

# NEURAL NETWORK TO IDENTIFY KEY POINTS OF THE HUMAN POSE IN PHYSIOTHERAPY REHABILITATION OF GAIT

Edição 123 JUN/23 SUMÁRIO / 15/06/2023

REGISTRO DOI: 10.5281/zenodo.8044461

R. B. Sanchez<sup>1,2,3,4</sup>,  
J. C. L. Fernandes<sup>1,4</sup>,  
M. A. Bissaco<sup>1</sup>,  
S. R. M. S. Boschi<sup>1</sup>,  
T. A. Scardovelli<sup>1</sup>,  
A. P. da Silva<sup>1</sup>,  
S. C. Martini<sup>1</sup>

*This work proposes a digital system with cloud neural network with video acquisition to obtain key points of the human pose in patients undergoing physiotherapy rehabilitation of lower limbs in gait mobility, allowing agility in the records, minimizing measurement errors, analysis clinical condition, and the exchange of information between professionals in the field, resulting in a shared knowledge base. A digital camera, 1080p and 60fps, used parallel to the ground, with height adjustment (0.6 to 2.4m), 1.5m from the*

*analysis field (3x3m to 3x10m) is used for measurement, rehabilitation or walking procedures. Values manually entered by the physiotherapist are applied for calibration and definition of limits for comparison with the data obtained by the system. In relation to the current ones, the advantage is given by digitally obtaining anatomical points and body segments while in manuals they allow greater error and too much time in this collection. The system is an auxiliary tool for the physiotherapist to supply the manual limitation and sharing, with agility among professionals, historical data and qualitative analyzes that accurately generate a patient profile and allow the adaptation of the procedures applied with agility, since it provides an immediate graphic analysis.*

*Introduction:* In Brazil, the “2010 Census Booklet for People with Disabilities” [1] by the Brazilian Institute of Geography and Statistics (IBGE), launched by the Human Rights Secretariat of the Presidency of the Republic in 2012, says that about 45,606,048 Brazilians, 23.9 % of the total population, have some type of disability, of which the visual represents 18.6% of the population, with the greatest impact, followed by the motor that reaches 7% of the population.

Motor or physical disability is a complete or partial limitation of the functioning of parts of the human body, easily identified in the lower and upper limbs and classified according to neurological or muscular origin, which requires from the health area several treatment methods characterized by the degree of severity of the disorder, the great challenge for physiotherapy being the solution to problems of functional mobility, with rehabilitation being a time-consuming procedure that performs several repetitive functions, however daily several new techniques appear in the rehabilitation process [2-6].

The human body is a single complex structure formed by different structures subject to being affected by different types and degrees of injuries [7], where individual treatments have a high operational cost [8] that has been replaced by collective sessions with groups of patients, however, biomechanics depends on experimental results, where the acquisition of values is a concern of the applied methodologies [9], which can be obtained with techniques such as kinematics, electromyography, dynamometry and anthropometry [10, 11], considering the complexity to obtain biomechanical parameters that allow defining anthropometric models that will be used by kinematics [12], in which two images are sufficient for one model, one in the sagittal plane and one in the frontal plane [13].

*Methods:* The digital system was developed with the resources shown in software and hardware implementation, in which it was applied for tests with volunteers present at the Polyclinic of the University of Mogi das Cruzes (UMC), according to authorization with a signed Authorization Term of Co-Participating Institution (TAIC) by the Polyclinic Coordinator Dr. Melquiades Machado Portela, in which they were separated into a Control Group (CG) with healthy participants and an Experimental Group (EG) of patients with reduced mobility of the lower limbs due to neurological and / or muscle (orthopedic) trauma, but that allow rehabilitation procedures. All participants complete the Free and Informed Consent Form (ICF) and the Term of Use of Voice Image and Sound.

The digital system with the digital camera in 1080p and 60fps used parallel to the ground, with height adjustment (0.6 to 2.4m), 1.5m from the analysis field (3x3m to 3x10m), performs the measurement of the patient standing or walking through the demarcated area, used for measurement, rehabilitation or walking procedures, through computer vision with machine learning identifying and recreating key points and body segments from pre-trained models that will be analyzed with a convolutional neural network without the need for markers on the skin

To enable the application and functionality of the developed digital system, two bases of pre-trained models were used, the MPII Human Pose Dataset which

produces 15 key points in the human body [14] and the COCO CAFFÉ model which produces 18 points. An important feature of these models is that they were trained with Convolutional Neural Networks (CNN), which favors the application, as it allows to recognize patterns [15], the machine learning being defined by the variation of the parameters associated with the variables and by its architecture and structure allow multi-layer training [16].

CNN are generally the most efficient for image and object recognition applications, allowing to apply in the recognition and classification of images of the human body [17], where VGGNet for human can recognition has advantages for having a deeper structure and small nuclei with an improved recognition rate in relation to other neural networks, including VGGNet was the CCN used in the MP11 and COCO Caffee models, extracting the confidence and affinity maps for key points [18-21].

In biomechanics, anthropometric measurements are standardized and referenced by anatomical points that allow better positioning of measuring instruments during anthropometric assessment [22, 23], where the measurements were obtained in the anatomical position of the human being, that is, the position of the body alive, standing upright, with arms across the body and palms facing forward.

In this way, the digital system with VGGNet makes the prediction of the two-dimensional Confidence Maps of the key points defined by the MP11 and COCO Caffee models [24] forming a data set for the estimation of the human pose with one or more people in the same frame allowing, in addition to jointly predicting the Affinity Fields to associate body points and segments by code.

*Research Ethics Committee:* The protocol applied in the Research Project according to the premises of Resolution 466/12 of the National Health Council, was submitted to the Ethics Committee of the University of Mogi das Cruzes (UMC) in its second version dated November 1 of 2019 and approved according to the consubstantiated opinion 3,693,711 issued by this CEP on November 8, 2019 under number CAAE 22125419.4.0000.5497.

*Software and hardware implementation:* The digital system considers an external webcam connected to the computer or notebook, allowing mobility and better positioning of the camera, for a better definition of digital images and video capture considers the minimum resolution of 1280×720 (ideal Full HD in 1920×1080) for better sharpness and minimum 30fps to avoid sudden movement and greater smoothness of movement. Tested the system with Logitech, model C922 HD Pro, with glass lens, recording in full HD 1080p and 30fps, 78° field of view.

The source code was developed in Python 3.6 with the OpenCV 4.2.1, Matplot and NumPy libraries, operating with Google Cloud architecture using Google Drive and in the Google Collaboratory environment that allows the creation of virtual machines based on Jupiter Notebook, stored in '.ipynb' format and making available via Google servers processing with NVIDIA® Tesla® K80 GPU acceleration and 13.7 gigabytes of available RAM to run in real time.

*Test of the developed digital system:* It is important that in the visual field of the camera is only the patient and / or the professional who will conduct the measurement procedures, as it prevents important points from being hidden. These obstructions in the visualization can be identified by the algorithm of the digital system that will estimate the movement and the position of the anatomical points, that is, of the key points defined for the human body, thus predicting the human pose.

Once the capture camera is positioned, the professional who conducts the session must manually start the system through the link made available to Google Colab establishing the connection to Google Drive through a single and encrypted validation, load the pre-trained models, perform the routines that they will access the folders to read and write the input and output files, access the camera to allow the capture of images and videos and prepare the storage of the digitally processed results.

The affinity map for the anatomical points processed by CNN considering the pre-training of the MPII and COCO CAFFEE will be generated allowing to virtually

restore the silhouette of the human skeleton by reconstructing the body segments based on the identified key points. Each identified point will be expressed on a scale on a heat map that indicates how much each point in the same region represents in affinity with the real point to be identified, expressing it with the best accuracy and precision, the output processed for the map of affinity and correlation with the confidence map is obtained from the code below.

```
ponto = 12
mapa_confianca = saida[0, ponto, :, :]
mapa_confianca = cv2.resize(mapa_confianca,
                             (imagem_largura, imagem_altura))
plt.figure(figsize = [14, 10])
plt.imshow(cv2.cvtColor(imagem, cv2.COLOR_BGR2RGB))
plt.imshow(mapa_confianca, alpha = 0.6)
plt.colorbar()
plt.axis("off")
```

Having identified the key points from the affinity map with the association of the MPII and COCO CAFFE models, the convolutional neural network will define the points found and the algorithm will express the location by the MPII output format, being:

- (0) Head, (1) Neck
- (2) Right shoulder, (3) Right elbow, (4) Right wrist• (5) Left shoulder, (6) Left elbow, (7) Left wrist
- (8) Right hip, (9) Right knee, (10) Right ankle
- (11) Left hip, (12) Left knee, (13) Left ankle
- (14) Chest
- (15) Background

Therefore, by executing the routines of the developed digital system, 15 points are obtained from the pre-trained models, where through the code snippet below it is possible to normalize their identification and define the affinities based on the real and physical structure of the human body.

```
pares_pontos = [[0,1],[1,2],[2,3],[3,4],[1,5],[5,6],[6,7],[1,14],[14,8],
                [8,9],[9,10],[14,11],[11,12],[12,13]]
```

The COCO CAFFE and MPII models through CNN allow to identify each x and y coordinate of each point by assigning a weight value to identify '0' for x and '1' for y.

```
for par in pares_pontos:
    #print(par)
    parteA = par[0]
    parte = par[1]
```

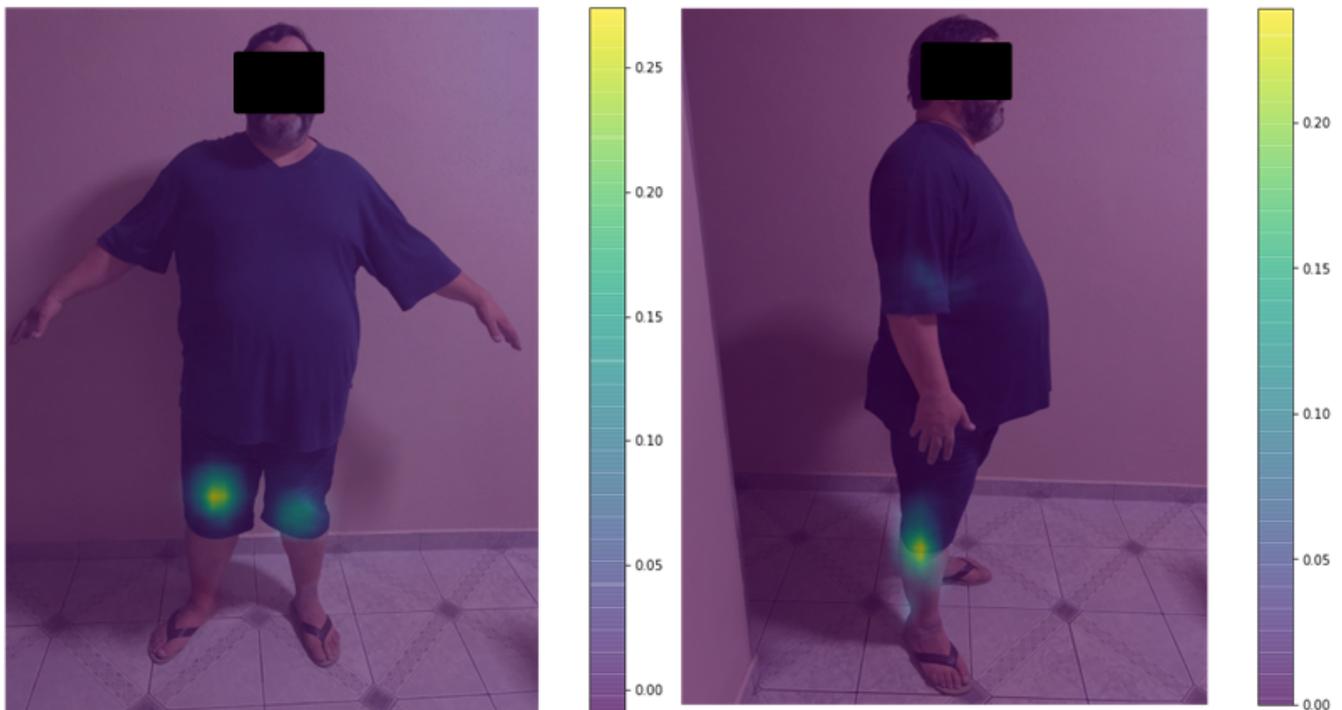
By means of the subroutine below, a mask with points and lines is generated to point out the visualization and the interconnection of each key point.

```
if pontos[parteA] and pontos[parteB]:
    cv2.line(imagem, pontos[parteA], pontos[parteB], cor_linha, 3)
    cv2.circle(imagem, pontos[parteA], 8, cor_ponto, thickness = -1,
               lineType = cv2.LINE_AA)
    cv2.line(mascara_mapa, pontos[parteA], pontos[parteB], cor_linha, 3)
    cv2.circle(mascara_mapa, pontos[parteA], 8, cor_ponto, thickness =
               -1, lineType = cv2.LINE_AA)
```

The reconstruction of the body segments occurs through the union of the identified points and as long as these are physically related, allowing the virtualization of the skeleton to be extracted as a mask superimposed on the image and on a black background.

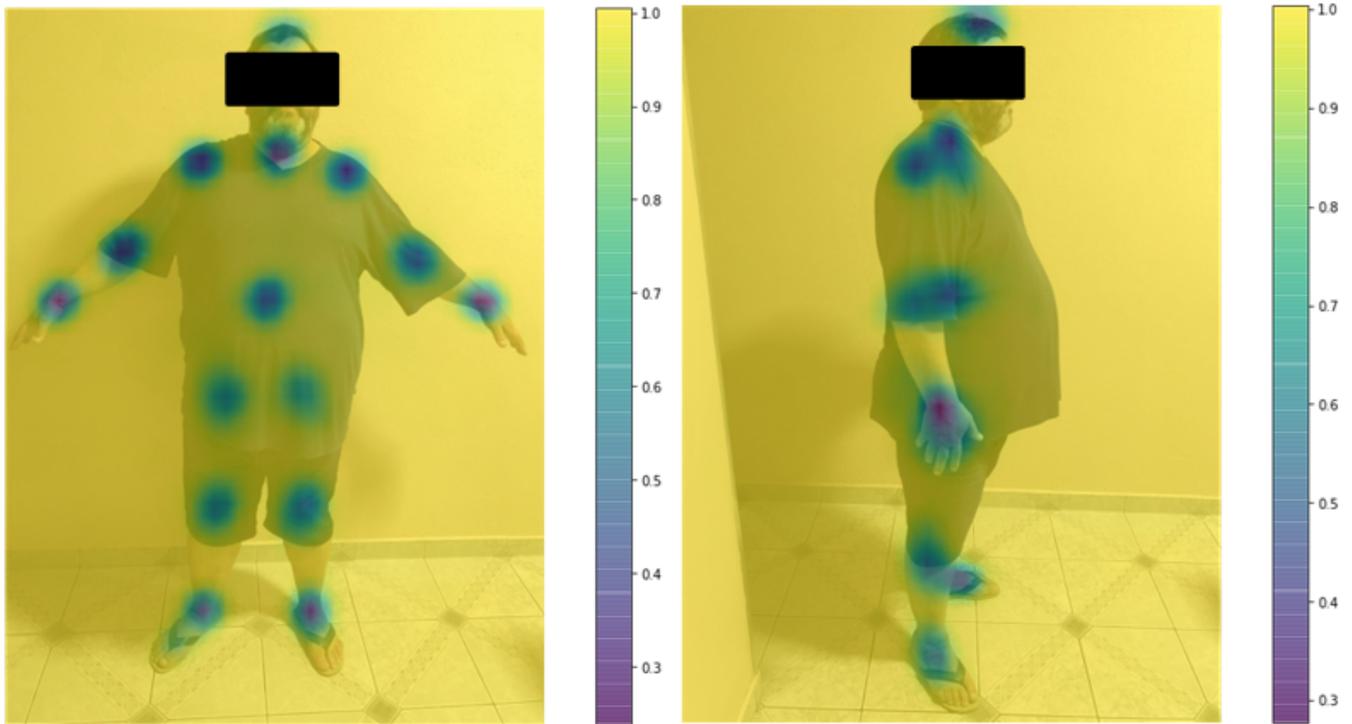
*Results and Discussion:* One of the advantages of the digital system stored in Google Drive and with Jupiter Notebook hosted on Google servers, occurs because it does not compromise the RAM memory and processing capacity of the Laptop, operating with generic processing directly from Google servers or its TPU and GPU, in tests with the GPU the processing was accelerated by making available with free access from Google Colab, hardware 'NVIDIA-SMI 450.66, Driver Version: 418.67 and CUDA Versions: 10.1' with 'Your runtime has 13.7 gigabytes of available RAM'.

As a first step towards fast processing, all input images and frames are scaled to 480p regardless of the original input size (ideal 1080p, minimum 720p), as this way it is possible to maintain the processing quality without losing image or frame characteristics. processed. In the preliminary tests of the digital system, images of an adult, male gender, chronological age of 54 years, body mass of 135kg and height of 1.79m were used, considered from the control group and healthy as to the mobility of the lower limbs, Fig 1 exposes an output with a scale 0 to 1 of the Affinity Map with focus, for example, on the right knee, this definition being important for the algorithm to refine the precision considering this point in the analysis of the lower limbs, in this case specifically for the lower limb right, although this point can be generalized for both members when considering the pelvic point.



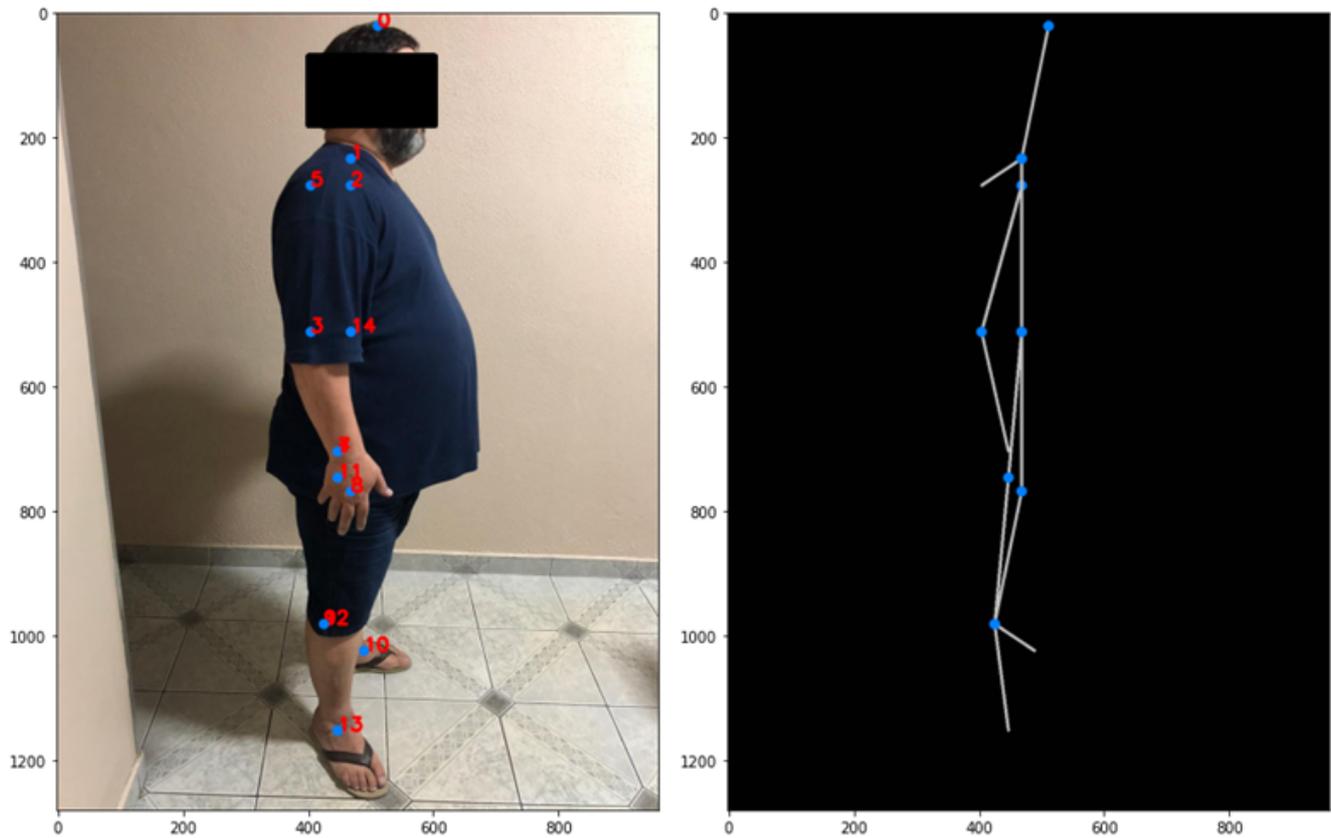
**Fig. 1** Images obtained in the sagittal and coronal planes with a focus on the right knee.

From the Affinity Map, the correspondence of the key points, with a scale from 0 to 1, of the MPII model, Fig. 2, was generated for the system test, where both pre-trained models had been loaded and processed in the analysis using the CNN.



**Fig. 2** Heat map and affinity of key points in the individual.

The result of the identification of the key points allows finding the Cartesian coordinate pairs to generate the virtualized spatial position of the anatomical points and respective body segments as shown in Fig. 3, where it is possible to numerically identify the points for the MPII model and the virtualization on a black background, however, in the sagittal plane there is naturally an obstruction in the visualization to the right side of the body for both computer and professional vision, so the digital system predicts by CNN and the parameters extracted from trained models, the position of key points not visible, thus the reconstruction of body segments.



**Fig. 3** Key points and virtualization of the skeleton in partial occlusion.

This identification of the unseen point 6, the left elbow, is possible because the adjacent points 7 and 5, respectively of the wrist and shoulder, were discovered, so in the processing the segment can be virtualized connecting the points.

Table 1 shows the location data for the points obtained based on Fig. 3 according to CNN processing in relation to the pre-trained models Mpii and COCO CAFFE, where it presents the positions in pixels and millimeters, considering the constant of 1px as a parameter. for each 0.26458mm for a resolution of 96dpi and its multiples, which must be adjusted to calibrate the system according to each type of application.

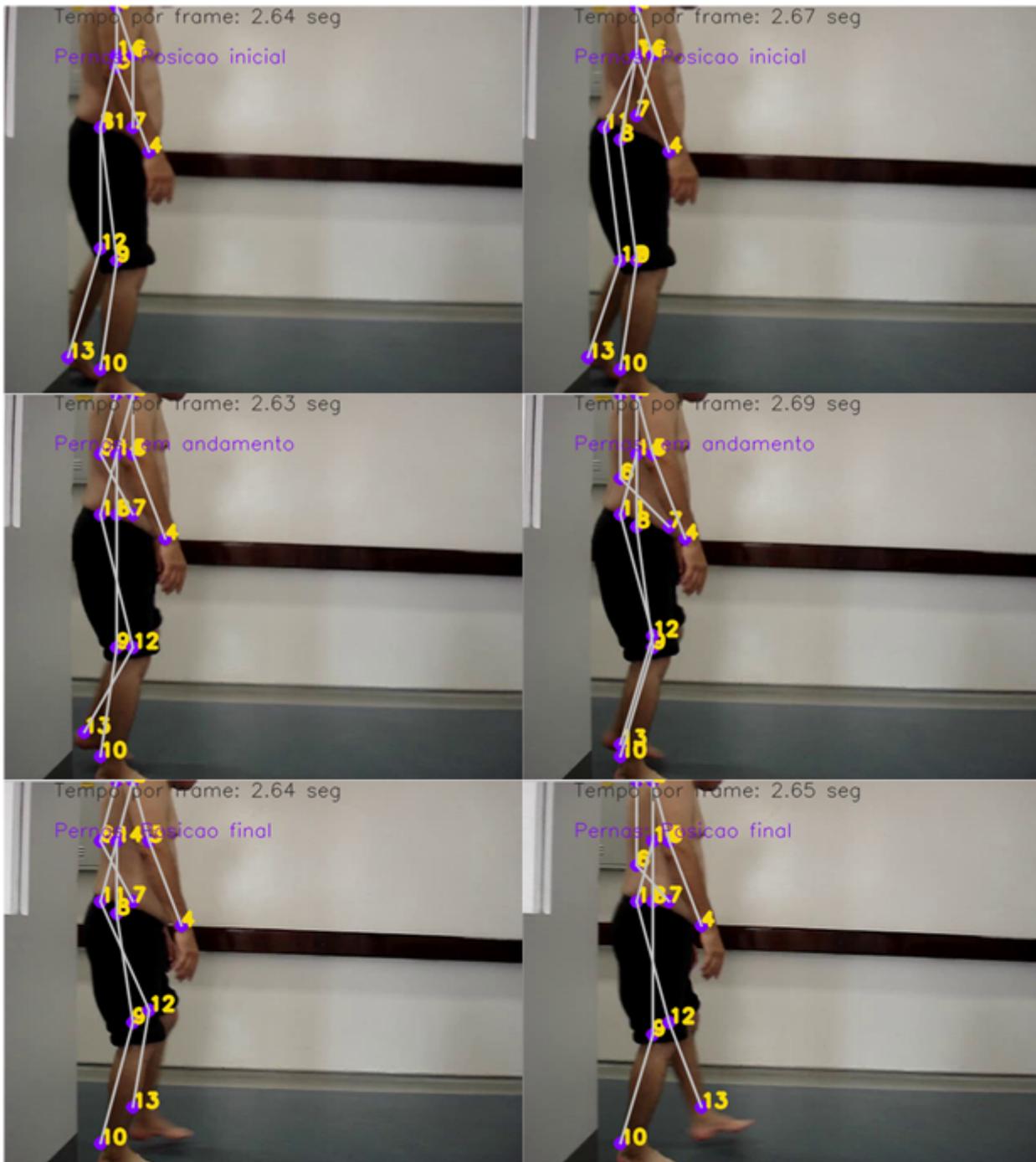
**Table 1:** List of the positions of the key points identified in Fig. 3.

KEY POINTS		CARTESIAN POSITION	
ID N°	DESCRIPTION	Fig. 3	
		px	mm

0	Head	(512 21)	(135,47 5,56)
1	Neck	(469 234)	(124,09 61,91)
2	Right shoulder	(469 277)	(124,09 73,29)
3	Right elbow	(405 512)	(107,16 135,47)
4	Right wrist	(448 704)	(118,53 186,27)
5	Left shoulder	(405 277)	(107,16 73,29)
6	Left elbow	<i>None</i>	<i>None</i>
7	Left wrist	(448 704)	(118,53 186,27)
8	Right hip	(469 768)	(124,09 203,20)
9	Right knee	(426 981)	(112,71 259,56)
10	Right ankle	(490 1024)	(129,65 270,93)
11	Left hip	(448 746)	(118,53 197,38)
12	Left knee	(426 981)	(112,71 259,56)
13	Left ankle	(448 1152)	(118,53 304,80)
14	Chest	(469 512)	(124,09 135,47)

Following the test, a video with a patient in rehabilitation was submitted to the digital system with gait dysfunction, that is, there is the incidence of pathological gait and the system performs, as well as in the previous adult, all processing with CNN and the MPII and COCO CAFFEE models, which will generate the Affinity and Trust Maps, Fig. 4, as well as plot a mask in the video frames with the identified key points and proceed with the virtualized connection of the pairs of points to generate the segments.

A second analysis to be seen in the video occurs with the identification by the algorithm with CNN that classifies the movement and its development, marking the condition of the execution in: initial, progress and final.



**Fig. 4** Execution of the analysis in pathological gait.

*Conclusion:* The digital computer vision system with the application of convolutional neural networks to identify the body pose opens a discussion regarding the current available applications and presents the coherent advantage in identifying the key points and the reproduction of the virtualized body segments from the capture of videos and images, regardless of the type

and model of digital camera used, maintaining the registration file both as a mask in the original frame and in reproduction with a black background exposing only the points and segments, predicting by the algorithm to predict the position of points outside the camera's visual field or the professional.

The system allows a low-cost application, as it does not depend on a specific camera and computer with high processing capacity or a specific computational environment, being flexible and allowing pairs of points to analyze sizes and distances, in addition to the ease of obtaining data to easily train models through machine learning and analyze angular and kinematic parameters in the future.

## References

1. CENSO, Cartilha do. Pessoas com deficiência. **Luiza Maria Borges Oliveira/Secretaria de Direitos Humanos da Presidência da República (SDH/PR)/Secretaria Nacional de Promoção dos Direitos da Pessoa com Deficiência (SNPD)/Coordenação-Geral do Sistema de Informações sobre a Pessoa com Deficiência**, 2010. Disponível em: <<http://www.sdh.gov.br/assuntos/pessoa-com-deficiencia/dados-estatisticos/arquivos/cartilha-do-censo-2010-pdf/view>>. Acesso em: 18 abr. 2018.
2. NEVES, Cristina; MOREIRA, Demóstenes. O exercício terapêutico no tratamento da lombalgia crônica: uma revisão da literatura. **Revista Brasileira de Ciência e Movimento**, v. 18, n. 4, p. 109-116, 2011.
3. ROSARIO, J. L.; MARQUES, Amelia P.; MALUF, A. S. Aspectos clínicos do alongamento: uma revisão de literatura. **Braz J Phys Ther**, v. 8, n. 1, p. 83-88, 2004.
4. SIMÕES, N. V. Lesões desportivas em praticantes de atividade física: uma revisão bibliográfica. **Braz. j. phys. ther.(Impr.)**, p. 123-128, 2005.
5. GRAY, Julie McLaughlin. Discussion of the ICDH-2 in relation to occupational therapy and occupational science. **Scandinavian Journal of Occupational Therapy**, v. 8, n. 1, p. 19-30, 2001.
6. BORGNETH, Livia. Considerações sobre o processo de reabilitação. **Acta fisiátrica**, v. 11, n. 2, p. 55-59, 2004.

7. GUIMARÃES, L. de S.; CRUZ, M. C. Exercícios terapêuticos: a cinesioterapia como importante recurso da fisioterapia. **Lato & Sensu**, v. 4, n. 1, p. 3-5, 2003.
8. FEUERWERKER, Laura. Modelos tecno assistenciais, gestão e organização do trabalho em saúde: nada é indiferente no processo de luta para a consolidação do SUS. **Interface-Comunicação, Saúde, Educação**, v. 9, n. 18, p. 489-506, 2005.
9. AMADIO, Alberto Carlos; DUARTE, Marcos. Fundamentos biomecânicos para a análise do movimento humano. 1996.
10. BAUMANN, W. Métodos de medição e campos de aplicação da biomecânica: estado da arte e perspectivas. In: **VI Congresso Brasileiro de Biomecânica. Brasília**. 1995.
11. AMADIO, A. C.; SERRÃO, J. C. Instrumentação em cinética. **Saad, M., Batistella, LR, Análise da Marcha. Manual do CAMO-SBMFR, São Paulo, Lemos Editorial**, p. 53-68, 1997.
12. MELO, Sebastião Iberes Lopes; SANTOS, Saray Giovana dos. Antropometria em biomecânica: características, princípios e modelos antropométricos. **Rev. bras. cineantropom. desempenho hum**, 2000.
13. REBELO, Francisco dos Santos. Sistema Digita–Aquisição de Dados Antropométricos Baseada em Técnicas Fotogramétricas para Aplicações em Ergonomia. **Manual Técnico. Lisboa, Portugal**, 2002.
14. ANDRILUKA, Mykhaylo et al. 2d human pose estimation: New benchmark and state of the art analysis. In: **Proceedings of the IEEE Conference on computer Vision and Pattern Recognition**. 2014. p. 3686-3693.
15. NIELSEN, Michael A. **Neural networks and deep learning**. San Francisco, CA, USA: Determination press, 2015.
16. BREDA, Vinícius Morais et al. Reconhecimento de gestos em vídeos utilizando modelos ocultos de Markov e redes neurais convolucionais aplicado a libras. 2018.
17. DE BRITO SANCHEZ, Renato et al. Artificial Intelligence to Detect Alzheimer's in Magnetic Resonances. In: **XXVI Brazilian Congress on Biomedical Engineering**. Springer, Singapore, 2019. p. 59-63.
18. JUN, He et al. Facial Expression Recognition Based on VGGNet Convolutional Neural Network. In: **2018 Chinese Automation Congress**

**(CAC)**. IEEE, 2018. p. 4146-4151.

19. SIMONYAN, Karen; ZISSERMAN, Andrew. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014.
  20. BAČANIN DŽAKULA, Nebojša et al. Convolutional Neural Network Layers and Architectures. In: **Sinteza 2019-International Scientific Conference on Information Technology and Data Related Research**. Singidunum University, 2019. p. 445-451.
  21. CAREY, Kevin et al. Comparison of skeleton models and classification accuracy for posture-based threat assessment using deep-learning. In: **Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II**. International Society for Optics and Photonics, 2020. p. 1141321.
  22. NORTON, Kevin; OLDS, Tim; ALBERNAZ, Nilda Maria Farias de. Antropométrica: um livro sobre medidas corporais para o esporte e cursos da área de saúde. In: **Antropométrica: um livro sobre medidas corporais para o esporte e cursos da área de saúde**. 2005. p. 398-398.
  23. BRENDLER, Clariana Fischer; TEIXEIRA, Fábio Gonçalves. Diretrizes para auxiliar na aplicação da Antropometria no desenvolvimento de Projetos de Produtos Personalizados. **Estudos em Design**, v. 21, n. 2, 2013.
  24. CAO, Zhe et al. Realtime multi-person 2d pose estimation using part affinity fields. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. 2017. p. 7291-7299
- 

<sup>1</sup>Núcleo de Pesquisa Tecnológica – NPT, Laboratório de Ambientes Virtuais e Tecnologia Assistiva – LAVITA, Universidade de Mogi das Cruzes – UMC, Mogi das Cruzes, Brazil

<sup>2</sup>E-mail: renatobritosanchez@gmail.com

<sup>3</sup>(Universidade Santo Amaro – UNISA, São Paulo, Brazil)

<sup>4</sup>(Núcleo de Pesquisa Eniac – NUPE, Centro Universitário Eniac – ENIAC, Guarulhos, Brazil)

← Post anterior

## RevistaFT

A **RevistaFT** é uma **Revista Científica Eletrônica Multidisciplinar Indexada de Alto Impacto e Qualis “B2” em 2023**. Periodicidade mensal e de acesso livre. Leia gratuitamente todos os artigos e publique o seu também [clikando aqui](#).



## Contato

**Queremos te ouvir.**

**WhatsApp:** 11 98597-3405

**e-Mail:** contato@revistaft.com.br

**ISSN:** 1678-0817

**CNPJ:** 48.728.404/0001-22

**CAPES** – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), fundação do Ministério da Educação (MEC), desempenha papel fundamental na expansão e consolidação da pós-graduação stricto sensu (mestrado e doutorado) em todos os estados da Federação.

## Conselho Editorial

**Editores Fundadores:**

Dr. Oston de Lacerda Mendes.

Dr. João Marcelo Gigliotti.

**Editor Científico:**

Dr. Oston de Lacerda Mendes

**Orientadoras:**

Dra. Hevellyn Andrade Monteiro

Dra. Chimene Kuhn Nobre

Dra. Edna Cristina

Dra. Tais Santos Rosa

**Revisores:**

Lista atualizada periodicamente em [revistaft.com.br/expediente](https://revistaft.com.br/expediente) Venha fazer parte de nosso time de revisores também!

Copyright © Editora Oston Ltda. 1996 - 2023

Rua José Linhares, 134 - Leblon | Rio de Janeiro-RJ | Brasil