

Assignment: Automatic Hand Tracking

Goal: Build an automatic pipeline that tracks hand movements in a video.

Tools: cv2, Google MediaPipe, SAM 2

Preliminaries

Data

Download the video [test.mp4](#).

Environment Set Up

- `conda create -n sam2 python=3.12`
- `conda activate sam2`
- `pip3 install torch torchvision torchaudio --index-url https://download.pytorch.org/whl/cu118`
 - May take a bit.
 - See [PyTorch installation](#) if this doesn't work.
- Follow the instructions for setting up SAM 2 [here](#)
- Install **mediapipe** with `pip install -q mediapipe`
- And **cv2** with `pip install opencv-python`

```

apple ~ ~/Documents/NYU
└ mkdir sam2-assign-ad

apple ~ ~/Documents/NYU
└ cd sam2-assign-ad

apple ~ ~/Documents/NYU/sam2-assign-ad
└ git init
hint: Using 'master' as the name for the initial branch. This default branch name
hint: is subject to change. To configure the initial branch name to use in all
hint: of your new repositories, which will suppress this warning, call:
hint:
hint:   git config --global init.defaultBranch <name>
hint:
hint: Names commonly chosen instead of 'master' are 'main', 'trunk' and
hint: 'development'. The just-created branch can be renamed via this command:
hint:
hint:   git branch -m <name>
Initialized empty Git repository in /Users/ad12/Documents/NYU/sam2-assign-ad/.git/
apple ~ ~/Documents/NYU/sam2-assign-ad
└ git remote add origin https://github.com/mannadamay12/sam2-assign-ad.git

apple ~ ~/Documents/NYU/sam2-assign-ad
└ git status
On branch master

No commits yet

nothing to commit (create/copy files and use "git add" to track)

apple ~ ~/Documents/NYU/sam2-assign-ad
└ code .

```

```

apple ~ ~/Documents/NYU/sam2-assign-ad
└ conda create -n sam2 python=3.12
Retrieving notices: ...working... done
Channels:
- defaults
Platform: osx-arm64
Collecting package metadata (repodata.json): done
Solving environment: done

## Package Plan ##

environment location: /opt/homebrew/Caskroom/miniconda/base/envs/sam2

added / updated specs:
- python=3.12

The following packages will be downloaded:

  package          |      build
  -----|-----
  ca-certificates-2024.12.31 | hca03da5_0      129 KB
  expat-2.6.4       | h313beb8_0      152 KB
  python-3.12.8     | h99e199e_0      13.9 MB
  tzdata-2025a      | h04d1e81_0      117 KB
  -----
                           Total:    14.3 MB

The following NEW packages will be INSTALLED:

  bzip2           pkgs/main/osx-arm64::bzip2-1.0.8-h80987f9_6
  ca-certificates pkgs/main/osx-arm64::ca-certificates-2024.12.31-hca03da5_0
  expat           pkgs/main/osx-arm64::expat-2.6.4-h313beb8_0
  libcxx           pkgs/main/osx-arm64::libcxx-14.0.6-h848a8c0_0
  libffi           pkgs/main/osx-arm64::libffi-3.4.4-hca03da5_1
  ncurses          pkgs/main/osx-arm64::ncurses-6.4-h313beb8_0
  openssl          pkgs/main/osx-arm64::openssl-3.0.15-h80987f9_0
  pip              pkgs/main/osx-arm64::pip-24.2-py312hca03da5_0
  python           pkgs/main/osx-arm64::python-3.12.8-h99e199e_0
  readline         pkgs/main/osx-arm64::readline-8.2-h1a28f6b_0
  setuptools       pkgs/main/osx-arm64::setuptools-75.1.0-py312hca03da5_0
  sqlite           pkgs/main/osx-arm64::sqlite-3.45.3-h80987f9_0

```

```
tk          pkgs/main/osx-arm64::tk-8.6.14-h6ba3021_0
tzdata     pkgs/main/noarch::tzdata-2025a-h04d1e81_0
wheel      pkgs/main/osx-arm64::wheel-0.44.0-py312hca03da5_0
xz          pkgs/main/osx-arm64::xz-5.4.6-h80987f9_1
zlib       pkgs/main/osx-arm64::zlib-1.2.13-h18a0788_1

Proceed ([y]/n)? y

Downloading and Extracting Packages:

Preparing transaction: done
Verifying transaction: done
Executing transaction: done
#
# To activate this environment, use
#
#   $ conda activate sam2
#
# To deactivate an active environment, use
#
#   $ conda deactivate

[?] ~ /~/Documents/NYU/sam2-assign-ad git master  ✓ 9s ✘ base
conda activate sam2

[?] ~ /~/Documents/NYU/sam2-assign-ad git master  ✓ sam2
pip3 install torch torchvision torchaudio --index-url https://download.pytorch.org/whl/cu118
Looking in indexes: https://download.pytorch.org/whl/cu118
ERROR: Could not find a version that satisfies the requirement torch (from versions: none)
ERROR: No matching distribution found for torch

[?] ~ /~/Documents/NYU/sam2-assign-ad git master  ✓ 1 ✘ 5s ✘ sam2
conda install pytorch::pytorch torchvision torchaudio -c pytorch
Channels:
- pytorch
- defaults
Platform: osx-arm64
Collecting package metadata (repodata.json): done
Solving environment: done
```

```
Downloading and Extracting Packages:

Preparing transaction: done
Verifying transaction: done
Executing transaction: done

[?] ~ /~/Documents/NYU/sam2-assign-ad git master ?1
git clone https://github.com/facebookresearch/sam2.git && cd sam2
Cloning into 'sam2'...
remote: Enumerating objects: 1070, done.
remote: Total 1070 (delta 0), reused 0 (delta 0), pack-reused 1070 (from 1)
Receiving objects: 100% (1070/1070), 134.70 MiB | 36.31 MiB/s, done.
Resolving deltas: 100% (376/376), done.

[?] ~ /~/Documents/NYU/sam2-assign-ad/sam2 git main
pip install -e .
Obtaining file:///Users/ad12/Documents/NYU/sam2-assign-ad/sam2
  Installing build dependencies ... done
    Checking if build backend supports build_editable ... done
    Getting requirements to build editable ... done
    Preparing editable metadata (pyproject.toml) ... done
Requirement already satisfied: torch>=2.5.1 in /opt/homebrew/Caskroom/miniconda/base/envs/sam2/lib/python3.12/site-packages (from SAM-2==1.0) (2.5.1)
Requirement already satisfied: torchvision>=0.20.1 in /opt/homebrew/Caskroom/miniconda/base/envs/sam2/lib/python3.12/site-packages (from SAM-2==1.0) (0.20.1)
Requirement already satisfied: numpy>=1.24.4 in /opt/homebrew/Caskroom/miniconda/base/envs/sam2/lib/python3.12/site-packages (from SAM-2==1.0) (2.0.1)
Collecting tqdm>=4.66.1 (from SAM-2==1.0)
  Using cached tqdm-4.67.1-py3-none-any.whl.metadata (57 kB)
Collecting hydra-core>=1.3.2 (from SAM-2==1.0)
  Downloading hydra_core-1.3.2-py3-none-any.whl.metadata (5.5 kB)
Collecting iopath>=0.1.10 (from SAM-2==1.0)
  Downloading iopath-0.1.10.tar.gz (42 kB)
  Preparing metadata (setup.py) ... done
Requirement already satisfied: pillow>=9.4.0 in /opt/homebrew/Caskroom/miniconda/base/envs/sam2/lib/python3.12/site-packages (from SAM-2==1.0) (11.0.0)
Collecting omegaconf<2.4,>=2.2 (from hydra-core>=1.3.2->SAM-2==1.0)
  Downloading omegaconf-2.3.0-py3-none-any.whl.metadata (3.9 kB)
Collecting antlr4-python3-runtime==4.9.* (from hydra-core>=1.3.2->SAM-2==1.0)
  Downloading antlr4-python3-runtime-4.9.3.tar.gz (117 kB)
  Preparing metadata (setup.py) ... done
Collecting packaging (from hydra-core>=1.3.2->SAM-2==1.0)
  Using cached packaging-24.2-py3-none-any.whl.metadata (3.2 kB)
```

```

Building wheel for iopath (setup.py) ... done
Created wheel for iopath: filename=iopath-0.1.10-py3-none-any.whl size=31528 sha256=9aabffcce3f0a46263f9e680b714802b42df1d7a6ce7c68ef
92df52fb83934d5
Stored in directory: /Users/ad12/Library/Caches/pip/wheels/7c/96/04/4f5f31ff812f684f69f40cb1634357812220aac58d4698048c
Successfully built SAM-2 antlr4-python3-runtime iopath
Installing collected packages: antlr4-python3-runtime, tqdm, sympy, portalocker, packaging, omegaconf, fsspec, iopath, hydra-core, SAM-2
Attempting uninstall: sympy
  Found existing installation: sympy 1.13.3
  Uninstalling sympy-1.13.3:
    Successfully uninstalled sympy-1.13.3
Successfully installed SAM-2-1.0 antlr4-python3-runtime-4.9.3 fsspec-2024.12.0 hydra-core-1.3.2 iopath-0.1.10 omegaconf-2.3.0 packaging-24.2 portalocker-3.1.1 sympy-1.13.1 tqdm-4.67.1

```

```

cd ..
pip install -q mediapipe

pip install opencv-python
Collecting opencv-python
  Downloading opencv_python-4.11.0.86-cp37abi3-macosx_13_0_arm64.whl.metadata (20 kB)
Requirement already satisfied: numpy>=1.21.2 in /opt/homebrew/Caskroom/miniconda/base/envs/sam2/lib/python3.12/site-packages (from opencv-python) (1.26.4)
  Downloading opencv_python-4.11.0.86-cp37abi3-macosx_13_0_arm64.whl (37.3 MB)
    37.3/37.3 MB 35.5 MB/s eta 0:00:00
Installing collected packages: opencv-python
Successfully installed opencv-python-4.11.0.86

```

Part 1: Detect Hands in the First Frame

References: [Google MediaPipe](#), [Google Colab Sample](#)

Using the references above, write code that can output information on hand location(s) in an image (represented however you wish - e.g., numpy array). This instruction is intentionally vague. Your goal is to return an output that can be used as a SAM 2 prompt (e.g., clicks, bounding boxes). Feel free to look into [SAM 2 Video Predictor Example](#) for example prompts.



Part 2: Use Part 1 and SAM 2 to Track Hands

References: [SAM 2 Repo](#), [SAM 2 Video Predictor Example](#)

Write a function that uses SAM 2 and the results from Part 1 to generate masks for every frame of the video.

- There can be multiple masks per frame.
- The results from Part 1 are your *input prompts* into SAM 2.
- Ideally, the parameters of this function are an input and output path. The function will write a new video with the masks to the output path.

Deliverables

- 1) **Link to Github code:** code style and organization are important and will be evaluated!
Include a README / pip requirements file for set up.
- 2) **Output video demo:** an output video with the masked hands