A Feature Importance Study in Ballet Pose Recognition with OpenPose

Margaux Fourie and Dustin van der Haar(✉)

University of Johannesburg, Kingsway Avenue and University Rds,
Auckland Park, Johannesburg, South Africa
{margauxf,dvanderhaar}@uj.ac.za

Abstract. Movements of the human body can finally be recognised and analysed using computer vision technology. Ballet is an activity that involves various movements and specific poses of the body, making it an attractive candidate for computer vision applications. This paper proposes a feature importance study for determining which body parts play the most significant role in ballet pose recognition. The study is based on the use of OpenPose for feature extraction together with Support Vector Machine, Random Forest and Gradient Boosted Tree classifiers. Recognition accuracies above 95% suggest that the methods are not only feasible but exhibit excellent results. The results also indicate that the body parts that were the most significant for the classification of ballet poses were those situated at the extremities of the body such as the wrists and feet. The study addresses challenges within the ballet domain as it relates to both training and choreography. Furthermore, the study confirms that as technology expands into all areas of life, it is worthwhile to explore the possibilities within artistic fields.

Keywords: Computer vision · Feature importance · Ballet · OpenPose

1 Introduction

An ever-increasing amount of application fields are using computer vision to assist and enhance activities within those domains. Human body movements and poses can accurately be recognised using recent advancements in graphics processing technologies and computer vision algorithms [1,2]. Ballet is a human activity that is especially attractive for computer vision due to the well-codified poses and the limited automated approaches that exist in the environment. The automatic identification of body parts that are significant for the recognition of different ballet poses becomes a relevant research problem considering the challenges that are present in ballet training and choreography.

Ballet has developed over multiple centuries, and its various established poses have become foundational elements of the art form [3]. It is, therefore, a frequent task in a ballet training environment for teachers and students to recognise and correct the poses being performed. To avoid bad training habits and injuries,

dancers need to be aware of the proper placement of different parts of the body when performing ballet poses [4]. There is a need for additional forms of training critique to avoid the development of flawed technique, which often results in injuries [4].

Ballet choreography is another area within ballet where pose recognition is relevant. Choreographers are responsible for creating dance pieces that are constructed by using sequences of poses. There is a need for the most important poses and body parts used in a choreographed piece to be determined and documented to reproduce created works with future generations effectively [5]. Both the areas of training and choreography in ballet have the potential to benefit from technological solutions to assist in correct training and the proper documentation of ballet choreography.

This paper proposes an approach that is based on a previous study [6] that has been completed for ballet pose recognition. Once distinct poses are recognised using computer vision methods, it is possible to determine the most important features used during classification. In turn, the particular parts of the body that played the most prominent part for the classification of the pose can be identified. The paper first provides information on the problem background along with current related work. The experiment setup is presented next, followed by the model. The results are then provided, and the paper ends with a discussion on future work and a conclusion.

2   Problem Background

Ballet has a vibrant historical dimension that reaches back to the 16th century [7]. Every step in a ballet class is, therefore, ingrained with centuries of traditions and adaptations [8]. In addition, ballet has also formed the basis for many other forms of dance [9]. Due to the well-established technique that is prescribed in ballet, it is a relevant application area for computer vision.

Ballet technique is a term which is used to describe the essential ingredients that enable a dancer to achieve the aesthetic appearance of poses and movements. Technique in ballet is mainly concerned with the proper placement of the different parts of the body. It involves concepts such as turnout, which is the outward rotation of the legs for a more appealing view of the legs and feet [3,10]. Alignment is another aspect of technique that refers to the vertical and horizontal lines of the shoulders and hips. In addition, stretched legs and feet are always emphasized in the building of a strong ballet technique [3].

Ballet training usually takes place in a studio with a class of students that are instructed by a ballet teacher. Advancement in ballet training has always relied on the verbal passing on of expertise from teachers to younger generations [11]. However, the traditional classroom approach presents the challenge of a lack of one-on-one attention that the students receive from the teacher [12]. There is a need for guided one-on-one training and correction, which has the potential to help dancers improve their skills. Furthermore, additional forms of direct feedback may create a better awareness of correct placement and prevent injuries caused by incorrect technique [4].

Choreography is another aspect that plays an important role within the ballet domain. It involves the construction of dance sequences that consist of a series of codified poses and movements [13]. The challenge that choreographers face is the accurate documentation of created works in order to ensure its preservation [14]. There is, therefore, an opportunity within the choreographic domain to explore the most significant poses and body parts that are used in dance pieces.

Since ballet technique is largely concerned with the placement of different body parts, the automated approaches that exist in this environment focus on the extraction of key body information. The various related research efforts in the ballet and technological domain are discussed in tue following section.

3   Related Work

A related area of technology that has been investigated as it applies to the ballet domain is wearable garments. An approach by Gupta et al. focused on the instruction of beginner adult ballet dancers who had a teacher demonstrate ballet movements wearing a full-body garment [15]. The garment would light up the essential body parts being used by a teacher during the demonstration. This study had the advantageous effect of enabling students to focus on the most important key-points instead of complex technicalities. Some of the limitations that this system had include the high cost of such a wearable garment and the restriction it placed on movements [15]. These limitations indicate that there is a gap for systems that are cost-effective and less restrictive.

Research has also been conducted in the area of ballet choreography by Dancs et al., which aimed to automate the recognition and recording of a choreographer's movement [16]. The study used Microsoft's Kinect sensor in order to detect the different joints of the body. Furthermore, the study used classification algorithms such as Nearest Neighbor as well as Support Vector Machine (SVM) methods, which produced promising results with over 90% for accuracy. The success of the approach by Dancs et al. indicates that it is worthwhile to explore how computer vision methods may contribute to addressing challenges in ballet choreography.

Related research that closely links to the work of this paper includes a fairly recent posture recognition system by Saha et al., which included 20 ballet poses as primitives [17]. The system made use of pre-processing methods involving skin color segmentation to arrive at minimised skeletons of the initial images. The mathematical Radon transform method was used to calculate line integral plots and ultimately match images to specific primitives for recognition. The system produced a promising recognition rate of 91.35%. Furthermore, Saha et al. indicated that the area of ballet pose recognition is fairly young with a variety of opportunities for future research [17]. There is, in particular, an opportunity to build on this work by looking at improved and recent ways to extract skeleton key-point features such as utilising an OpenPose approach [18].

This study further relates to the optimisation of classifiers. When it comes to image classification, techniques that are often used to tune hyper-parameters and

ultimately improve accuracy include random search, grid search and Bayesian optimisation [19]. Feature selection is another method that is used to improve classifiers by identifying which features are the most relevant to a particular problem [20]. A feature importance study is, therefore, a valuable step towards better recognition accuracies.

4   Experiment Setup

The feature importance study proposed in this paper is based on previous research [6] completed by the authors on the recognition of ballet poses. The experiment setup for this paper, which is concerned with feature importances, is therefore similar to the setup that was used for the completed pose recognition study.

The study compiled a primary dataset containing thirty classically trained ballet dancers as subjects that were captured performing eight distinct ballet poses. These poses included Demi-Plié, Second Position, Tendu, Sussous, Retiré, Développé, Arabesque, and Penché. The dancers were captured using a Microsoft Kinect sensor and a GoPro camera, which enabled the collection of video, depth and image data.

The dataset for this study consisted of Microsoft Kinect images that were captured at a 640 by 480 resolution. The data of about 7200 images was split into a training set consisting of 80% of the images and a testing set containing 20% of the images. An even distribution among different classes was used in both the training and testing sets. The authors focused on using the collected image data for this work and will, therefore, utilise the depth and video data in future work.

The success of collecting quality data for this study required certain constraints to be in place. It was first important that capturing should occur in a dance studio with an appropriate dance surface such as ballet mats. Mirrors or any clutter in the capture space were to be removed to minimise noise in the background. Furthermore, the lighting had to be at a suitable level. Concerning the participants, a role constraint of the study included that they had to be advanced level dancers that could execute each of the poses with sound technique. Lastly, standard black ballet attire had to be worn by participants to avoid unnecessary variations in the gathered image data.

Once the data has been collected, it can be used to apply various computer vision methods. The next section will unpack the model, which makes use of the captured dataset as the starting point to achieve pose recognition and the determination of feature importances.

5   Model

The model of this study is shown as a pipeline in Fig. 1, which consists of four separate stages, namely capturing, feature extraction, classification and feature

importance. Each of these phases involves a set of methods or actions that need to occur before moving on to the next stage.

The model has the captured dataset as a starting point from which features are extracted during the second stage. The feature extraction method used in this model is known as OpenPose [18]. OpenPose is a recent and useful feature extraction approach that uses a multi-stage Convolutional Neural Network (CNN) for extracting human skeleton key-point data from images.
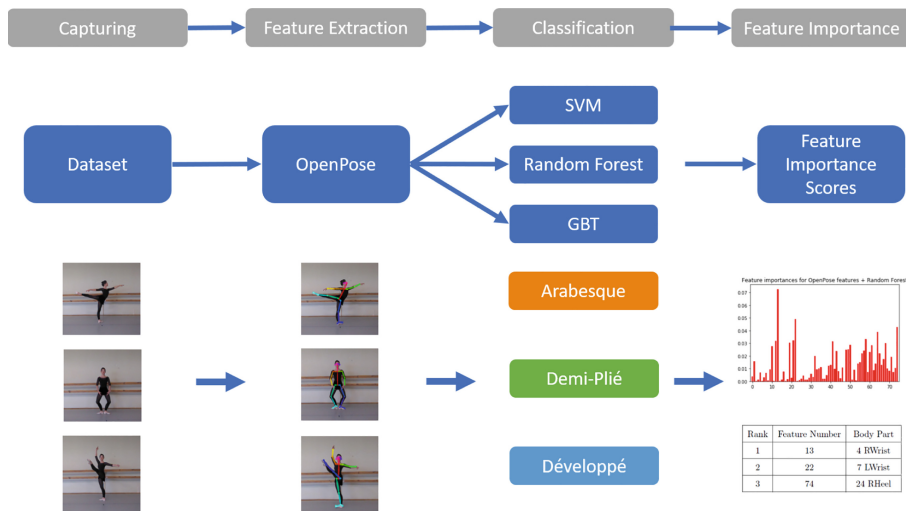

Fig. 1. Model for determining the most significant OpenPose features for different computer vision algorithms

The skeleton key-points can be seen in Fig. 2 with key body parts represented as numbers. Once the features have been extracted, different classification algorithms are utilised to perform training and testing on the dataset of ballet poses. The different classification methods that form a part of the model include Support Vector Machine (SVM), Random Forest (RF), as well as Gradient Boosted Tree (GBT). When training is completed, a feature importance study determines which OpenPose features played the most significant role in identifying particular poses. For the SVM model, the feature importances are determined by analysing the subsequent weights produced after training. For the RF and GBT classifiers, Gini importance is used, which computes how much each feature contributes to a decrease in node impurity. From the determined OpenPose key-point features, valuable insights are then derived by mapping the most important features to parts of the dancer's body.
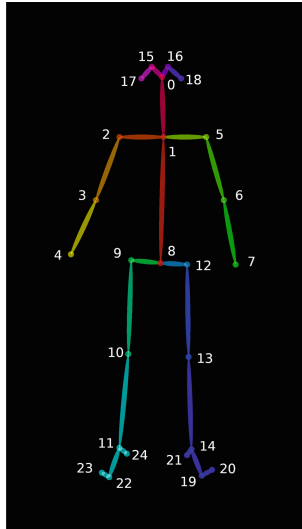
Fig. 2. Illustration of OpenPose skeleton key-point features [21]

6   Results

Before presenting the feature importance results from this study, a summary
of the recognition accuracies for each of the relevant pipelines are presented in
Table 1. The pipeline which achieved the best accuracy result was the OpenPose
and Random Forest variation with a score of 99.375%.

Table 1. Summary of the results obtained by the pose recognition study as percentages.

Pipeline variation	Accuracy
OpenPose + Random Forest	**99.375**
OpenPose + Gradient Boosted Trees	99.305
OpenPose + Support Vector Machine	99.097

Based on the accuracy scores achieved by the pose recognition study, it is
feasible to investigate how the OpenPose features impacted the identification of
different poses. There are a total of 75 features extracted by OpenPose when
an image is provided to the algorithm. The OpenPose features are based on 25
body parts which are represented in Fig. 2. The format in which the features
are output by OpenPose are: (x-coordinate, y-coordinate, confidence score) for
each of the 25 body parts which results in 75 features in total. It is therefore
possible to determine, based on the feature number, what type of feature it is

(x-coordinate, y-coordinate or confidence score) as well as what body part is associated with it. Table 2 shows the numbers and names associated with the OpenPose body parts.

Table 2. OpenPose body part numbers

Number	Body part	Number	Body part	Number	Body part
0	Nose	9	RHip	18	LEar
1	Neck	10	RKnee	19	LBigToe
2	RShoulder	11	RAnkle	20	LSmallToe
3	RElbow	12	LHip	21	LHeel
4	RWrist	13	LKnee	22	RBigToe
5	LShoulder	14	LAnkle	23	RSmallToe
6	LElbow	15	REye	24	RHeel
7	LWrist	16	LEye		
8	MidHip	17	REar		

For the extraction of meaningful information from the feature importance study, it is necessary to calculate a mapping between features and body parts. In order to determine which feature number is associated with which body part, the following algorithm has been constructed:

Input: f: feature number from 0 to 74
Output: b: the body part number that needs to be determined
Output: t: the feature type that needs to be determined
if $f \bmod 3 == 0$ **then**
$\quad | \quad b \leftarrow f \div 3$
$\quad | \quad t \leftarrow xCoordinate$
end
else if $f \bmod 3 == 1$ **then**
$\quad | \quad b \leftarrow (f-1) \div 3$
$\quad | \quad t \leftarrow yCoordinate$
end
else if $f \bmod 3 == 2$ **then**
$\quad | \quad b \leftarrow (f-2) \div 3$
$\quad | \quad t \leftarrow confidenceScore$
end

Algorithm 1: Mapping between features and body parts

6.1 Support Vector Machine Pipeline

The feature importances for the Support Vector Machine implementation is based on the feature weights that were produced and shown in Fig. 3. Valuable information can be extracted from the feature weight representation by making

use of Table 2 and the algorithm presented earlier in this section. The results from applying the mapping calculation for the SVM variation are presented in Table 3.
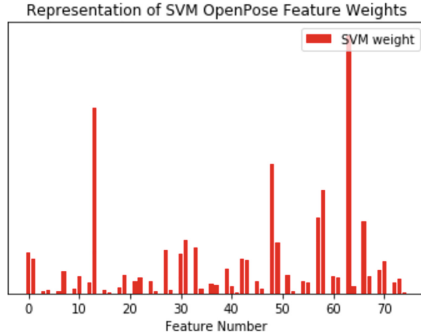

Fig. 3. Representation of feature weights for OpenPose + SVM

The top 10 most important body part features for the SVM pipeline are presented in Table 3. These results indicate that the three most significant body-parts for distinguishing between ballet poses using an SVM classifier are the left wrist, the right wrist and the left eye.

Table 3. Ranking for the top 10 important OpenPose features with SVM

Rank	Feature number	Body part	Feature type
1	63	21 LHeel	x-coordinate
2	13	4 RWrist	y-coordinate
3	48	16 LEye	x-coordinate
4	58	19 LBigToe	y-coordinate
5	57	19 LBigToe	x-coordinate
6	66	22 RBigToe	x-coordinate
7	31	10 RKnee	y-coordinate
8	49	16 LEye	y-coordinate
9	33	11 RAnkle	x-coordinate
10	27	9 RHip	x-coordinate

6.2   Random Forest Pipeline

The Random Forest Pipeline of this study made use of Gini importance to determine the most important features which are visible in Fig. 4. The most

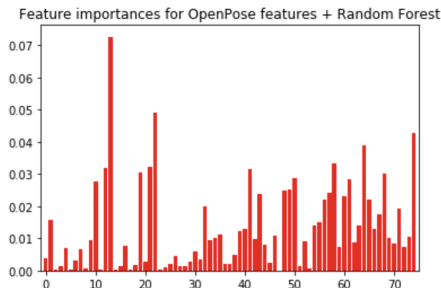noteworthy body parts that have been calculated for this pipeline are presented in Table 4.

Fig. 4. Feature importance representation for OpenPose + Random Forest

From Table 4 the top three most important body parts that play a role in the recognition task of the Random Forest classifier were the right wrist, the left wrist and the right heel. Each of these body parts is situated towards the ends of the limbs, which may indicate that parts that are further away from the body are more crucial for identifying poses using a Random Forest.

Table 4. Ranking for the top 10 important OpenPose features with Random Forest

Rank	Feature number	Body part	Feature type
1	13	4 RWrist	y-coordinate
2	22	7 LWrist	y-coordinate
3	74	24 RHeel	confidence score
4	64	21 LHeel	y-coordinate
5	58	19 LBigToe	y-coordinate
6	21	7 LWrist	x-coordinate
7	12	4 RWrist	x-coordinate
8	41	13 LKnee	confidence score
9	19	6 LElbow	y-coordinate
10	68	22 RBigToe	confidence score

6.3 Gradient Boosted Tree Pipeline

The Gradient Boosted Tree classifier also made use of Gini importance to extract important features as shown in Fig. 5. The body parts that had the highest importance for this GBT pipeline are presented in Table 5.
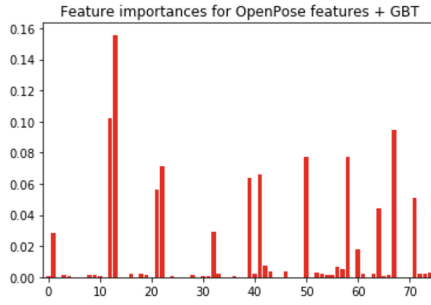
Fig. 5. Feature importance representation for OpenPose + GBT

From Table 5 it is clear that the right wrist as well as the right big toe had an important role to play in the distinction between ballet poses using a GBT classifier. This result gives an indication that body parts situated far away from the core of the body contribute the most when performing recognition using a GBT.

Table 5. Ranking for the top 10 important OpenPose features with GBT

Rank	Feature number	Body part	Feature type
1	13	4 RWrist	y-coordinate
2	12	4 RWrist	x-coordinate
3	67	22 RBigToe	y-coordinate
4	50	16 LEye	confidence score
5	58	19 LBigToe	y-coordinate
6	22	7 LWrist	y-coordinate
7	41	13 LKnee	confidence score
8	39	13 LKnee	x-coordinate
9	21	7 LWrist	x-coordinate
10	71	23 RSmallToe	confidence score

The results of the study indicate that the objective of determining the most significant body parts for recognising ballet poses with computer vision has been achieved. Generally, the skeleton points that had higher overall importance for distinguishing between different poses included the wrists as well as points situated in the feet or head. There is value in knowing which parts of the body carry a higher weight in determining how ballet poses may be classified as technology expands into this domain. It is especially relevant for ballet training where body parts with a higher feature importance score may need more emphasis during training to ensure the correct and improved execution of relevant poses.

For choreographers, a feature importance study has the potential to provide insight into which body parts play the most significant part in making the used poses and movements distinct from one another.

7 Future Work and Conclusion

This study has shown that it is possible to derive meaningful information from feature importance data gathered on different classifiers. The use of OpenPose key-point data contributed to effectively determine which parts of the body play a noteworthy role in the recognition of the poses chosen for this study. Some key findings in this paper indicated that the body parts situated further away from the centre of the body played a more significant role in the identification of different poses.

Future work for this study has the potential to contribute further to both the training and choreographic areas of ballet. For the ballet training environment, it would be valuable to build on the current study and consider the automatic correction of poses. On the choreographic side, this research may work towards the automated recording of dance sequences. Other computer vision techniques that are of interest for future work include various Convolutional Neural Network approaches, N-shot learning as well as Recurrent Neural Networks. Future work may also make use of video data in order to expand from static pose recognition to movement-based tracking and recognition.

The ballet field, with its deep historical roots and artistic associations, may seem to be in direct contrast with the modern, ever-growing field of technology. Despite the differences between the two fields, ballet's concrete underlying structure in terms of poses make it an ideal field to explore in conjunction with today's technological advancements [22]. Furthermore, technology enables the automation of tasks that were previously only performed by humans and it enables the discovery of new insights into the ballet art form. The field of ballet may, therefore, find value in the growth of automation through technology as it has the potential to serve as an enhancement tool for the improvement of current practices.

References

1. Nishani, E., Çiço, B.: Computer vision approaches based on deep learning and neural networks: Deep neural networks for video analysis of human pose estimation. In: 2017 6th Mediterranean Conference on Embedded Computing (MECO), pp. 1–4. IEEE (2017)
2. Yao, B., Hagras, H., Alhaddad, M.J., Alghazzawi, D.: A fuzzy logic-based system for the automation of human behavior recognition using machine vision in intelligent environments. Soft Comput. **19**(2), 499–506 (2014). https://doi.org/10.1007/s00500-014-1270-4

3. Royal Academy of Dancing. The Foundations of Classical Ballet Technique. Royal Academy of Dancing (1997)
4. Bronner, S., Ojofeitimi, S., Spriggs, J.: Occupational musculoskeletal disorders in dancers. Phys. Ther. Rev. **8**(2), 57–68 (2003)
5. Sperling, J.: How do you write down choreography?—jody sperling dance blog (2010)
6. Fourie, M., van der Haar, D.: Computer vision for the ballet industry: a comparative study of methods for pose recognition (2020)
7. Di Orio, L.: Ballet: Method to method. Dance Informa American Edition (2013)
8. Butterworth, J.: Dance Studies: The Basics. Routledge, Abingdon (2011)
9. New York Film Academy. Ballet and modern dance: Using ballet as the basis for other dance techniques (2014)
10. Kassing, G., Jay, D.M.: Teaching beginning ballet technique. Human Kinetics (1998)
11. Trajkova, M., Cafaro, F.: E-ballet: designing for remote ballet learning. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct - UbiComp 2016, pp. 213–216 (2016)
12. Dance Spirit. Working one-on-one: What to expect from private lessons (2014)
13. Speck, S., Cisneros, E.: Ballet for Dummies. Wiley, Hoboken (2003)
14. Snyder, A.F.: Securing Our Dance Heritage: Issues in the Documentation and Preservation of Dance. Council on Library and Information Resources, Washington DC (1999)
15. Gupta, M., Hallam, J., Keen, E., Lee, C., McKenna, A.: Ballet hero: building a garment for memetic embodiment in dance learning. In: Proceedings of the 2014 ACM International Symposium on Wearable Computers Adjunct Program - ISWC 2014 Adjunct, pp. 49–54 (2014)
16. Dancs, J., Sivalingam, R., Somasundaram, G., Morellas, V., Papanikolopoulos, N.: Recognition of ballet micro-movements for use in choreography. In: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1162–1167 (2013)
17. Saha, S., Konar, A.: Topomorphological approach to automatic posture recognition in ballet dance. IET Image Proc. **9**(11), 1002–1011 (2015)
18. Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., Sheikh, Y.: OpenPose: real-time multi-person 2D pose estimation using part affinity fields. arXiv preprint arXiv:1812.08008 (2018)
19. Bergstra, J., Bengio, Y.: Random search for hyper-parameter optimization. J. Mach. Learn. Res. **13**(Feb), 281–305 (2012)
20. Kuo, B.-C., Ho, H.-H., Li, C.-H., Hung, C.-C., Taur, J.-S.: A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **7**(1), 317–326 (2013)
21. Zhe, C., Hidalgo, G., Simon, T., Wei, S.-E., Sheikh, Y.: CMU-perceptual-computing-lab/openpose (2019)
22. LaViers, A., Chen, Y., Belta, C., Egerstedt, M.: Automatic sequencing of ballet poses. IEEE Robot. Autom. Mag. **18**(3), 87–95 (2011)