Article

Smart Boxing Glove "RD *α*": IMU Combined with Force Sensor for Highly Accurate Technique and Target Recognition Using Machine Learning

Dea Cizmic [1,†], Dominik Hoelbling [1,*,†], René Baranyi [1,2], Roland Breiteneder [1] and Thomas Grechenig [1,2]

1   Research Group for Industrial Software (INSO), Vienna University of Technology, 1040 Vienna, Austria
2   Research Industrial Systems Engineering (RISE), 2320 Schwechat, Austria
*   Correspondence: dominik.hoelbling@inso.tuwien.ac.at
†   These authors contributed equally to this work.

Abstract: Emerging smart devices have gathered increasing popularity within the sports community, presenting a promising avenue for enhancing athletic performance. Among these, the Rise Dynamics Alpha (RD *α*) smart gloves exemplify a system designed to quantify boxing techniques. The objective of this study is to expand upon the existing RD *α* system by integrating machine-learning models for striking technique and target object classification, subsequently validating the outcomes through empirical analysis. For the implementation, a data-acquisition experiment is conducted based on which the most common supervised ML models are trained: decision tree, random forest, support vector machine, k-nearest neighbor, naive Bayes, perceptron, multi-layer perceptron, and logistic regression. Using model optimization and significance testing, the best-performing classifier, i.e., support vector classifier (SVC), is selected. For an independent evaluation, a final experiment is conducted with participants unknown to the developed models. The accuracy results of the data-acquisition group are 93.03% (striking technique) and 98.26% (target object) and for the independent evaluation group 89.55% (striking technique) and 75.97% (target object). Therefore, it is concluded that the system based on SVC is suitable for target object and technique classification.

Keywords: machine learning; time-series classification; smart gear; martial arts; combat sports; boxing

1. Introduction

Martial arts and combat sports, such as boxing, kickboxing, karate, and kung fu, seem to constantly increase in acceptance and popularity worldwide [1], not solely in professional sports but even more in modern pop culture (e.g., movies, magazines, posters) and as a useful tool to gain physical fitness [2], which is why they have become appealing to a broader variety of people. Although this trend, in general, can be viewed as a positive healthy development, it also involves certain challenges, in particular, injury risks [3] or overtraining [4]. This is because the majority of combat sports techniques, even in the absence of direct body contact, entail substantial impacts with the potential to cause harm to anatomical structures such as bones, joints, tendons, ligaments, and muscles, particularly when executed incorrectly or when performed under conditions of pronounced fatigue [5,6]. Nonetheless, mastering combat sports techniques demands a substantial dedication of time and guidance, typically spanning several years, potentially leading to considerable expenses [7]. One assisting service to conquer these challenges can be provided by supporting technologies, such as the new prototype of a smart boxing glove, called "RISE Dynamics Alpha" (RD *α*) [8,9].

These smart boxing gloves are in an advanced prototypical state and consist of a novel validated [10] force sensor (validation data not yet publicly available), as well as an inertial measurement unit (IMU) for measuring acceleration and angular velocity. Besides the RD

α gloves, there exist other commercial products for data quantification. The products of FightCamp [11], Hykso [12], Rooq [13], StrikeTec [14], and Move It [15] all consist of wearable IMU sensors combined with a software application that connects to sensors, collects data and displays different statistics, like punch speed and frequency or sometimes the techniques, of a workout session. However, these products solely measure the acceleration and angular change and calculate their derivatives. The RD α smart boxing gloves, on the other hand, allow the direct measurement of the punching force. With its corresponding mobile software application, it calculates further hit-specific data, such as impact, peak force, maximum speed, and maximum acceleration (both independent from direction) but also provides the full force, speed, and acceleration curve as "punch details" during every target contact.

Linking this information to a specific striking technique as well as to a specific target object should provide further analysis and quantification possibilities for athletes and instructors alike. For example, tailored exercises with corresponding, individualized body-strain plans, can be designed for technique improvement or injury prevention.

1.1. Automatic Human Body Movement Recognition

The recognition of human body movement and activity in various sport domains is an active research topic in the field of data science and machine learning. There have been several research projects where, based on movement data collected from sensors, machine-learning (ML) approaches have been applied to solve the recognition of human body movement. For example, Perri et al. [16] focused in their study on the tennis-specific stroke and movement classification using machine learning based on data from a wearable sensor containing a tri-axial accelerometer, gyroscope, and magnetometer. Similarly, Kautz et al. [17] evaluated the application of deep neural networks for the recognition of volleyball-specific movement data collected by a tri-axial wearable sensor. In addition, Cust et al. [18] provides a systematic review of 52 studies on the topic of sport-specific movement recognition using machine and deep learning.

Among numerous other sporting disciplines, research projects have also been conducted in the realm of martial arts and combat sports, specifically examining and implementing machine-learning approaches. Therefore, the researchers aimed at solving certain classification tasks, like motion classification, with the goal to implement and deploy models that are capable of correctly recognizing e.g., a certain striking technique based on collected sensor data. These systems usually either make use of a wearable IMU sensor [19–21], depth images [22], or a 3D motion-capturing system based on video data combined with IMU sensors [23]. In addition, Lapkova et al. [24] used a stationary strain gauge sensor to measure the force and use the data as input for striking and kicking technique recognition.

As with the referenced projects, the RD α gloves also incorporate an IMU sensor but additionally include a force measurement unit. Based on the biomechanical characteristics of a striking technique, the presumption was established that the data of the sensors can be used to train an ML model for the classification of striking techniques as well as target objects. This assumes that it is possible to identify patterns in the sensor data that are common among the samples for each striking technique, based on the physical characteristics like limb trajectory and acceleration.

1.2. Hand-Striking Techniques

For this research, only striking techniques that can be executed wearing boxing gloves were considered. This is due to the composition of the RD α system since the exact hand and finger positions cannot be identified like it is necessary for martial arts and combat sports such as kung fu [25] or karate [26] to differentiate between techniques. Thus, for the striking technique recognition, techniques from Boxing and Kickboxing were considered, which can be described as follows according to their rulebooks [27,28]:

- Straight (Jab/Punch) [29] (see Figure 1a,b): Executed with the leading (Jab) or rear hand (Punch/Cross) in a straight line from the guard position towards the target object.
- Hook [29] (see Figure 1c): Executed either with the lead or rear hand from the guard position by extending and then rotating the arm. In the end position, the upper and lower arm build approximately a 90-degree angle and are parallel to the ground.
- Uppercut [30] (see Figure 1d): Executed either with the lead or rear hand from the guard position by slightly lowering down and rotating the hand and subsequently moving it upwards with the help of the upper body.
- Backfist [31] (see Figure 1e): Most often used in pointfighting, which is a sub-discipline of kickboxing. The backfist is executed with the leading hand by extending the arm towards a target but with the intention to hit the target with the backside of the fist [31].
- Ridge hand [31] (see Figure 1f): The hand is rotated and moved in an arc trajectory with the intention to hit the target object with the back of the hand.
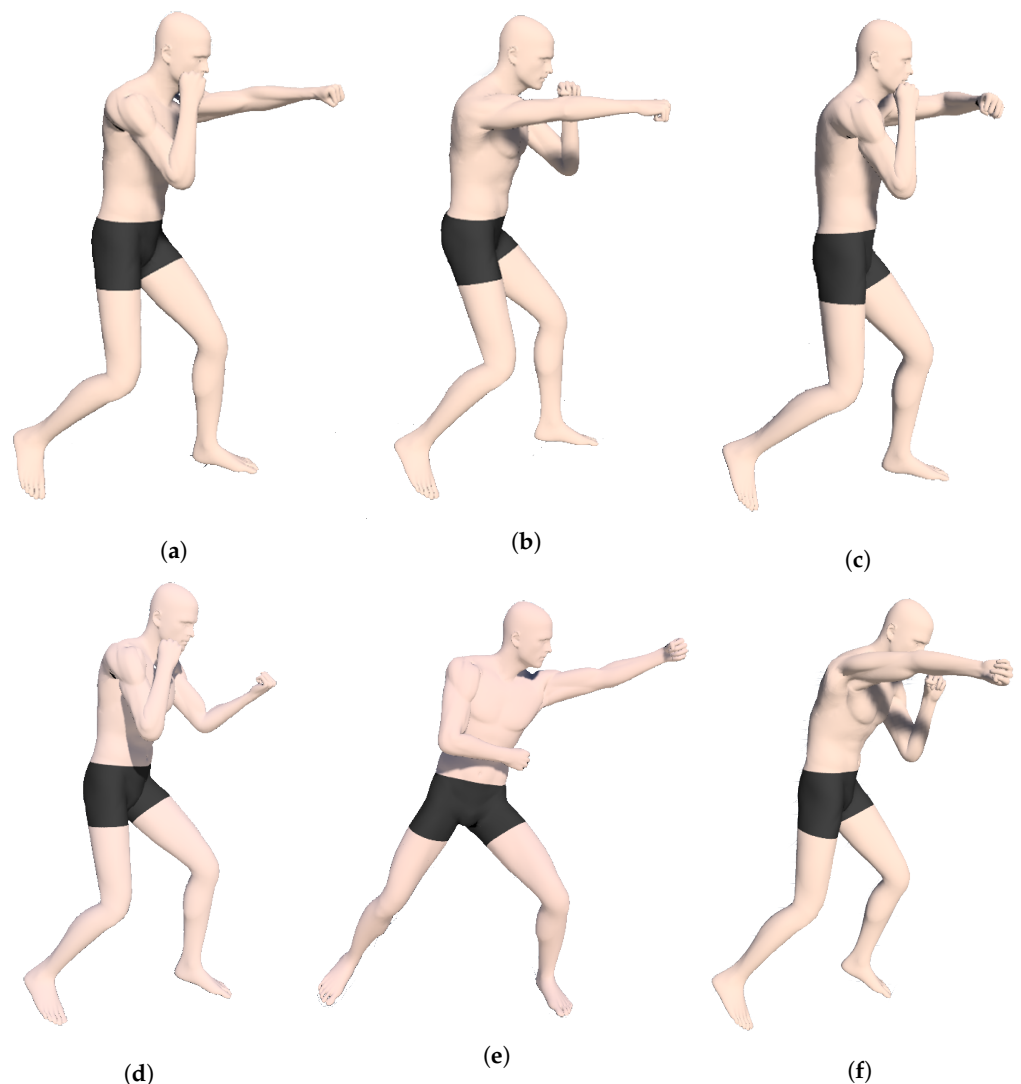

Figure 1. Illustration of the selected striking techniques. Images created using Magic Poser [32]. (**a**) Straight/Jab. (**b**) Punch/Cross. (**c**) Hook. (**d**) Uppercut. (**e**) Backfist. (**f**) Ridge hand.

All striking techniques can generally be split into three phases, the segment acceleration, the target contact, and the restoring phase [33]. For this paper, only the two former phases are relevant. The applied impact of a striking technique is the product of the effec-

tive mass of/behind the strike (*m*), the (negative) acceleration (*a*) during target contact, and the trajectory of the involved segments, during the target contact, see Equation (1).

$$p = \int_{start}^{end} F = \int_{start}^{end} m * a \tag{1}$$

Naturally, the negative acceleration of the glove is highly affected by the target (e.g., a comparison between the concrete wall and a soft punching bag). Furthermore, for each striking technique, the trajectory of the boxing glove differs, also resulting in different force profiles. These characteristics render the velocity before the target contact, the trajectory (acceleration and relative angle in all three directions), as well as the impact and potential peak force as crucial factors to determine the technique and target. Figure 2 shows the line chart of each feature over time of a jab. For comparison, five jab instances are plotted at once to depict the similarities between the curves. However, due to the positioning of the IMUs directly within the gloves, a solely rule-based differentiation between the techniques is not possible, as the rest-body movement is not known. This is why an ML-based approach might provide a solution to this problem.
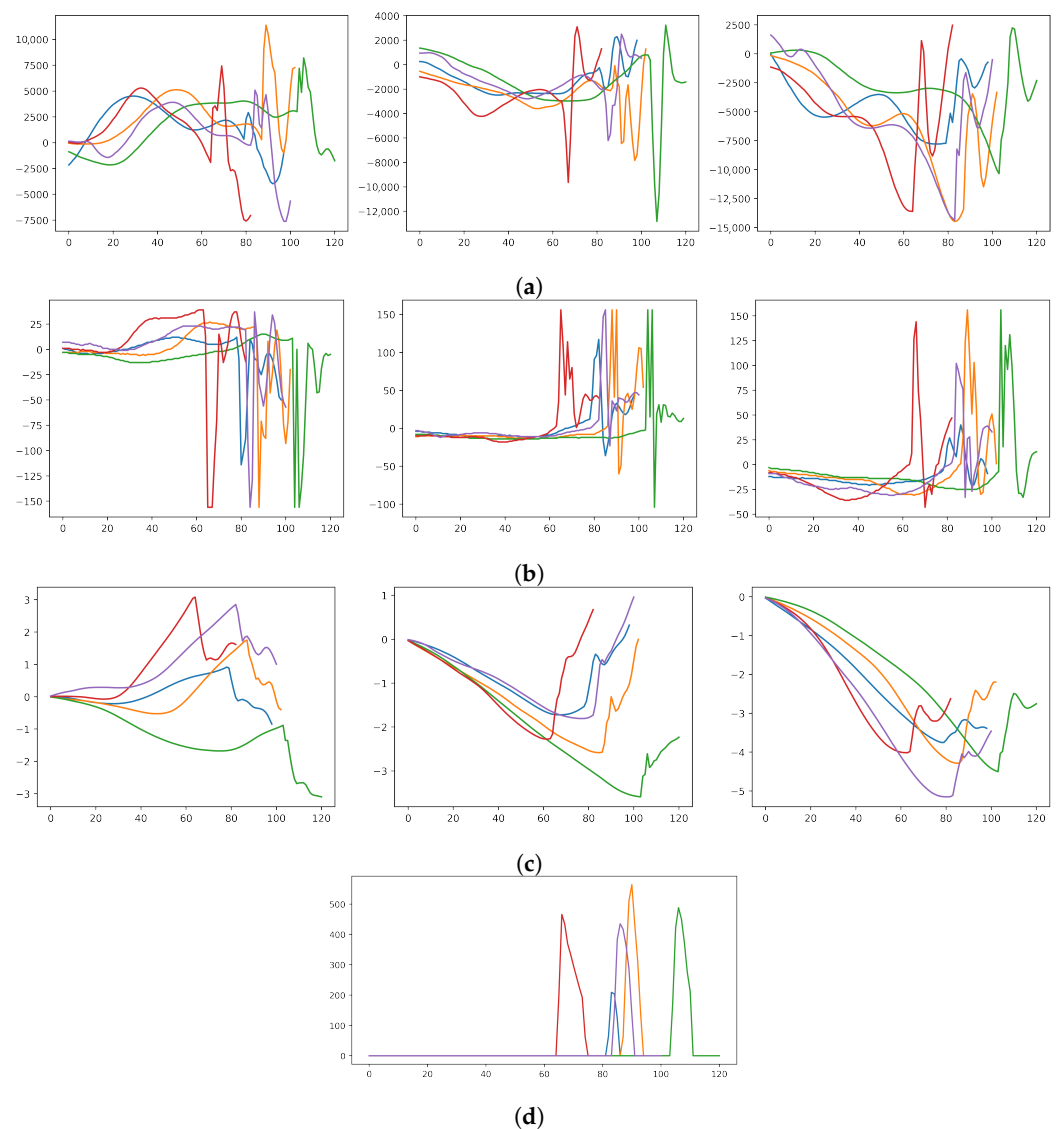

Figure 2. Illustration of five measurements of a jab technique depicting the sensor values over time. (**a**) Angular velocity. (**b**) Acceleration. (**c**) Velocity. (**d**) Force.

1.3. Aims and Novelty

This study outlines the development process of an ML-based extension for the RD α system. The primary objective was to create an ML-based system capable of accurately recognizing striking techniques and identifying the target object using data obtained from sensors embedded in the smart boxing gloves. To achieve this, ML models were implemented for the classification of various striking techniques, described in Section 1.2 and the differentiation between valid targets (e.g., punching bag, punch bag, gloves) and invalid targets (e.g., concrete wall) on which the striking techniques are executed.

To implement the classification models, it was necessary to identify the most suitable supervised machine-learning algorithms for the sensor data derived from the aforementioned sensors. Moreover, an assessment was carried out to determine which features and feature representations of the multi-variate time-series sensor data could be utilized for training and testing the supervised machine-learning models. Additionally, datasets were constructed for training and testing the implemented models, achieved through the conduction of two experiments involving mutually exclusive participant groups. Ultimately, the objective was to evaluate whether the developed classification approach, in conjunction with the data from the RD α system, could achieve a predictive accuracy of 85%. This value was chosen as a baseline since related work [19,20,23,24] showed the feasibility of achieving this accuracy rate. It was estimated that reaching this value would verify the proof of concept of the ML-based classification system and that the implementation of even higher accuracy can be achieved afterward through adaptations.

2. Materials and Methods

To create a structured and reproducible process for the implementation of the desired classification system, a customized ML workflow was designed according to common models and practices from the data science domain [34–36]. The workflow consists of the following steps:

1. Data acquisition
2. Data understanding and processing
3. Model training, testing, and optimization
4. Model comparison and selection
5. Model integration and final evaluation experiment

2.1. Data-Acquisition Experiment

Due to the novelty of the RD α smart gloves, it was necessary to create a dataset that would serve as a basis for model implementation. Thus, to construct a labeled dataset, a data-acquisition experiment was conducted. Members from local kickboxing gyms and the national team (currently active athletes of different sex, age, height, and experience level) were invited to repeatedly perform either individual striking techniques or a set of technique sequences, each consisting of four individual techniques executed immediately one after the other, with approximately 5 s break between techniques and 30 s breaks between sets. This approach was selected to include as much variation as possible in the data since the execution of a strike can vary when it is executed in sequence compared to when it is executed in isolation. Also, this best represents strike instances produced in a real-world scenario, e.g., a training session or combat since both options (individual and sequence) can occur. For the same reason, distance and stances or exact execution patterns were not predetermined by the researchers. The previously described striking techniques were all executed on four different target types (with approximately 5 min break to subjective physical recovery, between the executions on each target): punching bag (commercial standing bag), punch pads (held by an experienced coach with the instructions to "hold for but not hit against" each technique), gloves (used as punch pads) and a concrete wall which were further divided into two groups (valid and invalid target) for the classification. The valid targets are therefore selected to represent the targets most commonly used in a conventional training session. During the data-acquisition experiment,

each of the strike samples was labeled unanimously by two domain experts (>10 years of coaching in international events and minimum second level of national trainer education) with the respective label for the striking technique and the target object. Labeling is a crucial part of data acquisition since it allows the application of supervised machine learning which is more reliable and suitable for classification tasks than unsupervised machine learning that does not require labeled data. It is noted that due to the predetermined order of techniques, they predominantly represented an additional measure of data quality. The data-acquisition experiments' (n = 13) inclusion criteria allowed for healthy participants who at least practiced kickboxing or boxing for at least three months (evaluated through a questionnaire). Some athletes also had experience in other combat sports disciplines. The characteristics of the participants are shown in Table 1. The data of the data-acquisition experiment was used as input for the classification algorithms to train, optimize and test the classification models. Figure 3 shows the setup of the experiment depicting one participant executing the jab technique and the respective components.

Figure 3. System setup with the punching bag target. The RD α gloves send the sensor data via Bluetooth to the mobile app. The data are subsequently processed via Python for model implementation.

Table 1. Participant (n = 13) statistics of the data-acquisition experiment (Anthropometric measurements according to ISAK guidelines).

Attribute	Mean	Standard Deviation	Min.	Max.
Height [cm]	175.38	8.18	164	191
Arm length [cm]	64.31	4.1	59	70
Experience [years]	7.3	6.71	2	25

2.2. Data Understanding and Processing

After the data-acquisition experiment, the data were composed into a dataset that was then analyzed using statistical metrics (mean, standard deviation, minimum, maximum) and visualizations like box plots and histograms. Based on these findings, the data were scaled using the standardization method, and data-cleaning was performed by eliminating outliers regarding the feature duration, distance, and force measurement. Subsequently, a feature set was derived using the following statistical metrics: minimum, maximum,

arithmetic mean, standard deviation, skew, and kurtosis per each 3D axis of each sensor measurement and additionally the pairwise Pearson correlation coefficient between the x, y, and z-axis.

2.3. Model Training, Evaluation, and Optimization

Using the derived feature set, a baseline implementation of the following classifiers was implemented using Python 3.9 [37] and scikit-learn 1.0.2 [38] with its default configurations: decision tree (DT), random forest (RF), support vector machine (SVC), k-nearest neighbor (kNN), naive Bayes (NB), perceptron, multi-layer perceptron (MLP) and logistic regression (LR). For the optimization phase, the four best-performing classifiers were selected to be optimized using hyper-parameter tuning and grid search (using GridSearchCV provided by scikit-learn). Therefore, the accuracy (number of correctly classified instances compared to the total number of instances) was used as the main metric for assessing the generalization performance of the models along with the F1-score (harmonic mean between precision and recall). To prevent data leaking from the training into the test dataset, the three-way holdout method [39] was used to establish three different datasets: one for training, one for optimizing the models, and one for assessing the final predictive performance of the models [40].

2.4. Model Comparison and Selection

To identify the significance of the difference between the implemented models, a significance test was performed. For this, a paired t-Test (parametric) combined with 5×2 cross-validation [41,42] was conducted to assess the significance of the difference between all the established models. Therefore, the significance level α was set to 0.05 so that a p-value over 0.05 means no significant difference, indicating that the results of the two tested models are samples drawn from the same distribution. This means that the difference between the accuracy results is not due to an overall better-performing model but may be due to statistical anomalies in the data. Based on the results of the significance test as well as on the results from the test set, the best-performing models for the striking technique and target object classification were selected.

2.5. Model Integration and Final Evaluation Experiment

To integrate the final models into the RD α system and make a final evaluation, a classification API was implemented using Python and Flask [43]. Finally, the evaluation experiment was performed to evaluate the real-world predictive performance of the selected classification models based on unseen data from new participants during another experiment.

The evaluation experiment (n = 8) was performed in the same setup with the difference that the collected data were immediately classified using the classification API. Therefore, the ground truth was recorded, and the predictive accuracy of the classification was calculated at the end of the experiment. The participants for the evaluation experiment were specially selected according to certain criteria in Table 2 regarding sex, height, and experience level to cover a broad range of the characteristics within the general population. Furthermore, the two groups (data-acquisition group; independent evaluation group) were mutually exclusive. This was carried out to evaluate the generalization performance of the classification models based on unseen data from participants unknown to the models.

Table 2. Thresholds between participants of the experiment group to provide high diversity (anthropometric measurements according to ISAK guidelines). The eight chosen representative participants were above or below these guideline values (e.g., woman/tall/experienced representative = female, >170 cm, >4 years of experience; man/short/inexperienced representative = male, <181 cm, <3 years of experience).

Sex	Height u.t.	Height l.t.	Experience u.t.	Experience l.t.
Female	170 cm	165 cm	4 years	3 years
Male	186 cm	181 cm	4 years	3 years

(l.t. = lower threshold; u.t. = upper threshold).

3. Results

The results of this research can be divided into the products of the different ML workflow steps described in Section 2:

- Final content of the data-acquisition dataset for model training, optimization, and testing
- First baseline evaluation results
- Model optimization and testing results
- Significance test results and model selection
- Final evaluation experiment results

3.1. Dataset

After cleaning and pre-processing the data of the data-acquisition experiment, a total of 3453 strike samples were composed into a dataset with 66 features and additional labels for the striking technique as well as the target object. The following 64 features were derived from the IMU sensor measurements of the RD α gloves:

- Gyroscope: minimum, maximum, mean, standard deviation, kurtosis, and skew for each 3D axis and additionally the pairwise Pearson correlation coefficient between the axes.
- Acceleration: minimum, maximum, mean, standard deviation, kurtosis, and skew for each 3D axis and additionally the pairwise Pearson correlation coefficient between the axes.
- Velocity: minimum, maximum, mean, standard deviation, kurtosis, and skew for each 3D axis.
- Force: maximum and average force
- Distance: maximum distance covered by a striking instance
- Duration: total duration of a striking instance

Additionally to these features, a feature denoting the stance (southpaw/orthodox) of an athlete as well as one denoting the hand (left or right) were added as features totaling a feature set of 66 features. As for the independent evaluation group, a total of 1951 strike samples were produced including the same 66 features and two labels.

3.2. Baseline Evaluation

The results of the baseline evaluation using the default classifier configuration of scikit-learn 1.0.2 are shown in Table 3 for the striking technique classification and in Table 4 for the target object classification. This initial evaluation was made on the training and optimization portion of the dataset (using 10-fold cross-validation) from the data-acquisition experiment and served the purpose of identifying the performance differences between the classifiers as well as assessing the suitability of the designed feature set.

Table 3. Striking technique baseline classification results computed using 10-fold cross-validation on the training and optimization dataset.

Classifier	Accuracy	*F*1-Score (Weighted)
DT	77.53 (±2)%	77.50 (±2)%
RF	90.93 (±2)%	90.88 (±2)%
NB (Gaussian)	73.72 (±3)%	72.99 (±3)%
kNN	90.73 (±2)%	90.67 (±2)%
MLP	90.88 (±2)%	90.84 (±2)%
SVC	91.60 (±1)%	91.57 (±1)%
Perceptron	76.49 (±3)%	76.29 (±3)%
LR	82.28 (±3)%	82.17 (±3)%

(DT = Decision tree; NB = Naive Bayes; RF = Random forest; SVC = Support vector classifier; kNN = k-nearest neighbor; MLP = Multi-layer perceptron; LR = Logistic Regression).

Table 4. Target object baseline classification results computed using 10-fold cross-validation on the training and optimization dataset.

Classifier	Accuracy	*F*1-Score (Weighted)
DT	90.33 (±1)%	90.46 (±1)%
RF	93.99 (±1)%	92.45 (±1)%
NB (Gaussian)	84.79 (±2)%	86.94 (±2)%
kNN	96.05 (±1)%	95.69 (±1)%
MLP	96.78 (±1)%	96.67 (±1)%
SVC	95.69 (±1)%	95.01 (±1)%
Perceptron	91.42 (±1)%	91.19 (±1)%
LR	93.99 (±1)%	93.42 (±1)%

(DT = Decision tree; NB = Naive Bayes; RF = Random forest; SVC = Support vector classifier; kNN = k-nearest neighbor; MLP = Multi-layer perceptron; LR = Logistic Regression).

3.3. Model Optimization and Testing

Derived from the baseline results, the following four classifiers with the best-performing models (of both classification tasks) were selected for the optimization phase: random forest (RF), k-nearest-neighbor (kNN), support vector classifier (SVC) and multi-layer perceptron (MLP). For the optimization phase, the scikit-learn implementation of grid search [44] was utilized for each classifier with a predefined grid to establish the best-performing hyper-parameter configuration. For model comparison and final selection, 10-fold cross-validation was applied to the test portion of the dataset from the data-acquisition experiment to establish comparable accuracy results for the baseline and optimized configuration of each of the four classifiers. The results are shown in Table 5 for the striking technique and in Table 6 for the target object classification.

Table 5. Predictive performance for the striking technique classification computed using 10-fold cross-validation on the test dataset.

Classifier	Baseline Accuracy Result	Optimized Accuracy Result
RF	91.76%	92.08%
kNN	90.49%	92.87%
MLP	92.71%	92.71%
SVC	92.71%	93.03%

(RF = Random forest; SVC = Support vector classifier; kNN = k-nearest neighbor; MLP = Multi-layer perceptron).

Table 6. Predictive performance for the target object classification computed using 10-fold cross-validation on the test dataset.

Classifier	Baseline Accuracy Result	Optimized Accuracy Result
RF	94.36%	94.50%
kNN	95.37%	96.96%
MLP	97.97%	97.97%
SVC	95.80%	98.26%

(RF = Random forest; SVC = Support vector classifier; kNN = k-nearest neighbor; MLP = Multi-layer perceptron).

3.4. Model Comparison and Selection

As mentioned before, a significance test was performed to determine which of the models performs best. The results of the significance tests are shown in Table 7 for the striking technique and Table 8 for the target object classification. The results for the striking technique classification models show that there is only a significant difference between the models of the RF and SVC as well as the SVC and MLP models with a significance level of $\alpha = 0.05$. Regarding the results of the significance test for the target object classification models, slightly more models seem to be significantly different with the significance level of $\alpha = 0.05$. In particular, the following models differ: RF and SVC, RF and kNN, and RF and MLP. Relating these results to the results of the test set, it is evident that those models that have a small difference in accuracy also seem not to be significantly different. Combining the insights of all results, the conclusion is that the kNN, MLP, and SVC models all produce similarly good results. Thus, to make a final decision for the final models, the best accuracy result for the striking technique and strike target object was taken into account. Hence, the SVC model was selected for both classification tasks as the final model.

Table 7. Significance test results of the model comparison for the striking technique classification.

Classifier 1	Classifier 2	*p*-Value	*t*-Statistics
RF	SVC	0.028	−3.061
RF	kNN	0.156	−1.668
RF	MLP	0.111	−1.932
kNN	SVC	0.922	0.103
kNN	MLP	0.209	1.442
SVC	MLP	0.031	2.961

(RF = Random forest; SVC = Support vector classifier; kNN = k-nearest neighbor; MLP = Multi-layer perceptron).

Table 8. Significance test results of the model comparison for the target object classification.

| Classifier 1	Classifier 2	*p*-Value	*t*-Statistics
RF	SVC	0.002	−5.700
RF	kNN	0.008	−4.260
RF	MLP	0.001	−6.462
kNN	SVC	0.105	−1.977
kNN	MLP	0.486	−0.751
SVC	MLP	0.143	1.735

(RF = Random forest; SVC = Support vector classifier; kNN = k-nearest neighbor; MLP = Multi-layer perceptron).

3.5. Final System Evaluation

For the dataset of the evaluation experiment with 1951 punch samples of eight athletes, the results were as follows: accuracy of 89.55% for the striking technique and 75.97% for the binary target object classification. Besides the overall accuracy, it was also evaluated which striking techniques were recognized most precisely. For this, the confusion matrix (see Figure 4) of the striking techniques was computed together with the classification report per striking technique (Table 9). Therefore, the confusion matrix can be viewed in addition to the accuracy results to assess the distribution of the true positive and false positive instances indicating which techniques were most often confused by the model.

Table 9. Classification report including precision, recall, *F*1-score and support (number of samples) per class for the striking technique classification results.

Class	Precision	Recall	*F*1-Score	Support
Straight	0.98	0.92	0.95	776
Hook	0.81	0.83	0.82	434
Uppercut	0.85	0.95	0.90	475
Backfist	0.85	0.95	0.89	139
Ridge Hand	0.96	0.74	0.84	127
		Overall		
macro avg	0.89	0.88	0.88	1951
weighted avg	0.90	0.90	0.90	1951
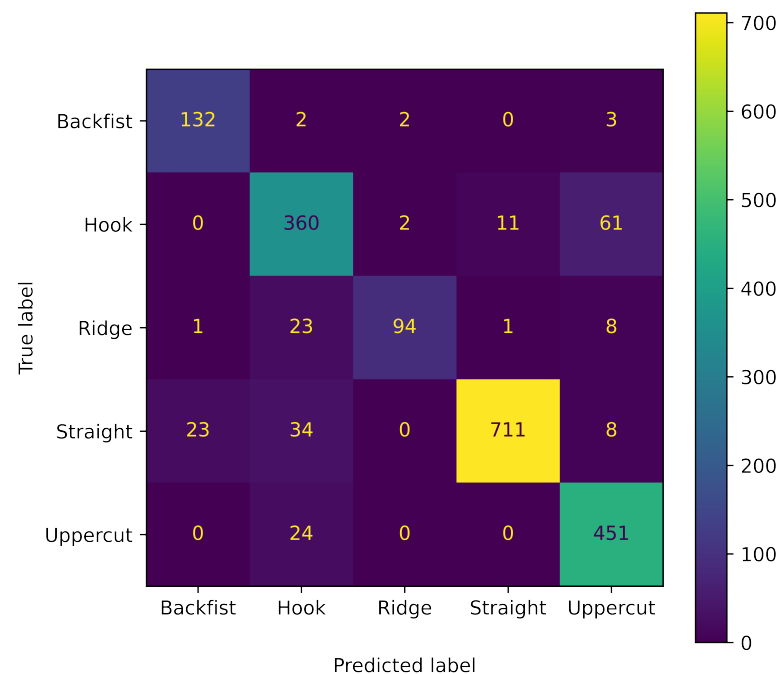
Figure 4. Confusion matrix for the striking technique classification evaluation showing the number of true positive and false positive instances.

In addition, it was assessed if there was a difference between the individual participant groups by computing the accuracy, averaged over the participants for each group (Figure 5).
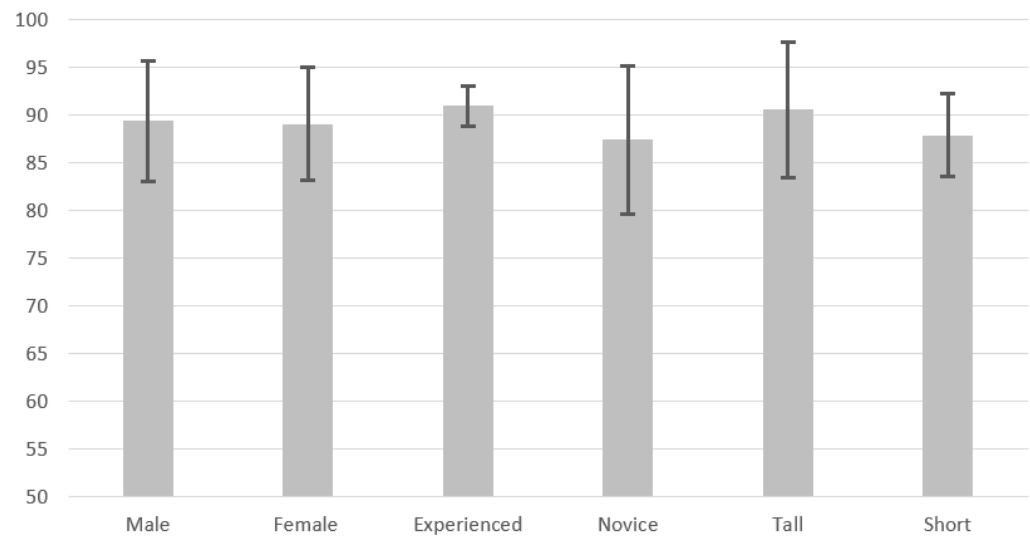
Figure 5. Average accuracy (in percent) including standard deviation per participant group (Experienced: >5 years of training and competition athlete; Novice <4 years of training and no competition experience).

4. Discussion

The results confirm that it is possible to implement a classification model for the striking technique classification based on IMU and force measurement sensors.

4.1. Classification of Sticking Technique and Target Objects

The implemented models show very high predictive performance results of over 93% for athletes known to the model and a robust result of over 89% for athletes unknown to the model. Regarding the target object classification, a similarly good result was produced for the athletes known to the model with over 98%. However, when tested with data of athletes unknown to it, the model only provided an accuracy of 75.97%, indicating that this model is slightly less resilient to data variation, although this issue can be corrected by e.g., including more diverse training data or by collecting more data from additional athletes. This is because better results are achieved when the training data show higher similarity to the testing data so including as much diverse data as possible in the training set increases the chances of including data similar to the real-world data.

4.2. Comparison of Technique Classification

The classification outcomes of the striking techniques in this study are comparable to other projects, such as Worsey et al. [19] (overall mean accuracy of 0.90 ± 0.12 and 0.87 ± 0.09 for two different sensor placements), Soekarjo et al. [23] (approximately 86% accuracy), Hanada et al. [45] (91.1% accuracy for person-independent evaluation), and Khasanshin [21] (ranging from $87.2 \pm 5.4\%$ to $95.33 \pm 2.51\%$).

Importance of Divers Participant Groups

However, the achieved accuracy in this study is slightly lower than the results reported by Wagner et al. [20] (98.68%). It is important to note that a direct comparison is not feasible since the researchers only classified three types of striking technique and both Wagner et al. [20] as well as Worsey et al. [19] did not evaluate the model using participants and data unknown to the model. It must be considered that using the same group of athletes on which the models were trained may introduce an optimistic bias and affect the model's performance on other athletes. Developing ML models requires establishing robustness and ensuring good generalization, which means accurate classifications for unseen data. Therefore, to realistically estimate the model's performance, evaluating it with new data

becomes crucial. In this research, deliberate efforts were made to introduce diversity within the training group and maximize it within the testing group, enhancing the models' generalizability, while slightly lowering the accuracy measures. Furthermore, the study explored the classification of two techniques, backfist and ridge hand, not previously included in related research, which increased the models' susceptibility to errors due to the added data variation that needed to be distinguished. Despite these additional challenges, the implemented models demonstrated good generalization performance compared to the related projects.

4.3. Comparison of Target Object Classification

In terms of target object classification, direct comparisons are not feasible since there are no similar experiments where the specific goal was to classify the target object in combat sports. Another crucial distinction between the mentioned projects and the RD α gloves described in this study is the availability of force measurement data. This not only offers an additional key feature but also allows for a more precise striking technique segmentation, as the exact point in time of contact with the target object is known.

4.4. Limitations and Outlook

For the herein described experiments, the prototype version 4.0.1 of the RD α smart gloves was used which included other IMU sensors than the current prototype and featured a higher degree of hardness which may have had an impact on the execution of the striking techniques, especially regarding the different target types. This is because due to their hardness, they reduce the level of wearing comfort which may have impacted the intensity of some striking techniques. Thus, producing data with the improved prototype (with more accurate IMU sensors and lower degree of hardness) and re-training the models may lead to even better accuracy results. Furthermore, due to COVID-19 regulations during the experiment, it was not feasible to recruit more participants at the time. Collecting more data, in general, should be performed as future work to improve the generalization performance of the models. In addition, it might also be possible to use the sensor as a serious gaming input device to capture additional movements for rehabilitation aspects. An already established serious game for hand (wrist and finger) movement, gesture, and touch exercises [46,47] could be enhanced with the usage of the gloves as well. The ML would need additional training to capture and interpret possible movements, but it should be feasible, at least for this serious game's hand movement aspects.

5. Conclusions

Overall, this research demonstrates the suitability of ML models as effective tools for technique and target object classifications, especially when combined with other measurement units like force sensors. The findings provide valuable insights that can serve as a foundation for addressing other classification challenges, such as assessing the quality of striking techniques. However, it is essential to acknowledge that achieving more robust results in quality classification would require a larger and more diverse data set with an appropriate preliminary quality assessment. The successful implementation of mutually exclusive data-acquisition and evaluation groups, deliberate diversity introduction, and exploration of new techniques showcases the effort made to enhance the generalizability and accuracy of the models.

Author Contributions: Conceptualization, D.C. and D.H.; Investigation, D.C.; Methodology, D.H.; Project administration, D.H. and T.G.; Resources, T.G.; Software, D.C.; Supervision, D.H., R.B. (René Baranyi), R.B. (Roland Breiteneder) and T.G.; Validation, D.C.; Writing—original draft, D.C. and D.H.; Writing—review and editing, D.H., R.B. (René Baranyi), R.B. (Roland Breiteneder) and T.G. All authors have read and agreed to the published version of the manuscript.

Abbreviations

The following abbreviations are used in this manuscript:

API	Application programming interface
DT	Decision tree
IMU	Inertial measurement unit
kNN	k-nearest neighbor
LR	Logistic Regression
ML	Machine Learning
MLP	Multi-layer perceptron
NB	Naive Bayes
RD α	Rise Dynamics Alpha
RF	Random forest
SVC	Support vector classifier
SVM	Support vector machine
ISAK	Society for the Advancement of Kinanthropometry

References

1. Green, T.; Svinth, J. Martial Arts of the World: An Encyclopedia of History and Innovation [2 volumes]: An Encyclopedia of History and Innovation. In *Martial Arts of the World: An Encyclopedia of History and Innovation*; ABC-CLIO: Santa Barbara, CA, USA, 2010.
2. Klein, C. Martial arts and the globalization of US and Asian film industries. *Comp. Am. Stud. Int. J.* **2004**, *2*, 360–384. [CrossRef]
3. Noble, C. Hand injuries in boxing. *Am. J. Sport. Med.* **1987**, *15*, 342–346. [CrossRef]
4. Nemček, D.; Dudíková, M. Self-Perceived Fatigue Symptoms After Different Physical Loads in Young Boxers. *Acta Fac. Educ. Phys. Univ. Comen.* **2022**, *62*, 123–133. [CrossRef]
5. Zetaruk, M.N.; Violan, M.A.; Zurakowski, D.; Micheli, L.J. Injuries in martial arts: A comparison of five styles. *Br. J. Sport. Med.* **2005**, *39*, 29–33. [CrossRef] [PubMed]
6. Pieter, W. Martial arts injuries. *Epidemiol. Pediatr. Sport. Inj.* **2005**, *48*, 59–73.
7. Biernat, E.; Krzepota, J.; Sadowska, D. Martial arts as a form of undertaking physical activity in leisure time analysis of factors determining participation of poles. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1989. [CrossRef]
8. Baldinger, A.; Ferner, T.; Hölbling, D.; Wohlkinger, W.; Zillich, M. Device for Detecting the Impact Quality in Contact Sports. Patent WO2020041806A1, 30 August 2018.
9. Hölbling, D.; Breiteneder, R.; Christoph, L. System Zur Automatisierten Wertungsvergabe bei Kampfsportarten. Patent A 50619/2021, 26 July 2021.
10. Hölbling, D. Verfahren Zur Kalibrierung Eines Schlaghandschuhes. Patent 31172-AT, 24 February 2023.
11. FightCamp. Available online: https://joinfightcamp.com/ (accessed on 25 May 2023).
12. Hykso. Available online: https://shop.hykso.com/ (accessed on 25 May 2023).
13. ROOQ Box. Available online: https://rooq-shop.com/ (accessed on 13 March 2022).
14. StrikeTec. Available online: https://striketec.com (accessed on 13 March 2022).
15. Move It. Available online: https://move-it.store/ (accessed on 13 March 2022).
16. Perri, T.; Reid, M.; Murphy, A.; Howle, K.; Duffield, R. Prototype Machine Learning Algorithms from Wearable Technology to Detect Tennis Stroke and Movement Actions. *Sensors* **2022**, *22*, 8868. [CrossRef] [PubMed]
17. Kautz, T.; Groh, B.; Hannink, J.; Jensen, U.; Strubberg, H.; Eskofier, B. Activity recognition in beach volleyball using a Deep Convolutional Neural Network. *Data Min. Knowl. Discov.* **2017**, *31*, 1678–1705. [CrossRef]

18. Cust, E.E.; Sweeting, A.J.; Ball, K.; Robertson, S. Machine and deep learning for sport-specific movement recognition: A systematic review of model development and performance. *J. Sport. Sci.* **2019**, *37*, 568–600. [CrossRef] [PubMed]

19. Worsey, M.T.O.; Espinosa, H.G.; Shepherd, J.B.; Thiel, D.V. An Evaluation of Wearable Inertial Sensor Configuration and Supervised Machine Learning Models for Automatic Punch Classification in Boxing. *IoT* **2020**, *1*, 21. [CrossRef]

20. Wagner, T.; Jäger, J.; Wolff, V.; Fricke-Neuderth, K. A machine learning driven approach for multivariate timeseries classification of box punches using smartwatch accelerometer sensordata. In Proceedings of the Innovations in Intelligent Systems and Applications Conference (ASYU), Izmir, Turkey, 31 October–2 November 2019; pp. 1–6. [CrossRef]

21. Khasanshin, I. Application of an Artificial Neural Network to Automate the Measurement of Kinematic Characteristics of Punches in Boxing. *Appl. Sci.* **2021**, *11*, 1223. [CrossRef]

22. Kasiri, S.; Fookes, C.; Sridharan, S.; Morgan, S. Fine-grained action recognition of boxing punches from depth imagery. *Comput. Vis. Image Underst.* **2017**, *159*, 143–153. [CrossRef]

23. Soekarjo, K.M.W.; Orth, D.; Warmerdam, E.; van der Kamp, J. Automatic Classification of Strike Techniques Using Limb Trajectory Data. In *Machine Learning and Data Mining for Sports Analytics*; Brefeld, U., Davis, J., Van Haaren, J., Zimmermann, A., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 131–141.

24. Lapková, D.; Kominkova Oplatkova, Z.; Pluhacek, M.; Senkerik, R.; Adamek, M. Analysis and Classification Tools for Automatic Process of Punches and Kicks Recognition. In *Pattern Recognition and Classification in Time Series Data*; IGI Global: Hershey, PA, USA, 2017. [CrossRef]

25. Fuchs, P.X.; Lindinger, S.J.; Schwameder, H. Kinematic analysis of proximal-to-distal and simultaneous motion sequencing of straight punches. *Sport. Biomech.* **2017**, *17*, 512–530. [CrossRef] [PubMed]

26. Rinaldi, M.; Nasr, Y.; Atef, G.; Bini, F.; Varrecchia, T.; Conte, C.; Chini, G.; Ranavolo, A.; Draicchio, F.; Pierelli, F. Biomechanical characterization of the Junzuki karate punch: Indexes of performance. *Eur. J. Sport Sci.* **2018**, *18*, 796–805. [CrossRef]

27. World Association of Kickboxing Organizations (WAKO). WAKO Kickboxing Rules. 2022. Available online: https://wako.sport/wp-content/uploads/2022/10/WAKO-Rules-25.10.2022.-revision-3.pdf (accessed on 20 June 2023).

28. International Boxing Association (IBA). IBA Rulebook. 2021. Available online: https://www.iba.sport/wp-content/uploads/2022/02/IBA-Technical-and-Competition-Rules_20.09.21_Updated_.pdf (accessed on 20 June 2023).

29. Gatt, I.; Allen, T.; Wheat, J. Quantifying wrist angular excursion on impact for Jab and Hook lead arm shots in boxing. *Sport. Biomech.* **2021**, 1–13. [CrossRef]

30. Dinu, D.; Louis, J. Biomechanical Analysis of the Cross, Hook, and Uppercut in Junior vs. Elite Boxers: Implications for Training and Talent Identification. *Front. Sport. Act. Living* **2020**, *2*, 598861. [CrossRef]

31. Mudrić, R.; Ranković, V. Analysis of Hand Techniques in Karate. *Sport. Sci. Pract.* **2016**, *6*, 47–74.

32. Wombat Studio. Magic Poser. Available online: https://magicposer.com/ (accessed on 24 May 2023).

33. Meinel, K.; Schnabel, G. *Bewegungslehre—Sportmotorik: Abriss Einer Theorie der Sportlichen Motorik unter Pädagogischem Aspekt*; Meyer & Meyer: Aachen, Germany, 2007.

34. Fayyad, U.; Piatetsky-Shapiro, G.; Smyth, P. From Data Mining to Knowledge Discovery in Databases. *AI Mag.* **1996**, *17*, 37–54.

35. Chapman, P.; Clinton, J.; Kerber, R.; Khabaza, T.; Reinartz, T.; Shearer, C.R.; Wirth, R. *CRISP-DM 1.0: Step-by-Step Data Mining Guide*; SPSS Inc.: Chicago, IL, USA, 2000.

36. SAS Institute Inc. Introduction to SEMMA. 2017. Available online: https://documentation.sas.com/?docsetId=emref&docsetTarget=n061bzurmej4j3n1jnj8bbjjm1a2.htm&docsetVersion=14.3&locale=en (accessed on 8 February 2021).

37. Foundation, P.S. Python. Available online: https://www.python.org/ (accessed on 28 March 2021).

38. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

39. Raschka, S. Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning. *arXiv* **2018**, arXiv:1811.12808.

40. Burkov, A. *The Hundred-Page Machine Learning Book*; OCLC: Dublin, OH, USA, 2019.

41. Dietterich, T.G. Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. *Neural Comput.* **1998**, *10*, 1895–1923. [CrossRef] [PubMed]

42. Raschka, S. MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack. *J. Open Source Softw.* **2018**, *3*, 638. [CrossRef]

43. Flask. Available online: https://flask.palletsprojects.com/en/2.0.x/ (accessed on 25 February 2022).

44. Buitinck, L.; Louppe, G.; Blondel, M.; Pedregosa, F.; Mueller, A.; Grisel, O.; Niculae, V.; Prettenhofer, P.; Gramfort, A.; Grobler, J.; et al. Tuning the Hyper-Parameters of an Estimator. Available online: https://scikit-learn.org/stable/modules/grid_search.html (accessed on 23 October 2021).

45. Hanada, Y.; Hossain, T.; Yokokubo, A.; Lopez, G. BoxerSense: Punch Detection and Classification Using IMUs. In *Sensor- and Video-Based Activity and Behavior Computing*; Ahad, M.A.R., Inoue, S., Roggen, D., Fujinami, K., Eds.; Springer: Singapore, 2022; pp. 95–114.

46. Baranyi, R.; Czech, P.; Walcher, F.; Aigner, C.; Grechenig, T. Reha@ Stroke-A mobile application to support people suffering from a stroke through their rehabilitation. In Proceedings of the 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH), Kyoto, Japan, 5–7 August 2019; pp. 1–8.

47. Baranyi, R.; Czech, P.; Hofstätter, S.; Aigner, C.; Grechenig, T. Analysis, Design, and Prototypical Implementation of a Serious Game Reha@ Stroke to Support Rehabilitation of Stroke Patients with the Help of a Mobile Phone. *IEEE Trans. Games* **2020**, *12*, 341–350. [CrossRef]