



Article

Automated Detection of Ear Tragus and C7 Spinous Process in a Single RGB Image—A Novel Effective Approach

Ivanna Kramer ^{1,2,*} , Sabine Bauer ^{2,†} and Anne Matejcek ¹¹ Institute for Computational Visualistics, University Koblenz-Landau, 56070 Koblenz, Germany; amatejcek@uni-koblenz.de² Institute of Medical Technology and Information Processing, University Koblenz-Landau, 56070 Koblenz, Germany; bauer@uni-koblenz.de

* Correspondence: ivannamyckhal@uni-koblenz.de

† These authors contributed equally to this work.

Abstract: Biophotogrammetric methods for postural analysis have shown effectiveness in the clinical practice because they do not expose individuals to radiation. Furthermore, valid statements can be made about postural weaknesses. Usually, such measurements are collected via markers attached to the subject's body, which can provide conclusions about the current posture. The craniovertebral angle (CVA) is one of the recognized measurements used for the analysis of human head–neck postures. This study presents a novel method to automate the detection of the landmarks that are required to determine the CVA in RGBs. Different image processing methods are applied together with a neuronal network Openpose to find significant landmarks in a photograph. A prominent key body point is the spinous process of the cervical vertebra C7, which is often visible on the skin. Another visual landmark needed for the calculation of the CVA is the ear tragus. The methods proposed for the automated detection of the C7 spinous process and ear tragus are described and evaluated using a custom dataset. The results indicate the reliability of the proposed detection approach, particularly head postures.

**Citation:** Kramer, I.; Bauer, S.;

Matejcek, A. Automated Detection of Ear Tragus and C7 Spinous Process in a Single RGB Image—A Novel

Effective Approach. *Biomedinformatics***2022**, *2*, 318–331. [https://doi.org/](https://doi.org/10.3390/biomedinformatics2020020)

10.3390/biomedinformatics2020020

Academic Editor: Pentti Nieminen

Received: 22 April 2022

Accepted: 2 June 2022

Published: 8 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: automated detection; craniovertebral angle; neuronal network; Openpose; RGB; C7; tragus

1. Introduction

The number of smart device users is increasing every year. The year 2013 recorded a total of more than 4.01 billion smartphone users, which increased to 5.07 billion users in 2019 [1]. It is forecast that the subscriptions associated with smartphones will continue to rise and will reach 7.7 billion in 2027 [2,3]. Moreover, tablet computer usage has increased dramatically in recent years, in terms of both the number of users and the type of applications [4].

With such increased use of smart devices, the number of people who report neck pain has also risen. Scientific studies show that more and more physiological and musculoskeletal complaints are being reported that can be traced to the use of smart devices [5–8]. The bent body position that is often taken when using a smart device can lead to an abnormal head posture and also affects the postural apparatus and thus the spinal structures. To investigate these effects on the human body, researchers evaluate the biomechanics of the head–neck system and the associated impairments such as dizziness or headaches associated with a variety of conditions.

In order to quantify these postural alterations, the neck flexion and forward head position (FHP) can be determined with the help of, e.g., the craniovertebral angle (CVA), head tilt angle (HTA), and shoulder angle (SHA).

The research in [9–11] showed that prolonged smartphone use has a direct effect on the HTA, SHA, and FHP as determined by the CVA.

In order to gain insights into the tendency towards scoliosis, the naked dorsal surface of 98 volunteers was scanned with Microsoft Kinect to determine the shoulder angle [12]. The position of the C7 spinous process was estimated using the method of pixel-shade difference. The method used to determine the C7 spinous process was not presented in detail, and it was not checked for validity. The aim of the study presented in [13] was the assessment of the relative angles between vertebral adjacent segments during gait using IMUs. The proposed method proved the usability of inertial sensors for the assessment of spinal posture. Statistically significant differences were shown with regard to the influence of gender, speed, and imposed cadence. In addition to the analysis of the head position, Ormos et al. [14,15] measured the range of motion of the cervical spine using a goniometer and determined the isometric strength of the neck muscles with a dynamometer for flexion, extension, and head tilt on both sides. People with FHP showed a lower pressure pain threshold (PPT) in all locations except for the upper trapezius and scalenus medius muscles. They also showed less extension and right-rotation range of motion. The objectives of the study in [16,17] were to quantify the neck posture using CVA and fatigue in neck muscles. Both studies used markers attached to the participant's tragus and C7 to measure the CVA.

Furthermore, different studies investigated the relation between a bent posture and neck pain, headache, and spinal deformations. For example, Ref. [18] investigated the effects of smartphone use for less than and more than four hours per day between two groups. They found a significant difference in forward head angle and the intensity of neck pain after prolonged use of smartphones. To explore the working mechanism of manual therapy, Ref. [19] investigated whether aspects of cervical spine function, such as cervical ROM, neck flexor endurance, and FHP, were mediators of the effect of manual therapy on headache frequency. The effect whereby FHP causes spinal deformation, which increases scapula deformation, lordosis of the cervical vertebra, and kyphosis of the upper thoracic vertebra, was confirmed by [20,21]. The craniocervical angle was manually measured on the basis of markers using a lateral digital photograph with a digital camera. In addition, the authors of [22] quantified neck postures using electrogoniometers. It was observed that the ergonomic loads were increased when compact and slate computers were used, especially when used in non-traditional work environments.

The study by [23] is a prospective, cross-sectional, observational investigation, evaluating 3D quantitative standing posture proprioceptive perception through an instinctive self-correction maneuver in nonspecific chronic and sub-acute patients with lower back pain. To measure the subjects' 3D whole-skeleton pose, a non-ionizing 3D optoelectronic stereophotogrammetric approach with 27 passive retro-reflective markers was used.

The methods reported in the previous publications are based on marker-based approaches or semi-automatic recording, or they need additional recording devices to determine the body position. To the best of our knowledge, no markerless, fully-automatic detection method for head and neck angles in a single RGB image exists. Therefore, the focus of our study was to propose a fully automated approach to analyze cervical spine posture using RGB images, without using additional markers as landmarks, to detect defined key points such as the tragus of the ear or the spinous process of C7. By means of this automatic landmark detection method, the craniovertebral angle (CVA) is an established indicator of the severity of forward head position [10,15,16,24–26], is determined.

2. Materials and Methods

The CVA [27,28] is determined using lateral view images of the head–neck postures as the angle formed between the horizontal and a line drawn from the midpoint of the ear tragus to the point on the skin overlying the tip of the spinous process of the seventh cervical vertebra, C7 (see Figure 1). In order to automatically calculate the CVA, two anatomical landmarks, i.e., the ear tragus and the spinous process of C7, must be detected first.

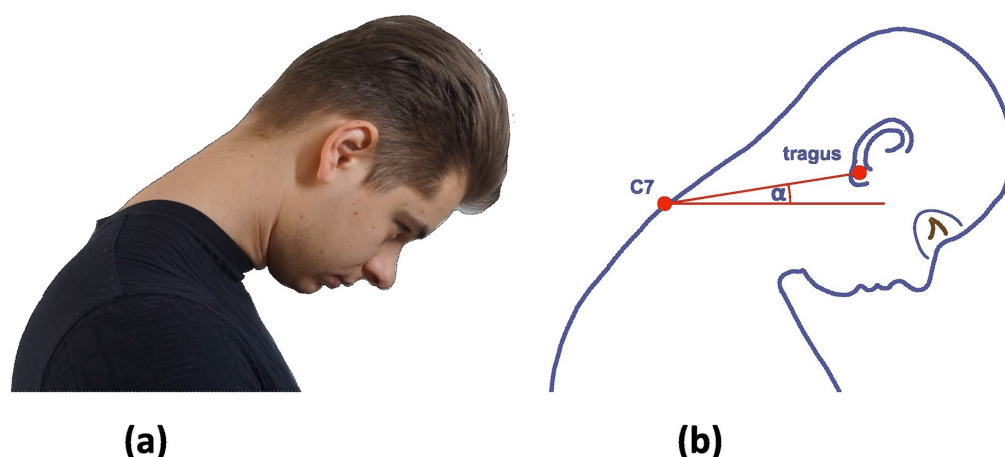


Figure 1. Image of a subject with flexed head position: (a) craniovertebral angle α was defined as the angle formed between the line drawn from the tragus of the ear to the skin tip location of the spinous process C7 and the horizontal line (b).

2.1. Determination of Spinous Process

The spinous process of the seventh human vertebra is an important landmark in the analysis of head–neck postures. When lowering or stretching the head forward, the spinous process of C7 forms a clear bulge in the skin on the back of the neck. In order to find a 2D pixel position of the skin tip of the spinous process in a single RGB image, a method based on the approximation of the line to the neck contour is proposed.

Openpose is a neural network that can recognize the 2D positions of landmarks on the human body and facial structures in an RGB image. In the context of this study, Openpose was used to determine a region of interest (RoI) surrounding the landmarks in the neck region. The proposed method, depicted in Figure 2, is to take a photograph of the subject and extract 18 key body points using the real-time skeleton detection model of Openpose [29] on it. The determination of the RoI is an essential part of the proposed approach. If the neck curvature is outside of the RoI, the C7 spinous process cannot be detected. Otherwise, if the RoI was too large, the image filters applied in the further steps could extract too many features in the background, which could lead to errors in the landmark detection.

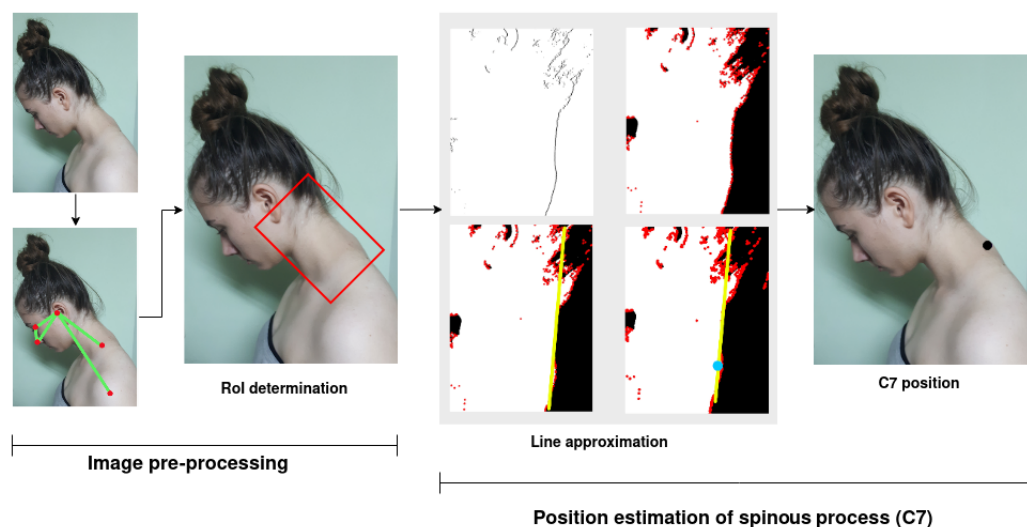


Figure 2. Determination of the position of the skin bulging through the C7 spinous process in an RGB image: based on 2D body key points (marked red) detected in the image, the RoI was determined. Using computer vision methods, corner detection was performed. The 2D position of the spinous process of C7 was estimated through the line (marked yellow) approximated to the neck curvature.

The RoI was calculated based on the 2D coordinates of the selected body parts in the Openpose output such as the x and y pixel positions of the nose and the ear (see Figure 3):

$$upper_{left} = nose_x + \frac{ear_x - nose_x}{2}, \quad nose_y + \frac{ear_y - nose_y}{2}, \quad (1)$$

$$upper_{right} = ear_x + (ear_x - nose_x), \quad ear_y + (ear_y - nose_y), \quad (2)$$

$$lower_{left} = upper_{left_x} - ((ear_y - nose_y) \cdot 2), \quad upper_{left_y} + ((ear_x - nose_x) \cdot 2), \quad (3)$$

$$lower_{right} = upper_{right_x} - ((ear_y - nose_y) \cdot 2), \quad upper_{right_y} + ((ear_x - nose_x) \cdot 2). \quad (4)$$

The input image was cropped to only the relevant area where the neck curvature and C7 spinous process bulge were visible, and at the same time, the area of the background was kept as small as possible.

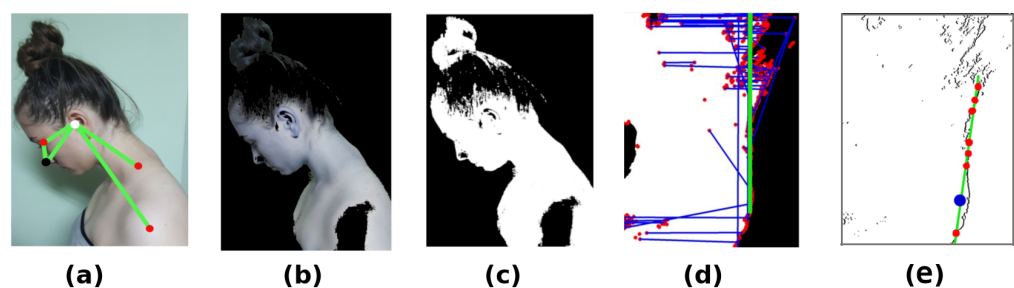


Figure 3. Body positions of the nose (marked black) and ear (marked white), which are used for the determination of RoI, are extracted using a pre-trained Openpose model (a). For the separation of the body from the background, color segmentation (b) is applied to the image. Then the body is marked white and the background black (c). In the binary image, the 2D corner-point coordinates are detected. For each of the corner points (marked red) and its neighbor, a straight line (marked blue) is generated. Subsequently, the number of corner points lying on each line is determined and compared. The line with the most corner points located on it is selected as the approximation line of the neck contour (green line). (Note: for clarity, not all lines are visualized in the image) (d). In order to calculate the position of the intersection, points (marked red) between the approximation line (marked green) and the neck contour (black) are used (e). The C7 spinous process is located in the middle of the adjacent intersection points with the maximum distance from each other.

Before computer vision methods for the edge extraction could be applied to the image, a color segmentation was used in order to separate the human body from the background. A color interval was first determined in the HSV color space. The output of this step is shown in Figure 3b. In the next step, the binary image was created (see Figure 3c). All pixels whose color value was inside the color interval of the skin were set as white, and every other pixel was set as black. This binary image was then passed to the Sobel operator [30] to select prominent contours in the RoI. The Sobel filter is a classical edge detector, which is commonly used in image processing. It extracts edges by performing the gradient on the image and emphasizes edges in the vertical or horizontal direction. A binary image favors the use of the Sobel filter. When using an RGB image instead, results can vary widely due to different lighting and color values. Using a binarized image, the Sobel operator can easily find the edges between the person and the background, since factors such as light and shadow are no longer taken into account. For a better interpretation of the edge detector results, corner points were calculated on each found contour using the Harris operator [31]. The 2D coordinates of the calculated corners points were then used to determine a straight line that approximated the neck contour. In order to determine this approximation line, the corner points found were first sorted according to their y -coordinate. For each corner point and its successor, a straight line was calculated. We then checked how many other corner points lay on each of the lines. Finally, the line with the most corner points was chosen to be the approximation line for the neck contour. The reason that the number

of corner points on the line was also taken as a criterion was that the number of corner points on the neck contour was always significantly higher than the number of vertices on, for example, hair contours.

Once the approximation line was determined, its intersection points with the neck contour were calculated utilizing the edges extracted by the Sobel operator in the neck region. In Figure 3e, it is shown that the estimated position of the C7 spinous process is located between two adjacent intersection points with a distance greater than a pre-defined threshold.

2.2. Estimation of the Ear Tragus

The ear tragus was the second most important reference point for the CVA determination. The pipeline for the calculation of the ear tragus position was similar to the method proposed in Section 2.1 for the detection of C7. The RoI with the ear tragus was determined using the same pre-trained Openpose model as the RoI used for the C7 spinous process. However, in this case, rather than the neck contour, the ear should be mainly visible. For the RoI extraction, the Euclidean distance $dist_{NE}$ between the nose's key point and the ear was calculated. The vertices of the RoI were determined in such a way that the ear was located in the center:

$$upper_{left} = ear_x - \frac{dist_{NE}}{2}, \quad ear_y - \frac{dist_{NE}}{2} \quad (5)$$

$$upper_{right} = ear_x + \frac{dist_{NE}}{2}, \quad ear_y - \frac{dist_{NE}}{2} \quad (6)$$

$$lower_{left} = ear_x - \frac{dist_{NE}}{2}, \quad ear_y + \frac{dist_{NE}}{2} \quad (7)$$

$$lower_{right} = ear_x + \frac{dist_{NE}}{2}, \quad ear_y + \frac{dist_{NE}}{2}. \quad (8)$$

For the localization of the ear tragus, the ear contour was first detected by approximating an ellipse model on the outer shape of the ear, and then the pixel intensities inside the selected ellipse were analyzed. For this purpose, the Canny algorithm [32] was applied to the ear's RoI in order to detect edges. The Canny edge detector was chosen since the ear had many complex structures that could be missed by the Sobel operator. Afterward, the edges that were close to each other were connected to the contours, which were used to fit an ellipse to the ear's outer boundaries. The ellipses were generated in such a way that they maximally surrounded the found contours in the image (see Figure 4). In the last step, the ellipse whose center was the closest to the ear landmark point found by Openpose was selected to build a bounding ellipse of the outer ear contours.

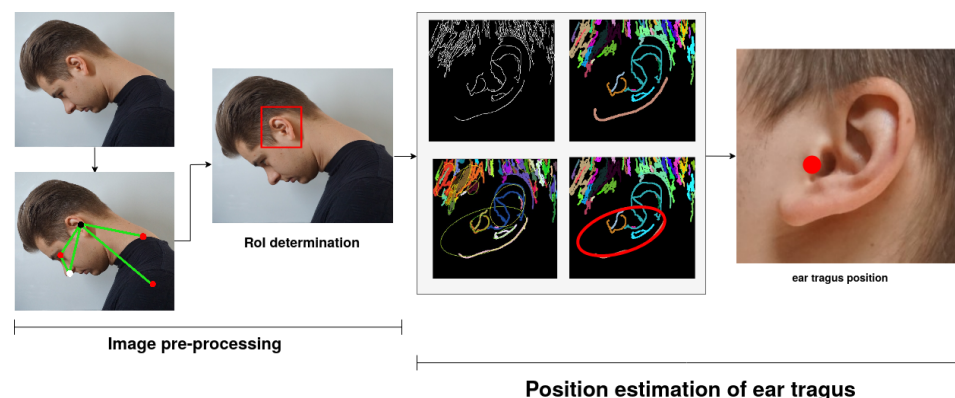


Figure 4. Estimation of the ear tragus in the lateral-view image of the subject: in the first step, image pre-processing was performed, which includes the extraction of the RoI (marked by red frame) based on the Openpose output. In the further processing steps, image features such as edges and contours were detected in the RoI. An ellipse was associated with each contour. The ellipse positioned closest to the Openpose ear point was then selected. Intensities were analyzed inside of the chosen ellipse. The ear tragus was estimated to be located in the direction of the smallest distance from the ear channel to the ear border represented through an ellipse.

In the next step, the area inside the ear ellipse was analyzed. The goal was to find the region with the darkest pixels, since this is where the auditory canal is most likely to be found. To do so, the minimum intensity value of all points within the ellipse was determined. The position of the ear tragus can be derived by applying an edge detector in the direction of the closest distance to the bounding ellipse that surrounds the ear.

2.3. Data Acquisition

In order to validate the proposed method for the automated calculation of the CVA in RGB images, a custom dataset was generated with a total of 79 subjects, of whom 45 were male and 34 were female.

For each test participant, the following additional data, required for the CVA analysis, were collected: demographic data such as age and gender, spinal disorders, and average daily duration of smartphone usage. Their mean age was 26.6 years, with the youngest person being 21 years old and the oldest being 61 years old. The height of the test subjects ranged from 163 cm to 196 cm. Of the 79 participants, 6 had been diagnosed with cervical spine disorders. The estimated daily duration of smartphones usage among the test subjects was 3.16 h on average, in a range from 0.3 h to 7 h.

The images in the dataset showed each subject performing four head–neck postures: straight neck, maximal head flexion, forward head posture, and head-down position (see Figure 5). The dataset resulted in a total of 316 images. The aim of the recorded dataset was twofold: on the one hand, the validation of the proposed methods for the detection of the anatomical points needed for calculation of the CVA and, on the other hand, the analysis of the CVA in four different head–neck positions.

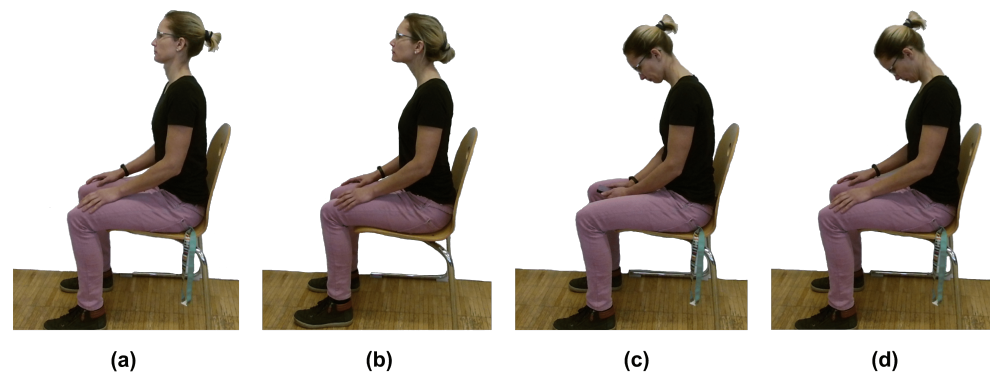


Figure 5. Sample images of different head–neck postures: straight-neck (a), forward-head (b), and head-down position during smartphone use (c) and maximal head flexion (d).

The images were recorded with a widely available Microsoft Kinect v2 3D camera, which captured 2D RGB images, as well as the depth data with a separate depth sensor [33]. The resolution of the RGB camera was 1920×1080 pixels, and the depth sensor had a resolution of 512×424 pixels. The camera and the data recorded in the dataset are depicted in Figure 6a. Although the dataset was intended for the evaluation of the automated calculation of the craniocervical angle in a single RGB image, depth maps were captured for each RGB frame and can be used for future research projects.

The photographs were taken in sitting positions in a left lateral view with the same background. For each recording, the camera was located at a distance of 1.3 m away from the subject on a tripod at a height of 1.2 m from the floor. In particular, to be able to determine the position of C7's skin bulge, the area of the neck bulge through C7 should not be covered by clothing or hair. The participants were asked to wear suitable clothing to keep the region of the head and neck uncovered and tie their hair back. In addition, all participants looked in the same direction and followed the instructions given to the participants before recording. Some of the subjects were asked to cover their neck, and 48 images were obtained, which could be used to validate whether the method could correctly predict the negative classes or not.



Figure 6. Microsoft Kinect v2 camera with included sensors (a) and illustration of three modalities captured in the dataset (b): RGB image, depth map, and 3D point cloud.

The labeling of the captured photographs was performed by three experts on the 2D data. The experts were asked to mark the posterior position of the C7 spinous process, the ear tragus, and the ear lobula (see Figure 7). In this study, two points, the C7 spinous process and the ear tragus, were used for analysis and validation. In order to evaluate the quality of the annotation performed by the experts, inter-rater reliability (IRR) [34,35] was chosen as a validation metric. To calculate the IRR for three ratings, Krippendorff's α_K coefficient [35] was applied, which measures the extent of the agreement between multiple raters. A α_K value that approaches 1.0 indicates high confidence in the labeling accuracy and results in the high-quality dataset. The calculated reliability coefficient for the given data showed a high inter-rater agreement with an α_K of 0.991186.

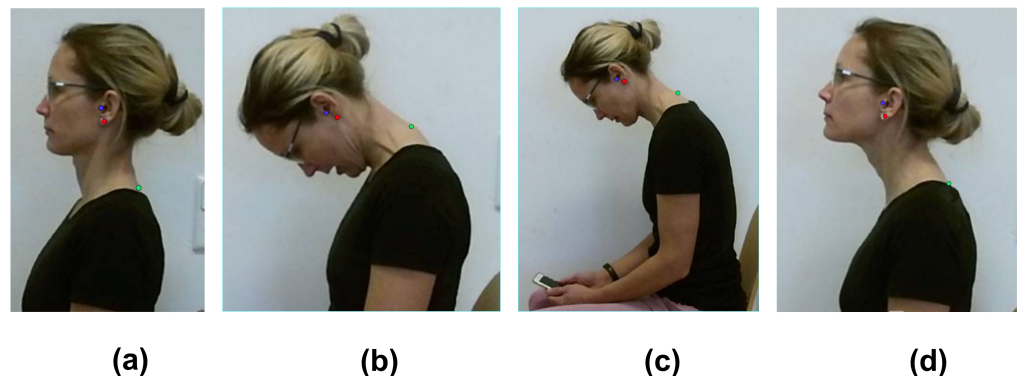


Figure 7. Labeled sample images of recorded head-neck postures: straight posture (a), maximal flexed head (b), using smartphone (c), and forward head position (d). The C7 spinous process is marked green, the ear tragus is blue, and the ear lobula is shown in red.

3. Results

In order to validate the performance of the proposed methods, the general detection accuracy was determined by calculating the ratio of the correct outcomes to all possible method responses:

$$acc = \frac{TP + TN}{TP + FP + FN + TN} \quad (9)$$

where TP denotes true positive, TN denotes true negative, FP denotes false positive, and FN denotes false negative predictions. The detection accuracy for two detected anatomical points is shown in Table 1.

Table 1. Detection accuracy determined for C7 spinous process and ear tragus.

	C7 Spinous Process	Ear Tragus
Detection accuracy (acc)	80%	83%

The performance of the proposed detection methods was analyzed by considering the detected and observed values of the x and y coordinates for the chosen landmarks. The relationships of the observed and detected x and y coordinates for the C7 spinous process and ear tragus are visualized in Figure 8. From the presented diagrams, it can be seen that the proposed detection approaches performed well for the most of the images in the dataset. In general, the detected points are aligned to the line of the optimal fit in all of the sub-plots.

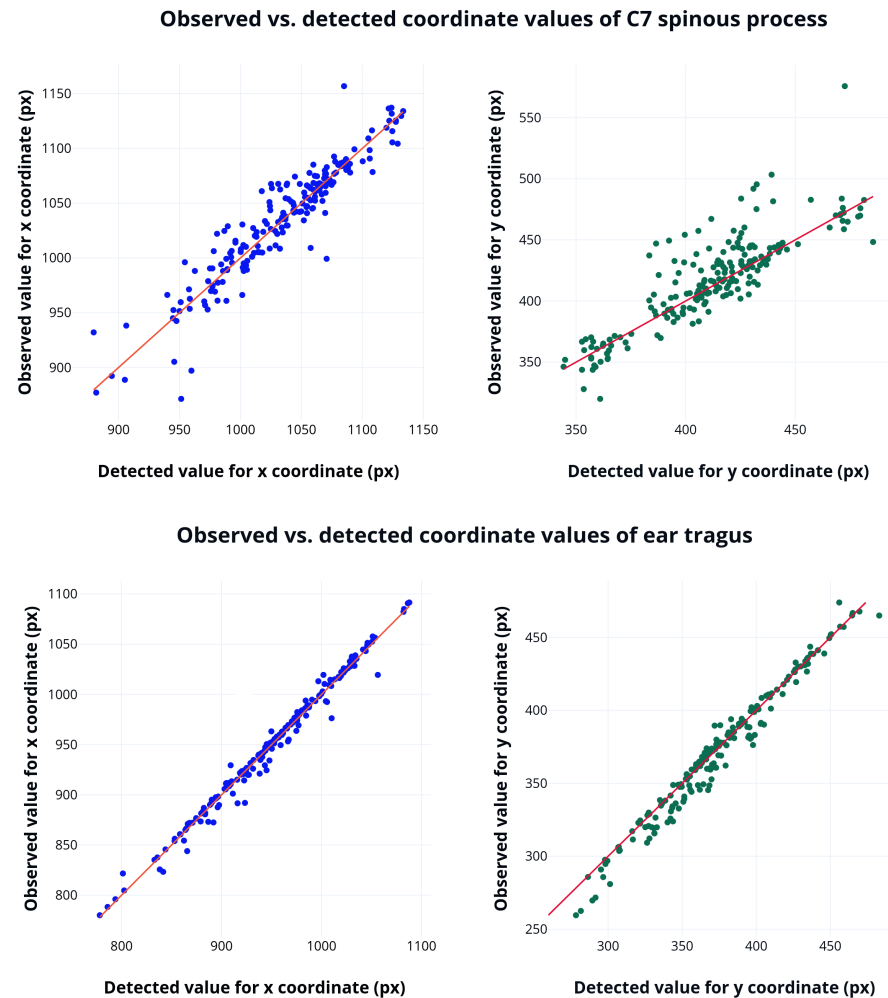


Figure 8. Comparison of the observed and detected values of the 2D coordinate of C7 spinous process and ear tragus. Values of x and y coordinates for the corresponding anatomical landmarks are visualized in the separate sub-plots. The line for the optimal fit is marked in red in each diagram.

By comparing the determined points for the C7 spinous process to the detected points of the ear tragus, it can be seen that the detection of the ear tragus provided more accurate results. In some datapoints, the detector for the C7 spinous process computed a slightly higher y coordinate value compared to its ground truth. However, the proposed method for ear tragus detection showed a relatively small number of incorrect outcomes.

In order to quantify the performance of the proposed detectors, PRESS statistics were calculated for each coordinate of the corresponding landmark points by using the following equation:

$$PRESS_{coordinate} = \sum_{i=1}^N (v_{coordinate,i} - \hat{v}_{coordinate,i})^2, \quad (10)$$

where v_i is the observed value of the corresponding coordinate, $\hat{v}_i$ is its predicted value, and N is the number of images in the dataset. Additionally, the coefficient of determination R^2 was calculated:

$$R_{coordinate}^2 = \frac{PRESS}{\sum_{i=1}^N (v_{coordinate,i} - \bar{v}_{coordinate,i})^2}, \quad (11)$$

The results of the PRESS statistics are presented in Table 2. High values of R_x^2 and R_y^2 for the C7 spinous process and ear tragus indicate strong correlation between predicted datapoints and the corresponding ground truth. While comparing proposed detection approaches for the C7 spinous process and for the ear tragus much lower PRESS statistics are shown for the ear tragus demonstrating high predictive ability of this method.

Table 2. PRESS statistics and coefficient of determination R^2 calculated for x and y coordinates of C7 spinous process and ear tragus.

Measure	C7 Spinous Process	Ear Tragus
$PRESS_x$	68,760.81	11,094.77
$PRESS_y$	80,683.63	13,316.21
R_x^2	0.99	0.99
R_y^2	0.99	0.99

To assess the success of the methods graphically, the residuals for the x and y coordinates of the determined landmarks are depicted in the scatter-plots in Figure 9. The upper sub-plots representing residual values for the determined point of the C7 spinous process show that most residuals for the x and y coordinates are distributed symmetrically between 20 px and -20 px , and the only exceptions are a few outliers, which demonstrate a difference of up to 80 px from the ground truth value in the x coordinate and up to -60 px in the y coordinate. The residual datapoints for ear tragus detection are located mostly between 0 px and 5 px for both coordinates of the calculated landmark. However, in the validation dataset used, some incorrect detections were performed by the ear tragus detector, which are shown as outliers in the corresponding diagram. The minimum residual of the x coordinate was -36 px , while the minimum residual for the y coordinate was -22 px .

Considering the detection error of the proposed methods, the mean distance error $\bar{e}$ between the predicted point p_i and the corresponding ground truth $\hat{p}_i$ point for each posture class was determined. The mean distance error was calculated by applying the Euclidean distance in each of the images i :

$$\bar{e}_{class} = \frac{1}{N} \sqrt{\sum_{i=1}^N (p_{class,i} - \hat{p}_{class,i})^2}, \quad (12)$$

where N is the total number of images in the dataset. The results of the mean detection errors and of the standard deviation of the detection for the C7 spinous process and ear tragus are depicted in Figure 10. It can be seen that the smallest detection error was indicated for the posture class smartphone use, where $\bar{e}_{C7}$ for the C7 spinous process was $15.75 \pm 12.06\text{ px}$, and $\bar{e}_{ear}$ was $6.7 \pm 7.4\text{ px}$ for the ear tragus and maximal flexed where we observed the values of $14.18 \pm 8.03\text{ px}$ for the C7 spinous process and $7.02 \pm 7.07\text{ px}$ for the ear tragus. For the C7 spinous process, the maximum mean distance error of $22.01 \pm 18.06\text{ px}$ was calculated for the straight posture. The highest mean distance error for the ear tragus detection of $9.27 \pm 7.34\text{ px}$ was indicated in the posture class of head forward.

The overall mean distance error for the detection of the C7 spinous process $\bar{e}_{C7}$ was calculated to be $18.07 \pm 13.67\text{ px}$, and the overall mean distance error for the ear tragus $\bar{e}_{ear}$ was $7.96 \pm 7.45\text{ px}$.

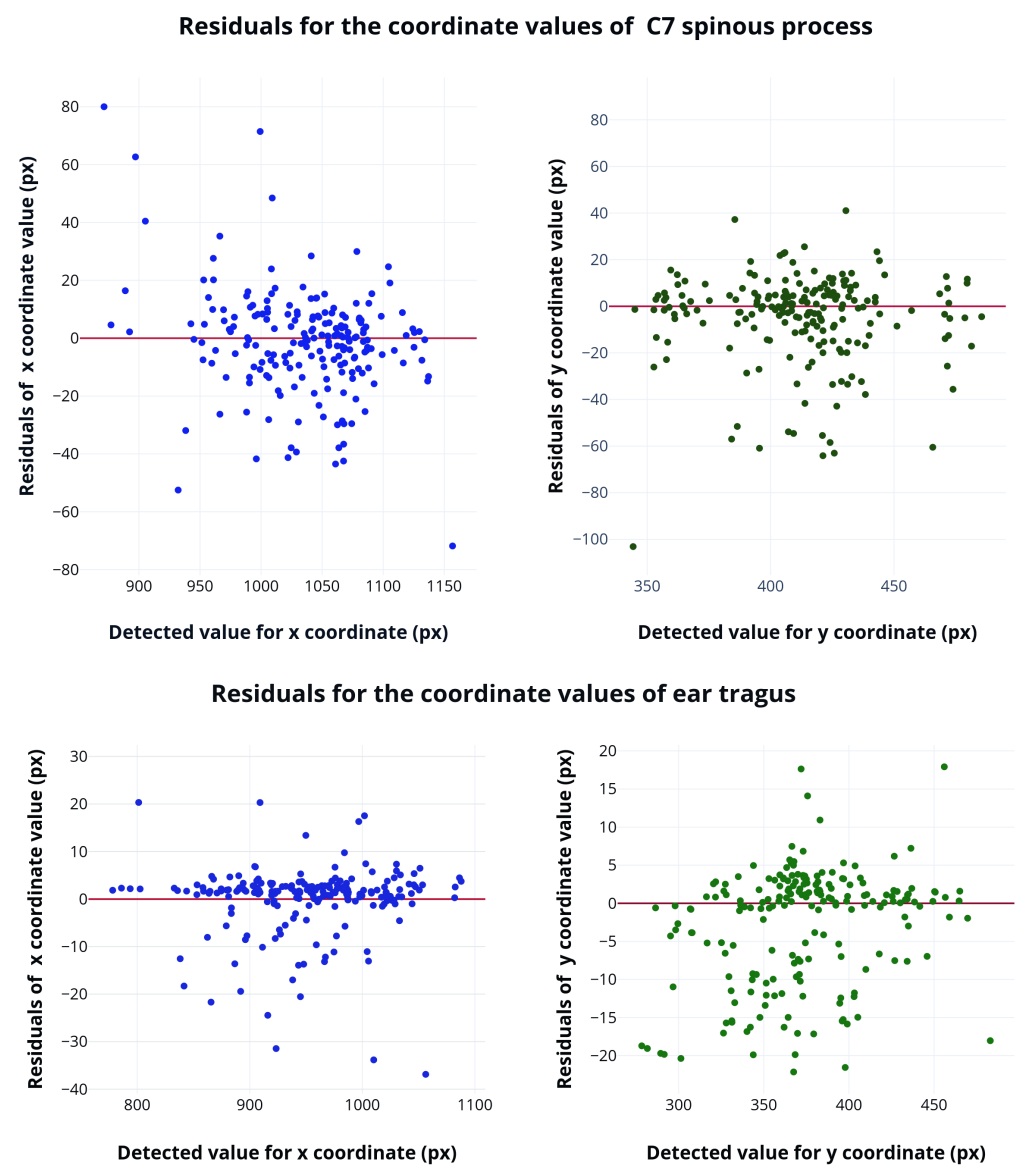


Figure 9. Residuals calculated for x and y coordinates of the detected landmarks: the zero-line is marked in red, the points for the x coordinates are marked in blue, and the green points represent residuals for the y coordinate.

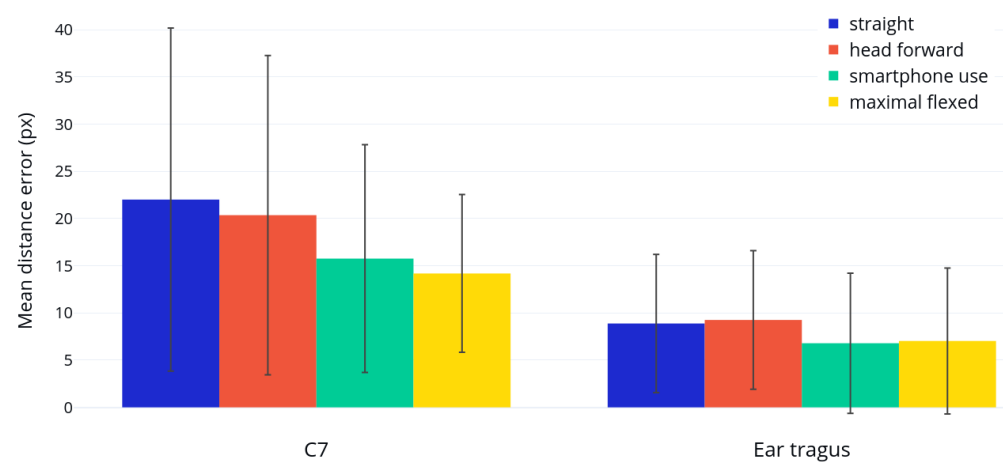


Figure 10. Mean distance error calculated for the detected landmarks. The distance error together with standard deviation is visualized separately for each posture class.

Figure 11 shows sample images from the recorded validation dataset and detected landmark points for the C7 spinous process as well as for the ear tragus. For the subject depicted in Figure 11a, the determined point of the C7 spinous process is slightly offset from the observed point in the horizontal direction.

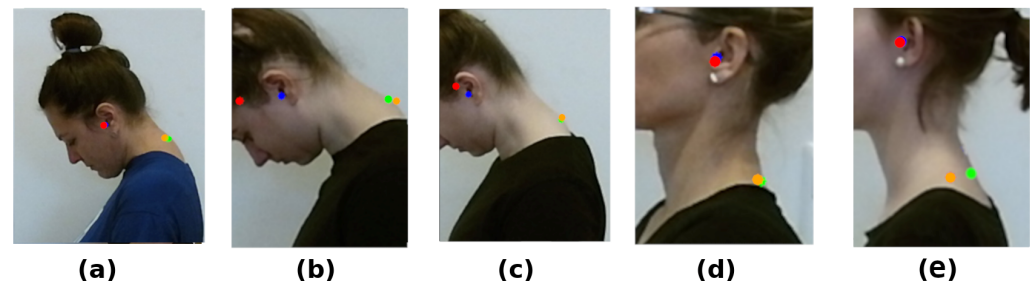


Figure 11. Sample predictions for different postures: minor deviation between the predicted and labeled points for the C7 spinous process and the ear tragus (a), the wrongly detected ear tragus (b), the found point for the ear tragus is located on the outer ear border (c), good prediction sample (d), the found position of the C7 spinous process with the slight deviation from the corresponding ground truth point (e). Green marks the positions of the C7 spinous process and blue indicates the ear tragus annotated by experts; predicted positions are marked using red for the ear tragus and orange for the C7 spinous process.

A higher displacement can be seen for the tragus point in Figure 11b. It is falsely positioned outside of the ear border, in contrast to the detection in Figure 11c, where it is still localized in the ear region, but in the wrong position. Moreover, as shown in Figure 11b, the detector estimated the position of the C7 spinous process over the skin tip. An example of a small shift for both the tragus point and the C7 spinous process point is shown in Figure 11d. There is also the possibility that one of the two key points is found well, but there are deviations for the other. While the key point of the ear tragus is well estimated, the key point for the C7 spinous process is localized in the middle of the neck (Figure 11e).

There is also a possibility that one of the two key points may be found successfully while the other landmark is miscalculated, such as the case shown in Figure 11e, where the point of the C7 spinous process is detected in the middle of the neck and the ear tragus is located correctly.

4. Discussion

The CVA is an important characteristic in head–neck postural assessment. However, the automatic detection of the head–neck landmarks, such as the C7 spinous process and ear tragus, which are required for the calculation of the CVA, is a challenging task.

The methods proposed in this study utilize image filters in order to detect landmarks required for CVA calculation in a single RGB image. In the first step of the proposed approach, the RoI was extracted based on the output from a body pose prediction model. For the localization of the C7 spinous process, line approximation was applied to the neck curvature. The intersection points between the neck contour and the approximation line were extracted in the subsequent step, and the distance for each pair of adjacent intersection points was calculated. Finally, the position of the C7 spinous process was determined as the midpoint of the vector built from the intersection points with the maximum distance.

For the ear tragus detection, the ear RoI was first extracted similarly to that of the C7 spinous process. Using a classical edge detector, the contours were extracted inside the RoI. From the extracted contours, an ellipse was fitted to contour of the ear. The ear tragus was then determined by analyzing the intensities inside the associated ellipse.

In general, the presented results demonstrated the capability of the developed method to detect the desired landmarks in an RGB image. The detector for the C7 spinous process as well as for the ear tragus showed the best performance for the subjects performing

the smartphone use and maximal flexed postures. In these postures the bulge of the C7 spinous process was prominently expressed on the skin for most of the subjects, so the values of the detected C7 spinous process points remained within the error-tolerance range. The ear tragus detection showed good prediction results in the images of these postures as well. In some rare cases, the ear detector falsely located the position of the ear tragus shifted to the left side of the determined RoI, as shown in Figure 11b. The reason for this miscalculation was the incorrect estimation of the ear contour. In this case, multiple edges in the hair region given by the Sobel filter were taken into account during the determination of the ear contours. Subsequently, the wrong ellipse was associated with the ear, and the ear tragus was found in the wrong position. In future work, we can overcome this issue by filtering the hair region out from the edge detection.

From the high values of the PRESS statistics as well as from the graphical representation of the residuals, relatively large deviations between the detected and labeled points of the C7 spinous process could be observed for some datapoints. Furthermore, based on the mean detection error, it can be stated that the most mismatched points were found for the subjects performing the straight head and forward head postures. In general, the morphological conditions of humans are subject to high natural variance. While in one person the spinous process of the C7 vertebra might be clearly visible, it may hardly be seen in another subject. In addition, when the head is held upright or positioned forward, the spinous process generally does not stand out as prominently as when the head is bent.

5. Conclusions

This study aimed to present and evaluate a novel approach for the automated detection of the body landmarks, namely the C7 spinous process and the ear tragus, required for the determination of the CVA. The proposed methods take a single RGB image and localize the 2D position in pixel coordinates for the desired key points using simple but effective computer vision methods. Both detection methods utilize the Sobel edge detector for their core calculations.

The proposed detectors demonstrated robust detection results for the smartphone use and maximal flexed posture groups; however, they showed some discrepancy in the detection of the straight and forward head posture classes. In order to improve the detection of the C7 spinous process in these particular poses, machine learning approaches can be used. In these approaches, the models learn to localize the pre-defined landmarks in the image from the data provided.

A limitation of the method is that the background where the person is recorded needs to be homogeneous. Otherwise, the color segmentation can detect some false-positive regions. Another limitation of this approach is that landmark occlusion is not considered in the method; i.e., landmarks covered by hair or clothing cannot be determined. To overcome this problem in the future, a machine learning model will be trained using the recorded dataset in order to detect the C7 spinous process.

Another future task is to extend the current method's implementation to the automatic determination and analysis of CVAs for different postures and compare them with the data recorded in the dataset.

A particularly attractive feature of this method is that the current posture can be automatically deduced without attaching additional physical markers. This rules out the possibility that a changed posture, which is caused by potentially interfering markers, will unknowingly influence the actual posture and thus falsify the measurement results. Especially when using physical markers during movement, there is the possibility that these markers will move away from their initial positions and shift with the skin. This implies that the unintentionally changed local position of the markers can lead to incorrect positions of joints.

Furthermore, markerless analysis of the head-neck postures can be beneficial in different experimental scenarios, especially if a study with specific vulnerable participants needs to be carried out. The proposed methods are a promising approach that can be

extended to include the detection of other prominent landmarks on the human body. Therefore, this approach could also be of interdisciplinary interest—for example, to dentists, physical therapists, speech therapists, and other professionals, since it may assist them in their clinical practice and in the context of scientific research.

All of the presented results so far relate to detection in 2D pixel space. However, using depth maps recorded in the dataset, a projection of the detected points in the 3D world is possible. This will be addressed in future works.

Author Contributions: Conceptualization, I.K. and S.B.; methodology, S.B.; software, A.M.; validation, I.K.; formal analysis, A.M.; investigation, S.B.; data curation, I.K.; writing—original draft preparation, S.B. and I.K.; writing—review and editing; visualization, I.K.; supervision, S.B.; project administration, S.B. and I.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CVA	craniovertebral angle
FHP	forward head position
HTA	head tilt angle
RoI	region of interest
IRR	inter-rater reliability
PRESS	predicted residual error sum of squares

References

1. Abdulla, M.; Smaeel, A. Providing Information through Smart Platforms: An Applied Study on Academic Libraries in Saudi Universities. *J. Educ. Soc. Behav. Sci.* **2019**, *30*, 1–24. [\[CrossRef\]](#)
2. Saksena, R.; Lu, D.; Celik, I. Ericsson Mobility Report. Report, Ericsson. 2021. Available online: <https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/november-2021> (accessed on 15 December 2021).
3. Number of Smartphone Users from 2016 to 2021. Website. 2021. Available online: <https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/> (accessed on 15 December 2021).
4. Vasavada, A.N.; Nevins, D.D.; Monda, S.M.; Hughes, E.; Lin, D.C. Gravitational demand on the neck musculature during tablet computer use. *Ergonomics* **2015**, *58*, 990–1004. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Jacobs, K.; Hudak, S.; McGiffert, J. Computer-Related posture and musculoskeletal discomfort in middle school students. *Work* **2009**, *32*, 275–283. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Siu, D.C.H.; Tse, L.A.; Yu, I.T.S.; Griffiths, S.M. Computer products usage and prevalence of computer related musculoskeletal discomfort among adolescents. *Work* **2009**, *34*, 449–454. [\[CrossRef\]](#)
7. Palmer, K.; Ciccarelli, M.; Falkmer, T.; Parsons, R. Associations between exposure to Information and Communication Technology (ICT) and reported discomfort among adolescents. *Work* **2014**, *48*, 165–173. [\[CrossRef\]](#)
8. Park, J.; Kim, J.; Kim, J.; Kim, K.; Kim, N.; Choi, I.; Lee, S.; Yim, J. The effects of heavy smartphone use on the cervical angle, pain threshold of neck muscles and depression. *Adv. Sci. Technol. Lett.* **2015**, *91*, 12–17. [\[CrossRef\]](#)
9. Shousha, T.; Hamada, H.; Abo-zaid, N.; Abdelsamee, M.; Behiry, M. The effect of smartphone use on neck flexion angle and hand grip power among adolescents: Cross-sectional study. *J. Hum. Sport Exerc.* **2021**, *16*, 883–891. [\[CrossRef\]](#)
10. Shaghayeghfard, B.; Ahmadi, A.; Maroufi, N.; Sarrafzadeh, J. Evaluation of forward head posture in sitting and standing positions. *Eur. Spine J.* **2016**, *25*, 3577–3582. [\[CrossRef\]](#)
11. Kang, J.H.; Park, R.Y.; Lee, S.J.; Kim, J.Y.; Yoon, S.R.; Jung, K.I. The effect of the forward head posture on postural balance in long time computer based worker. *Ann. Rehabil. Med.* **2012**, *36*, 98–104. [\[CrossRef\]](#)
12. Castro, A.; Pacheco, J.; Lourenço, C.; Queirós, S.; Moreira, A.; Rodrigues, N.; Vilaça, J. Evaluation of spinal posture using Microsoft Kinect™: A preliminary case-study with 98 volunteers. *Porto Biomed. J.* **2017**, *2*, 18–22. [\[CrossRef\]](#)

13. Digo, E.; Pierro, G.; Pastorelli, S.; Gastaldi, L. Evaluation of spinal posture during gait with inertial measurement units. *Proc. Inst. Mech. Eng. Part H J. Eng. Med.* **2020**, *234*, 1094–1105. [[CrossRef](#)] [[PubMed](#)]
14. Ormos, G. Survey of neck posture, mobility and muscle strength among schoolchildren. *Man. Med.* **2016**, *54*, 156–162. [[CrossRef](#)]
15. Martinez-Merinerio, P.; Nuñez-Nagy, S.; Achalandabaso-Ochoa, A.; Fernandez-Matias, R.; Pecos-Martin, D.; Gallego-Izquierdo, T. Relationship between Forward Head Posture and Tissue Mechanosensitivity: A Cross-Sectional Study. *Ann. Rehabil. Med.* **2012**, *36*, 98–104. [[CrossRef](#)]
16. Choi, K.H.; Cho, M.U.; Park, C.W.; Kim, S.Y.; Kim, M.J.; Hong, B.; Kong, Y.K. A Comparison Study of Posture and Fatigue of Neck According to Monitor Types (Moving and Fixed Monitor) by Using Flexion Relaxation Phenomenon (FRP) and Craniovertebral Angle (CVA). *Int. J. Environ. Res. Public Health* **2020**, *17*, 6345. [[CrossRef](#)]
17. Lee, S.; Choi, Y.H.; Kim, J. Effects of the cervical flexion angle during smartphone use on muscle fatigue and pain in the cervical erector spinae and upper trapezius in normal adults in their 20s. *J. Phys. Ther. Sci.* **2017**, *29*, 921–923. [[CrossRef](#)] [[PubMed](#)]
18. Marina Samaan, E.E.; Elnahhas, A.; Hendawy, A. Effect of prolonged smartphone use on cervical spine and hand grip strength in adolescence. *Int. J. Multidiscip. Res. Dev.* **2018**, *5*, 49–53.
19. Castien, R.; Blankenstein, N.; van der Windt, D.; Heymans, M.; Dekker, J. The Working Mechanism of Manual Therapy in Participants With Chronic Tension-Type Headache. *J. Orthop. Sport. Phys. Ther.* **2013**, *43*, 693–699. [[CrossRef](#)]
20. Shin, Y.; Kim, W.; Kim, S. Correlations among visual analogue scale, neck disability index, shoulder joint range of motion, and muscle strength in young women with forward head posture. *J. Exerc. Rehabil.* **2017**, *13*, 413–417. [[CrossRef](#)]
21. Singla, D.; Veqar, Z. Association Between Forward Head, Rounded Shoulders, and Increased Thoracic Kyphosis: A Review of the Literature. *J. Chiropr. Med.* **2017**, *16*, 220–229. [[CrossRef](#)]
22. Werth, A.J.; Babski-Reeves, K. Effects of portable computing devices on posture, muscle activation levels and efficiency. *Appl. Ergon.* **2014**, *45*, 1603–1609. [[CrossRef](#)]
23. Kinel, E.; Roncoletta, P.; Pietrangelo, T.; D’Amico, M. 3D Stereophotogrammetric Quantitative Evaluation of Posture and Spine Proprioception in Subacute and Chronic Nonspecific Low Back Pain. *J. Clin. Med.* **2022**, *11*, 546. [[CrossRef](#)] [[PubMed](#)]
24. Brunton, J.; Brunton, E.; Mhuiri, A.N. Reliability of measuring natural head posture using the craniovertebral angle. In *Proceedings of the Irish Ergonomics Society Annual Conference*; Irish Ergonomics Society: Limerick, Ireland, 2003; pp. 37–40.
25. van Niekerk, S.M.; Louw, Q.; Vaughan, C.; Grimmer-Somers, K.; Schreve, K. Photographic measurement of upper-body sitting posture of high school students: A reliability and validity study. *BMC Musculoskelet. Disord.* **2008**, *9*, 113. [[CrossRef](#)] [[PubMed](#)]
26. Salahzadeh, Z.; Maroufi, N.; Ahmadi, A.; Behtash, H.; Razmjoo, A.; Gohari, M.; Parnianpour, M. Assessment of forward head posture in females: Observational and photogrammetry methods. *J. Back Musculoskelet. Rehabil.* **2013**, *27*, 131. [[CrossRef](#)] [[PubMed](#)]
27. Weber, P.; Corrêa, E.C.R.; Milanesi, J.M.; Soares, J.C.; Trevisan, M.E. Cranio-cervical posture: Cephalometric and biophotogrammetric analysis. *Braz. J. Oral Sci.* **2012**, *11*, 416–421.
28. Iunes, D.; Bevilacqua-Grossi, D.; Oliveira, A.; Castro, F.; Salgado, H. Comparative analysis between visual and computerized photogrammetry postural assessment. *Braz. J. Phys. Ther.* **2009**, *13*, 308–315. [[CrossRef](#)]
29. Osokin, D. Real-time 2d multi-person pose estimation on cpu: Lightweight openpose. *arXiv* **2018**, arXiv:1811.12004.
30. Vincent, O.R.; Folorunso, O. A descriptive algorithm for sobel image edge detection. *Proc. Informing Sci. Educ. Conf. (InSITE)* **2009**, *40*, 97–107.
31. Derpanis, K.G. *The Harris Corner Detector*; York University: Toronto, ON, Canada, 2004; Volume 2.
32. Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, PAMI-8, 679–698. [[CrossRef](#)]
33. Brancati, R.; Cosenza, C.; Niola, V.; Savino, S. Experimental Measurement of Underactuated Robotic Finger Configurations via RGB-D Sensor. *Advances in Service and Industrial Robotics. RAAD 2018. Mech. Mach. Sci.* **2019**, *67*, 533. [[CrossRef](#)]
34. Viera, A.J.; Garrett, J.M. Understanding interobserver agreement: The kappa statistic. *Fam. Med.* **2005**, *37*, 360–363.
35. Krippendorff, K. *Content Analysis: An Introduction to Its Methodology*, 4th ed.; Sage Publications, International Educational and Professional Publisher: Thousand Oaks, CA, USA; London, UK; New Delhi, India, 2018; pp. 221–285.