# Computer Vision

Project 2023-24

Identification of American Sign Language Letters

Motivation

Sign languages are often used for communication with deaf people. The objective of these project is the identification of some letters of the American sign language using images of the hand of a person that is communicating using this kind of language.

From American Sign Language
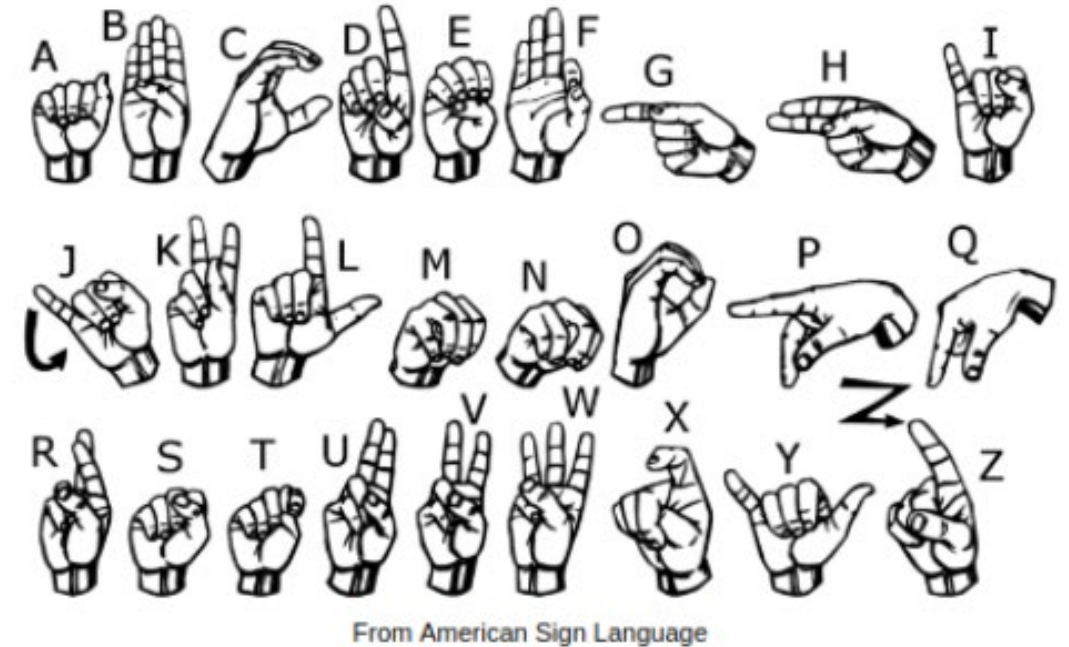
The project is divided into three tasks:

- In the first task, given an image showing one hand (and background and possibly part of the arm), the goal is the detection of the region of the interest (ROI) that contains the hand, also known as "Bounding Box" (BB).
- The second task consists in the identification of the letter that is being shown in an image, but considering only a subset of images showing 7 of the possible (26) letters.
- The third task is similar to the second one, but using now the bounding-boxes obtained in task 1 instead of the GT bounding-boxes.
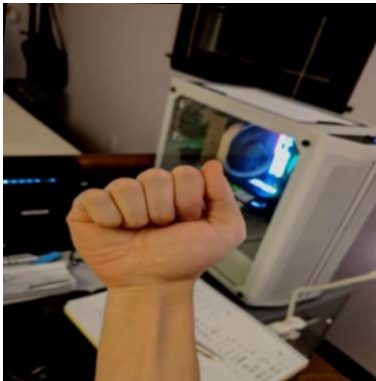
Project datasets were adapted from:
American Sign Language Letters Dataset (roboflow.com)
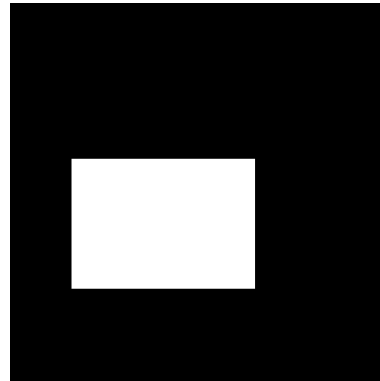
Computer Vision 2023-24
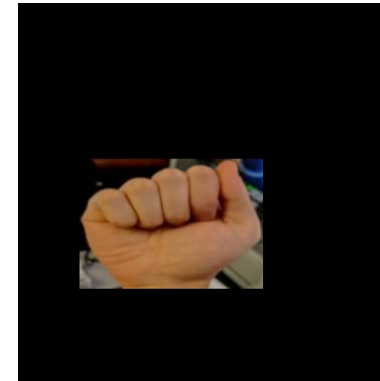
Task 1

Goal:

Development of an algorithm for the detection of the hand shown in the image and create a rectangular region of interest (Bounding Box) for the segmented area.

Original image   **GT (task1)**   **Input to Task2**

Datasets

A training dataset with 38 images is available for the development of task 1.

The ground truth (GT) dataset is formed by 38 binary images (masks) containing white rectangles corresponding to the positions of the hands in the training dataset.

Task 1

Evaluation

The result to obtain for each image is a rectangular area (bounding box) with the position of the hand as detected by the algorithm.

Using these results, and for the whole dataset, the following performance measures should be obtained.

- The number of true positives (TP):

    A segmented area is considered a TP if the corresponding Jaccard index is equal or larger than 0.5; the number of true positives is identical to the number of correctly detected hands for the dataset;

- The number of false positives (FP):

    A segmented area is considered a false positive if the corresponding Jaccard index is lower than 0.5 but greater than zero;

- The number of false negatives (FN):

    A false negative is a bounding box from the GT that does not correspond to any segmented region;

- The Recall, Precision and F_1-measure (R, P, F_1) for the whole dataset.

Task 2

Goal:

Development of an algorithm to identify a subset of letters of the American sign language. The letters to be identified are: letter A, letter B, letter C, letter I, letter L, letter V and letter W.

Dataset

The dataset to be used in this task is obtained from the corresponding datasets in task 1, as the input images are now the images showing only the hand areas as defined by the original images and the corresponding GT masks (example in slide 3).

The results of the classification are available in the "LetterClassification_GT" excel file.

Evaluation

The result to obtain for each image is the classification in one of the seven classes (A, B , C, I, L, V, W).

Using these results, and for each dataset, a confusion matrix should be presented containing the results of the classification process.

The Recall, Precision and F_1-measure (R, P, F_1) for the dataset classification results should also be reported.

Task 3

Goal:

Identical to task 2, but the input images are now the combination of the original images with the bounding-box masks obtained in task 1.

Evaluation

The results to obtain are the same as in task 2.

Some references

[1] Kakumanu, P.; Makrogiannis, S.; Bourbakis, N. A survey of skin-color modeling and detection methods. Pattern Recognition. 2007, 40, 1106–1122.

[2] V. Vezhnevets, V. Sazonov, A. Andreeva, A survey on pixel-based skin color detection techniques, GRAPHICON03, 2003, pp. 85–92.

[3] Buza, E.; Akagic, A.; Omanovic, S. Skin detection based on image color segmentation with histogram and K-means clustering. In Proceedings of the 10th International Conference on Electrical and Electronics Engineering (ELECO), Bursa, Turkey, 30 November–2 December 2017; pp. 1181–1186.

[4] Leite, M.; Parreira, W.D.; Fernandes, A.M.d.R.; Leithardt, V.R.Q. Image Segmentation for Human Skin Detection. Appl. Sci. 2022, 12, 12140. https://doi.org/10.3390/app122312140.

Students are encouraged to find their own references.

Major guidelines and Deliverables

- Students should form groups of **3-4 students**; however, upon request, a group can be formed by just one student. Groups with just one student are only required to develop Task 1.

- The delivery deadline for the report, code and contributions is May 17, 2024.

- The delivery deadline for the material of the presentation is May 19, 2024.

Deliverables

Each group is required to deliver:

- A four A4 page report in IEEE Transactions article format describing the methodology and obtained results, both qualitatively and quantitatively;

- The MATLAB code developed during the project;

- The material prepared for the public presentation of the project (15 minutes maximum for groups with 3-4 students; 10 minutes maximum for individual groups);

- A brief description of the contribution of each element of the group to the project.