

ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ МАШИННОГО ЗРЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ ОДИНОЧНЫХ ПРЯМЫХ УДАРОВ В БОКСЕ

И.Я. Хасаншин, Д.С. Уткин, Д.Н. Дербин

Финансовый университет при Правительстве Российской Федерации, Москва, Россия

Аннотация

Целью работы было проведение экспериментальных исследований по распознаванию одиночных ударов в боксе на основе фреймворка MediaPipe, включающего в себя модели сверточных нейронных сетей BlazePose для определения позы человека.

Методы и организация исследования. Для распознавания движений и вида удара применялся фреймворк с открытым исходным кодом MediaPipe, в который встроена сверточная нейронная сеть для определения движений и позы человека BlazePose. Исследования проводились с участием 14 спортсменов с тренировочным опытом 2-3 года. В контрольной группе было 14 мужчин, возраст – 21 ± 3 года, вес – 70 ± 12 кг, рост – 175 ± 11 см. Были исследованы только прямые удары, которые наносились справа и слева со сменой стоек. Классификация ударов производилась на основе алгоритма k-ближайших соседей (англ. k-nearest neighbors algorithm, k-NN).

Результаты исследования. Исследования показали, что точность распознавания ударов передней левой рукой (0,86) и правой рукой (0,87) ниже, чем такие же показатели для перекрестных ударов – 0,95 и 0,96. Средняя точность по всем классам составила 0,93. Оценка точности полученных результатов классификации была произведена на основе F-метрики.

Заключение. Простота и эффективность методики позволяют легко внедрить ее в тренерскую практику, а также делают ее перспективной для дальнейших исследований других видов ударов.

Ключевые слова: компьютерное зрение, сверточная нейронная сеть, распознавание удара в боксе, BlazePose, MediaPipe.

APPLICATION OF MACHINE VISION TECHNOLOGIES FOR RECOGNITION OF SINGLE DIRECT PUNCHES IN BOXING

I.Ya. Khasanshin, e-mail: iykhasanshin@fa.ru, ORCID: 0000-0003-3809-8624

D.S. Utkin, e-mail: 190713@edu.fa.ru, ORCID: 0000-0002-1996-0234

D.N. Derbin, e-mail: 190763@edu.fa.ru, ORCID: 0000-0002-5700-6927

Financial University under the Government of the Russian Federation, Moscow, Russia

Abstract

The aim of the research was to conduct experimental studies on the recognition of single punches in boxing based on the MediaPipe framework, which includes models of convolutional neural networks BlazePose to determine a person's posture.

Methods and organization of the research. To recognize movements and the type of a punch, the open source MediaPipe framework was used, which has a built-in convolutional neural network BlazePose to determine the movements and a posture of a person. The studies were conducted with the participation of 14 athletes with training experience of 2-3 years. There were 14 men in the control group, age – 21 ± 3 years old, weight – 70 ± 12 kg, height – 175 ± 11 cm. Only straight punches that were applied to the right and left with a change of positions were investigated. The recognition of punches was carried out on the basis of the k-nearest neighbors algorithm (k-NN).

The research results. Studies have shown that the accuracy of recognition by the left hand (0.86) and the right hand (0.87) is lower than the same indicators for cross-counters – 0.95 and 0.96. The average accuracy for all classes was 0.93. The accuracy of the classification results obtained was evaluated on the basis of the F-score.

Conclusion. The simplicity and effectiveness of the technique allow implementing it into coaching practice, and make it promising for further research of other types of punches.

Keywords: computer vision, convolutional neural network, punch recognition in boxing, BlazePose, MediaPipe.

ВВЕДЕНИЕ

Компьютерные технологии начинают играть одну из ведущих ролей в подготовке спортсменов. Внедрение машинного зрения, скоростной видеосъемки, датчиков скорости, веса, силы и т.д. произвело настоящую революцию в анализе техники движений, совершенствовании сбора и обработки спортивных параметров.

Машинное зрение, компьютерный анализ видеоданных активно применяются для распознавания движений и анализа техники в различных видах спорта: фехтовании [12], теннисе [13], футболе [2], каратэ [8], боксе [10].

В статье [8] авторами Emad и др. ставилась задача исследования движений формального комплекса каратэ (ката) хэйаншодан на основе ИК-камеры (Kinect 2 for Windows) и препроцессинга данных при помощи модели F-DTW (англ. fast dynamic time warping – алгоритм динамической трансформации временной шкалы). Авторами [8] на основе анализа данных движений вырабатывались сообщения о том, насколько правильно выполнялись движения, каковы были ошибки, а также давались рекомендации, как можно улучшить движения и удары. Точность разработанной системы составила 91,07%. В работе [10] для классификации ударов использовались алгоритмы: метод опорных векторов (англ. SVM, support vector machine) и Random forest (с англ. – случайный лес). Экспериментальные результаты показали точность 97,3% для спортсменов, участвовавших в набо-

ре данных, и 96,2% для спортсменов контрольной группы.

МЕТОДЫ И ОРГАНИЗАЦИЯ ИССЛЕДОВАНИЯ

Для распознавания движений и вида удара применялся фреймворк с открытым исходным кодом MediaPipe [8], который представляет собой кроссплатформенное решение создания мультимодальных конвейеров машинного обучения для обработки таких данных, как аудио, потоковое видео. Платформа MediaPipe написана на языках C++, Java и Obj-C и состоит из различных API (англ. Application Programming Interface – программный интерфейс приложения) для вывода данных, расчета и построения графов. Решение MediaPipe для определения движений и позы человека базируется на модели BlazePose [1]. Модель BlazePose основана на архитектуре сверточной нейронной сети.

Blaze Pose состоит из двух моделей машинного обучения: детектора и оценщика. Детектор «вырезает» область с фигурой человека из входного изображения, в то время как оценщик принимает изображение обнаруженного человека с разрешением 256x256 в качестве входных данных и выводит 33 ключевые точки, топология которых показана на рисунке 1 [8].

Пример результирующего изображения с определенными ключевыми точками показан на рисунке 2.

Исследования проводились с участием 14 спортсменов с тренировочным

опытом 2-3 года. В контрольной группе было 14 мужчин, возраст – 21 ± 3 года, вес – 70 ± 12 кг, рост – 175 ± 11 см.

Эксперименты проводились только для прямых ударов, которые наносились справа и слева со сменой стоек. Каждый удар был нанесен спортсменами 150 раз. Таким образом, датасет для каждого спортсмена состоял из 600 изображений конечных фаз ударов каждого вида из левосторонней и правосторонней стоек, всего 8400 изображений. Так как боксер в стойке обычно держит переднюю руку полусогнутой, то в датасет также были включены изображения боксеров в стойке, чтобы модель не распознала эти положения тела как тот или иной удар.

Таким образом, классификация проводилась по следующим категориям:

- левый прямой удар передней рукой, то есть рукой той же стороны, что и нога, которая располагается впереди, в английской терминологии – jab;
- правый прямой удар задней рукой или

правый перекрестный удар, в английской терминологии – cross;

- правый прямой удар передней рукой;
- правый прямой удар задней рукой;
- категория «стойка».

Фотографии ударов были выполнены с разрешением 3024x4032. Распознавание ударов производилось на основе алгоритма k-ближайших соседей (англ. k-nearest neighbors algorithm, k-NN) [8].

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И ИХ ОБСУЖДЕНИЕ

Для использования в классификаторе k-NN изображения конечных фаз ударов были конвертированы при помощи ресурса Colab [7].

На рисунке 3 показаны боксеры с размеченными ключевыми точками.

Оценка точности полученных результатов классификации была произведена на основе F-метрики (англ. – F-score). Данная метрика была проведена при помощи библиотеки sklearn [9]. В многоклассовой классификации эти

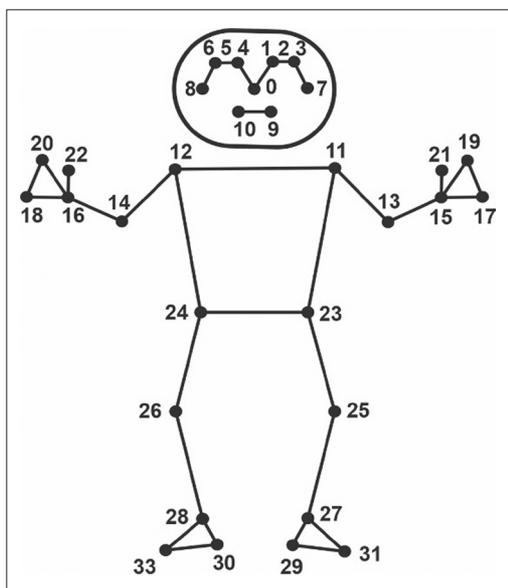


Рисунок 1 – Топология ключевых точек, определяемых моделью BlazePose
Figure 1 – Pose Landmark Model BlazePose



Рисунок 2 – Изображение человека с обозначенными ключевыми точками [8]
Figure 2 – An image of a person with pose landmarks [8]

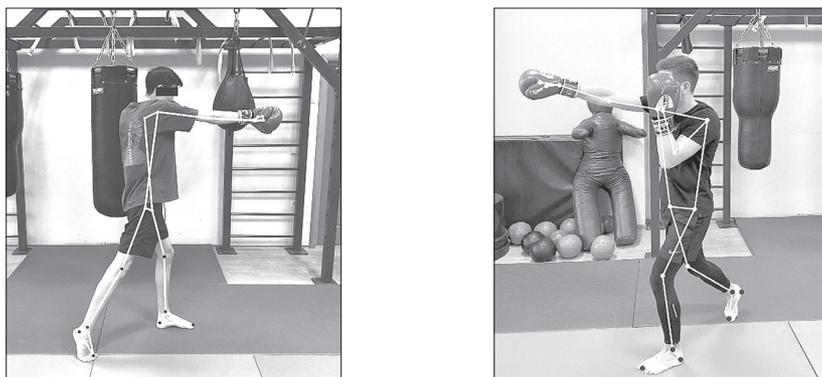


Рисунок 3 – Фотографии боксеров в конечной фазе ударов с обозначенными ключевыми точками
Figure 3 – Photos of boxers in the final phase of punches with pose landmarks

показатели вычисляются для каждого класса отдельно (за исключением точности) и F-метрика имеет вид F1-метрики:

$$F1 = (2 \cdot (PR \cdot RC)) / (PR + RC), \quad (1)$$

где PR – точность модели, находится по (2), RC – полнота или чувствительность модели (3).

Точность – это доля верных ответов модели:

$$PR = TP / (TP + FP), \quad (2)$$

где TP – количество случаев, верно отнесенных к данному классу, FP – количество случаев, когда случаи были отнесены к данному классу, но такими не являлись. Полнота определяет количество определенных истинно положительных случаев, к примеру, отнесенных к классу «удар левой передней рукой», среди всех меток класса, которые были определены как «удар левой передней рукой»:

$$RC = TP / (TP + FN), \quad (3)$$

где TP – количество случаев, верно отнесенных к данному классу, FN – количество случаев, когда случаи не были отнесены к данному классу, но такими являлись.

Результаты экспериментальных исследований показаны в таблице, где виды удара обозначены: левый прямой удар передней рукой – JBL, правый прямой удар передней рукой – JBR, левый прямой удар задней рукой – CBL, правый прямой удар задней рукой – CBR, класс «стойка» – None.

Точность распознавания ударов передней левой рукой и правой рукой ниже, чем такие же показатели для перекрестных ударов, так как модель в отдельных случаях все же «путала» эти удары с выставленной вперед рукой в стойке (класс «стойка» – None). С. Касири и др. [10], которые вели исследования, касающиеся распознавания ударов в боксе, на основе дальностной ИК-камеры и метода опорных векторов (англ. – Support Vector Machines), получили среднюю точность 0,938. Однако нужно отметить, что исследования проводились для всех основных видов ударов и применялась 3d-камера. В [10] также для ударов передней рукой точность распознавания была ниже, чем для ударов задней рукой. В работе

Таблица – Результаты распознавания ударов
Table – Results of punch recognition

Класс/Class	Точность /Precision	Полнота/Recall	F1-метрика/F1-score	Средняя точность/ Average Precision
JBL	0,86	0,88	0,87	0,93
JBR	0,95	0,96	0,95	
CBL	0,87	0,88	0,87	
CBR	0,96	0,93	0,94	
None	0,93	0,94	0,93	

Ф. Малавски, Б. Кволек [12] для распознавания ударов в фехтовании была получена точность до 0,98 для алгоритма динамической трансформации временной шкалы (англ. – *dynamic time warping*). В данной работе была использована стереокамера, а также акселерометры, которые крепились на запястьях спортсменов.

ЗАКЛЮЧЕНИЕ

Были проведены исследования по распознаванию одиночных ударов в боксе на основе машинного зрения, видеозахвата изображений ударов. Применялась модель сверточных нейронных сетей для

определения позы человека BlazePose, инкапсулированная во фреймворк MediaPipe. Классификация проводилась по пяти категориям – класс «стойка» (без ударов), удары передней и задней рукой из правосторонней и левосторонней стоек. Оценка точности по каждому классу проводилась на основе F-метрики, средняя точность по всем категориям – 0,93, что показывает высокую эффективность технологии.

Простота и эффективность методики позволяют легко внедрить ее в тренерскую практику, а также делают ее перспективной для дальнейших исследований других видов ударов.

ЛИТЕРАТУРА

1. Bazarevsky V., Kartynnik Y., Vakunov A., Raveendran K., Grundmann M. BlazeFace: Sub-millisecond Neural Face Detection on Mobile GPUs. doi = 10.48550/ARXIV.1907.05047.
2. Bhargava, Neha & Cuzzolin, Fabio. (2020). Challenges and Opportunities for Computer Vision in Real-life Soccer Analytics. Conference: AAAI 2020 - Workshop on AI in Team Sports (AITS 2020)At: New York, USA.
3. Campos, Guilherme O.; Zimek, Arthur; Sander, Jörg; Campello, Ricardo J. G. B.; Mícenková, Barbora; Schubert, Erich; Assent, Ira; Houle, Michael E. (2016). «On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study». *Data Mining and Knowledge Discovery*, 30 (4): 891–927. doi:10.1007/s10618-015-0444-8. ISSN 1384-5810. S2CID 1952214
4. D. Dinu, B. Millot, J. Slawinski, J. Louis. An Examination of the Biomechanics of the Cross, Hook and Uppercut between Two Elite Boxing Groups. *Proceedings, 2020; Volume 49*, pp. 61. <https://doi.org/10.3390/proceedings2020049061>.
5. D. K. Kimm, D. Thiel. Hand Speed Measurements in Boxing. *Procedia Engineering*, 2015; Volume 112, pp. 502-506. <https://doi.org/10.1016/j.proeng.2015.07.232>.
6. Emad, Bassel & Atef, Omar & Shams, Yehya & El-Kerdany, Ahmed & Shorim, Nada & Nabil, Ayman & Atia, Ayman. (2020). *iKarate: Karate Kata Guidance System*. *Procedia Computer Science*. 175. 149-156. 10.1016/j.procs.2020.07.024.
7. https://colab.research.google.com/drive/19txHpN8exWhstO6WVkfYVVC6uug_oVR
8. <https://mediapipe.dev/>
9. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html
10. Kasiri, Soudeh & Fookes, Clinton & Sridharan, Sridha & Morgan, Stuart. (2017). Fine-grained action recognition of boxing punches from depth imagery. *Computer Vision and Image Understanding*. 159. 10.1016/j.cviu.2017.04.007.
11. M. Zago, M. Codari, F. Marcello Iaia & C. Sforza Multi-segmental movements as a function of experience in karate, *Journal of Sports Sciences*, 2017; Volume 35, Issue 15, pp. 1515-1522, DOI: 10.1080/02640414.2016.1223332.
12. Malawski, Filip & Kwolek, Bogdan. (2016). Classification of basic footwork in fencing using accelerometer. 51-55. 10.1109/SPA.2016.7763586.
13. Messelodi, Stefano & Modena, Carla & Ropele, V. & Marcon, S. & Sgro, Michael. (2019). A Low-Cost Computer Vision System for Real-Time Tennis Analysis. 106-116. 10.1007/978-3-030-30642-7_10.
14. N. Haralabidis, D.J. Saxby, C. Pizzolato, L. Needham, D. Cazzola, C. Minahan. Fusing Accelerometry with Videography to Monitor the Effect of Fatigue on Punching Performance in Elite Boxers. *Sensors 2020; Volume 20*, pp. 5749. <https://doi.org/10.3390/s20205749>.

REFERENCES

1. Bazarevsky V., Kartynnik Y., Vakunov A., Raveendran K., Grundmann M. BlazeFace: Sub-millisecond Neural Face Detection on Mobile GPUs. doi = 10.48550/ARXIV.1907.05047.
2. Bhargava, Neha & Cuzzolin, Fabio. (2020). Challenges and Opportunities for Computer Vision in Real-life Soccer Analytics. Conference: AAAI 2020 - Workshop on AI in Team Sports (AITS 2020)At: New York, USA.
3. Campos, Guilherme O.; Zimek, Arthur; Sander, Jörg; Campello, Ricardo J. G. B.; Mícenková, Barbora; Schubert, Erich; Assent, Ira; Houle, Michael E. (2016). «On the evaluation of unsupervised

- outlier detection: measures, datasets, and an empirical study». *Data Mining and Knowledge Discovery*. 30 (4): 891–927. doi:10.1007/s10618-015-0444-8. ISSN 1384-5810. S2CID 1952214
4. D. Dinu, B. Millot, J. Slawinski, J. Louis. An Examination of the Biomechanics of the Cross, Hook and Uppercut between Two Elite Boxing Groups. *Proceedings*, 2020; Volume 49, pp. 61. <https://doi.org/10.3390/proceedings2020049061>.
 5. D. K. Kimm, D. Thiel. Hand Speed Measurements in Boxing. *Procedia Engineering*, 2015; Volume 112, pp. 502-506. <https://doi.org/10.1016/j.proeng.2015.07.232>.
 6. Emad, Bassel & Atef, Omar & Shams, Yehya & El-Kerdany, Ahmed & Shorim, Nada & Nabil, Ayman & Atia, Ayman. (2020). iKarate: Karate Kata Guidance System. *Procedia Computer Science*. 175. 149-156. 10.1016/j.procs.2020.07.024.
 7. https://colab.research.google.com/drive/19txHpN8exWhstO6WVkfYVVC6uug_oVR
 8. <https://mediapipe.dev/>
 9. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html
 10. Kasiri, Soudeh & Fookes, Clinton & Sridharan, Sridha & Morgan, Stuart. (2017). Fine-grained action recognition of boxing punches from depth imagery. *Computer Vision and Image Understanding*. 159. 10.1016/j.cviu.2017.04.007.
 11. M. Zago, M. Codari, F. Marcello laia & C. Sforza Multi-segmental movements as a function of experience in karate, *Journal of Sports Sciences*, 2017; Volume 35, Issue 15, pp. 1515-1522, DOI: 10.1080/02640414.2016.1223332.
 12. Malawski, Filip & Kwolke, Bogdan. (2016). Classification of basic footwork in fencing using accelerometer. 51-55. 10.1109/SPA.2016.7763586.
 13. Messelodi, Stefano & Modena, Carla & Ropele, V. & Marcon, S. & Sgro, Michael. (2019). A Low-Cost Computer Vision System for Real-Time Tennis Analysis. 106-116. 10.1007/978-3-030-30642-7_10.
 14. N. Haralabidis, D.J. Saxby, C. Pizzolato, L. Needham, D. Cazzola, C. Minahan. Fusing Accelerometry with Videography to Monitor the Effect of Fatigue on Punching Performance in Elite Boxers. *Sensors* 2020; Volume 20, pp. 5749. <https://doi.org/10.3390/s20205749>.

СВЕДЕНИЯ ОБ АВТОРАХ:

Хасаншин Ильшат Ядыкарович (Khasanshin Ilshat Yadikarovitch) – кандидат технических наук; Финансовый университет при Правительстве РФ; 125993 (ГСП-3), г. Москва, Ленинградский просп., 49; e-mail: iykhasanshin@fa.ru; ORCID: 0000-0003-3809-8624.

Уткин Даниил Сергеевич (Utkin Daniil Sergeevich) – научный сотрудник Лаборатории робототехники, интернета вещей и встраиваемых систем; Финансовый университет при Правительстве РФ; 125993 (ГСП-3), г. Москва, Ленинградский просп., 49; e-mail: 190713@edu.fa.ru; ORCID: 0000-0002-1996-0234.

Дербин Дмитрий Николаевич (Dmitriy Derbin Nikolaevich) – научный сотрудник Лаборатории робототехники, интернета вещей и встраиваемых систем; Финансовый университет при Правительстве РФ; 125993 (ГСП-3), г. Москва, Ленинградский просп., 49; e-mail: 190763@edu.fa.ru; ORCID: 0000-0002-5700-6927.

Поступила в редакцию 04 мая 2021 г.

Принята к публикации 18 мая 2022 г.

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Хасаншин, И.Я. Применение технологий машинного зрения для распознавания одиночных прямых ударов в боксе / И.Я. Хасаншин, Д.С. Уткин, Д.Н. Дербин // *Наука и спорт: современные тенденции*. – 2022. – Т. 10, № 2. – С. 43-48. DOI: 10.36028/2308-8826-2022-10-2-43-48

FOR CITATION

Khasanshin I.Ya., Utkin D.S., Derbin D.N. Application of machine vision technologies for recognition of single direct punches in boxing, *Science and sport: current trends*, 2022, vol. 10, no.2, pp. 43-48 (in Russ.) DOI: 10.36028/2308-8826-2022-10-2-43-48