A MACHINE LEARNING/DEEP LEARNING HYBRID FOR AUGMENTING TEACHER-LED ONLINE DANCE EDUCATION

Catherine Hung

Palo Alto Senior High School, Palo Alto, USA

ABSTRACT

For online dancers, learning a dance move properly without the feedback of a live instructor can be challenging because it is difficult to determine whether a move is done correctly. The lack of proper guidance can result in doing a move incorrectly, causing injury. In this work – we explore the use of a hybrid Deep Learning/Machine Learning approach to classify dance moves as structurally correct or incorrect. Given a video clip of the dancer doing a move, such as the grand plie, the algorithm should detect the correctness of the movement. To capture the overall movement, we proposed various methods to process data, starting with deep learning techniques to convert video frames into landmarks. Next, we investigate several approaches to combining landmarks from multiple frames and training machine learning algorithms on the dataset. The distinction between correct and incorrect grand plies achieved accuracies of over 98%.

KEYWORDS

Deep Learning, Machine Learning, Classification, Online Dance Education

1. INTRODUCTION

Dancing attracts many people due to its numerous health benefits ranging from enhancing cognitive function to improving balance [1]. Additionally, many individuals dance to improve their mood and overall well-being [2]. In 2021, around 24.71 million people in the US, or around 7.5% of the US population, took advantage of the benefits of dance [3].

While many attend in-person dance classes, due to the COVID-19 pandemic and lockdown in 2020, dance studios were required to stop in-person learning. Thus, many resorted to online dance instruction and dance apps. STEEZY, which has over one million downloads, is an app that teaches users how to dance in various styles ranging from K-pop to ballet for novices, intermediate, and advanced dancers [4]. Even after the lockdown, many still attend online dance classes because they allow for more flexibility in one's schedule, are typically cheaper than in-person classes, and provide people in remote areas the opportunity to participate in classes and learn from teachers all around the world [5]. Yet, despite these advantages, without a teacher physically present to correct them, dancers may incorrectly learn dance moves. Also, it can be hard for dance teachers to correct many students online without some physical contact, and it is difficult for inexperienced new dancers to self-correct [6].

Incorrectly doing even a basic dance move by using incorrect technique and poorly aligning one's body can injure dancers, preventing them from furthering their dance education and learning more complicated moves [7].

David C. Wyld et al. (Eds): CCNET, AIMLA, CICS, IOTBS, NLTM, COIT - 2023
pp. 29-40, 2023. CS & IT - CSCP 2023 DOI: 10.5121/csit.2023.130403

Ballet is a classical style of dance whose movements and technique serve as the foundation of other dance styles [8]. In ballet, a grand plie, an elementary move, helps develop the balance and stability needed for other moves [9]. For this reason, it is important to be able to do this move correctly. It involves the bending of the knees up until the thighs are horizontal to the floor while maintaining a straight back. While this move seems straightforward in theory, there are several ways in which dancers can incorrectly perform this move, such as by incorrectly bending one's spine or abruptly moving from one pose to another after a long pause. This can result in injuries such as muscle strain [10]. Because of this, it is critical for online dancers to learn how to perform each dance move correctly.

AI has the potential to detect and differentiate incorrect dance moves from correct dance moves given its demonstrated effectiveness across a range of sectors including the detection of correct mask-wearing and sports [11]. This paper explores how deep learning and machine learning techniques can be leveraged in the dance teaching domain. Dance teachers can use the feedback on the correctness of a move to determine areas of growth for students.

To build an AI that accurately detects correct and incorrect grand plies, we gathered videos of dancers performing this move correctly and incorrectly and consulted an experienced dancer to label the videos for our dataset. We then used a deep learning technique to extract the physical landmarks of each video frame. Subsequently, we processed these landmarks in a variety of ways to assess different aspects of the motion, such as speed and smoothness. After that, we used our processed landmarks to train three well-known and robust machine learning algorithms: K-Nearest Neighbors, Random Forest Classifier, and MLP. We then evaluated the results.

After using different techniques to process our data and train our models, we were eventually able to achieve accuracies over 98% and occasionally 100%. Overall, we achieved similar results for all three of the machine learning algorithms we ran our processed data on.

Our paper outline is as follows. Section 2 discusses related work and our novel contributions. Section 3 examines how our solution works, Section 4 covers our results and analysis, Section 5 discusses our observations, and Section 6 notes our conclusions and future work.

2. RELATED WORK

Improving dance education is a field of interest to many. Multiple studies have looked into ways to implement technology into correcting dance movements. For example, past studies have used physical sensors that are attached to dancers to track the movements of their body parts [5,12]. Other studies have collected data through cameras and analyzed the data with machine learning frameworks, such as OpenPose [13,14,15].

Similarly, some studies are focused on detecting the movement of multiple dancers at a time to possibly be utilized in a class with multiple dancers [16,14]. Other studies are focused on tracking the movement of a single dancer or athlete [12, 13, 17]. For example, Woah.AI is an app that teaches users how to do modern TikTok dances and uses AI to provide feedback [13].

Having data on the body movements of dancers can be used to benefit different aspects of dance. One study is using this data to analyze how dance-related injuries develop over time in order to find methods to prevent them [12]. Whereas, another study is concerned with generating dances that are more realistic, creative, and appealing based on specific guidelines such as composition, performance, and evaluation [18].

Our study differs from the related work above in several ways. First, our goal is to detect and correct dance moves. This is unlike other research that is intended to generate new dances or track how dance injuries develop over time. Additionally, unlike how many researchers are detecting dance moves through physical sensors, which may be uncomfortable to dance with, we are using non-invasive techniques such as cameras to detect dance moves. We expect our approaches to be more practical for use in dance education. Also, other researchers are looking at how to detect the dance moves of multiple dancers in a room. We, however, are only detecting the dance moves of a single dancer since our research is aimed toward online dancers.

Online dance classes are challenging for both students and teachers. One study that analyzed virtual dance education found that while it is progressing promisingly, there is still improvement needed to be done [19]. In online streaming platforms like Zoom, because students are represented by tiny squares, the teacher is unable to see an individual student in detail and has difficulty tracking the issues of multiple students at the same time. Our application has the potential to overcome these limitations since each video feed can be analyzed separately.

3. SOLUTIONS

We begin by outlining the requirements for an effective solution.

3.1. Requirements

A. Correct movement requires the integration of the full body. Therefore, in order to properly assess the movement, the algorithm should analyze a full-body video clip of the dancer doing a grand plie.

B. Given the full-body video clip, the algorithm should be able to accurately detect whether the dancer is correctly or incorrectly doing the grand plie.

C. Smoothness and continuity are important elements of a correct grand plie [10]. Therefore, the algorithm should be able to leverage the data to assess the smoothness and continuity of a given grand plie.

D. In addition to smoothness, the algorithm should also be able to detect other essential factors when classifying the grand plie as correct or incorrect. One important factor is the placement of the hips. In a correct grand plie, the hips are between and in line with the knees. Whereas, in an incorrect grand plie, dancers place their hips behind or below their knees. Another factor is the placement of the heels. Correct grand plies involve the heels naturally and slightly lifting up as dancers are bending their knees. However, dancers who incorrectly do grand plies forcibly push their heels up high [10].

3.2. How Our Solution Works

Next, we outline our algorithmic approach and our dataset.

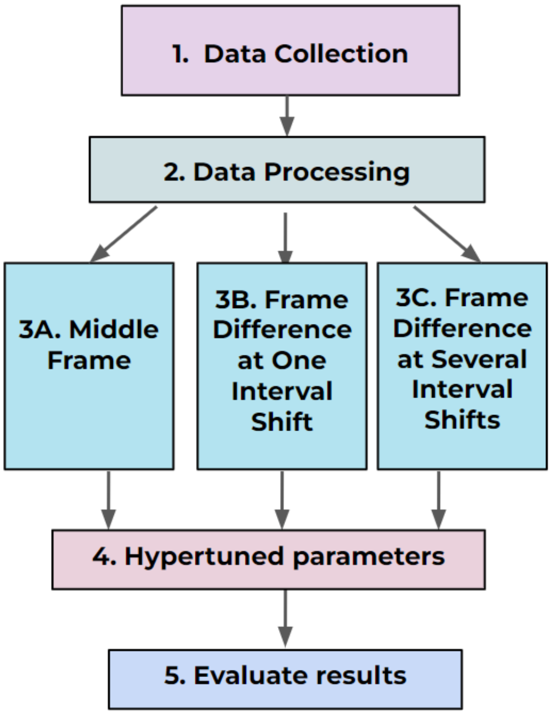

Figure 1. Experimental Process

Figure 1 describes our overall process in creating our model that detects whether or not a dancer is doing a grand plie correctly. For our first step, we gathered videos of correct and incorrect grand plies. Then, in Step 2, we processed these videos using deep learning algorithms to generate landmarks. Next, in Steps 3A, 3B, and 3C, we combined the landmarks from different frames using three different methods. Then, in Step 4, we trained our data on machine learning algorithms. Finally, we analyzed our findings in Step 5.

3.2.1. Data

In total, we collected 42 video clips of dancers correctly doing a grand plie and 34 video clips of dancers incorrectly doing a grand plie. These videos came from various YouTube videos, Instagram posts, and live dancers. These video clips were labeled by a trained dancer and we had permission to use them.

3.2.2. Data Processing

To convert the video clips into features, we used MediaPipe Pose, a machine learning solution for tracking the locations of body parts [20]. Given a full-body RGB multi-frame video of a person, MediaPipe Pose uses BlazePose, a built-in neural network solution, to locate the person of interest and identify the locations of the 33 landmarks, such as the left leg, right hip, and left ankle [21]. Within each landmark, the x-component, y-component, the landmark depth (z), and the likelihood of a landmark being visible in the image (v) are given as numerical values. Thus, for each video frame, 132 data points are collected. The label of the video is one of two values — whether the move is correctly or incorrectly executed. Additionally, using MediaPipe Pose also allows our approach to be deployed in a variety of delivery vehicles, such as in a mobile app or website.

Figure 2. MediaPipe Pose for tracking body position. The image was provided by a volunteer who consented to its use.

3.2.3. Processing the MediaPipe Pose Output

We experimented with processing the MediaPipe Pose output in three ways – listed as 3(A-C) below. The specific calculations are listed in the equations section.

3A. Initially, we used one frame per video for our dataset. We theorized that with a captured moment in the middle frame of each video clip, the AI algorithm could detect flaws in body placement, such as the intentional lifting of the heels. We define the middle frame as the frame in the exact middle of the video. The middle of the grand plie is when dancers' errors such as the sticking out of hips and the intentional lifting of the heels are the most prominent [10].

While our first method could potentially capture snapshots of error, such as the intentional lifting of the heels, it fails to consider the overall smoothness of the movement. When incorrectly doing the grand plie, dancers tend to pause for a while when their legs are bent the most, whereas correct grand plies involve gradual, nonstop movement [10]. To capture this movement, we took the difference in the landmark coordinates between two specified frames at certain distances from the middle frame of the video. If these two frames have some difference, the dancer could be gradually moving throughout the grand plie. Little to no change in the difference between these two frames indicates that the dancer is pausing at a pose. Furthermore, a large difference shows that the dancer could be abruptly moving from one move to another. This approach is illustrated in 3B and 3C.

3B. We initially calculated the difference in landmark positions between two frames, one frame that was twenty frames before the midpoint of the video and another frame that was twenty frames after the video midpoint. Our new dataset is comprised of taking the frame difference at one interval shift per video.

3C. To increase the size of our dataset, we calculated the landmark difference of two frames at a total of three different interval shifts. We define an interval shift as a time location or frame shifted by an interval relative to the middle of the video. We extended method 3B by adding two additional interval shifts. We added frame differences between a frame that was 15 frames before the midpoint and a frame that was 25 frames after the midpoint. We also calculated the difference between a frame 25 frames before the midpoint and another one 15 frames after the midpoint. Our dataset for 3C thus consisted of triple the number of samples in 3B.

3.3. Equations

We name the equation parameters as below.

Videos (*1..N*)
Frames - *F*. F_i *(frames in video i)*
Frame midpoint $= M_i = Floor(F_i/2)$
Left_bookend = *Lb*
Right bookend = *Rb*
Length $L = Rb - Lb$

3.3.1. Middle Frame (3A)

$$M_i \tag{1}$$

Each row has 132 features for each frame in the video. We use the frame *Mi* that is in the middle of each video. What this means is that the first group of data has *N* samples. Each sample represents one video which contains the middle frame of each video.

3.3.2. Frame Difference at One Interval Shift (3B)

We used Length = 40. The Left bookend Lb and Right bookend Rb are as follows.

$$Lb = M_i - 20 \tag{2}$$
$$Rb = M_i + 20 \tag{3}$$

We convert each feature to the difference between that feature's value in Frame *Rb* and the value in Frame *Lb*.

For example:

$$\text{Nose\_x} = \text{Frame }(Rb)\text{\_Nose\_x} - \text{Frame }(Lb)\text{\_Nose\_x} \tag{4}$$

What this means is that the second group of data has *N* samples. Each sample represents one video. Each feature is the difference of that feature's value between Frames *Lb* and *Rb* where the sample has Left and Right bookends as calculated in equations (2) and (3).

3.3.3. Frame Difference at Several Interval Shifts (3C)

We used Length = 40. The Left bookend Lb and Right bookend Rb are as follows.

$$Lb = M_i - 25 \tag{5}$$
$$Rb = M_i + 15 \tag{6}$$
$$Lb = M_i - 25 \tag{7}$$
$$Rb = M_i + 15 \tag{8}$$

What this means is that the third group of data has $3N$ samples. Each video produces three different samples. Each feature is the difference of that feature's value between Frames Lb and Rb where the first sample has (2) and (3), the second sample has (5) and (6), and the third sample has (7) and (8).

In response to the unbalanced dataset with more correct grand plies, we utilized SMOTE to oversample incorrect grand plies [22]. SMOTE, also known as Synthetic Minority Oversampling Technique, aims to create a balanced data set by replicating some of the data from the minority class to make the size of the minority class the same as that of the majority class.

3.4. Final Algorithm

We used three different ML algorithms — K-Nearest Neighbors, Random Forest Classifier, and MLP — to test our model and hypertuned parameters within each algorithm to obtain the best accuracy.

- Random Forest Classifier is a classification method composed of many decision trees. Each decision tree, which is built through random methods, outputs a class prediction and the Random Forest Classifier outputs the most popular class prediction [23]. With Random Forest Classifier, the data iterated through an interval of the number of estimators ranging from 1 to 13 and through maximum depths ranging from 10 to 40.
- K-Nearest Neighbors, also known as KNN, is another classification method that presumes that similar data points with the same class are located close together. To classify a data point, it looks at the classes of the other close data points, or neighbors, and determines the most popular class. With KNN, the data iterated through an interval of the number of neighbors ranging from 1 to 10 neighbors [24].
- Multilayer Perceptron, or MLP, is a basic neural network that consists of multiple layers including an input layer, several hidden layers, and an output layer [25]. With MLP, the training iterated through different learning rates ranging from 0.01 to 0.1 and maximum epochs ranging from 20 to 125. Early stopping was enabled and the tolerance was set to 0.00001.

4. RESULTS

Overall, throughout the three different methods of processing data, algorithms with and without SMOTE resulted in similar maximum accuracies. In the cases with a difference in accuracies, using an algorithm with SMOTE slightly lowered the highest accuracy. For example, this was the case in Fig. 3; MLP had an accuracy of 0.7826 without SMOTE and an accuracy of 0.7391 with SMOTE. Given that there were more correct samples than incorrect samples, the algorithms could have been biased to predict that a given video is correct without SMOTE. Additionally, all three algorithms had the highest accuracy when given a dataset with frame differences at several interval shifts per video clip. In Fig. 5, all maximum accuracies were above 90% with Random Forest Classifier even achieving an accuracy of 100%. The three algorithms had the lowest accuracy when a frame difference was only calculated at one interval shift per video. Fig. 4 demonstrates that the single frame difference approach (3B) did not generate good accuracies.

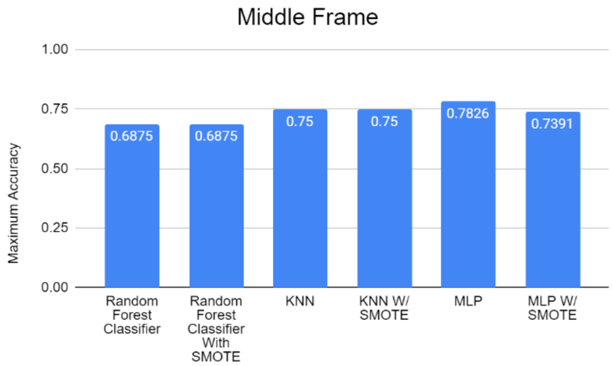

Figure 3. Middle Frame Method (3A)

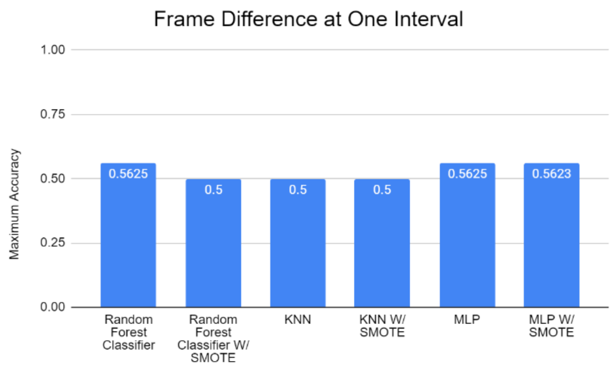

Figure 4. Frame Difference at One Interval Method (3B)

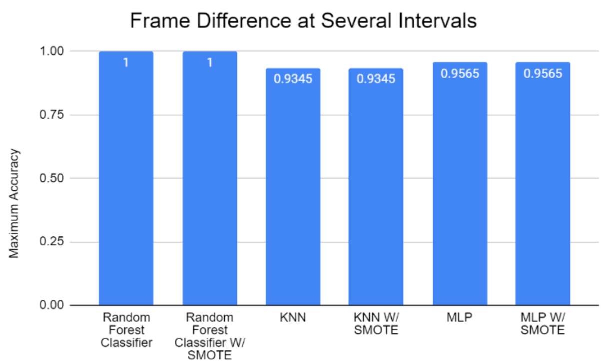

Figure 5. Frame Difference at Several Intervals Method (3C)

Overall, these three algorithms performed roughly the same, and there was not an algorithm that reached a maximum accuracy that was drastically different from the other algorithms. In Fig. 5, Random Forest Classifier performed slightly higher for frame differences at several interval shifts, achieving a maximum accuracy of 100%.

The tables below display our results. Each table consists of the hyperparameters for each algorithm as well as the highest and lowest accuracy that resulted with hyperparameter tuning. There is a table for each of our 3 methods: 3A, 3B, and 3C.

Table 1. Results of Using Middle Frame Method (3A)

Algorithms	Hyperparameters	Accuracy
Random Forest Classifier	Estimators (1-13), Max Depth (10-40)	0.5625-0.6875
KNN	#Neighbors (1-10)	0.5625-0.75
MLP	Learning rate (0.01-0.1), Max Iteration (20 - 125)	0.3478 - 0.7826

Based on Table 1, while MLP achieved the highest accuracy (0.7826) than the other two algorithms, the lowest accuracy was 0.3478. Whereas, even though the highest accuracies from KNN and Random Forest Classifier are lower than that of MLP, they both had the lowest accuracies of 0.5625. Therefore, MLP required more tuning to get to a higher accuracy than the other two algorithms.

Table 2. Results of Using Frame Difference at One Interval Method (3B)

Algorithms	Hyperparameters	Accuracy
Random Forest Classifier	Estimators (1-13), Max Depth (10-40)	0.3125-0.6875
KNN	#Neighbors (1-10)	0.125-0.5
MLP	Learning rate (0.01-0.1), Max Iteration (20 - 125)	0.3158 - 0.5625

Referring to Table 2, Random Forest Classifier and KNN both had the same difference between the highest and lowest accuracy (0.375), whereas MLP had a slightly smaller range (0.2467). Therefore, for this solution, MLP required less tuning than Random Forest Classifier and KNN.

Table 3. Results of Using Frame Difference at Several Intervals Method (3C)

Algorithms	Hyperparameters	Accuracy
Random Forest Classifier	Estimators (1-13), Max Depth (10-40)	0.7826 - 1
KNN	#Neighbors (1-10)	0.4783 - 0.9348
MLP	Learning rate (0.01-0.1), Max Iteration (20 - 125)	0.4783 - 0.9565

Referring to Table 3, Random Forest Classifier had the lowest accuracy range (0.2174) compared to KNN and MLP, which had accuracy ranges of both 0.4565 and 0.4782 respectively. These ranges are both more than double that of Random Forest Classifier. Therefore, for this solution, Random Forest Classifier required the least amount of tuning.

5. DISCUSSION

The three different algorithms – KNN, MLP, and Random Forest Classifier – all had around the same accuracy for each method of data processing. For all three algorithms, taking the frame difference at several different interval shifts of each video resulted in the highest accuracy, which was above 90% according to Fig. 4. Whereas, taking the frame difference at only one interval of each video resulted in the lowest accuracy (50 - 68.75%) for each of the algorithms which can be seen in Fig. 3.

Having the frame difference at three different interval shifts per video tripled the sample size to 228 and also increased the test size to around 46. There could have been a higher accuracy because there was more training data (around 182) to practice on and more samples to test on. Having more frame shifts at different time locations of the movement provides more insight on correct versus incorrect grand plies. In Fig. 4, the algorithms performed the same with and without SMOTE. Even though there were 24 more sample videos of correct grand plies than incorrect grand plies, the sample sizes for each, both over 100, could have been large enough for the algorithm to not be favored toward correct grand plies.

A limitation of the study was the small sample size. There were only 76 total video clips – 42 of dancers correctly doing a grand plie and 34 of dancers incorrectly doing grand plies. Even though a difference of 8 samples would not be much for a larger dataset, for a smaller sample size like 76, this difference could be impactful enough to introduce bias favored toward correct grand plies, which could have been the case in Fig. 2 and Fig. 3. For these charts, Random Forest Classifier and KNN performed slightly worse with SMOTE where the number of correct and incorrect samples is the same. Additionally, due to a small sample size, the test set, which was 20% of the sample size or around 15, was also small, which did not give as much variation in the accuracies. However, increasing the test set would simultaneously decrease the training set, so the algorithm may not have many samples.

6. Conclusions and Future Work

To allow more online dancers to be successful and to combat the possible injuries that could arise with online dance classes, we built an AI system that detects correct and incorrect grand plies to a high level of accuracy.

SMOTE was used to account for the imbalance of correct and incorrect videos. We then ran it through three algorithms – Random Forest Classifier, KNN, and MLP — and hypertuned parameters. Taking the frame difference of frames at multiple interval shifts was the most successful. Capturing the position of the dancer at three different time locations possibly increased the robustness of the AI system performance. The three different algorithms performed similarly in terms of accuracy for the three different methods.

In the future, it would be interesting to combine our method of capturing the frame intervals with capturing the middle frame of the videos to determine if the results differ. Further work could also be done to identify the ideal frame intervals and the number of frame shifts for each dance move.

This method could also be tested to detect the correctness of more complicated dance moves, such as a turn, a leap, or even a series of dance moves, such as a grand plie and a turn in one video clip. Furthermore, given that MediaPipe Pose permits this, our solution could be deployed as a mobile app or a website that could be compatible with Zoom and other virtual school platforms.

References

[1] M. Alricsson, K. Harms-Ringdahl, K. Eriksson, and S. Werner, "The effect of dance Training on Joint Mobility, Muscle Flexibility, Speed and Agility in Young Cross-country Skiers - a Prospective Controlled Intervention Study." *Scandinavian Journal of Medicine & Science in Sports*, vol. 13, no. 4, 14 July 2003, pp. 237-43. *National Library of Medicine*, https://doi.org/10.1034/j.1600-0838.2003.00309.x. Accessed 6 July 2022.

[2] A. Maraz, O. Király, R. Urbán, M. D. Griffiths, and Z. Demetrovics, "Why Do You Dance? Development of the Dance Motivation Inventory (DMI)." *PLOS ONE*, vol. 10, no. 3, 24 Mar. 2015, p. e0122866. *National Library of Medicine*, https://doi.org/10.1371/journal.pone.0122866. Accessed 6 July 2022.

[3] D. Lange. "Dance, step, and other choreographed exercise participants US 2021." *Statista*, February 2022, https://www.statista.com/statistics/756629/dance-step-and-other-choreographed-exercise-participants-us/. Accessed 6 July 2022.

[4] "Steezy Studio: Reach Your Dance Goals." *STEEZY Studio | Reach Your Dance Goals*, https://www.steezy.co/#levels.

[5] "Pros and cons of doing dance classes online - go&dance." *Go&Dance*, https://www.goandance.com/en/blog/lifestyle/72-pros-and-cons-of-doing-dance-classes-online. Accessed 6 July 2022.

[6] Bellerose, Samantha. "Should you enroll your child in online virtual dance lessons?" *Dance Parent 101*, https://danceparent101.com/should-you-enroll-your-child-in-online-virtual-dance-lessons/. Accessed 6 July 2022.

[7] A. Malkogeorgos, F. Mavrovouniotis, G. Zaggelidis, and C. Ciucurel, "Common dance related musculoskeletal injuries." *Journal of Physical Education and Sport*. *Research Gate*, www.researchgate.net/publication/287785890_Common_dance_related_musculoskeletal_injuries. Accessed 6 July 2022.

[8] H. Kantilaftis. "Ballet and modern dance: using ballet as the basis for other dance techniques." *New York Film Academy*, 5 August 2014, https://www.nyfa.edu/student-resources/ballet-and-modern-dance/. Accessed 6 July 2022.

[9] T. Tellier. "Plié power." *Everyday Ballet*, 10 June 2017, https://www.everydayballet.com/plie-power/. Accessed 6 July 2022.

[10] I. Levendosky. "The ups and downs of demi and grand plies." *Front Range Classical Ballet Academy*, 14 October 2018, https://frcballet.com/blog/2018/10/11/the-ups-and-downs-of-demi-and-grand-plies. Accessed 6 July 2022.

[11] J. Tomás, A. Rego, S. Viciano-Tudela, and J. Lloret. "Incorrect facemask-wearing detection using convolutional neural networks with transfer learning." *Healthcare*, vol. 9, no. 8, 16 Aug. 2021, p. 1050, https://doi.org/10.3390/healthcare9081050.

[12] UNT Health Center. "Using 3D motion technology to prevent injuries in professional dancers." *The DO*, 8 October 2018, https://thedo.osteopathic.org/2018/10/using-3d-motion-technology-to-prevent-injuries-in-professional-dancers/. Accessed 6 July 2022.

[13] "Woah! You Can Really Dance!" *Woah*, https://www.woah.ai/.

[14] M. Trajkova. "Designing AI-Based Feedback for Ballet Learning." *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (2020): n. pag.

[15] Z. Cao, G. Hidalgo, T. Simon, S. Wei, and Y. Sheikh. "OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields." *Arxiv*. *Cornell University*, https://doi.org/10.48550/arXiv.1812.08008. Accessed 19 Jan. 2023.

[16] A. Faridee, S. R. Ramamurthy, H. M. S. Hossain, and N. Roy, "HappyFeet: recognizing and assessing dance on the floor." *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications* (2018): n. pag.

[17] A. Luna, et al. "Artificial Intelligence application versus physical therapist for squat evaluation: a randomized controlled trial." *Scientific Reports*, vol. 11, no. 1, 13 Sept. 2021, https://doi.org/10.1038/s41598-021-97343-y.

[18] S. Ibrahimi. "Composition, performance and evaluation: a dance education framework for AI systems." International Conference on Computational Creativity 2021, 2021.

[19] You, Yuhui. "Online Technologies in Dance Education (China and Worldwide Experience)." *Research in Dance Education*, 14 Oct. 2020, pp. 1-17. *Taylor and Francis Online*, https://doi.org/10.1080/14647893.2020.1832979. Accessed 9 Sept. 2022.

[20] "Pose-Mediapipe."*Google*, https://google.github.io/mediapipe/solutions/pose.html. Accessed 6 July 2022.

[21] V. Bazarevsky, et al. "BlazePose: on-device real-time body pose tracking." *arXiv e-prints*, https://doi.org/10.48550/arXiv.2006.10204. Accessed 6 July 2022.

[22] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, *SMOTE: Synthetic Minority Over-sampling Technique*. ArXiv, 2011. *Arxiv*, https://doi.org/10.48550/arXiv.1106.1813.

[23] "sklearn.ensemble.RandomForestClassifier—scikit-learn 1.1.1 documentation." *Scikit-learn*, https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html. Accessed 7 July 2022.
[24] "sklearn.neighbors.KNeighborsClassifier — scikit-learn 1.1.1 documentation." *Scikit-learn*, https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html. Accessed 7 July 2022.
[25] "sklearn.neural_network.MLPClassifier — scikit-learn 1.1.1 documentation." *Scikit-learn*, https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html. Accessed 7 July 2022

AUTHORS

Catherine Hung is a senior at Palo Alto High School in Palo Alto, California. Her research interests are artificial intelligence and machine learning.