Implementing Computational Geometry for Finger Counting in Color Images: A Non-Machine Learning Approach

Thato Senoamadi

7 June 2023

1 Introduction

The objective of this project is to devise an image processing methodology that does not rely on machine learning, but instead leverages computational geometry and image processing principles to identify and count the number of fingers visible in a color image of a cropped hand. The central algorithm driving this project is the convex hull algorithm, a fundamental tool widely utilized in both computational geometry and image processing. Our aim is to establish a reliable system capable of accurately discerning the number of raised fingers in an image. This system could hold significant potential for numerous applications, including gesture recognition, sign language interpretation, and beyond.

2 Dataset

The dataset harnessed for this project is procured from Kaggle, initially composed of more than 18,000 cropped images displaying both the right and left hand with varying numbers of fingers raised. For the purposes of this project, a subset of 300 images, equally distributed among the number of fingers raised from 0 to 5, has been selected. The choice of a smaller dataset is justified by the non-learning nature of the approach used in this study, which negates the need for a vast collection of images. Nevertheless, to maintain generalizability and avoid bias, the chosen images exhibit variation in size and the angle at which the hand is depicted.

3 Methodology

In this project, the Python programming language has been utilized, in conjunction with the OpenCV and Numpy libraries. The purpose of this section is to delineate the methodology employed, providing a detailed, step-by-step overview of the adopted approach.

3.1 Pre-processing

Upon arranging and categorizing the chosen images based on the count of raised fingers, where the name of each folder signifies the number of raised fingers in its contained pictures, we proceed with preprocessing our dataset. This involves a series of steps: Firstly, we convert the color image to grayscale. This simplifies further operations and reduces the image's dimensionality. Secondly, we implement a Gaussian filter to blur the image, aiming to minimize noise interference. Lastly, a binary threshold is applied to distinctively segregate the hand from the background. These steps are encapsulated in a function that yields the processed image alongside the binary image.

3.2 Identifying the Hand Contour

The subsequent stage of our methodology involves extracting the contour of the hand from the thresholded image. Given the structure of our dataset, the largest contiguous region in the image is anticipated to represent the hand. Therefore, we consider the contour with the maximum area as the hand contour.

The critical purpose behind this step is twofold:

1. Finger Detection: We utilize the hand contour to discern raised fingers by identifying the convexity defects within it. This identification is achieved through the application of the Convex Hull algorithm.

2. Hand Area Calculation: The hand contour is also instrumental in calculating the hand area. This computation is particularly significant to distinguish between images of a hand with no fingers raised and one with a single finger raised. The hand area provides a quantitative measure that aids in this differentiation.

3.3 Estimation of Palm Area

Recognizing the distinction between a closed fist and a hand with a single finger raised is crucial because the hand's overall shape and contour don't alter significantly between these two states. This similarity could lead to misclassification of closed fists. Nonetheless, the hand's area is expected to exhibit a significant change when moving from a closed fist to a single raised finger. This difference in area forms a reliable criterion for differentiation. The key obstacle lies in determining the precise threshold to use for this differentiation as it can vary greatly due to factors such as the camera's distance from the subject, variations in hand sizes, and other considerations. Given that hands with more than one finger raised generally have a substantially larger area than closed fists, our focus is on distinguishing between zero and one raised finger scenarios. To establish this threshold, the following steps are undertaken:

1. The average hand area across images featuring a closed fist is computed.
2. Subsequently, the average hand area where a single finger is raised is determined.
3. The threshold is then established at the midpoint between these two calculated average areas.

3.4 Finger Detection

The ultimate phase of our approach involves determining the count of raised fingers in the image. This operation is founded on the principles of convexity defects and hand area computation, which we've previously outlined. Following the computation of the threshold area, we first evaluate whether the area of the hand in the image exceeds this threshold. If this condition holds, we initially presume the presence of a single raised finger. This presumption stems from the method we use to compute the area threshold, which is the mean of areas corresponding to a closed hand and a single raised finger. Therefore, any hand area surpassing this threshold is indicative of at least one finger being raised. For recognizing the existence of further raised fingers, we analyze the convexity defects present. Each valid defect, satisfying predetermined depth and angle criteria, signifies an additional raised finger.

4 Assessment of Methodology

The proposed methodology was evaluated on a dataset comprising of 300 images, distributed evenly across six classes representing the number of raised fingers ranging from zero to five. The obtained results are as follows:

- Zero fingers raised: 78.00% accuracy
- One finger raised: 70.00% accuracy
- Two fingers raised: 78.00% accuracy
- Three fingers raised: 96.00% accuracy
- Four fingers raised: 100.00% accuracy
- Five fingers raised: 76.00% accuracy

The overall average accuracy of the method was found to be 83.00%, indicative of a promising approach. Notably, the methodology exhibits perfect accuracy for images displaying four raised fingers. This is attributed to the distinct convexity defects formed by each of the four extended fingers, satisfying the depth and angle conditions required for valid defect classification. Therefore, the algorithm can accurately detect the presence of each finger.

On the other hand, the method's least accurate performance is evident in cases where a single finger is raised. The low accuracy can potentially be ascribed to the visual similarities between images of a hand with one finger raised and other classes, such as a closed hand.

A notable constraint of the proposed methodology is its reliance on the quality of the input images. Factors such as insufficient isolation of the hand from the background, excessive variance in images, or a high level of noise can potentially diminish the performance of the system.

5 Reflective Analysis

The methodology implemented in this project hinges on the employment of non-learning based strategies, integral to Digital Image Processing. These techniques, including Image Segmentation, thresholding, and Convex sets, were applied to discern and enumerate the fingers visible in a colour image. The selection and execution of various steps throughout the approach were governed by considerations of problem nature, simplicity, and operational efficiency. Significant aspects of the methodology include:

1. Conversion to Grayscale: The decision to transition to grayscale was rooted in the desire for procedural simplicity and a reduction in data dimensionality. The grayscale conversion allows for the manipulation of a single channel as opposed to three (Red, Green, Blue), which streamlines subsequent processing and analysis stages.
2. Application of Gaussian Blur: The inclusion of Gaussian blur was purposed for noise mitigation. This operation smoothens the image and

minimizes minor distortions that could potentially disrupt the detection of the hand's contour and finger placement.

3. Implementation of the Convex Hull Algorithm: The Convex Hull algorithm is an integral component of the approach. It offers a straightforward and efficient means to identify hand contours and detect fingers via convexity defects. Given its capacity to create a convex polygon encompassing the hand, the algorithm aligns well with the natural structure of a spread hand, making it particularly suited to our problem.

6 Conclusion

The proposed method showcased the effectiveness of leveraging traditional Digital Image Processing techniques for finger detection and counting in hand images. The combination of grayscale conversion, Gaussian blur, and the Convex Hull algorithm provided an intuitive and efficient means of tackling the problem, achieving an average accuracy of 83.00% across various hand postures.

Despite this promising outcome, there remains potential for further enhancement. A major area of improvement lies in refining the image segmentation process to better cope with diverse lighting conditions and noise levels. Moreover, distinguishing between a closed fist and a hand with one finger raised represents another challenge due to their inherent visual similarities.

Looking ahead, the integration of machine learning approaches could contribute to overcoming these hurdles and boost overall accuracy. Such a blend of traditional and AI-powered techniques might afford the system greater adaptability, enabling it to cope with a wider range of hand sizes, image quality, and lighting conditions.

To conclude, this project has demonstrated a significant step towards robust finger counting solutions with potential applications in numerous fields. The exploration and refinement of this approach is, therefore, a worthwhile pursuit for future research.