). In the case of SL processing, one needs on-line classification within the entire alphabet of gestures. To eliminate this contradiction, we have proposed a modification of the few-shot method presented in Fig. 1.

To make few-shot learning model applicable to the given task, we propose a range of modifications, our contribution is three fold.

1) Standard approach to teaching where we have access to training examples and information about to which class each sample belongs to. We propose a new distance based training pipeline by few-shot architecture with an output dimension equal to the number of classes. During each training step our method learns to output probability for each class in dataset.

2) (top image) Traditional few-shot learning used when there are a large number of classes and a small number of examples within each class. The dataset is divided into episodes, consisting of examples (train) shots and queries (test) shots. Based on examples, the model "remembers" how each class looks like and, when receiving some queries (test) shots, it is able to carry out a classification. The task of the few shot learning model in this setting is to learn to distinguish different classes by building separable representation of each class in episode and compare test examples with built representation. Formally, we have function, f_φ, with learnable parameters φ and some squared Euclidean distance d between outputs of this function:

$$f_\varphi \colon \mathbb{R}^D \to \mathbb{R}^M.$$

$\mathbb{R}^D$ is the D-dimensional feature vector, $\mathbb{R}^M$ is M-dimensional representation vector. Our neural networks produce a distribution over classes, p_φ, for an inputs sample x based on a Softmax function over distances between representations, built for each class:

$$p_\varphi(y = k|\mathbf{x}) = \frac{\exp(-d(f_\varphi(\mathbf{x}), \mathbf{c}_k))}{\sum\limits_{k'} \exp(-d(f_\varphi(\mathbf{x}), \mathbf{c}_{k'}))}$$

$\mathbf{c}_k$ corresponds to a prototype of a class k, similarly.

In our experiment, we have selected computationally inexpensive metric-based models, presented in the Background and Related Works section, being the most common in the field, namely, Matching Networks [32] and Prototypical Networks [24]. We used a typical implementation of the random forest classifier from the scikit-learn1 package, all parameters are set by default.
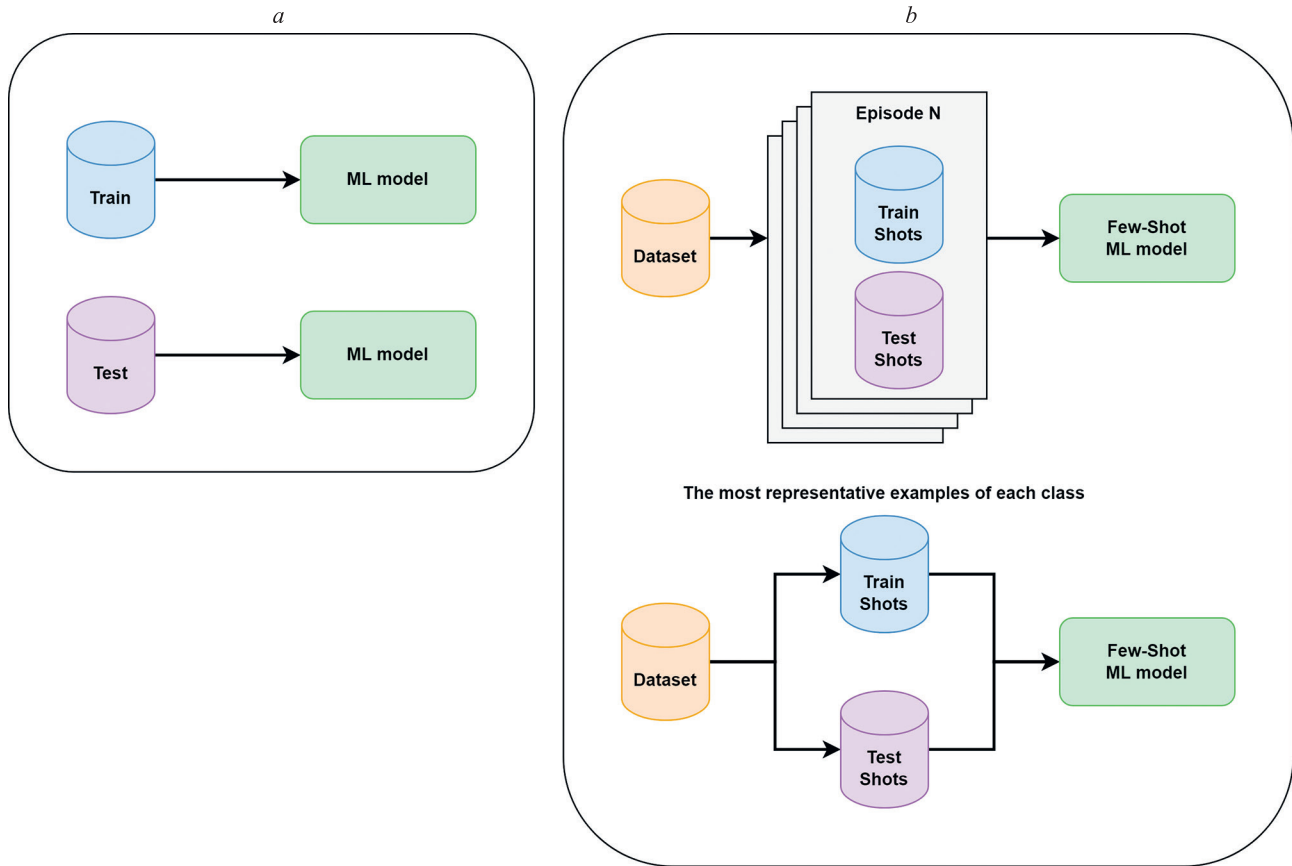
Научно-технический вестник информационных технологий, механики и оптики, 2022, том 22, № 3
Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2022, vol. 22, no 3

561

Fig. 1. Different approaches to model training

For investigations, we used a putEMG dataset [33] containing the EMG signal records of the corresponding hand gesture. The dataset includes records of 8 basic gestures performed by 44 volunteers. The total size of the dataset was 160,000 rows with a fairly even distribution across classes (Fig. 2).

To take the data, 3 Myo Thalmic bracelets [7] were used which were sequentially attached to the arm: the first one was approximately 1/4 of the forearm length from the elbow, each with a shift of approximately 1/5 of the forearm length. For our experiments, we used data from 30 volunteers and formed the following sub-datasets from them: 3 datasets containing data from each bracelet (1, 2,

and 3); 3 datasets containing data for pairs of bracelets (1 and 2, 2 and 3, 1 and 3); 1 dataset containing data for all three bracelets. Each of the constructed sub-datasets was grouped into 7 classes into new datasets, from which 400 rows were taken. As a result, each of the 7 classes contains 30 instances of 400 rows each, of which data for 20 volunteers were used to create a sup-porting set, and the remaining 10 were used to create queries.

Results

The results of the experiment are presented in the Fig. 3. The Fig. 3, *a*, *b* use notation typical for few-shot methods, namely: 'k-way' is the number of classes; 'n-shot' is the number of samples elicited from each class; and Q is the number of examples per class. In all figures, as an *y* axis value, we use a measure of the effectiveness of the classification of gestures CA (categorical efficiency), i.e., the frequency of the correct class definition, indicated, respectively, for each bracelet or their combination (e.g., CA1 stands for categorical efficiency acquired using only the first bracelet, CA12 stands for categorical efficiency acquired using the first and the second one, and so forth). Fig. 3, *c* shows the value of accuracy (*y* axis) obtained by a Random Forest classifier at the optimal number of splits.

Discussion

To compare achieved results, it is better to use Fig. 3, since they offer a visual approach. First of all, we should
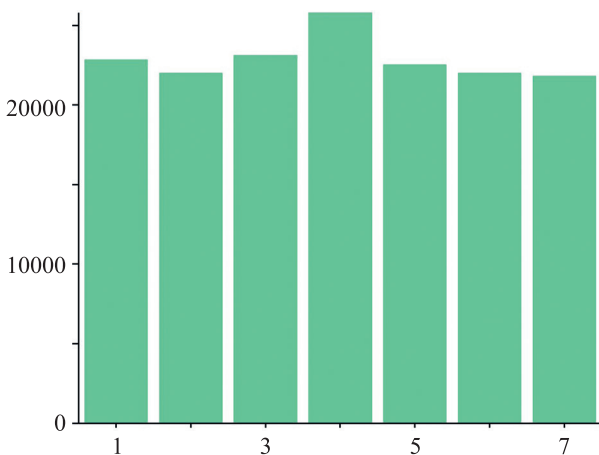
Fig. 2. Distribution of dataset elements by classes

562

Научно-технический вестник информационных технологий, механики и оптики, 2022, том 22, № 3
Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2022, vol. 22, no 3

Fig. 3. ProtoNets results (*a*); Matching Networks results (*b*); Random Forest results (*c*).

k-way is the number of classes; n-shot is the number of samples elicited from each class; Q is the number of examples per class

Научно-технический вестник информационных технологий, механики и оптики, 2022, том 22, № 3
Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2022, vol. 22, no 3

563

notice that using the few-shots methods resulted in a better ability to distinguish samples in comparison to the random forest algorithm, see Fig. 3, *c*. Secondly, we found out that the difference between various combinations of bracelets does not have the significant impact on the result of a classification. We were not able to establish any kind of cor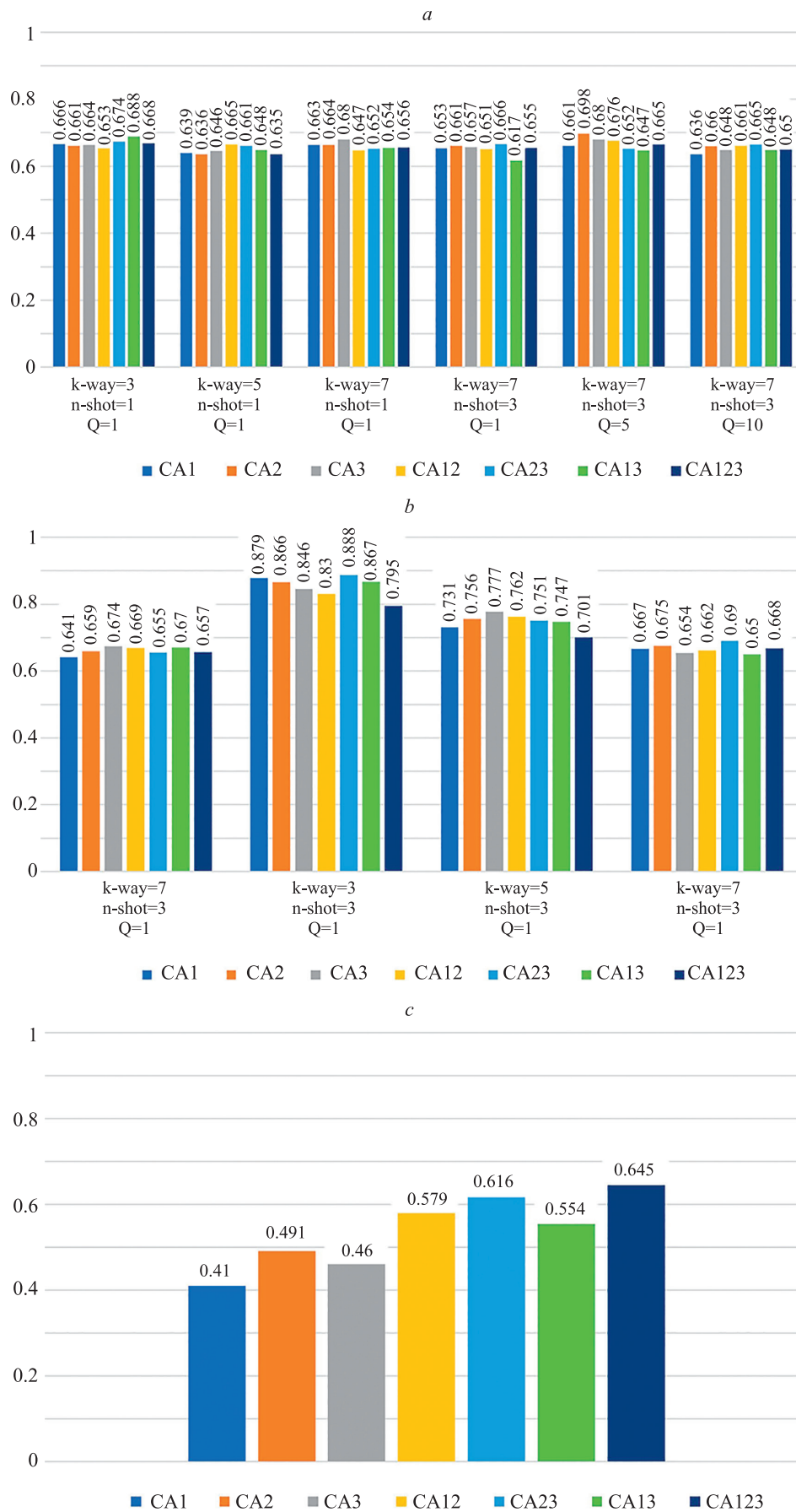relation between used bracelet (or their combinations) and the calculated metric. Considering those two facts, we concluded that it is more suitable to use the few-shot approach and that there is no major impact on a result depending on used bracelets.

Conclusion

In our paper, we explored various scenarios for model training. We tested the range of few-shot models, with different settings for few datasets and their combinations. We carried out the grid search over the hyper-parameters of a few-shot model for dataset; we adjusted the different values of LR and the optimizer: n-shot, k-way, q-queries. Results are presented in Fig. 3, *a*, *b*. Our results show that training with few-shot neural nets on 56k samples we can achieve better results than training on random forest with 160k samples, see Fig. 3, *c*.

Our results show that increasing the number of features via combining different datasets does not significantly increase the accuracy of the model. In our opinion, this behavior is primarily associated with an increase in the dataset noise after the dataset combination. The noisiness of the data is often the cause of difficulties in predictions using neural networks. The lack of strong patterns in the data obtained from the different band also affects the accuracy of neural networks.

Finally, combining datasets increase the dimension of the input data. For a single dataset, the sample size is (8, 20, 20), which equals to 3200 features characterizing a certain action. In the case of combining datasets the sample size is (24, 20, and 20), the number of features is 9600, this makes model training more complicated and requires using complex deep learning techniques. Using deep models has its draw-backs in speed of inference, slowing down the inference can affect the quality of the tool application in real-time cases.

Thus, the use of a minimum number of sensors in combination with few-shot signal processing techniques provides the possibility of organizing quick and cost-effective interaction with people with hearing and speech disabilities.

References

1. Agrawal S.C., Jalal A.S., Tripathi R.K. A survey on manual and non-manual sign language recognition for isolated and continuous sign. *International Journal of Applied Pattern Recognition*, 2016, vol. 3, no. 2, pp. 99–134. https://doi.org/10.1504/ijapr.2016.079048
2. Bragg D., Koller O., Bellard M., Berke L., Boudrealt P., Braffort A., Caselli N., Huenerfauth M., Kacorri H., Verhoef T., Vogler C., Morris M.R. Sign language recognition, generation, and translation: An interdisciplinary perspective. *Proc. of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS)*, 2019, pp. 16–31. https://doi.org/10.1145/3308561.3353774
3. Kamal S.M., Chen Y., Li S., Shi X., Zheng J. Technical approaches to Chinese sign language processing: A review. *IEEE Access*, 2019, vol. 7, pp. 96926–96935. https://doi.org/10.1109/ACCESS.2019.2929174
4. O'Connor T.F., Fach M.E., Miller R., Root S.E., Mercier P.P., Lipomi D.J. The Language of Glove: Wireless gesture decoder with low-power and stretchable hybrid electronics. *PLoS ONE*, 2017, vol. 12, no. 7, pp. e0179766. https://doi.org/10.1371/journal.pone.0179766
5. Song Y., Lee S., Choi Y., Han S., Won H., Sung T.-H., Choi Y., Bae J. Design framework for a seamless smart glove using a digital knitting system. *Fashion and Textiles*, 2021, vol. 8, no. 1, pp. 6. https://doi.org/10.1186/s40691-020-00237-2
6. Zhou Z., Chen K., Li X., Zhang S., Wu Y., Zhou Y., Meng K., Sun C., He Q., Fan W., Fan E., Lin Z., Tan X., Deng W., Yang J., Chen J. Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays. *Nature Electronics*, 2020, vol. 3, no. 9, pp. 571–578. https://doi.org/10.1038/s41928-020-0428-6
7. Bernhardt P. *Myo SDK Beta 7*. Available at: https://developerblog.myo.com/myo-sdk-beta-7/ (accessed:10.02.2022).
8. Abreu J.G., Teixeira J.M., Figueiredo L.S., Teichrieb V. Evaluating sign language recognition using the Myo armband. *Proc. of the 18th Symposium on Virtual and Augmented Reality (SVR)*, 2016, pp. 64–70. https://doi.org/10.1109/SVR.2016.21
9. Wang Y., Yao Q., Kwok J., Ni L.M. Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys*, 2020, vol. 53, no. 3, pp. 63. https://dl.acm.org/doi/10.1145/3386252
10. Wang F., Zhao S., Zhou X., Li C., Li M., Zeng Z. An recognition–verification mechanism for real-time Chinese sign language recognition based on multi-information fusion. *Sensors (Basel)*, 2019, vol. 19, no. 11, pp. 2495. https://doi.org/10.3390/s19112495

Литература

1. Agrawal S.C., Jalal A.S., Tripathi R.K. A survey on manual and non-manual sign language recognition for isolated and continuous sign // International Journal of Applied Pattern Recognition. 2016. V. 3. N 2. P. 99–134. https://doi.org/10.1504/ijapr.2016.079048
2. Bragg D., Koller O., Bellard M., Berke L., Boudrealt P., Braffort A., Caselli N., Huenerfauth M., Kacorri H., Verhoef T., Vogler C., Morris M.R. Sign language recognition, generation, and translation: An interdisciplinary perspective // Proc. of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS). 2019. P. 16–31. https://doi.org/10.1145/3308561.3353774
3. Kamal S.M., Chen Y., Li S., Shi X., Zheng J. Technical approaches to Chinese sign language processing: A review // IEEE Access. 2019. V. 7. P. 96926–96935. https://doi.org/10.1109/ACCESS.2019.2929174
4. O'Connor T.F., Fach M.E., Miller R., Root S.E., Mercier P.P., Lipomi D.J. The Language of Glove: Wireless gesture decoder with low-power and stretchable hybrid electronics // PLoS ONE. 2017. V. 12. N 7. P. e0179766. https://doi.org/10.1371/journal.pone.0179766
5. Song Y., Lee S., Choi Y., Han S., Won H., Sung T.-H., Choi Y., Bae J. Design framework for a seamless smart glove using a digital knitting system // Fashion and Textiles. 2021. V. 8. N 1. P. 6. https://doi.org/10.1186/s40691-020-00237-2
6. Zhou Z., Chen K., Li X., Zhang S., Wu Y., Zhou Y., Meng K., Sun C., He Q., Fan W., Fan E., Lin Z., Tan X., Deng W., Yang J., Chen J. Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays // Nature Electronics. 2020. V. 3. N 9. P. 571–578. https://doi.org/10.1038/s41928-020-0428-6
7. Bernhardt P. Myo SDK Beta 7 [Электронный ресурс]. URL: https://developerblog.myo.com/myo-sdk-beta-7/ (дата обращения: 10.02.2022).
8. Abreu J.G., Teixeira J.M., Figueiredo L.S., Teichrieb V. Evaluating sign language recognition using the Myo armband // Proc. of the 18th Symposium on Virtual and Augmented Reality (SVR). 2016. P. 64–70. https://doi.org/10.1109/SVR.2016.21
9. Wang Y., Yao Q., Kwok J., Ni L.M. Generalizing from a few examples: A survey on few-shot learning // ACM Computing Surveys. 2020. V. 53. N 3. P. 63. https://dl.acm.org/doi/10.1145/3386252
10. Wang F., Zhao S., Zhou X., Li C., Li M., Zeng Z. An recognition–verification mechanism for real-time Chinese sign language recognition based on multi-information fusion // Sensors. 2019. V. 19. N 11. P. 2495. https://doi.org/10.3390/s19112495

564

Научно-технический вестник информационных технологий, механики и оптики, 2022, том 22, № 3
Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2022, vol. 22, no 3

11. Kim S., Kim J., Ahn S., Kim Y. Finger language recognition based on ensemble artificial neural network learning using armband EMG sensors. *Technology and Health Care*, 2018, vol. 26, suppl. 1, pp. 249–258. https://doi.org/10.3233/THC-174602

12. Paudyal P., Lee J., Banerjee A., Sandeep K.S. A comparison of techniques for sign language alphabet recognition using armband wearables. *ACM Transactions on Interactive Intelligent Systems*, 2019, vol. 9, no. 2-3, pp. 1–26. https://doi.org/10.1145/3150974

13. Tateno S., Liu H., Ou J. Development of sign language motion recognition system for hearing-impaired people using electromyography signal. *Sensors*, 2020, vol. 20, no. 20, pp. 5807. https://doi.org/10.3390/s20205807

14. Sheng X., Lv B., Guo W., Zhu X. Common spatial-spectral analysis of EMG signals for multiday and multiuser myoelectric interface. *Biomedical Signal Processing and Control*, 2019, vol. 53, pp. 101572. https://doi.org/10.1016/j.bspc.2019.101572

15. Zhang L., Shi Y., Wang W., Chu Y., Yuan X. Real-time and user-independent feature classification of forearm using EMG signals. *Journal of the Society for Information Display*, 2019, vol. 27, no. 2, pp. 101–107. https://doi.org/10.1002/jsid.749

16. Das P., Paul S., Ghosh J., Palbhowmik S., Neo-Gi B., Ganguly A. An approach towards the representation of sign language by electromyography signals with fuzzy implementation. *International Journal of Sensors, Wireless Communications and Control*, 2017, vol. 7, no. 1, pp. 26–32. https://doi.org/10.2174/2210327907666170222093839

17. Côté-Allard U., Fall C.L., Drouin A., Campeau-Lecours A., Gosselin C., Glette K., Laviolette F., Gosselin B. Deep learning for electromyographic hand gesture signal classification using transfer learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2019, vol. 27, no. 4, pp. 760–771. https://doi.org/10.1109/TNSRE.2019.2896269

18. Tsinganos P., Cornelis B., Cornelis J., Jansen B., Skodras A. Data augmentation of surface electromyography for hand gesture recognition. *Sensors*, 2020, vol. 20, no. 17, pp. 4892. https://doi.org/10.3390/s20174892

19. Li W., Shi P., Yu H. Gesture recognition using surface electromyography and deep learning for prostheses hand: state-of-the-art, challenges, and future. *Frontiers in Neuroscience*, 2021, vol. 15, pp. 621885 https://doi.org/10.3389/fnins.2021.621885

20. Rahimian E., Zabihi S., Asif A., Farina D., Atashzar S.F., Mohammadi A. FS-HGR: Few-shot learning for hand gesture recognition via electromyography. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2021, vol. 29, pp. 1004–1015. https://doi.org/10.1109/TNSRE.2021.3077413

21. Finn C., Abbeel P., Levine S. Model-agnostic meta-learning for fast adaptation of deep networks. *Proceedings of Machine Learning Research*, 2017, vol. 70, pp. 1126–1135.

22. Lee Y., Choi S. Gradient-based meta-learning with learned layerwise metric and subspace. *Proc. of the 35th International Conference on Machine Learning (ICML)*. V. 7, 2018, pp. 4574–4586.

23. Koch G., Zemel R., Salakhutdinov R. Siamese neural networks for one-shot image recognition. *ICML Deep Learning Workshop*. V. 2, 2015.

24. Snell J., Swersky K., Zemel R. Prototypical networks for few-shot learning. *Advances in Neural Information Processing Systems*, 2017, pp. 4077–4087.

25. Vaezi Joze H.R., Koller O. MS-ASL: A large-scale data set and benchmark for understanding American sign language. *Proc. of the 30th British Machine Vision Conference (BMVC)*, 2019.

26. De Coster M., Van Herreweghe M., Dambre J. Sign language recognition with transformer networks. *Proc. 12th International Conference on Language Resources and Evaluation (LREC)*, 2020, pp. 6018–6024.

27. Pigou L., Van Herreweghe M., Dambre J. Sign classification in sign language corpora with deep neural networks. *Proc. of the International Conference on Language Resources and Evaluation (LREC)*, Workshop, 2016, pp. 175–178.

28. Pradhan A., He J., Jiang N. Performance optimization of surface electromyography based biometric sensing system for both verification and identification. *IEEE Sensors Journal*, 2021, vol. 21, no. 19, pp. 21718–21729. https://doi.org/10.1109/JSEN.2021.3079428

29. Young A.J., Hargrove L.J., Kuike T.A. Improving myoelectric pattern recognition robustness to electrode shift by changing interelectrode distance and electrode configuration. *IEEE Transactions on*

11. Kim S., Kim J., Ahn S., Kim Y. Finger language recognition based on ensemble artificial neural network learning using armband EMG sensors // Technology and Health Care. 2018. V. 26. S. 1. P. 249–258. https://doi.org/10.3233/THC-174602

12. Paudyal P., Lee J., Banerjee A., Sandeep K.S. A comparison of techniques for sign language alphabet recognition using armband wearables // ACM Transactions on Interactive Intelligent Systems. 2019. V. 9. N 2-3. P. 1–26. https://doi.org/10.1145/3150974

13. Tateno S., Liu H., Ou J. Development of sign language motion recognition system for hearing-impaired people using electromyography signal // Sensors. 2020. V. 20. N 20. P. 5807. https://doi.org/10.3390/s20205807

14. Sheng X., Lv B., Guo W., Zhu X. Common spatial-spectral analysis of EMG signals for multiday and multiuser myoelectric interface // Biomedical Signal Processing and Control. 2019. V. 53. P. 101572. https://doi.org/10.1016/j.bspc.2019.101572

15. Zhang L., Shi Y., Wang W., Chu Y., Yuan X. Real-time and user-independent feature classification of forearm using EMG signals // Journal of the Society for Information Display. 2019. V. 27. N 2. P. 101–107. https://doi.org/10.1002/jsid.749

16. Das P., Paul S., Ghosh J., Palbhowmik S., Neo-Gi B., Ganguly A. An approach towards the representation of sign language by electromyography signals with fuzzy implementation // International Journal of Sensors, Wireless Communications and Control. 2017. V. 7. N 1. P. 26–32. https://doi.org/10.2174/2210327907666170222093839

17. Côté-Allard U., Fall C.L., Drouin A., Campeau-Lecours A., Gosselin C., Glette K., Laviolette F., Gosselin B. Deep learning for electromyographic hand gesture signal classification using transfer learning // IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2019. V. 27. N 4. P. 760–771. https://doi.org/10.1109/TNSRE.2019.2896269

18. Tsinganos P., Cornelis B., Cornelis J., Jansen B., Skodras A. Data augmentation of surface electromyography for hand gesture recognition // Sensors. 2020. V. 20. N 17. P. 4892. https://doi.org/10.3390/s20174892

19. Li W., Shi P., Yu H. Gesture recognition using surface electromyography and deep learning for prostheses hand: state-of-the-art, challenges, and future // Frontiers in Neuroscience. 2021. V. 15. P. 621885 https://doi.org/10.3389/fnins.2021.621885

20. Rahimian E., Zabihi S., Asif A., Farina D., Atashzar S.F., Mohammadi A. FS-HGR: Few-shot learning for hand gesture recognition via electromyography // IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2021. V. 29. P. 1004–1015. https://doi.org/10.1109/TNSRE.2021.3077413

21. Finn C., Abbeel P., Levine S. Model-agnostic meta-learning for fast adaptation of deep networks // Proceedings of Machine Learning Research. 2017. V. 70. P. 1126–1135.

22. Lee Y., Choi S. Gradient-based meta-learning with learned layerwise metric and subspace // Proc. of the 35th International Conference on Machine Learning (ICML). V. 7. 2018. P. 4574–4586.

23. Koch G., Zemel R., Salakhutdinov R. Siamese neural networks for one-shot image recognition // ICML Deep Learning Workshop. V. 2. 2015.

24. Snell J., Swersky K., Zemel R. Prototypical networks for few-shot learning // Advances in Neural Information Processing Systems. 2017. P. 4077–4087.

25. Vaezi Joze H.R., Koller O. MS-ASL: A large-scale data set and benchmark for understanding American sign language // Proc. of the 30th British Machine Vision Conference (BMVC). 2019.

26. De Coster M., Van Herreweghe M., Dambre J. Sign language recognition with transformer networks // Proc. 12th International Conference on Language Resources and Evaluation. LREC. 2020. P. 6018–6024.

27. Pigou L., Van Herreweghe M., Dambre J. Sign classification in sign language corpora with deep neural networks // Proc. of the International Conference on Language Resources and Evaluation (LREC), Workshop. 2016. P. 175–178.

28. Pradhan A., He J., Jiang N. Performance optimization of surface electromyography based biometric sensing system for both verification and identification // IEEE Sensors Journal. 2021. V. 21. N 19. P. 21718–21729. https://doi.org/10.1109/JSEN.2021.3079428

29. Young A.J., Hargrove L.J., Kuike T.A. Improving myoelectric pattern recognition robustness to electrode shift by changing interelectrode distance and electrode configuration // IEEE Transactions on

Научно-технический вестник информационных технологий, механики и оптики, 2022, том 22, № 3
Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2022, vol. 22, no 3

565

Biomedical Engineering, 2012, vol. 59, no. 3, pp. 645–652. https://doi.org/10.1109/TBME.2011.2177662

30. Benatti S., Farella E., Gruppioni E., Benini L. Analysis of robust implementation of an EMG pattern recognition based control. *Proc. of the Analysis of Robust Implementation of an EMG Pattern Recognition Based Control (BIOSIGNALS)*, 2014, pp. 45–54. https://doi.org/10.5220/0004800300450054

31. George J.A., Neibling A., Paskett M.D., Clark G.A. Inexpensive surface electromyography sleeve with consistent electrode placement enables dexterous and stable prosthetic control through deep learning. *arXiv*, 2003, arXiv:2003.00070. https://doi.org/10.48550/arXiv.2003.00070

32. Vinyals O., Blundell C., Lillicrap T., Kavukcuoglu K., Wierstra D. Matching networks for one shot learning. *Advances in Neural Information Processing Systems*, 2016, pp. 3637–3645.

33. Kaczmarek P., Mańkowski T., Tomczyński J. putEMG–A surface electromyography hand gesture recognition dataset. *Sensors*, 2019, vol. 19, no. 16, pp. 3548. https://doi.org/10.3390/s19163548

Biomedical Engineering. 2012. V. 59. N 3. P. 645–652. https://doi.org/10.1109/TBME.2011.2177662

30. Benatti S., Farella E., Gruppioni E., Benini L. Analysis of robust implementation of an EMG pattern recognition based control // Proc. of the Analysis of Robust Implementation of an EMG Pattern Recognition Based Control (BIOSIGNALS). 2014. P. 45–54. https://doi.org/10.5220/0004800300450054

31. George J.A., Neibling A., Paskett M.D., Clark G.A. Inexpensive surface electromyography sleeve with consistent electrode placement enables dexterous and stable prosthetic control through deep learning // arXiv. 2003. arXiv:2003.00070. https://doi.org/10.48550/arXiv.2003.00070

32. Vinyals O., Blundell C., Lillicrap T., Kavukcuoglu K., Wierstra D. Matching networks for one shot learning // Advances in Neural Information Processing Systems. 2016. P. 3637–3645.

33. Kaczmarek P., Mańkowski T., Tomczyński J. putEMG–A surface electromyography hand gesture recognition dataset // Sensors. 2019. V. 19. N 16. P. 3548. https://doi.org/10.3390/s19163548

Authors

Grigory F. Shovkoplias — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57222048908, https://orcid.org/0000-0001-7777-6972, gfshovkoplias@itmo.ru

Dmitriy A. Strokov — Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0000-0003-1924-0621, dmitry.strokov@mail.ru

Daniil V. Kasantsev —Senior Laboratory Assistant, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0000-0001-7974-0922, kazantsev@itkvar.ru

Aleksandra S. Vatian — Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57191870868, https://orcid.org/0000-0002-5483-716X, alexvatyan@gmail.com

Arip A. Asadulaev — Assistant, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0000-0002-2581-935X, aripasadulaev@itmo.ru

Ivan V. Tomilov — Senior Laboratory Assistant, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0000-0003-1886-2867, ivan-tomilov3@yandex.ru

Anatoly A. Shalyto — D. Sc., Full Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57222048908, https://orcid.org/0000-0002-2723-2077, anatoly.shalyto@gmail.com

Natalia F. Gusarova — PhD, Senior Researcher, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57162764200, https://orcid.org/0000-0002-1361-6037, natfed@list.ru

Авторы

Шовкопляс Григорий Филиппович — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57222048908, https://orcid.org/0000-0001-7777-6972, gfshovkoplias@itmo.ru

Строков Дмитрий Андреевич — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0000-0003-1924-0621, dmitry.strokov@mail.ru

Казанцев Даниил Владимирович — старший лаборант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0000-0001-7974-0922, kazantsev@itkvar.ru

Ватьян Александра Сергеевна — доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57191870868, https://orcid.org/0000-0002-5483-716X, alexvatyan@gmail.com

Асадулаев Арип Амирханович — ассистент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0000-0002-2581-935X, aripasadulaev@itmo.ru

Томилов Иван Вячеславович — старший лаборант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0000-0003-1886-2867, ivan-tomilov3@yandex.ru

Шалыто Анатолий Абрамович — доктор технических наук, профессор, профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57222048908, https://orcid.org/0000-0002-2723-2077, anatoly.shalyto@gmail.com

Гусарова Наталия Федоровна — кандидат технических наук, старший научный сотрудник, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57162764200, https://orcid.org/0000-0002-1361-6037, natfed@list.ru

566

Научно-технический вестник информационных технологий, механики и оптики, 2022, том 22, № 3
Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2022, vol. 22, no 3