On the Impact of Computer Vision Algorithms on Sport Training Automation: Proof of Concept for Shadow Boxing Virtual Instructor

Ilya Makarov[1,2], Stanislav Petrov[1]

[1]HSE University, Moscow, Russia, Pokrovsky boulevard 11, 109028 Moscow, Russian Federation

[2]Artificial Intelligence Research Institute (AIRI), Moscow, Russia, Nizhny Susalny lane 5 p. 19, 105064 Moscow, Russian Federation

Abstract

We overview several XR applications of deep convolutional neural networks to the opportunity for creating an automated sports training process. From smart soccer and basketballs to automated fitness training programs, the progress of computer vision methods combined with personalized recommender systems and specific data science algorithms allows one to train non-contact training processes in a semi-automated manner with virtual mirrors and XR devices. We overview modern progress in this area and also present our own prototype of Shadow Boxing simulator for Virtual Mirror aiming to match part of boxer training with automated control and gamification process. We show that the trend of automating training instructors in sport leads to a positive shift in sportsmen and trainees' view of artificial intelligence in common life.

Keywords

Computer Vision, Mixed Reality, Automated Sports Instructor, Boxing Simulator

1. Introduction

Nowadays, mixed reality (MR) and artificial intelligence (AI) allow humans to engage in different types of sport life activities, which previously requires human-to-human interaction, but now can be substituted with training in virtual, augmented, or mixed reality. The idea of that is quite simple: usually, the instructor repeats the sequences of commands, which are limited in number and length, and such a small world of possible actions allows AI specialists to train the model for sport and fitness exercises, while automatically personalize the process of enriching training program based on exact person progress. The overview of the existing application is presented in Table 1. Although there are commercial products in this field, it is still open for market changes and new startups, while waiting for significant improvements on computer vision and life-long learning algorithms supporting the systems of automated sports training.

AISMA-2021: International Workshop on Advanced in Information Security Management and Applications, Stavropol, Krasnoyarsk, Russia, October 1, 2021

iamakarov@hse.ru (I. Makarov); stasdp@mail.ru (S. Petrov)

0000-0002-3308-8825 (I. Makarov)

Table 1
Table of Smart Fitness Simulators based on Computer Vision (CV) or Data Science (DS) techniques.

Name	CV/DS	Product/ Prototype	Reference
Smartspot (fitness virtual mirror)	CV	Prototype	[1]
D.Gym (gym tracking and analytics)	CV	None	[2]
Fitbod (recommending exercises)	DS	Product	[3]
Phormatics (exercise pose evaluation)	CV	Prototype	[4]
Perch (barbell exercises)	CV	Prototype	[5]
EPAM Smart Gym (barbell)	CV	Prototype	[6]
Smart Balls by DU (soccer/basketball training)	CV	Product	[7]
Just Dance (dancing video game)	CV	Product	[8]

2. Core Concepts in Automated Instructors

In order to implement a working prototype of virtual avatar training sportsman one need to solve the following problems:

- to develop RGBD body and skeleton motion segmentation [9] and tracking [10], usually used for aerobics/dance exercises (see [8][1], in the latter video upper left corner contains Human Figure Segmentation based on which each body part movement is scored and then visualized in-game);
- to implement visualization in computer graphics engines, for example, Unreal Engine 4 or Unity [11];
- to analyze wearable smart devices for exercises on gym equipment, for example, [5], and track heart-beat, velocity, weight (see simple prototype[2]), and visualize right and wrong points using human pose estimation[3] [12, 13, 14];
- to train several tracks recommendation system similar to [3] [4] allowing to personalize training experience [15].

In what follows, we describe one of such systems (work-in-progress) designed for a particular type of boxing training, called 'Shadow Boxing', which requires the imagination of a virtual enemy and working on boxing techniques without haptic interaction with a partner/instructor.

3. Proposal of Shadow Boxing MR App

Every boxer uses shadow boxing during training. Beginners should start warming up with shadow boxing for at least 15 minutes. Usually, it takes 30 minutes for common boxers and 60 minutes for professionals.

[1]https://www.youtube.com/watch?v=PvZA8NKgrBI
[2]https://youtu.be/3CSzFhabZ3g
[3]https://github.com/cbsudux/awesome-human-pose-estimation
[4]https://www.youtube.com/watch?v=yOizW6130QY

Figure 1: Concept of Boxing in VR

One of the main conditions for evaluation of Shadow Boxing is the fact that sportsmen should not be getting tired when shadowboxing. Another reason for training using shadow boxing is the development of weak back muscles, the low endurance of which may be crucial if the boxer is weakly adapted to many missing hits in a real fight after training with a bag. The main idea is that a boxer has to elaborate on improving their imagination of fight, developing so-called "muscle memory" and working on certain goals during shadow boxing, which can be done anywhere at anytime.

The main objectives of Shadow Boxing can be divided into an improvement of several key parameters:

1. boxing technique, offense, and defense (complex):
   a) correct hits (tracking, CV)
   b) balance (tracking, CV)
   c) combinations of hits as simple techniques (tracking, CV)

2. overall fighting abilities:
   a) strength (sensor)
   b) power (acceleration, CV)
   c) speed (speed, CV)
   d) endurance (time series, CV & DS)
   e) footwork (tracking, CV)
   f) rhythm (complex CV, unknown)

4. Related Work

4.1. Existing Boxing Applications

In the first part, we overview existing products and prototypes in the area of Boxing simulators.

4.1.1. Patents and Existing Prototypes in Boxing

Robotic boxing punching bags were presented in [16] and real-world prototype [17]. Both ideas keep insisting on interaction with a physical bag to hit. The scientific explanation of suggested models can be found in the corresponding patent and punch measurement work [18].

4.1.2. Sensor application using electrodes for imitating hit feeling

In [19], the similar idea of haptic feedback was used for reconstruction of hand and finger movement based on myogram, tracking muscle activities in the middle between hand and elbow.

4.1.3. VR applications

Currently, there are two popular video games: [20][5] and [21][6]. The first provides the experience of virtual punching bug, order of hits, slow-motion, and other gamified moments of simulating boxing match; the second is the fist mitt hitting training without a partner. Both applications lack the realism of punching training and interaction with both virtual enemy and virtual trainer. Moreover, the entertainment model of immersion in video games was prioritized over the training aspect of professional and novice boxers via human-computer interaction in VR.

4.2. Computer Vision Research on Boxing Recognition

Existing research on boxing action recognition and related topics are mostly connected to human pose estimation and segmentation, and also boxing hit recognition.

In [22], the authors presented a punch recognition system based on prior knowledge of boxing punches. The authors of [23] presented a robust framework to recognize fine-grained boxing punches from specifically posed depth images from head/ceiling view.

An overview of the application of deep learning to RGB-D-based motion recognition: RGB-based, depth-based, skeleton-based, and RGB+D-based, can be found in [24]. The game and sports action recognition dataset most close to our task was presented in [25]. A survey on deep learning for sport-specific movement recognition performance describes advances in the field of sports action recognition [26].

Shortly speaking, there are datasets allowing recognition of certain types of the simplest punches and human movements, while also estimating parameters of such punches, such as speed, acceleration, and exact three-D locations of joints. However, there is no complex study and available dataset of boxing techniques, combinations, and parameters of such training. This is the situation, in which a combination of boxing instructors and CV specialists will be required for synergy between technologies and expert domain knowledge.

[5] https://www.youtube.com/watch?v=7zuKLu7dOng

[6] https://www.youtube.com/watch?v=n51DqCpqGSU

5. Our Vision of Shadow Boxing Application

We will in detail describe our vision of Shadow Boxing gamification and implementation of human-computer interaction with Virtual Boxing Instructor.

We aim to create a simple Boxer avatar visualized on a virtual mirror that authorizes the person (possibly based on biometry). At first, it is required to calibrate the system while asking the user to make simple movements, single hits, and short combinations in order to adapt hit score and tracking systems for the user.

We visualize the animation of correct combination and user's performance showing the differences and mistakes in the process of training as follows. An animated avatar shows movements and a user repeats them. After a certain sequence is executed, the sportsman sees his score, and a slow-motion replay of his action is compared with a virtual avatar executing the same motion sequence. The differences between recorded movements of the avatar and the user are visualized on the screen and interpreted in advises how to overcome this difference.

Depending on boxer level, to master certain combinations we may divide the sequence into well-done and the problem moves, which may be trained separately in slower motion to later combine them in one sequence after each sub-part is mastered. Each simple and combined boxing sequence should be defined together with the boxing instructor or be based on some boxing training manual.

Every task is considered as a sequence of simple moves and punches initialized by timer/sound, and visualized as a sequence of cheat sheets during boxer performing this sequence. For example, the combination "hit left - hit left - jab right - evade left - move right" is evaluated based on position, speed, acceleration, and depth parameters of each movement in parallel with recognition of each movement from one or several RGBD cameras, which can be placed in front, around or above sportsman. The collection of a new dataset will be required depending on the complexity of the prototype and different boxing punches modalities.

Based on boxer data we can collect his performance, evaluate his parameters (endurance, speed, balance, transitions from defense to offense, etc.), and recommend he train certain sequences in which he lacks skills. In the first step, it will be hardcoded rules based on boxing theory ideas, later we will use collected data to use data mining techniques provided self-supervised recommendations.

As for implementation, we start with person tracking and identification and training deep learning models for human 3D skeleton reconstruction (pose estimation) so it will be coherent with boxer movement on video and show avatar repeating these actions. Then we deal with boxing hit recognition (hook, jab, etc., on which there are no available open datasets) from RGBD video and evaluation of speed and power parameters of how the hits should be done, so that the system will be able to recognize simple hits and their combinations, score the hit based on inner parameters and show how it should have been properly done. Finally, we aim to develop a virtual enemy avatar for which the training boxer should develop his counter-measure hits and movement, first with virtual hints, and further, just by observing his movements. The resulting application will support these two modes of training as repeating hits and knowing what hits to throw against the virtual enemy.

6. Development and Release Description for Shadow Boxing Training Application

Below, we describe the structure of input information, problems in AI and CV to be solved, and consistent increase in functionality of the Shadow Boxing application. We also envision the development of consequent updates to Shadow Boxing Training based on increased use of existing boxing fights datasets and MR head-mounted displays.

6.1. Input Channels

The combinations of different inputs should be presented for a successful Demo integrating into a working prototype. The user has to be properly led to the current goals, see the results of training, collect the feedback, and do not be disappointed by mistakes of recognition module, which requires very robust and precise algorithms to be developed applied to the boxing domain.

We consider the combination of the following input channels:

- Graphical avatar, tracking the position of human, and providing personal training based on input parameters of training mode and current level;
- Sound input into headphones for instructor commands, especially during combination movements, during which boxer could not properly see the whole picture, but has to respond fast to possible motion errors or react to virtual enemy actions;
- AR interface to visualize virtual enemy and make the training of real fight possible;
- Possible electrode-based sensors for simulating recoil feedback and sensory feelings of touching/punching may be tested.

6.2. Problems Pipeline to be Solved

We formulate core concepts of computer vision and machine learning fields to be solved in order to make the prototype of Shadow Boxing training sufficient for evaluation in a real-world scenarios with professional boxing sportsmen, together with complexity estimation based on state-of-the-art results in the related fields.

1. Computer Vision
    a) Person Identification (simple)
    b) Person Tracking, including reidentification in occluded scenarios (middle)
    c) Depth Person Segmentation (middle)
    d) Depth Person Skeleton Reconstruction (middle-hard)
    e) Boxer Action Recognition (middle-hard)
    f) Generating Avatar Movement based on collected data (hard)
    g) Parsing boxing TV videos and Enemy reconstruction (very hard)
    h) Programming game AI for Boxing sparring (state-of-art)

2. Machine Learning
    a) Robust measurement of punch parameters (simple)

b) Measuring human performance during series of training (middle)
c) Recommending simple moves to work on (simple)
d) Recommending long sequences for skills improvement (middle)
e) Boxing Technique learning as a list of sequences to be learned combined with proper evaluation of sub-part movements (hard, state-of-art)

In addition, Reinforcement Learning methods may be applied to reverse engineer agents' behavior, thus training a virtual agents to perform actions by receiving rewards for correct actions recognized by computer vision encoder similar to environments for training agents in 3D shooter games [27, 28, 29, 30, 31, 32] or 2D interaction games [33, 34, 35].

6.3. AR Helmet Training (Demo v1.5)

As a continuation of the previous work, we aim to work on a prototype of a Mixed Reality application, in which the boxer will wear an optical see-through head-mounted display allowing him to see visual enemy or training targets in Augmented Reality. We plan to add the following elements based on the previous description.

- Implementation of simplest training exercises similar to VR Creed Boxing game: hitting certain areas in front of the boxer in the corresponding sequence, hitting virtual bag to hit other targets, evade, speed in-/out- from the enemy zone, etc.
- Virtual BOTs playing box against a person with health bar measures and training certain boxing techniques to counter them, such as training mode against Southpaw.

6.4. TV highlighting mode (Demo v2.0)

Standalone application to recognize and parse boxing matches, visualize highlights and recognize the sequence of actions to later map them to predefined short avatar scripted scenes for learning in AR boxing demo. It could be used for TV parsing and slow-motion reconstructions of highlighted events.

The p

6.5. Personalized enemy for Shadow Boxing Sparring in AR (Demo v2.5)

After we learn how to parse TV videos of boxing matches, we could collect data on the specific person and implement his skills and techniques into AI in-game prototype, which later will be used as an opponent in AR for training against exact type of opponent with known skills, parameters, and techniques.

This last task is challenging and at the current moment, nobody tries to do this, but the current progress in neural networks and human action recognition allow us to state that we could solve this task in Machine Learning and Game Artificial Intelligence fields.

7. Example of implementation from scratch

One of the main goals is to create a neural network that could capture actions that differs a little from each other. Also, the domain in our work – boxing, is a very narrow area for action recognition. Thus, there is no dataset with different boxing punches where the boxer stays right in front of the camera.

We aim to create a domain-specific network that could help boxers in their training. Given that we can assume some limitations to model usage:

- The presence of only one person in the video
- This person is 2-3 meters away in front of the camera
- Boxing is performed in direction of the camera

In this work, we wanted to create a dataset with as many as possible features. There are several ways of doing that, such as recording video and using different neural networks to extract pose, mask, depth, etc., or to record video using a special camera that writes depth, IR, pose, and mask data. The first approach is straightforward and accurate – using video and modern neural networks, we can collect pose estimation, depth, and mask of the human, but this approach is computationally expensive for detailed streaming data annotation despite usefulness of the methods proposed in [36, 37]. It can be efficiently done without solving the mask problem if only depth information is required [38, 39, 40, 41, 42, 43, 44, 45, 46, 47]. In addition, pose estimation may be efficiently solved by graph neural networks, which are of great use for extracting and preserving structural dependencies. Our works in the domain of graph feature engineering can be found in [48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63].

The latter approach is more expensive than the former but gives a human pose, the mask of a person, and a depth map in near real-time and does not require wearing additional sensors from the sportsman. As the balance for our task, we stop on the latter approach.

For dataset collection purposes, we have ORRBEC Astra Pro. It can capture RGB image data in 1280×780 resolution, depth image 640×480 with 30fps both. It uses USB 2.0 to connect to a computer.

ORRBEC developed Astra SDK for body tracking. It was used to develop a program that in different threads read and write to correspondent files: RGB video of the action, Depth video, Mask of the person body with a mask of the floor, Pose keypoints.

For this work, we collected video representations of four basic boxer actions: Defense, Cross punch, Hook, Uppercut. Each class contains 200 samples of the author boxing 2-3 meters away in front of the camera. Actions captured in one location, different clothes, and different techniques were used(with wrist rotation and without it). All punches was performed with the right hand.

Data preparation could be split into 3 stages:

- Input video splinted manually in 32 video pieces representing one boxing punch
- Splinted video cut to 480×480 resolution from the right and left sides
- Cut video is reshaped to 224×224 resolution video Data augmentation.

There were 3 types of data augmentation were performed. On the second step of the data, the preparation video was randomly cropped from right between 60- and 100-pixel range Temporal augmentation. For each 32-frame cut video clip randomly was taken bias in a range from -8 to -4 and from 4 to 8. This bias served to produce another 32-frame but using starting point plus bias.

In fact, Astra Camera recorded video with some freezes and not all samples were captured with 30fps, but as actions were performed with the different speeds it had no effect on the collected data.

In this work we stopped our choice on Multi streams I3D network[36]: I3D is a good performer on Kinetics dataset, is not enormous as C3D, and does not need sampling frames from the video. We omitted the use of optical flow as an input, instead of using depth flow, and do not take into account the pretraining encoder on ImageNet. The training was performed with the size of batch = 2 in 1 epoch. Test dataset contained 30% of all data. The size of the train dataset was 2750. The training was end to end with Cross-Entropy loss on the sum of the logits of all streams. We trained the network in 3 setups: all four streams; depth, pose and mask streams; depth and pose streams. After training, we get average accuracy on trimmed videos around 95% on test and validation sets. That could be explained by overfitting on small-size dataset.

Results demonstrate that the I3D net successfully classified all actions from the test data set. To improve the results, an augmented extended dataset was recorded and tested providing more robust and believable results for shadow boxing simulation [64].

8. Conclusion

We overview the state-of-the-art research and production projects in the field of automating fitness and sports training. We showed that despite the current advances in the deep learning field, technological aspects of creating a production-ready prototype are still quite challenging tasks.

We further discuss the particular problem of developing a computer vision based simulator for Shadow Boxing, and the problem arising during this development. We envision the future progress of such a simulator based on current deep learning and computer vision techniques and suggest the exact scenario of increasing complexity and usability of such systems in the boxer training process.

We aim to show the first version of the prototype during the conference and collect feedback both, in terms of user study and algorithm accuracy impacting the whole simulator. We also aim to measure human perception on whether such a system may substitute a human boxing instructor in non-contact sports training.

References

[1] Smartspot, https://smartspot.io, 2017.
[2] D.Gym, http://dohyungkim.com/dgym/, 2017.
[3] Fitbod, https://fitbod.me, 2018.
[4] Phormatics, https://github.com/jrobchin/phormatics, 2018.
[5] Perch, https://www.perch.fit/, 2018.
[6] EPAM, https://bit.ly/2uWmxUm, 2019.
[7] Smart Balls, https://buy.dribbleup.com/collection/smart-balls, 2019.
[8] Just Dance, https://www.ubisoft.com/en-us/game/just-dance-2019, 2019.

[9] L. L. Presti, M. La Cascia, 3d skeleton-based human action classification: A survey, Pattern Recognition 53 (2016) 130–147.
[10] Z.-Q. Cheng, Y. Chen, R. R. Martin, T. Wu, Z. Song, Parametric modeling of 3d human body shape—a survey, Computers & Graphics 71 (2018) 88–100.
[11] C. George, M. Spitzer, H. Hussmann, Training in ivr: investigating the effect of instructor design on social presence and performance of the vr user, in: Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, ACM, ACM, New York, USA, 2018, p. 27.
[12] D. Mehta, S. Sridhar, O. Sotnychenko, H. Rhodin, M. Shafiei, H.-P. Seidel, W. Xu, D. Casas, C. Theobalt, Vnect: Real-time 3d human pose estimation with a single rgb camera, ACM Transactions on Graphics (TOG) 36 (2017) 44.
[13] H. Fang, S. Xie, Y.-W. Tai, C. Lu, Rmpe: Regional multi-person pose estimation, in: The IEEE International Conference on Computer Vision (ICCV), volume 2, IEEE, New York, USA, 2017, pp. 1–10.
[14] G. Varol, D. Ceylan, B. Russell, J. Yang, E. Yumer, I. Laptev, C. Schmid, Bodynet: Volumetric inference of 3d human body shapes, arXiv preprint arXiv:1804.04875 arXiv:1804.04875 (2018) 1–27.
[15] T. T. Tran, J. W. Choi, C. Van Dang, G. SuPark, J. Y. Baek, J. W. Kim, Recommender system with artificial intelligence for fitness assistance system, in: 2018 15th International Conference on Ubiquitous Robots (UR), IEEE, IEEE, New York, USA, 2018, pp. 489–492.
[16] Boxing buddy system, https://patents.google.com/patent/US9586120B1/en, 2014.
[17] BotBoxer, http://botboxer.com/, 2018.
[18] S. Chadli, N. Ababou, A. Ababou, A new instrument for punch analysis in boxing, Procedia Engineering 72 (2014) 411 – 416. URL: http://www.sciencedirect.com/science/article/pii/S187770581400589X. doi:https://doi.org/10.1016/j.proeng.2014.06.073, the Engineering of Sport 10.
[19] P. Lopes, A. Ion, P. Baudisch, Impacto: Simulating physical impact by combining tactile stimulation with electrical muscle stimulation, in: Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology, UIST '15, ACM, New York, NY, USA, 2015, pp. 11–19. URL: http://doi.acm.org/10.1145/2807442.2807443. doi:10.1145/2807442.2807443.
[20] Creed: Rise to Glory, https://survios.com/creed/, 2018.
[21] The Fastest Fist, https://store.steampowered.com/app/544540/The_Fastest_Fist/, 2018.
[22] S. Kasiri-Bidhendi, C. Fookes, S. Morgan, D. T. Martin, S. Sridharan, Combat sports analytics: Boxing punch classification using overhead depthimagery, in: 2015 IEEE International Conference on Image Processing (ICIP), IEEE, New York, USA, 2015, pp. 4545–4549. doi:10.1109/ICIP.2015.7351667.
[23] S. Kasiri, C. Fookes, S. Sridharan, S. Morgan, Fine-grained action recognition of boxing punches from depth imagery, Computer Vision and Image Understanding 159 (2017) 143 – 153. URL: http://www.sciencedirect.com/science/article/pii/S1077314217300668. doi:https://doi.org/10.1016/j.cviu.2017.04.007, computer Vision in Sports.
[24] P. Wang, W. Li, P. Ogunbona, J. Wan, S. Escalera, Rgb-d-based human motion recognition with deep learning: A survey, Computer Vision and Image Understanding 171 (2018) 118 – 139. URL: http://www.sciencedirect.com/science/article/pii/S1077314218300663.

doi:https://doi.org/10.1016/j.cviu.2018.04.007.
[25] V. Bloom, D. Makris, V. Argyriou, G3d: A gaming action dataset and real time action recognition evaluation framework, in: Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on, IEEE, IEEE, New York, USA, 2012, pp. 7–12.
[26] E. E. Cust, A. J. Sweeting, K. Ball, S. Robertson, Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance, Journal of Sports Sciences 0 (2018) 1–33. doi:10.1080/02640414.2018.1521769, pMID: 30307362.
[27] I. Makarov et al., First-person shooter game for virtual reality headset with advanced multi-agent intelligent system, in: Proceedings of the 24th ACM international conference on Multimedia, 2016, pp. 735–736.
[28] I. Makarov, M. Tokmakov, L. Tokmakova, Imitation of human behavior in 3d-shooter game, AIST'2015 Analysis of Images, Social Networks and Texts (2015) 64.
[29] I. Makarov et al., Modelling human-like behavior through reward-based approach in a first-person shooter game, in: EEML Proceedings, 2016, pp. 24–33.
[30] I. Makarov, P. Polyakov, R. Karpichev, Voronoi-based path planning based on visibility and kill/death ratio tactical component, in: Proceedings of AIST, 2018, pp. 129–140.
[31] I. Makarov, P. Polyakov, Smoothing voronoi-based path with minimized length and visibility using composite bezier curves, in: AIST (Supplement), 2016, pp. 191–202.
[32] I. Makarov, O. Konoplia, P. Polyakov, M. Martynov, P. Zyuzin, O. Gerasimova, V. Bodishtianu, Adapting first-person shooter video game for playing with virtual reality headsets., in: FLAIRS Conference, 2017, pp. 412–416.
[33] I. Makarov, D. Savostyanov, B. Litvyakov, D. I. Ignatov, Predicting winning team and probabilistic ratings in "dota 2" and "counter-strike: Global offensive" video games, in: Proceedings of AIST, Springer, 2017, pp. 183–196.
[34] I. Kamaldinov, I. Makarov, Deep reinforcement learning in match-3 game, in: Proceedings of CoG, IEEE, 2019, pp. 1–4.
[35] I. Kamaldinov, I. Makarov, Deep reinforcement learning methods in match-3 game, in: Proceedings of AIST, Springer, 2019, pp. 51–62.
[36] J. Hong, B. Cho, Y. W. Hong, H. Byun, Contextual action cues from camera sensor for multi-stream action recognition, Sensors 19 (2019) 1382.
[37] K. Lomotin, I. Makarov, Automated image and video quality assessment for computational video editing, in: International Conference on Analysis of Images, Social Networks and Texts, Springer, 2020, pp. 243–256.
[38] I. Makarov, V. Aliev, O. Gerasimova, Semi-dense depth interpolation using deep convolutional neural networks, in: Proceedings of the 2017 ACM on Multimedia Conference, ACM, 2017, pp. 1407–1415.
[39] I. Makarov, D. Maslov, O. Gerasimova, V. Aliev, A. Korinevskaya, U. Sharma, H. Wang, On reproducing semi-dense depth map reconstruction using deep convolutional neural networks with perceptual loss, in: Proceedings of the 27th ACM International Conference on Multimedia, 2019, pp. 1080–1084.
[40] D. Maslov, I. Makarov, Online supervised attention-based recurrent depth estimation from monocular video, PeerJ Computer Science 6 (2020) e317.

[41] I. Makarov, V. Aliev, O. Gerasimova, P. Polyakov, Depth map interpolation using perceptual loss, in: Mixed and Augmented Reality (ISMAR-Adjunct), 2017 IEEE International Symposium on, IEEE, 2017, pp. 93–94.
[42] A. Korinevskaya, I. Makarov, Fast depth map super-resolution using deep neural network, in: 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), IEEE, 2018, pp. 117–122.
[43] I. Makarov, A. Korinevskaya, V. Aliev, Fast semi-dense depth map estimation, in: Proceedings of the 2018 ACM Workshop on Multimedia for Real Estate Tech, ACM, ACM, NY, USA, 2018, pp. 18–21.
[44] I. Makarov, A. Korinevskaya, V. Aliev, Sparse depth map interpolation using deep convolutional neural networks, in: 2018 41st IC on Telecommunications and Signal Processing (TSP), IEEE, IEEE, NY, USA, 2018, pp. 1–5.
[45] I. Makarov, A. Korinevskaya, V. Aliev, Super-resolution of interpolated downsampled semi-dense depth map, in: Proceedings of the 23rd International ACM Conference on 3D Web Technology, ACM, ACM, NY, USA, 2018, p. 27.
[46] D. Maslov, I. Makarov, Fast depth reconstruction using deep convolutional neural networks, in: International Work-Conference on Artificial Neural Networks, Springer, 2021, pp. 456–467.
[47] I. Makarov, I. Guschenko-Cheverda, Learning loss for active learning in depth reconstruction problem, in: Proceedings of CINTI'21, IEEE, 2021, pp. 1–6.
[48] K. Tikhomirova, I. Makarov, Community detection based on the nodes role in a network: The telegram platform case, in: International Conference on Analysis of Images, Social Networks and Texts, Springer, 2020, pp. 294–302.
[49] I. Makarov, A. Savchenko, A. Korovko, L. Sherstyuk, N. Severin, A. Mikheev, D. Babaev, Temporal graph network embedding with causal anonymous walks representations, arXiv preprint arXiv:2108.08754 (2021).
[50] P. Zolnikov, M. Zubov, N. Nikitinsky, I. Makarov, Efficient algorithms for constructing multiplex networks embedding, in: Proceedings of EEML conference, 2019, pp. 57–67.
[51] I. Makarov, D. Kiselev, N. Nikitinsky, L. Subelj, Survey on graph embeddings and their applications to machine learning problems on graphs, PeerJ Computer Science (2021).
[52] I. Makarov, M. Makarov, D. Kiselev, Fusion of text and graph information for machine learning problems on networks, PeerJ Computer Science 7 (2021).
[53] I. Makarov, K. Korovina, D. Kiselev, Jonnee: Joint network nodes and edges embedding, IEEE Access (2021) 1–14.
[54] I. Makarov, O. Bulanov, L. E. Zhukov, Co-author recommender system, in: International Conference on Network Analysis, Springer, 2016, pp. 251–257.
[55] M. K. Rustem, I. Makarov, L. E. Zhukov, Predicting psychology attributes of a social network user, in: Proceedings of the Fourth Workshop on Experimental Economics and Machine Learning (EEML'17), Dresden, Germany, September 17-18, 2017, CEUR WP, 2017, pp. 1–7.
[56] I. Makarov, O. Bulanov, O. Gerasimova, N. Meshcheryakova, I. Karpov, L. E. Zhukov, Scientific matchmaker: Collaborator recommender system, in: International Conference on Analysis of Images, Social Networks and Texts, Springer, 2017, pp. 404–410.
[57] I. Makarov, O. Gerasimova, P. Sulimov, L. E. Zhukov, Recommending co-authorship via

network embeddings and feature engineering: The case of national research university higher school of economics, in: Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries, ACM, 2018, pp. 365–366.
[58] I. Makarov, O. Gerasimova, P. Sulimov, K. Korovina, L. E. Zhukov, Joint node-edge network embedding for link prediction, in: International Conference on Analysis of Images, Social Networks and Texts, Springer, 2018, pp. 20–31.
[59] I. Makarov, O. Gerasimova, P. Sulimov, L. E. Zhukov, Co-authorship network embedding and recommending collaborators via network embedding, in: International Conference on Analysis of Images, Social Networks and Texts, Springer, 2018, pp. 32–38.
[60] I. Makarov, O. Gerasimova, P. Sulimov, L. E. Zhukov, Dual network embedding for representing research interests in the link prediction problem on co-authorship networks, PeerJ Computer Science 5 (2019) e172.
[61] I. Makarov, O. Gerasimova, Predicting collaborations in co-authorship network, in: 2019 14th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), IEEE, 2019, pp. 1–6.
[62] I. Makarov, O. Gerasimova, Link prediction regression for weighted co-authorship networks, in: International Work-Conference on Artificial Neural Networks, Springer, 2019, pp. 667–677.
[63] I. Makarov, A. Oborevich, Network embedding for cluster analysis, in: Proceedings of CINTI'21, IEEE, 2021, pp. 1–6.
[64] A. Broilovskiy, I. Makarov, Human action recognition for boxing training simulator, in: International Conference on Analysis of Images, Social Networks and Texts, Springer, 2020, pp. 331–343.