Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 7, S. 163-170, 7, 2023

Araştırma Makalesi

IJANSER

https://as-proceeding.com/index.php/ijanser
ISSN: 2980-0811

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 7, pp. 163-170, 7, 2023

Research Article

Enhancing Human Movement Classification in Imbalanced Data with a Hybrid AI Approach

Kürşad UÇAR*[1]

[1]*Department of Electric and Electronic Engineering/Faculty of Technology, Selcuk University, Turkey*

(kucar@selcuk.edu.tr)

(Received: 16 August 2023, Accepted: 28 August 2023)

(1st International Conference on Recent and Innovative Results in Engineering and Technology ICRIRET 2023, August 16-18, 2023)

ATIF/REFERENCE: Uçar, K. (2023). Enhancing Human Movement Classification in Imbalanced Data with a Hybrid AI Approach. *International Journal of Advanced Natural Sciences and Engineering Researches*, 7(7), 163-170.

Abstract **–** The monitoring and classification of human movements have garnered significant importance across various domains, including hazardous situation detection and elderly care. In recent years, the application of Artificial Intelligence (AI) has emerged as a promising approach to predicting and analyzing human movements in such contexts. However, as the complexity and diversity of target movements increase, obtaining sufficient and high-quality data for AI methods becomes increasingly challenging. A particular obstacle researchers and practitioners encounter is the uneven distribution of movement data, where certain movements are frequently repeated and well-represented in the dataset, while others occur infrequently, leading to an imbalanced dataset. This data imbalance poses a critical issue as it hampers the effectiveness of AI models, particularly for those movements with limited available data. The scarcity of samples for certain classes can lead to biased learning, resulting in reduced accuracy and poor generalization of the model. To address this critical challenge, this research proposes a hybrid AI methodology specifically designed to enhance the classification accuracy of two important human movements: turning and walking. The experimental results demonstrate the efficacy of the proposed method. It significantly improves the overall classification accuracy compared to only one AI technique, especially for those movements characterized by limited data availability. By effectively addressing the data imbalance problem, the hybrid AI methodology ensures that the model's performance is not disproportionately skewed towards the overrepresented classes, yielding a more balanced and reliable human movement monitoring and classification system.

Keywords – Human Activity, Classification, Artificial Intelligence, Hybrid Model, Limited Data

I. Introduction

The precise monitoring and classification of human movements have garnered immense importance across an array of critical domains, ranging from child and elderly care to the detection of dangerous situations and the tracking of sequential tasks [1-3]. In these spheres, the meticulous analysis of people's movements, coupled with continuous observation, facilitates invaluable follow-ups such as health status assessments, organ functionality evaluations, and the identification of potentially harmful behaviors. Moreover, the application of movement tracking technologies enables the timely detection of violent acts, such as physical aggression, battering, or extortion, thereby aiding in the early recognition of suspicious individuals who might pose a threat to others [4, 5]. Additionally, the prediction and prevention of hazardous movements for employees in high-risk work environments and the accurate monitoring of sequential tasks are vital applications that have the potential to enhance workplace safety and productivity [6].

An instrumental tool in this pursuit is the ubiquitous camera, which operates tirelessly and serves as an accessible imaging network, particularly in the realm of security cameras. Harnessing the power of deep learning and artificial intelligence algorithms has further expanded the horizons of image analysis, facilitating swift and contextually relevant results [7, 8]. As such, the endeavor to comprehend and control human behavior in videos has evolved into a captivating and crucial subject, aimed at mitigating possible accidents and dangerous situations.

A foundational aspect of human movements lies in the sequence of positional changes within bodily organs. Consider, for instance, the act of eating, where the hands engage in a repetitive motion towards the mouth, or the phenomenon of walking, characterized by the rhythmic swinging of arms and coordinated bodily locomotion [9]. By accurately tracking these organs in images, it becomes possible to discern and interpret diverse human movements.

While it is indeed feasible to detect and classify movements from individual images, complexities arise when dealing with movements that involve opposing actions. For instance, the task of donning and doffing a jacket requires contrasting movements. Attempting to deduce this action from a single image of the jacket in motion can prove challenging, necessitating a more nuanced approach that entails the classification of movements based on the sequential flow of organ positions.

At the core of the pursuit of effective movement classification lies the collection and labeling of motion images of individuals. However, this process is frequently confronted by its time-consuming and labor-intensive nature, especially as the scope of target movements expands. The issue is further compounded by the varied frequencies and durations of different movements, making it challenging to maintain a balanced dataset. For instance, walking, being a common activity, is frequently encountered and extends over a considerable duration, while standing up, on the other hand, is less frequent and brief in duration. Such data imbalance poses substantial obstacles for artificial intelligence, which struggles to learn movements characterized by limited data, ultimately affecting the training and classification efficacy [10, 11].

In light of these challenges, this study endeavors to present a novel and advanced hybrid artificial intelligence method devised for the classification of back-and-forth walking and turning movements, which contend with data imbalance. Initially, a sophisticated deep learning-based YOLOv7 [12] approach is employed to identify and precisely locate ten crucial human organs. Subsequently, the motion data from these localized regions in consecutive images are meticulously measured and effectively classified using state-of-the-art SVM and Decision Tree techniques. While traditional methods exhibit limited success in accurately classifying forward and backward movements, the proposed hybrid approach demonstrates significant enhancements, particularly in the case of rotational movements, where data availability is substantially limited.

II. MATERIALS AND METHOD

A. *YOLO (You Look Only Once)*

YOLO [13] is a popular and extremely powerful object detection algorithm used in computer vision. YOLOv7 [12] is an evolution of the YOLO series and stands for "You Only Look Once version 7". The primary purpose of YOLO is to detect and localize objects in images or videos. In doing so, it divides the input image into a grid and estimates the bounding boxes and class probabilities for each grid cell. Unlike most deep learning algorithms, YOLO stands out with its real-time performance and efficiency. This makes it suitable for applications such as driverless cars, surveillance systems and robotics. YOLOv7 may include improvements over its predecessors, such as changes in network architecture, training strategies, or additional features.

B. *DT (Decision Tree)*

DT [14] is a popular machine learning algorithm used for both classification and regression tasks. It works by recursively subdividing the attribute space based on the values of different attributes. Each segment creates a "node" in the tree and the process continues until a stopping criterion is met (e.g. maximum tree depth, minimum sample per leaf). In case of classification, each leaf node represents a class label and for regression the leaf nodes contain predicted continuous values. DTs are easy to interpret and visualize, making them valuable for understanding how decisions are made by the model. However, they can be prone to overfitting, especially when the tree becomes too deep or complex. Techniques such as pruning and using ensemble methods (e.g. Random Forest) can help alleviate this problem.

C. SVM (Support Vector Machine)

SVM [15] is also a machine learning algorithm used for classification and regression tasks. Its main purpose is to find a hyperplane that optimally separates data points from different classes in a multidimensional space. In the case of binary classification, SVM aims to maximize the margin (distance) between the two classes by providing a solid decision boundary. SVM is effective for both linearly separable and nonlinearly separable data, thanks to the use of kernel functions that transform the input space into a higher dimensional feature space from which the data can become separable. SVM is widely used in various fields such as image recognition, text classification and bioinformatics.

D. Proposed Method

Movement analysis is a dynamic process, evolving over time, and relying on a single image for estimating movements may lead to misleading or inadequate results. To address this challenge, the proposed method in this study is centered around organ detection in images and the subsequent measurement of organ movements in consecutive frames. Ten target organs were selected for tracking in this study, including the head, two arms (from shoulders to wrists), two hands, the chest, the abdomen, two legs, and two feet. The YOLO (You Only Look Once) algorithm was employed to detect these organs in images captured at intervals of 50 milliseconds. However, as YOLO does not inherently differentiate between left and right organs, additional checks were introduced to establish the correct positional relationships among the detected organs.

Several criteria were applied to validate the detected organ positions:

1) The head must consistently occupy the uppermost position, and the feet must remain at the bottom throughout the target movements.

2) The legs must be positioned below the chest, and the head must be above the torso at all times.

3) Sequential images should not exhibit abrupt jumps in organ positions, ensuring smooth transitions ($|x_t - x_{t-1}| < 0.1$, $|y_t - y_{t-1}| < 0.1$, where x_t and y_t are the midpoints of organs at time t.

4) Organs such as hands, feet, legs, and arms were categorized into right and left, and in cases where only one organ is detected in an image, the closest matching organ (right or left) from the previous image is assigned.

5) Over-detected organs were compared with their positions in the previous image and the closest match was considered for subsequent analysis.

6) In the absence of detected organs, their positions were recorded as 0.

7) If YOLO potentially fails to detect an organ, a comparison is made with the previous image. If the organ was not detected and its approximate location was not close to the image edges in the previous frame (approximate location > 0.1), the previous position is utilized as a reasonable approximation for the current position.

Once the organ detection process is completed, artificial intelligence techniques are employed to classify the movements. SVM and DT algorithms were utilized to analyze the positional changes of all organs and classify forward-backward walking and forward-backward rotational movements. As depicted in Table 1, due to variations in data amount, achieving overall satisfactory classification success becomes challenging. Specifically, rotational movements, which suffer from limited data compared to walking, exhibit inadequate classification results. Consequently, a two-stage classification approach was adopted. Initially, walking and turning movements without direction were classified, followed by the classification of movement direction in the second stage. At this stage, SVM and DT showed discrepancies in classifying movements and directions. Leveraging these differences, when the outputs from the two methods diverged, more definitive conclusions could be drawn, leading to increased classification success.

Table 1. Number of movements

Movement	Number of movements	Number of same movements
Walking for.	393	830
Walking back.	437	
Rotate for.	89	150
Rotate back	61	

The proposed method's block diagram is visually depicted in Fig. 1, showcasing the step-by-step process of organ detection, subsequent checks, and the classification of movements through artificial intelligence algorithms. By incorporating robust organ detection and applying a two-stage classification approach, this study seeks to enhance the accuracy and reliability of movement analysis, particularly in cases characterized by imbalanced data distribution. The methodology holds promising potential for revolutionizing movement tracking and classification applications, fostering safer environments, and optimizing performance across a wide range of domains.

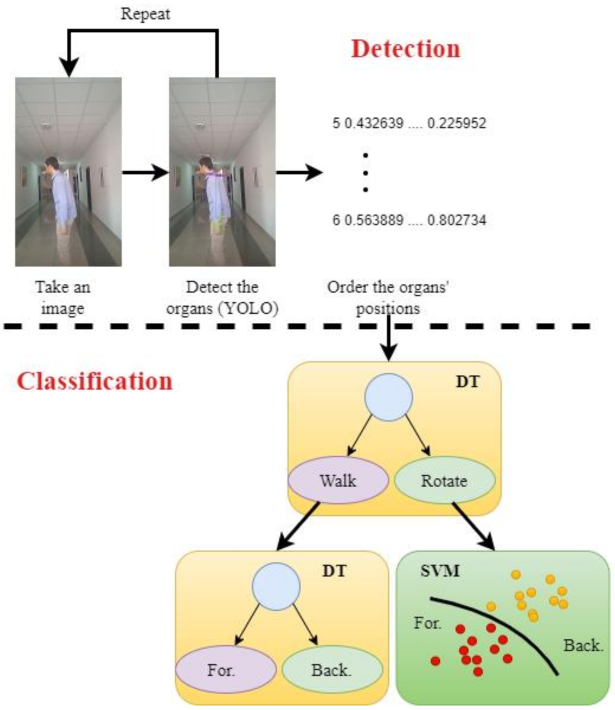


Fig. 1 Block diagram of method

III. DATASET

The process of motion classification in this study involves the use of two distinct datasets that are sequentially employed. The first dataset encompasses images utilized for training the YOLO model, responsible for detecting organs, while the second dataset includes the locations of the detected organs, obtained after the completion of YOLO training.

For the YOLO training and test sets, 980 images were extracted from a video lasting approximately 49 seconds, with an image captured every 50 milliseconds. The image dimensions were set at 720 x 1280 pixels. The objective was to detect ten specific organs as shown in Fig. 2, within these images. Subsequently, a random allocation was performed, dedicating 70% of the images to training, 10% for testing, and 20% for validation.

In the second stage, the detected organs' positions were determined within the images. However, evaluating the positions of organs within individual images in isolation would not provide meaningful insights into movements. To address this, the relative amount of movement for each organ's midpoint between successive images was quantified using Eq. (1).

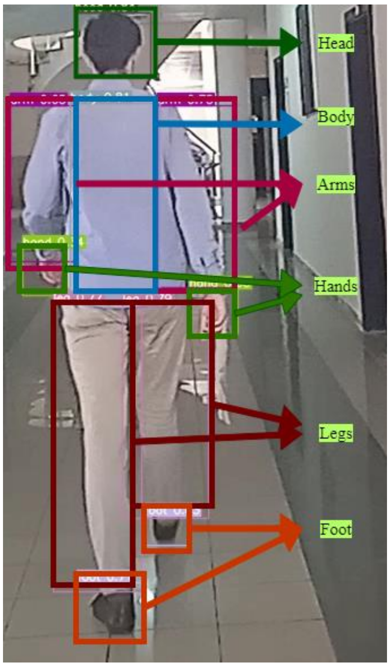


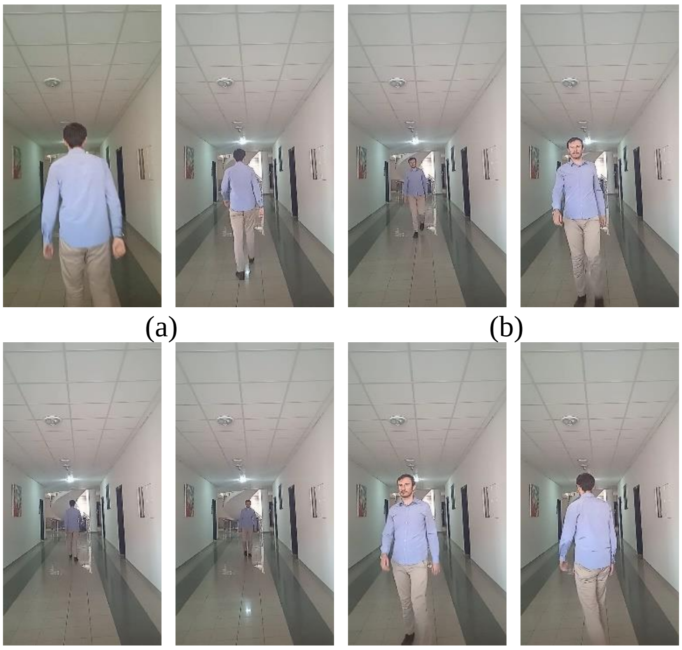
Fig. 2 Target organs

$$mo_{x,y}^{i} = o_{x,y}^{i} - o_{x,y}^{i-1} \tag{1}$$

Here, *mo* represents the amount of movement of the organ, *x* and *y* represent the position of the organ in the i^{th} image, and *o* it represents the organs. In the classifiers, 60% of the detection results of YOLO were used in training and 40% in the test. The movements in the images are as in Fig. 3 and their amounts are as in Table 1.

IV. Experiment and Results

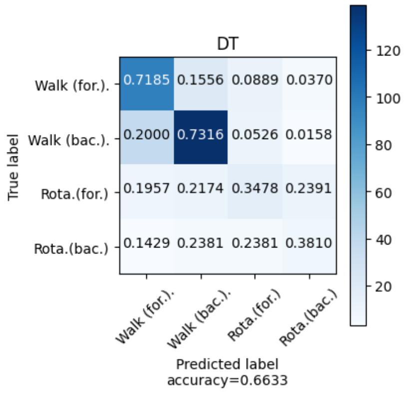
Two different experiments were carried out to classify the movements. In the first trials, four movements were tried to be classified simultaneously with DT and SVM. Classification results for four classes with DT and SVM were obtained as in Fig. 4. Despite all the mentioned corrections, it is obvious that these results are not satisfactory for motion classification. As a maximum, approximately 66.33% classification success was achieved with SVM. Approximately 36% of rotational movements are correctly classified. The effect of absolute values at this stage was not expected to be positive. Because the amount of movement of the organs in the forward and backward movements are of opposite signs.



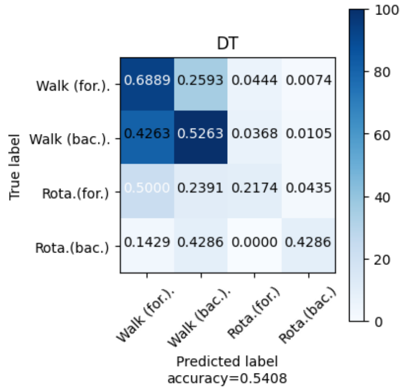
(a) (b) (c) (d)

Fig. 3 Movements, (a) walk forward, (b) walk backward (c) turn forward, (d) turn back

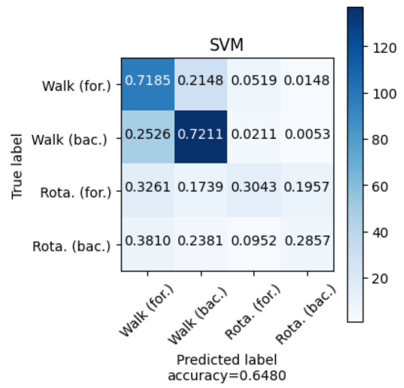
Since DT and SVM could not achieve high success in classifying these four movements, the movements were first divided into two groups without specifying the direction with the proposed method. In other words, movements as in the last column of Table 1 are classified as walking and turning movements. In fact, what is done here is to deal with the classification problem regardless of the direction of the transaction movements. Later, the movements were classified as forward-backward. When walking and turning movements are forward and backward, the amount of movement of the organs is close to each other. However, since these movements are opposite to each other, values with opposite signs are obtained in Eq. (1). Absolute values are used to eliminate this sign difference. Thus, it can be even more successful to classify turning and walking movements, regardless of whether they are forward or backward. This effect can also be seen in the turning and walking classification in Fig. 5. It has been possible to successfully classify over 90% of movements. Also, using absolute values here improves performance by 2-3%. While walking could be classified as 95.38% successfully, this rate was best 65.67% in turning.



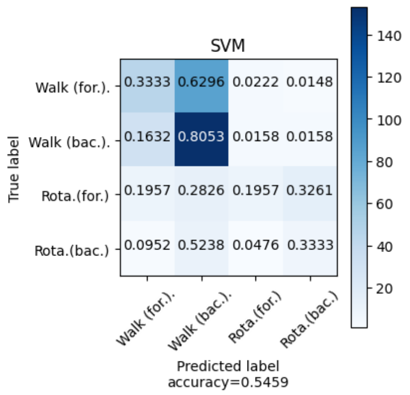
(a)



(b)

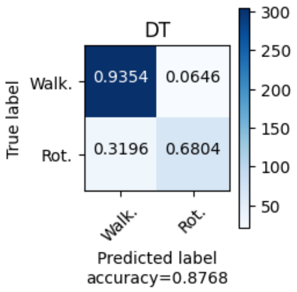


(c)

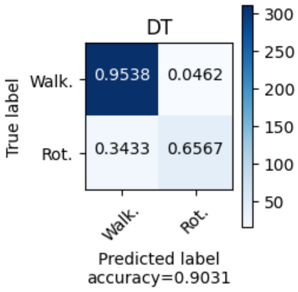


(d)

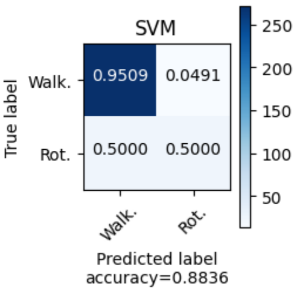
Fig. 4 Classification results for four activities, (a) DT result with original data, (b) DT result with absolute value, (c) SVM result with original data, (d) SVM result with absolute value



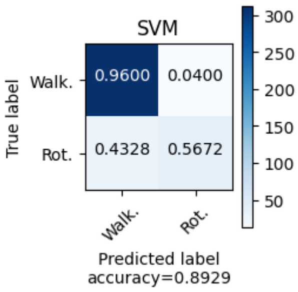
(a)



(b)



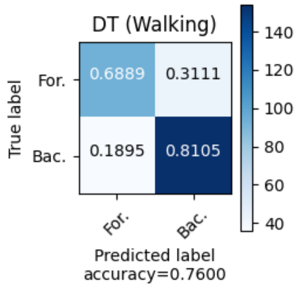
(c)



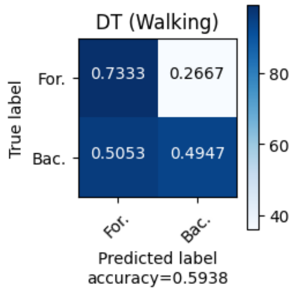
(d)

Fig. 5 Classification results for two activities, (a) DT result with original data, (b) DT result with absolute value, (c) SVM result with original data, (d) SVM result with absolute value

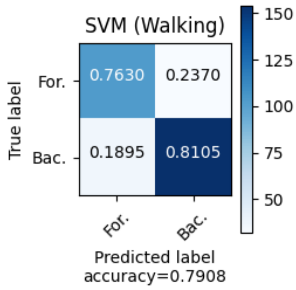
After the motion class was determined, the direction of the movements was determined by using two methods. The classification results of walking and turning directions are as in Fig. 6 and Fig. 7, respectively.



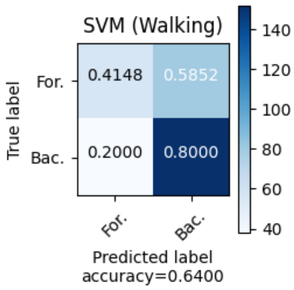
(a)



(b)

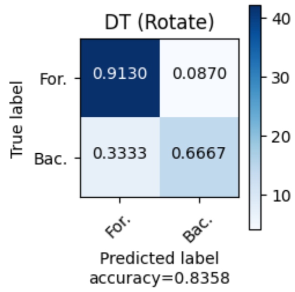


(c)

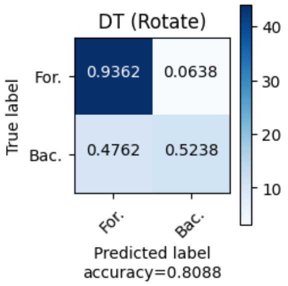


(d)

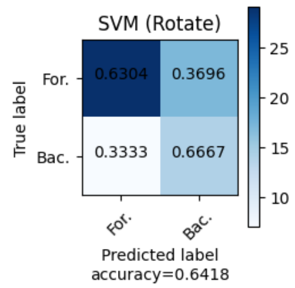
Fig. 6 Classification of walking direction, (a) DT result with original data, (b) DT result with absolute value, (c) SVM result with original data, (d) SVM result with absolute value



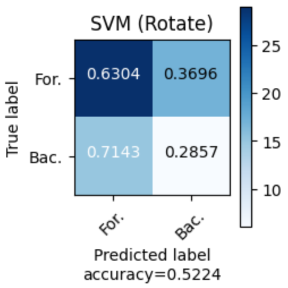
(a)



(b)



(c)



(d)

Fig. 7 Classification of rotate direction, (a) DT result with original data, (b) DT result with absolute value, (c) SVM result with original data, (d) SVM result with absolute value

In Fig 5, the best classification was obtained with absolute values. After the absolute values, the classification results for the direction of the movements were obtained as in Fig. 6 and Fig. 7. Absolute values had a negative effect on direction determination. In the two-stage classification method, the classification success obtained by using Eq. (2) for the classification of movements together with their directions was obtained as cr = 0.72.

$$cr = \frac{cr_w^1 \ x \ cr_w^2 \ x \ s_w + cr_r^1 \ x \ cr_r^2 \ x \ s_r}{s_w + s_r} \qquad (2)$$

Here, cr_w^1 and cr_w^2, cr_t^1 and cr_r^2 are the classification success of walking and turning movements in the first and second steps, respectively, and s_w and s_r are the sample numbers in walking and turning. In addition to the approximately 6% improvement in the classification success of this method, the most important point is its effect on the classification of rotational movements with little data. When the four movements were classified, the success of the rotational movements was around 36%, while it was 54.89% with the two-stage method applied.

When Fig. 6 and Fig. 7 are examined, it is seen that SVM is more successful in walking and DT is more successful in turning. For this reason, it is understood that the hybrid system developed by using different artificial intelligence methods while estimating the direction in different movements as in the proposed method can be used in such tasks.

V. DISCUSSION

This study presents a hybrid artificial intelligence approach designed to address the challenges of motion classification with unstable data numbers. By employing Decision Trees (DT) and Support Vector Machine (SVM) algorithms, we effectively classified forward-backward turning and forward-backward walking movements. Furthermore, the two-stage classification method significantly improved classification results when attempting to classify all four movements simultaneously.

The significance of this research lies in its successful handling of unbalanced data and limited samples per class, factors that often hinder the training of artificial intelligence models. With the proposed method, the adverse impact of data imbalance was mitigated, leading to notable improvements in overall classification performance, particularly for classes with sparse data.

VI. CONCLUSION

In conclusion, our hybrid artificial intelligence approach showcases promising results in motion classification, offering a robust solution for scenarios with unstable data numbers. The integration of DT and SVM algorithms, along with the two-stage classification methodology, effectively addresses the challenges posed by unbalanced data distribution and limited samples per class. The proposed method exhibits substantial improvements in classification accuracy, particularly for classes with sparse data, demonstrating its potential for real-world applications.

REFERENCES

[1] I. Seflek, Y. E. Acar, and E. Yaldiz, "Small motion detection and non-contact vital signs monitoring with continuous wave doppler radars," *Elektronika ir elektrotechnika*, vol. 26, no. 3, pp. 54-60, 2020.

[2] M. Wang, Y. D. Zhang, and G. Cui, "Human motion recognition exploiting radar with stacked recurrent neural network," *Digital Signal Processing*, vol. 87, pp. 125-131, 2019.

[3] M. Yildirim and A. Çinar, "A new model for classification of human movements on videos using convolutional neural networks: MA-Net," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 9, no. 6, pp. 651-659, 2021.

[4] E. Bermejo Nievas, O. Deniz Suarez, G. Bueno García, and R. Sukthankar, "Violence detection in video using computer vision techniques," in *Computer Analysis of Images and Patterns: 14th International Conference, CAIP 2011, Seville, Spain, August 29-31, 2011, Proceedings, Part II 14*, 2011: Springer, pp. 332-339.

[5] M. Ramzan *et al.*, "A review on state-of-the-art violence detection techniques," *IEEE Access*, vol. 7, pp. 107560-107575, 2019.

[6] K.-J. Wang and D. Santoso, "A smart operator advice model by deep learning for motion recognition in human–robot coexisting assembly line," *The International Journal of Advanced Manufacturing Technology*, pp. 1-20, 2022.

[7] E. Balouji and O. Salor, "Classification of power quality events using deep learning on event images," in *2017 3rd International conference on pattern recognition and image analysis (IPRIA)*, 2017: IEEE, pp. 216-221.

[8] K. B. Obaid, S. Zeebaree, and O. M. Ahmed, "Deep learning models based on image classification: a review," *International Journal of Science and Business*, vol. 4, no. 11, pp. 75-81, 2020.

[9] P. Sarcevic, Z. Kincses, and S. Pletl, "Online human movement classification using wrist-worn wireless sensors," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 1, pp. 89-106, 2019.

[10] L. Wang, M. Han, X. Li, N. Zhang, and H. Cheng, "Review of classification methods on unbalanced data sets," *IEEE Access*, vol. 9, pp. 64606-64628, 2021.

[11] J. Song, X. Lu, and X. Wu, "An improved adaboost algorithm for unbalanced classification data," in *2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery*, 2009, vol. 1: IEEE, pp. 109-113.

[12] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 7464-7475.

[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779-788.

[14] B. De Ville, "Decision trees," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 5, no. 6, pp. 448-455, 2013.

[15] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and their applications*, vol. 13, no. 4, pp. 18-28, 1998.