

One task. Two runs. A short report.

Use the ready-made course website. No account, coding, installation, deployment or model training is required for this assignment.

[Open the course lab](https://huggingface.co/spaces/ZevonLew/AI_Powered_Image_Recognition): https://huggingface.co/spaces/ZevonLew/AI_Powered_Image_Recognition

1. Choose one task and a built-in sample

Choose Face verification, Gesture recognition or Text OCR. You only need one task. If you are unsure, open Text OCR and choose Lab card. Use a supplied sample before trying your own permitted image.

Face verification compares two face images. Gesture recognition classifies a supported static hand pose. OCR reads printed English. Each task answers a different question.

2. Make the first run

Select Run recognition and wait for the result. First use may load model files. Save a screenshot showing the sample, the setting and the actual result. This is your baseline.

3. Change one setting and run again

Keep the same input. Choose just one of these routes:

OCR: use Lab card; compare One text block with Sparse text. Check the actual returned words against the card.

Face: use Same person; compare distance thresholds 0.60 and 0.40. A lower threshold is stricter. Read the distance and the match decision separately.

Gesture: use Open palm; compare score thresholds 0.60 and 0.90. A higher threshold is stricter. Read the candidate, score and accepted label separately.

Click Run recognition again after changing the setting. Save a second screenshot. The result may stay the same; this is a valid observation. You do not need to create a failure or prove an improvement.

4. Write four to six sentences

Name your task and sample. State the first setting and result, then the changed setting and second result. Explain a difference or no change. Add one limitation, such as testing only this one sample. You may write in English or Chinese.

Use your own results. A high score alone does not prove correctness. For OCR, compare characters with the source. For faces and gestures, explain the task-specific output rather than calling every number accuracy.

5. Submit one page

Put both screenshots and your four to six sentences in a one-page PDF or document. Check that the task, sample, two settings and results can be identified. A JSON export is optional extra evidence.

You do not need to submit source code, create a new website, register an account, train a model or cover all three tasks. Reading the papers and repositories is optional deeper learning.

Explore the expanded sample library

The website contains 50 cases: 7 face, 18 gesture and 25 OCR cases. Use the Sample group menu for starter examples, all seven named gesture poses, varied inputs, difficult cases and natural images. Select a case, then run it. Case names describe intended inputs, not guaranteed predictions.

Gesture categories are Open_Palm, Closed_Fist, Victory, Thumb_Up, Thumb_Down, Pointing_Up and ILoveYou. Try occlusion, two hands, a distant hand and no hand as well. Face cases include a mask, a group portrait and an empty scene. OCR cases include menus, notices, one line, sparse labels, similar characters, small print and handwriting.

In Text OCR, choose **Natural images**. This group contains the existing handwritten note and seven new scenes: a road sign, a waiting hall, a bus stop, shop opening hours, a sidewalk cafe menu, supermarket price labels and library directions. These are AI-generated teaching scenes with manually checked main text. Start with Sparse text for signs or separated labels; compare One text block when the words form a compact block. Check the words, numbers and reading order against the image. Background objects and arrows may produce extra characters.

In our actual local checks, some small or faint print succeeded. A new palm had a correct raw candidate but a score of 0.593, below 0.60; a new fist returned no detected hand. An identifier changed from ROOM 1002 to ROOM 1002. Inspect the actual output before naming an error.

Optional extensions: choose one after the core report

Extension 1 — compare recognizers (advanced): choose one task and the same checked inputs. For OCR, a separately configured recognizer such as PaddleOCR may be compared with Tesseract. A second recognizer is not integrated on this website.

Extension 2 — tune a parameter (easy): keep the input and backend fixed. Change only one existing threshold or layout setting. Explain the actual difference or lack of change.

Extension 3 — test a new domain (easy): choose one new condition such as a group portrait, occluded hand or photographed label. Record what changed and what was returned.

Extension 4 — analyze an error (easy): distinguish missing input, no detection, rejected candidate and incorrect prediction. Show the input and output, then propose one next test.

Extension 5 — evaluate a small set (easy): write reference answers, record predictions under fixed settings and count checked correct, partial and wrong outcomes. Do not infer general accuracy from this synthetic library.

Extension 6 — train or fine-tune (advanced bonus): use permitted labeled examples, keep a pretrained baseline and reserve an independent final test. Coding and a separate training workflow are needed; training is not required for the core report.

If a run cannot finish

Wait for loading, retry the same built-in sample and check the displayed message. If the public site remains unavailable, save the message and report the blocked run to your instructor. Do not invent an output or treat an error screen as a prediction.

Use and check the demo responsibly

Use synthetic examples or images you have permission to process. All three tasks run in your browser; uploaded images are not sent to an inference server. Gesture recognition requires browser WebGL2 support. Keep the input and device fixed between runs.

Face comparison is not identity proof. Static gestures are not sign-language translation. English OCR still needs checking against the image. One built-in example cannot prove general accuracy.

