

Article

Addressing Ergonomic Challenges in Agriculture through AI-Enabled Posture Classification

Siddhant Kapse^{1,2}, Ruoxuan Wu^{1,3} and Ornwipa Thamsuwan^{1,*} 

¹ Department of Mechanical Engineering, École de technologie supérieure, Montreal, QC H3C 1K3, Canada; siddhantkapse.iitkgp@gmail.com (S.K.); ruoxuanwu@outlook.com (R.W.)

² Department of Metallurgical and Material Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, West Bengal, India

³ School of Biomedical Sciences and Engineering, South China University of Technology, Guangzhou 511442, China

* Correspondence: ornwipa.thamsuwan@etsmtl.ca

Abstract: In this study, we explored the application of Artificial Intelligence (AI) for posture detection in the context of ergonomics in the agricultural field. Leveraging computer vision and machine learning, we aim to overcome limitations in accuracy, robustness, and real-time application found in traditional approaches such as observation and direct measurement. We first collected field videos to capture real-world scenarios of workers in an outdoor plant nursery. Next, we labeled workers' trunk postures into three distinct categories: neutral, slight forward bending and full forward bending. Then, through CNNs, transfer learning, and MoveNet, we investigated the effectiveness of different approaches in accurately classifying trunk postures. Specifically, MoveNet was utilized to extract key anatomical features, which were then fed into various classification algorithms including DT, SVM, RF and ANN. The best performance was obtained using MoveNet together with ANN (accuracy = 87.80%, precision = 87.46%, recall = 87.52%, and F1-score = 87.41%). The findings of this research contributed to the integration of computer vision techniques with ergonomic assessments especially in the outdoor field settings. The results highlighted the potential of correct posture classification systems to enhance health and safety prevention practices in the agricultural industry.

Keywords: ergonomics; posture; agriculture; computer vision; machine learning



Citation: Kapse, S.; Wu, R.; Thamsuwan, O. Addressing Ergonomic Challenges in Agriculture through AI-Enabled Posture Classification. *Appl. Sci.* **2024**, *14*, 525. <https://doi.org/10.3390/app14020525>

Academic Editor: Steven Davy

Received: 24 July 2023

Revised: 2 January 2024

Accepted: 3 January 2024

Published: 7 January 2024



Copyright: © 2024 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/>).

1. Introduction

Musculoskeletal disorders (MSDs) have been significant occupational health concerns, particularly in physically demanding industries such as agriculture. The prevalence of MSDs among agricultural workers has been alarming, with studies indicating a high prevalence of musculoskeletal symptoms and disorders worldwide [1]. These conditions not only affected the well-being and quality of life of workers but also resulted in economic burdens due to decreased productivity and increased healthcare costs [2].

In the agricultural sector, ergonomic interventions could play a vital role in minimizing the risk of MSDs and improving overall worker health and safety. Ergonomics, as the science of fitting work tasks to the capabilities and limitations of individuals, aims to optimize the interaction between workers and their work environment. Through ergonomic assessment [3], it is possible to identify and mitigate factors that contribute to non-neutral postures and repetitive movements, which are known risk factors for the development of MSDs [4].

Prolonged exposure to non-neutral postures, such as awkward trunk bending, and twisting places excessive stress on the musculoskeletal system, leading to tissue damage and the onset of musculoskeletal disorders [5]. Therefore, accurate and objective detection and characterization of postures are crucial for understanding ergonomic risks and implementing effective preventive measures.

This research aimed to apply existing yet novel technologies to assess and classify agricultural working posture into ergonomics-based categories. Particularly, Artificial Intelligence (AI) was utilized in ergonomic assessments to overcome the limitations of traditional methods to provide insights into the ergonomic risk factors among agricultural workers.

2. Related Works

Traditionally, postural analysis in ergonomic studies relied on manual observation, self-reporting, or subjective assessments, such as Ovako Working Posture Analysis System (OWAS) [6], Strain Index [7], Nordic Musculoskeletal Questionnaire [8], etc. However, these methods are subjective, time-consuming, and prone to bias.

Meanwhile, the field of ergonomics posture categorization has seen significant advancements in recent years, with researchers focusing on developing objective methods to assess and analyze human postures in various settings. Numerous studies have investigated different techniques and approaches for posture detection, aiming to improve ergonomics practices and promote better musculoskeletal health. One common approach in the literature involves the use of wearable sensors such as accelerometers [9,10] and inertial measurement units [11] (IMUs) to capture and analyze body movements and postures. Motion capture systems can be implemented based on a system of special markers, sensors or trackers located on the torso, hands and legs, combined into a system that can accurately record the movements of the entire body [12,13]. These sensors provide real-time data on joint angles, acceleration, and orientation, allowing for the quantification and classification of different postures.

The emergence of AI, particularly computer vision techniques [14], has provided new opportunities to automate and enhance the characterization of postures, including in the field of ergonomics. These methods utilize image and video processing algorithms to extract relevant features from visual data captured by cameras or depth sensors. Techniques such as image segmentation [15], posture estimation [16], and skeletal tracking [17] have been employed to identify and analyze body postures.

In the meantime, researchers have explored machine learning and deep learning to accurately classify and identify specific postures based on sensor data [18,19]. Specifically, transfer learning [20], a subfield of deep learning, has obtained attention for its ability to leverage pre-trained models to improve classification performance. Also, researchers have leveraged other deep learning architectures [21] and posture estimation models like OpenPose [22] and PoseNet [23] to achieve high accuracy and robustness in posture detection from image data.

Among these models, Movenet [24] stands out as a promising option for application in various scenarios, due to its strong ability to detect key joint features using image information. It has been proved that MoveNet can be applied to create a software for monitoring physical activities in the elderly [25], and can be extended to the classification of stroke patients based on videos captured by smartphones [26]. In scenarios where precise and complex measuring instruments are challenging to use for posture detection, MoveNet achieves more accurate results with simple image data alone, showcasing its immense application potential.

While significant progress has been made in posture detection research, there is still room for improvement. Challenges remain in dealing with variations in environmental conditions, clothing, and individual body characteristics. For instance, human activity classification may be confounded by the background of the image such as in a farm, and the clothes that a person wears, their gender role or ethnicity might be related to their occupation. Moreover, the practical implementation of posture detection systems in real-world agricultural settings requires considerations such as computational efficiency, system usability, and integration with existing agricultural practices.

3. Methodology

This section describes how we collected and prepared data then explored AI-based approaches like CNNs [27], transfer learning [20] and MoveNet [24] to improve the accuracy and enable real-time posture classification. Ultimately, we emphasized the significance of key anatomical landmarks to mitigate interference from background and other factors in the images by calculating geometric parameters and spatial relationships from these landmarks. This approach was expected to not only enhance accuracy but also reduce the model's input volume, leading to improving computational efficiency.

3.1. Data Collection

We conducted field visits to one plant nursery on three different days and captured images of workers engaged in various activities. Specifically, the videos were obtained from a group of eight workers present in an outdoor plant nursery. Among these workers, six were engaged in tasks such as weeding, plant trimming, and plant transportation. However, the remaining two workers served as mechanics in a workshop setting, and as a result, their videos were not included in the analysis. In general, all the workers participating in this study were usually exposed to extreme trunk flexion and lateral bending.

Each participant was assigned a researcher who followed them during their work tasks. We aimed to capture the workers in natural and unposed positions, reflecting real-world scenarios. Also, we paid attention to factors such as lighting conditions, camera angles, and framing. The researcher carried a video camera with a monopod to ensure stable footage and clear visibility. The recording sessions were conducted with the full consent of the workers, and they were fully informed about the purpose and use of the recorded videos. The research protocol was approved by the Research Ethics Board of the École de technologie supérieure (Reference H20211103, approved on 21 January 2022).

Additionally, to overcome the limitation of a small dataset, an open-source dataset was considered such as the MPII Human Pose dataset [28] and Open Data Commons [29], which contained a wide range of trunk posture images and environmental conditions. Their images that align with our specified classes were chosen, rather than the entire dataset. The combination of the datasets from fieldwork and these two benchmarking sources in our research were expected to enhance the generalization and robustness of our classification models. By incorporating external datasets, we aimed to introduce a diverse range of scenarios, ensuring that our models could effectively handle variations in posture categories, body sizes, and environmental conditions.

3.2. Data Labeling

To represent the diverse range of trunk postures, we defined and assigned three appropriate labels named “neutral”, “slight bend”, and “full bend” as shown in Figure 1. The choice of the three classes corresponded with the common Rapid Upper Limb Assessment [30] and Rapid Entire Body Assessment [31] methods, categorizing trunk postures in the forward or lateral bending ranges of 0–20 degrees, 20–60 degrees, and greater than 60 degrees, respectively. These labels enabled us to categorize and analyze the variations in trunk postures captured in the dataset. To our best knowledge, we ensured that no bias or skewed data were present in our model training and evaluation.

A total of 200 images were included in the study. The dataset was divided into two parts: the training dataset (85%) and the test dataset (15%). On the one hand, the training dataset was formed from images collected from the external data source. On the other hand, the test dataset was a mix of images from both the external data source and images collected from the agriculture field during our research.

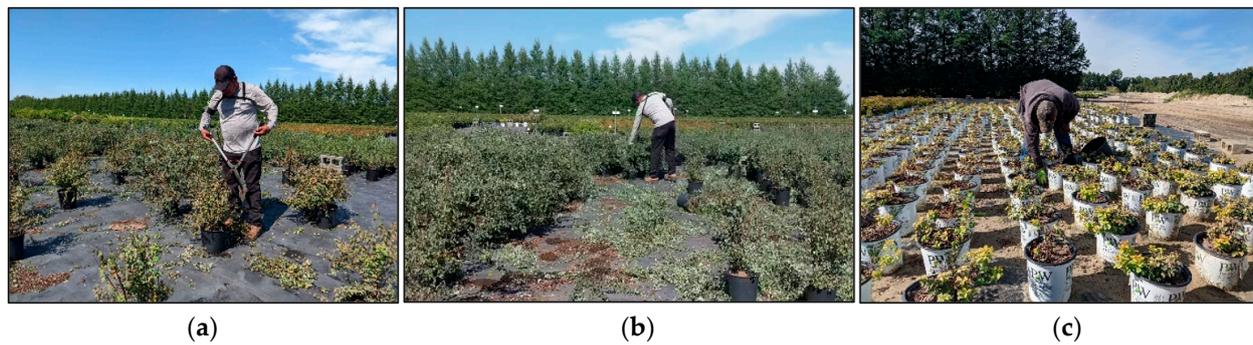


Figure 1. Sample images displaying different trunk posture classes—(a) Neutral Posture, (b) Half Bend Posture, and (c) Full Bend Posture.

3.3. Model Training

3.3.1. Convolutional Neural Network (CNN)

Initially, we employed a convolutional neural network (CNN) model [27], considering its widespread use and success in image classification. CNNs consist of (1) convolutional layers that apply filters to extract meaningful features from the input images, (2) pooling layers that reduce the spatial dimensions, preserving important features while reducing computational complexity, and (3) fully connected layers perform classification based on the extracted features. We trained a CNN model from scratch to establish a baseline performance for trunk posture classification. The model consists of three convolutional layers with increasing filter sizes (32, 64, and 128), each followed by a max-pooling layer. The output is then flattened and passed through two fully connected layers. We used ReLU activation functions in the convolutional layers and a SoftMax activation function in the output layer to classify the trunk posture categories. The model was trained using the Adam optimizer with a learning rate of 0.001 for 20 epochs.

Moreover, to mitigate overfitting and improve the model's generalization capability, we explored the application of data augmentation techniques. Data augmentation [32] involves applying various transformations, such as rotation, scaling, and flipping, to artificially expand the diversity of the training data.

3.3.2. Transfer Learning

Some popular pre-trained CNN models were explored, including VGG16, Inception, ResNet and MobileNet. VGG16 [33] is a CNN architecture known for its simplicity and effectiveness. It consists of 16 layers: 13 convolutional layers and 3 fully connected layers. Inception, also known as GoogLeNet [34], is characterized by its use of “inception modules” that efficiently capture information at multiple scales. Inception models employ various kernel sizes in parallel convolutional layers to capture both local and global features. This architecture has been successful in achieving high accuracy on image classification tasks while maintaining a relatively lightweight design. ResNet [35] is an architecture that introduced the concept of residual connections to address the problem of vanishing gradients in deep neural networks. By allowing information to bypass certain layers, ResNet enables the training of extremely deep networks. MobileNet [36] is a lightweight CNN aiming to provide efficient and accurate models that can be deployed on devices with limited computational resources and power constraints. MobileNet achieves this by utilizing depthwise separable convolutions, which separate the standard convolution into depthwise and pointwise convolutions. This reduces the number of parameters and computations required, making it more efficient while still maintaining good performance.

The pre-trained models were then fine-tuned on the collected dataset by adjusting the internal parameters and retraining the top layers to adapt them to the specific trunk posture classification task.

3.3.3. MoveNet Feature Extraction

MoveNet [37], renowned for its lightweight architecture and excellence in posture estimation, was utilized to extract key anatomical points from the collected images. The key points, as shown in Figure 2, included the positions and confidence scores of essential human anatomical points such as shoulder, elbow, wrist, hip, knee, ankle and more. These points provided valuable information about body joint positions, allowing for a more detailed analysis of trunk postures.



Figure 2. Output of MoveNet for Human Posture Estimation. The circles are anatomical key points, and the lines connected between some of the key points. The yellow lines are the visible ones and the purple lines are not visible but predicted via MoveNet.

We utilized the anatomical landmarks derived from MoveNet to generate a feature vector for the purpose of posture classification. This involved a series of steps, including centering the human in the image and scaling it to a uniform size. The resulting coordinate data were then transformed into a flattened feature vector.

3.3.4. Customized Feature Calculation and Selection

To enhance the accuracy of the classification task for agricultural bend detection, we calculated different features from MoveNet anatomical landmarks. These new features encompassed various aspects, including geometric properties and spatial relationships. We emphasized information of tilt angles and distances of different parts of the body. The tilt angles were calculated as the arctangent of the slope between the coordinate of two anatomical points (x_1, y_1) and (x_2, y_2) , i.e., the MoveNet landmarks as shown in Equation (1). The distances were calculated in terms of Euclidean distance between the pair of (x_1, y_1) and (x_2, y_2) coordinates using Equation (2):

$$\text{slope} = \left(\frac{y_1 - y_2}{x_1 - x_2} \right) \quad (1)$$

$$\text{distance} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (2)$$

To identify the most informative features and eliminate irrelevant ones, we employed the feature selection method of the Chi-Square Test, which is widely used in classification tasks [38–40]. This statistical method identifies features that are most likely to be indepen-

dent of the class labels and irrelevant for classification. The score of a chi-squared test was calculated to evaluate this relevance, according to Equations (3) and (4).

$$\text{Chi-square} = -\sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})}{E_{ij}} \quad (3)$$

where O_{ij} is the counts of samples with a certain value of features, and E_{ij} equals to:

$$E_{ij} = \frac{(n_{*j} n_{i*})}{n} \quad (4)$$

where n_{*j} is the counts of samples with a particular feature value, n_{i*} is the counts of samples belonging to a specific class, and n is the number of samples. All of these are calculated based on the assumption of independence [41]. A higher score indicates a higher difference between observed frequency and expected frequency, implying a stronger correlation between features and categories.

3.3.5. Classification of Trunk Posture

Based on the key points extracted using MoveNet; we employed various classification algorithms to accurately classify the trunk postures. The classification algorithms used in our study included Decision Trees (DT); Support Vector Machines (SVM); Random Forests (RF); and Artificial Neural Networks (ANN).

DT [42] is a simple yet powerful model that classifies based on a series of if-else conditions, by partitioning the data into different branches based on feature values. SVM [43] aims to find an optimal hyperplane to separate different classes in the feature space. RF [44] is an ensemble learning model that combines multiple DTs to improve classification accuracy. By aggregating the predictions of individual trees, RF can provide robust and reliable classifications. ANN [45], a deep learning model inspired by the structure of the human brain, consists of interconnected layers of artificial neurons. ANN can capture complex patterns and relationships in the data, making it suitable for tasks such as image classification.

Each algorithm was trained on the labeled key points, extracted from the MoveNet, to learn patterns and relationships between the key points and the corresponding trunk posture categories. The trained models were then used for classifying new instances of trunk postures.

3.4. Cross-Validation

Cross-validation [46] is a vital technique in machine learning to evaluate the performance of models on limited datasets and avoid overfitting. It involves partitioning the available data into multiple subsets, or “folds”, to train and test the model iteratively. By assessing the model’s performance on different subsets of the data, cross-validation provides more reliable performance metrics and enhances the model’s generalization capabilities.

In our research, we utilized stratified 3-fold cross-validation [47] to evaluate the effectiveness of our classification models. This ensured that each fold contained a representative distribution of the trunk posture classes, maintaining the original class proportions. This approach helps in preventing bias and ensures a more robust evaluation of the models’ performance across different trunk postures.

3.5. Evaluation Metrics

The performance of our model was evaluated using the numbers of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions. The evaluation metrics were accuracy, precision, recall or sensitivity, and F1-score, which provided a comprehensive assessment of both the model’s ability to correctly classify different postures and its overall performance.

Accuracy gives the proportion of the total number of predictions that were correct. Precision or the positive predictive value, is the fraction of positive values out of the total predicted positive instances. In other words, precision is the proportion of positive values that were correctly identified. Sensitivity, recall, or the TP rate is the fraction of positive values out of the total actual positive instances, i.e., the proportion of actual positive cases that are correctly identified). The F1 score is the harmonic mean of precision and sensitivity; it gives importance to both factors.

4. Results

4.1. CNN and Transfer Learning with Pre-Trained Models

During training the CNN model, we observed a training accuracy of 92.5% and a testing accuracy of 41.2%. These results indicate that the model suffered from significant overfitting and struggled to generalize well to unseen data. While data augmentation showed some improvement in reducing overfitting, the resulting performance enhancement was not substantial enough to meet our desired accuracy goals.

To improve the classification performance, we explored the use of transfer learning with pre-trained models. By fine-tuning the pre-trained models and replacing the last fully connected layer to match the number of classes in our problem, we achieved higher accuracies compared to the CNN from scratch model as shown in Table 1. These models demonstrated better generalization capabilities and achieved higher testing accuracies, but testing accuracies were still low.

Table 1. Performance metrics of transfer learning models.

| Pre-Trained Models | Training Results | | Test Results | | |
|--------------------|------------------|----------|--------------|--------|----------|
| | Accuracy | Accuracy | Precision | Recall | F1-Score |
| MobileNet | 99.77% | 65.56% | 67.96% | 65.56% | 63.11% |
| ResNet | 52.50% | 36.67% | 36.67% | 36.67% | 26.49% |
| Inception | 94.32% | 62.22% | 63.18% | 62.22% | 62.18% |
| VGG-16 | 99.77% | 60.00% | 62.96% | 60.00% | 58.98% |

4.2. MoveNet and Classification

Due to the limitations of the CNN models, we explored the use of MoveNet, a posture estimation model, for trunk posture detection. MoveNet provided us with key anatomical points representing body joint positions and angles relevant to trunk postures.

In the cross-validation phase, we observed the performance metrics of our classification models using stratified 3-fold cross-validation. Following the cross-validation phase, we tested the final trained models on our independent test dataset to evaluate their performance in real-world scenarios. Table 2 presents the accuracy obtained from the cross-validation for each model, as well as the accuracy, precision, recall, and F1 score achieved by each model on the test dataset.

Table 2. Performance metrics of the classification using MoveNet features.

| Pre-Trained Models | Training Results | | Test Results | | |
|--------------------|------------------|----------|--------------|--------|----------|
| | Accuracy | Accuracy | Precision | Recall | F1-Score |
| SVM | 82.91% | 69.05% | 68.05% | 67.15% | 66.64% |
| DT | 87.18% | 71.43% | 72.74% | 70.19% | 70.38% |
| RF | 91.45% | 71.43% | 71.88% | 70.67% | 70.87% |
| ANN | 94.02% | 80.49% | 80.61% | 78.93% | 79.92% |

The results demonstrated that MoveNet, combined with ANN, yielded improved accuracy in trunk posture classification compared to the other models. However, from the results, we can observe that the model has exhibited overfitting, indicating that we still need to find alternative methods to enhance the model's test performance.

4.3. Transformed and Selected Features and Classification

With the key anatomical points representing body joint positions relevant to trunk postures provided by MoveNet, we further obtained features like angles between segments that might be better related to trunk posture. We computed the Chi-Square Test scores for our calculated features, as shown in Figure 3. We selected several features with higher scores as representative input to models. Based on our experiments, it was proved that features such as the normalized planar angle between nose and hips, shoulders and hips, and the normalized distance between ears and ankles, hips and ankles, etc. were most helpful for model classification. We set a score threshold of 2.5, the 10 most significant features were selected for subsequent classification models.

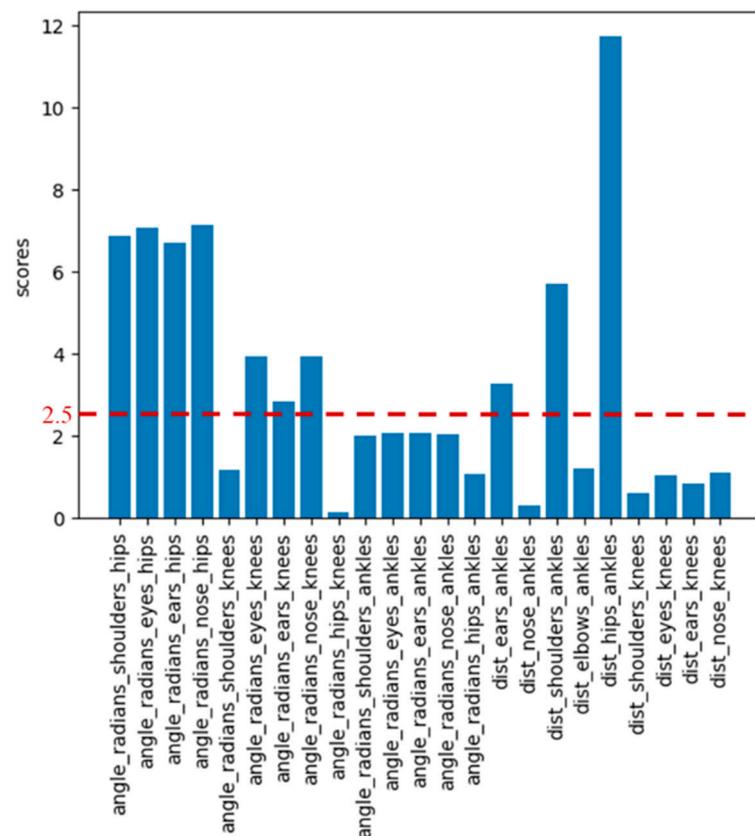


Figure 3. Bar chart showing chi-Square test scores (vertical axis) of all calculated features (horizontal axis), along with the cut-off score in the red line.

To gain a deeper understanding of the relationship between the selected 10 features and the targeted class labels, we generated violin plots for these selected features, as shown in Figure 4. Violin plots summarize statistics and density shape into a single plot, which provides an insightful visualization of the distribution of feature values concerning the class labels [48]. From the violin plots, we can see the frequency distribution of different eigenvalues. Also, the shapes of the violin plots listed in the figure were different between the three classes. In Figure 4, the violin plots are only shown for the normalized distance between ankles and hips, the normalized distance between shoulders and knees, the normalized planar angle between nose and hips and the normalized planar angle between nose and knees. The violin plots for the distance between ears and ankles exhibited a shape

akin to that of the distance between hips and ankles. Additionally, the planar angle between nose and hips, eyes and hips, shoulders and hips, ears and hips also shared a resemblance, indicating similar feature distributions. We did not delete similar features because some key parts (such as ears) would have been occluded in some agricultural activity pictures, particularly those captured from sideways angles or with individuals bent over.

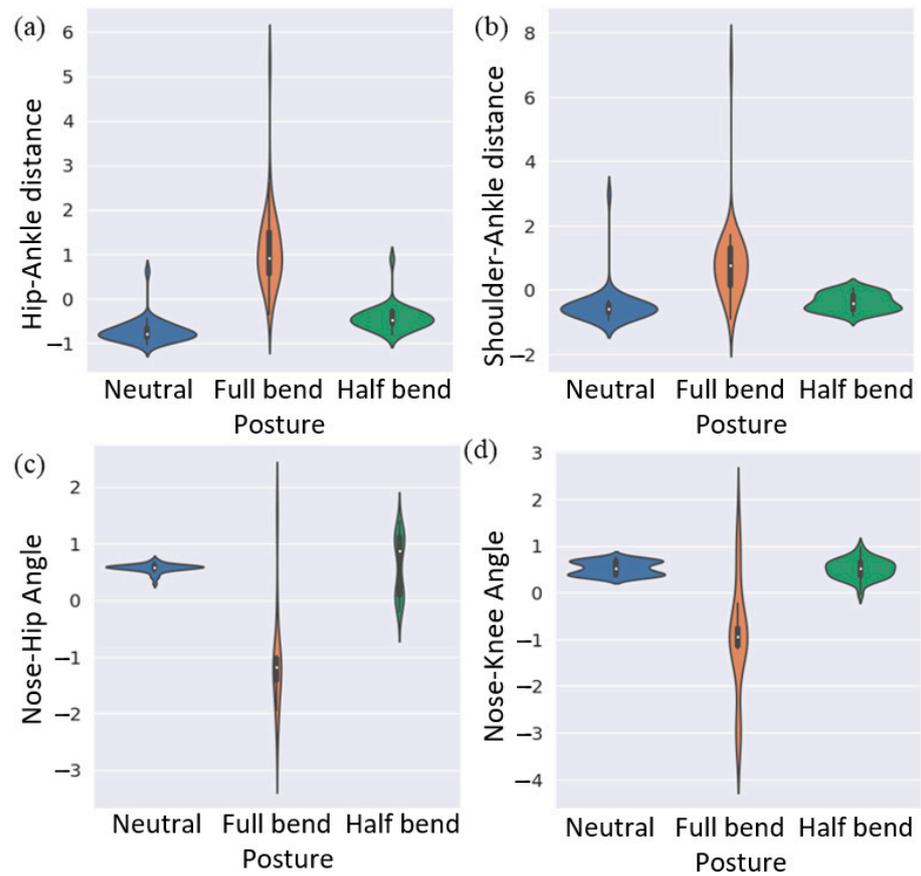


Figure 4. Violin plots for selected normalized features separated by trunk postures: (a) distance between hip and ankle, (b) distance between shoulder and ankle, (c) angle between the position of nose and hip, and (d) angle between the position of nose and knee.

With the nonlinear model that account for the selected features together like ANN, we could hopefully complete the classification task with better performance. We put the calculated features into the model and compared the result with applying the landmarks from Movenet directly. The new results are shown in Table 3 and confusion matrix in Figure 5.

Table 3. Performance metrics of the classification using new features.

| Pre-Trained Models | Training Results | | Test Results | | |
|--------------------|------------------|----------|--------------|--------|----------|
| | Accuracy | Accuracy | Precision | Recall | F1-Score |
| SVM | 85.47% | 80.48% | 81.78% | 80.48% | 80.83% |
| DT | 98.29% | 78.57% | 79.41% | 78.57% | 78.75% |
| RF | 98.29% | 85.36% | 87.92% | 85.36% | 86.53% |
| ANN | 94.44% | 87.80% | 87.46% | 87.52% | 87.41% |

Confusion Matrix of Posture Classification Model Using New Features

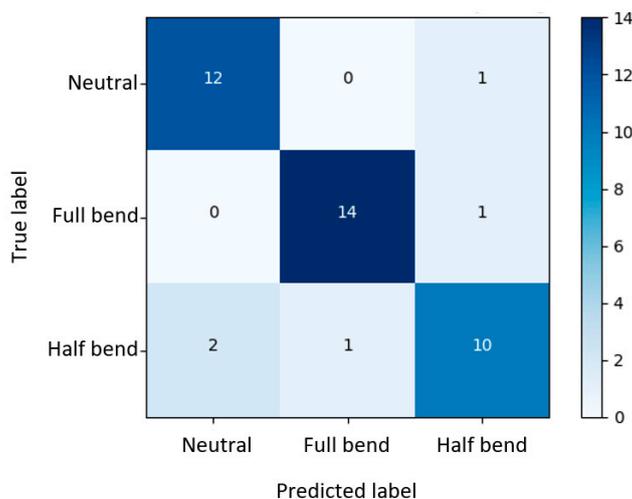


Figure 5. Confusion matrix of posture classification using the new slope and distance features and ANN model (the best model).

With the calculated and selected features that were most relevant to trunk bending, the model's test accuracy and F1 performance improved by almost 10%. From the results of the confusion matrix, it can be observed that the cases where predictions did not match the labels were primarily associated with the 'half bend' one, while 'fully bend' and 'neutral' were less prone to misclassification. This indicates that our model was capable of handling most of the posture classification tasks in this context, i.e., a setting in an outdoor plant nursery environment, with only minor enhancements in accuracy required. We also proved that the use of anatomical points and corresponding features allowed for a more comprehensive understanding of trunk postures, leading to more accurate classification results.

5. Discussion

We attempted to use CNNs and MoveNet for trunk posture classification within the domain ergonomics in agricultural field settings. Although not proposing a new CNN architecture, our study contributed to the understanding of the suitability of standard CNNs, and developed key features suitable for posture classification and pre-trained models like MoveNet for detecting and classifying postures in a unique environment. The application of these techniques offers insights and advantages in analyzing and evaluating trunk postures from images.

CNNs have been widely utilized for image classification and have demonstrated impressive performance in various domains. However, when applied to trunk posture classification in this study, i.e., the ergonomics context, CNNs encountered certain limitations. One of the main challenges was their susceptibility to variations in background image composition, lighting conditions, and occlusions in the outdoor settings. Apart from that, CNNs rely on large amounts of diverse and labeled training data to learn and generalize effectively; meanwhile the images obtained from our study were neither many nor diverse enough. This was due to the nature of ergonomic field study in a single location; that is, there was not much variety of tasks as compared to human daily activities. The complex nature of trunk postures, along with potential variations in clothing, body sizes, and environmental factors, could also lead to difficulties in accurately classifying different trunk posture categories using CNNs. Moreover, CNNs are generally computationally intensive, requiring significant computational resources and time for training and inference, which may limit their real-time application in agriculture field settings.

To address the limitations of CNNs, we adopted MoveNet for extracting original anatomical position features for trunk posture classification. MoveNet offered several

advantages in our research context. Firstly, MoveNet was specifically designed for posture estimation and could detect some key anatomical points, such as hips and shoulders. This made it well-suited for capturing detailed information about body joint positions and angles relevant to trunk postures. By leveraging MoveNet's capabilities, we could accurately identify and classify different trunk posture categories in images, providing valuable insights into ergonomic risk factors in the agricultural field.

Secondly, MoveNet, known for its lightweight architecture, made our image classification more computationally efficient. The model could process static images with fast inference times, enabling real-time trunk posture analysis even on resource-constrained devices commonly found in agricultural environments. This efficiency allows for practical deployment and usefulness of the system in real-world scenarios.

Additionally, MoveNet offered our study flexibility and scalability in the context of trunk posture detection. Meaning that, the model could be trained and fine-tuned on diverse datasets, encompassing a wide range of trunk postures in various image backgrounds including agricultural settings. Due to the aforementioned advantages, MoveNet has proven to be versatile and effective in various scenarios. However, as a general-purpose model, it gives all significant key points on the human body of which are redundant features that can interfere with the model's performance when applied. Therefore, to overcome these challenges, we further computed key features for posture classification based on the features extracted from MoveNet. For instance, we calculated planar angles between the nose and hips from the slope of the line between two points, among other features. We then employed feature selection methods to filter in only these crucial features for subsequent model training, resulting in improved classification outcomes. This adaptability ensured that the model could accurately classify trunk postures across various farming activities while accounting for individual variations, thus contributing to more comprehensive ergonomic evaluations and interventions.

It is worth noting that we also attempted to consider other recognized models, including YoloV5 [49], PoseNet [23] and OpenPose [22] in our study. However, Yolo required a substantial number of labeled images for effective training while the number of our dataset was small [50]. Given this limitation, MoveNet's posture estimation seemed a more feasible choice. Furthermore, while PoseNet specialized in whole-body posture estimation and OpenPose was quite comprehensive, they were computationally intensive for real-time applications. One of our key interests was the MoveNet's lightweight architecture, which enabled efficient real-time posture estimation. This lightweight nature was particularly valuable for our research involving dynamic agricultural scenarios, where the real-time analysis of workers' postures was crucial.

While our research provided valuable insights into posture detection in the agricultural field, there were certain limitations that should be acknowledged. Firstly, the number of images used or collected for our dataset was limited. This could potentially impact the generalizability of our model to other work tasks and environments outside the specific plant nursery setting.

Furthermore, we observed that there were fewer instances of the 'half bend' class in our dataset. This could be attributed to the nature of the work, where employees were primarily engaged in sustained bending (in 'full bend' class) and walking around (in 'neutral' class).

In addition, it is important to acknowledge that our research focused on a specific context and may have certain limitations in terms of its applicability to different work tasks and environments. Further research is needed to validate the effectiveness and generalizability of the approach in diverse agricultural settings.

6. Conclusions

This study attempted to employ various computer vision and machine learning techniques to characterize human trunk posture in ergonomic assessment. The limitations of CNNs in accurately classifying trunk postures from images prompted us to explore

alternative approaches. MoveNet, with its focus on posture estimation and lightweight architecture, emerged as a promising solution. However, some of the joint landmarks on the human body it computed, such as wrists and elbows, might not be suitable for our agricultural posture classification task and may interfere with the model's prediction results. By leveraging features relevant to targeted postures, we overcame the limitations of the CNNs and MoveNet and achieved accurate trunk posture classification in agricultural field ergonomics, which opens new possibilities for ergonomic assessments and musculoskeletal health improvement among agricultural workers. However, future research directions may involve constructing a more comprehensive feature library related to posture estimation in agricultural tasks, expanding the dataset to encompass a wider range of postures and environmental conditions, position feature reconstruction of occluded areas in pictures, and exploring the integration of complementary computer vision techniques to enhance the accuracy and robustness of trunk posture classification.

Author Contributions: Conceptualization, O.T.; methodology, S.K., R.W. and O.T.; software, S.K. and R.W.; validation, S.K., R.W. and O.T.; formal analysis, S.K. and R.W.; investigation, O.T.; resources, O.T.; data curation, S.K. and O.T.; writing—original draft preparation, S.K. and R.W.; writing—review and editing, S.K., R.W. and O.T.; visualization, S.K. and R.W.; supervision, O.T.; project administration, O.T.; funding acquisition, O.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by École de technologie supérieure start-up fund for new professor and Mitacs Globalink Research Internship program grant number IT33168. The publication fee was partially supported by the Natural Sciences and Engineering Research Council of Canada, Discovery Grant Program (RGPIN-2022-0327).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Research Ethics Committee of École de technologie supérieure (protocol code H20211103 and date of approval on 21 January 2022) for studies involving humans.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The image and video data collected in the field are not publicly available in order to protect the participants' privacy.

Acknowledgments: We are thankful to Union des producteurs agricoles, who helped us with participant recruitment, and the plant nursery, who allowed us to conduct the study onsite. We would also like to thank the research assistants, Rebeca Villanueva-Gomez, Amine Zougali and Mohamed Garouche, for their support during the data collection. Finally, we are grateful to all the participants for letting us observe them and for sharing their experience.

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

References

1. Barneo-Alcántara, M.; Díaz-Pérez, M.; Gómez-Galán, M.; Carreño-Ortega, Á.; Callejón-Ferre, Á.-J. Musculoskeletal disorders in agriculture: A review from web of science core collection. *Agronomy* **2021**, *11*, 2017. [\[CrossRef\]](#)
2. Naeini, H.S.; Karuppiyah, K.; Tamrin, S.B.; Dalal, K. Ergonomics in agriculture: An approach in prevention of work-related musculoskeletal disorders (WMSDs). *J. Agric. Environ. Sci.* **2014**, *3*, 33–51.
3. Westgaard, R.H.; Winkel, J. Ergonomic intervention research for improved musculoskeletal health: A critical review. *Int. J. Ind. Ergon.* **1997**, *20*, 463–500. [\[CrossRef\]](#)
4. Punnett, L.; Wegman, D.H. Work-related musculoskeletal disorders: The epidemiologic evidence and the debate. *J. Electromyogr. Kinesiol.* **2004**, *14*, 13–23. [\[CrossRef\]](#)
5. Keyserling, W.M.; Brouwer, M.; Silverstein, B.A. A checklist for evaluating ergonomic risk factors resulting from awkward postures of the legs, trunk and neck. *Int. J. Ind. Ergon.* **1992**, *9*, 283–301. [\[CrossRef\]](#)
6. Karhu, O.; Kansil, P.; Kuorinka, I. Correcting working postures in industry: A practical method for analysis. *Appl. Ergon.* **1977**, *8*, 199–201. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Garg, A.; Kapellusch, J.; Hegmann, K.; Wertsch, J.; Merryweather, A.; Deckow-Schaefer, G.; Malloy, E.J.; The WISTAH Hand Study Research Team. The Strain Index (SI) and Threshold Limit Value (TLV) for Hand Activity Level (HAL): Risk of carpal tunnel syndrome (CTS) in a prospective cohort. *Ergonomics* **2012**, *55*, 396–414. [\[CrossRef\]](#)

8. Kuorinka, I.; Jonsson, B.; Kilbom, A.; Vinterberg, H.; Biering-Sørensen, F.; Andersson, G.; Jørgensen, K. Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. *Appl. Ergon.* **1987**, *18*, 233–237. [[CrossRef](#)]
9. Thamsuwan, O.; Galvin, K.; Tchong-French, M.; Aulck, L.; Boyle, L.N.; Ching, R.P.; McQuade, K.J.; Johnson, P.W. Comparisons of physical exposure between workers harvesting apples on mobile orchard platforms and ladders, Part 2: Repetitive upper arm motions. *Appl. Ergon.* **2020**, *89*, 103192. [[CrossRef](#)]
10. Thamsuwan, O.; Galvin, K.; Tchong-French, M.; Aulck, L.; Boyle, L.N.; Ching, R.P.; McQuade, K.J.; Johnson, P.W. Comparisons of physical exposure between workers harvesting apples on mobile orchard platforms and ladders, Part 1: Back and upper arm postures. *Appl. Ergon.* **2020**, *89*, 103193. [[CrossRef](#)]
11. Sabatini, A.M. A review of wearable inertial sensors and algorithms for human motion pattern recognition. *Sensors* **2011**, *11*, 11556–11565.
12. Choo, C.Z.Y.; Chow, J.Y.; Komar, J. Validation of the Perception Neuron system for full-body motion capture. *PLoS ONE* **2022**, *17*, e0262730. [[CrossRef](#)] [[PubMed](#)]
13. Wu, Y.; Tao, K.; Chen, Q.; Tian, Y.; Sun, L. A Comprehensive Analysis of the Validity and Reliability of the Perception Neuron Studio for Upper-Body Motion Capture. *Sensors* **2022**, *22*, 6954. [[CrossRef](#)] [[PubMed](#)]
14. Seo, J.; Yin, K.; Lee, S. Automated Postural Ergonomic Assessment Using a Computer Vision-Based Posture Classification. *Constr. Res. Congr.* **2016**, *2016*, 809–818. [[CrossRef](#)]
15. Bulat, A.; Tzimiropoulos, G. Human Pose Estimation via Convolutional Part Heatmap Regression. *Lect. Notes Comput. Sci.* **2016**, *9911*, 717–732. [[CrossRef](#)]
16. Guler, R.A.; Neverova, N.; Kokkinos, I. DensePose: Dense Human Pose Estimation in the Wild. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7297–7306. [[CrossRef](#)]
17. Shotton, J.; Fitzgibbon, A.; Cook, M.; Sharp, T.; Finocchio, M.; Moore, R.; Kipman, A.; Blake, A. Real-time human pose recognition in parts from single depth images. In Proceedings of the CVPR 2011, Colorado Springs, CO, USA, 20–25 June 2011; pp. 1297–1304. [[CrossRef](#)]
18. Antwi-Afari, M.F.; Qarout, Y.; Herzallah, R.; Anwer, S.; Umer, W.; Zhang, Y.; Manu, P. Deep learning-based networks for automated recognition and classification of awkward working postures in construction using wearable insole sensor data. *Autom. Constr.* **2022**, *136*, 104181. [[CrossRef](#)]
19. Liaqat, S.; Dashtipour, K.; Arshad, K.; Assaleh, K.; Ramzan, N. A Hybrid Posture Detection Framework: Integrating Machine Learning and Deep Neural Networks. *IEEE Sens. J.* **2021**, *21*, 9515–9522. [[CrossRef](#)]
20. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]
21. Jose, J.; Shailesh, S. Yoga Asana Identification: A Deep Learning Approach. *IOP Conf. Ser. Mater. Sci. Eng.* **2021**, *1110*, 012002. [[CrossRef](#)]
22. Chen, W.; Jiang, Z.; Guo, H.; Ni, X. Fall Detection Based on Key Points of Human-Skeleton Using OpenPose. *Symmetry* **2020**, *12*, 744. [[CrossRef](#)]
23. Kendall, A.; Grimes, M.; Cipolla, R. PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015.
24. Zhu, J.; Cheng, C.; Shen, S.; Sun, L. MoveNet: Efficient Convolutional Neural Networks for Real-time Human Pose Estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021.
25. Dorado Chaparro, J.; Fernández-Bermejo Ruiz, J.; Santofimia Romero, M.J.; del Toro García, X.; Cantarero Navarro, R.; Bolaños Peño, C.; Llumiguano Solano, H.; Villanueva Molina, F.J.; Gonçalves Silva, A.; López, J.C. Phyx.io: Expert-Based Decision Making for the Selection of At-Home Rehabilitation Solutions for Active and Healthy Aging. *Int. J. Environ. Res. Public Health* **2022**, *19*, 5490. [[CrossRef](#)] [[PubMed](#)]
26. Feliandra, Z.B.; Khadijah, S.; Rachmadi, M.F.; Chahyati, D. Classification of Stroke and Non-Stroke Patients from Human Body Movements using Smartphone Videos and Deep Neural Networks. In Proceedings of the 2022 International Conference on Advanced Computer Science and Information Systems, Depok, Indonesia, 1–3 October 2022. [[CrossRef](#)]
27. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, *77*, 354–377. [[CrossRef](#)]
28. Andriluka, M.; Pishchulin, L.; Gehler, P.; Schiele, B. 2D Human Pose Estimation: New Benchmark and State of the Art Analysis. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014.
29. Miller, P.; Styles, R.; Heath, T. Open data commons, a license for open data. *Proc. LDOW* **2008**, *2008*, 369.
30. McAtamney, L.; Nigel Corlett, E. RULA: A survey method for the investigation of work-related upper limb disorders. *Appl. Ergon.* **1993**, *24*, 91–99. [[CrossRef](#)] [[PubMed](#)]
31. Hignett, S.; McAtamney, L. Rapid Entire Body Assessment (REBA). *Appl. Ergon.* **2000**, *31*, 201–205. [[CrossRef](#)]
32. Shorten, C.; Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*, 60. [[CrossRef](#)]
33. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
34. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.

35. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
36. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv* **2017**, arXiv:1704.04861.
37. Jo, B.; Kim, S. Comparative Analysis of OpenPose, PoseNet, and MoveNet Models for Pose Estimation in Mobile Devices. *Trait. Du Signal* **2022**, *39*, 119–124. [[CrossRef](#)]
38. Jin, X.; Xu, A.; Bie, R.; Guo, P. Machine Learning Techniques and Chi-Square Feature Selection for Cancer Classification Using SAGE Gene Expression Profiles. In *Data Mining for Biomedical Applications*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 106–115. [[CrossRef](#)]
39. Thaseen, I.S.; Kumar, C.A.; Ahmad, A. Integrated Intrusion Detection Model Using Chi-Square Feature Selection and Ensemble of Classifiers. *Arab. J. Sci. Eng.* **2019**, *44*, 3357–3368. [[CrossRef](#)]
40. Zhai, Y.; Song, W.; Liu, X.; Liu, L.; Zhao, X. A Chi-Square Statistics Based Feature Selection Method in Text Classification. In Proceedings of the 2018 IEEE 9th International Conference on Software Engineering and Service Science, Beijing, China, 23–25 November 2018; pp. 160–163. [[CrossRef](#)]
41. Rachburee, N.; Punlumjeak, W. A comparison of feature selection approach between greedy, IG-ratio, Chi-square, and mRMR in educational mining. In Proceedings of the 2015 7th International Conference on Information Technology and Electrical Engineering, Chiang Mai, Thailand, 29–30 October 2015; pp. 420–424. [[CrossRef](#)]
42. Quinlan, R. Induction of decision trees. *Mach. Learn.* **1986**, *1*, 81–106. [[CrossRef](#)]
43. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
44. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
45. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
46. Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*; O'Reilly Media: Sebastopol, CA, USA, 2019.
47. Hand, D.J.; Gneiting, T.; Raftery, A.E. Comparing predictive accuracy. *J. Bus. Econ. Stat.* **2001**, *19*, 321–333.
48. Hintze, J.L.; Nelson, R.D. Violin Plots: A Box Plot-Density Trace Synergism. *Am. Stat.* **1998**, *52*, 181–184. [[CrossRef](#)]
49. Jocher, G. Ultralytics YOLOv5. *Zenodo* **2020**. [[CrossRef](#)]
50. Jiang, P.; Ergu, D.; Liu, F.; Cai, Y.; Ma, B. A Review of YOLO Algorithm Developments. *Procedia Comput. Sci.* **2022**, *199*, 1066–1073. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.