

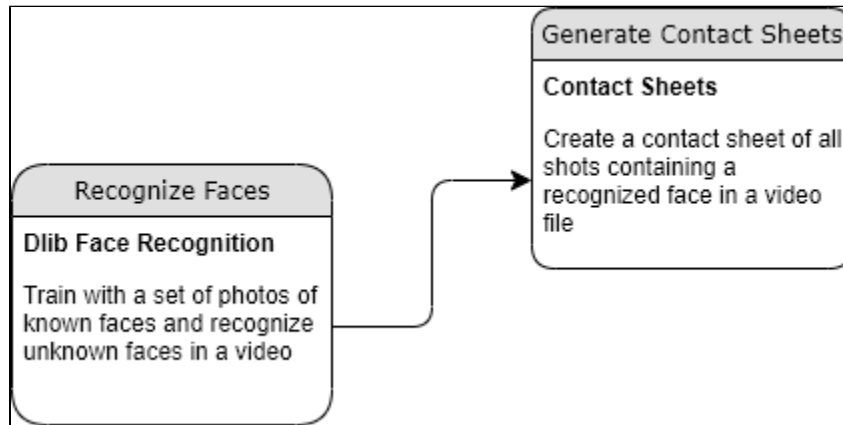
MGM - Facial Recognition

- [Category description and use cases](#)
- [Output standard](#)
- [Recommended tool\(s\)](#)
 - [Python face_recognition](#)
- [Evaluation summary](#)

Category description and use cases

To allow collection managers to locate known persons in collections materials. If, for example, a collection has many images of someone important to their institution and suspects they appear in video footage but would like to confirm, or would like to know where in a video the person appears, a face recognition tool would provide that information.

Workflow example:



Output standard

Summary:

Element	Datatype	Obligation	Definition
media	object	required	Wrapper for metadata about the source media file.
media.filename	string	required	Filename of the source file.
media.duration	string	required	The duration of the source file.
media.frameRate	number	required	The frame rate of the video, in FPS.
media.numFrames	number	required	The number of frames in the video.
media.resolution	object	required	Resolution of the video.
media.resolution.width	number	required	Width of the frame, in pixels.
media.resolution.height	number	required	Height of the frame, in pixels.
frames	array	required	List of frames containing identified faces.
frames[*]	object	optional	A frame containing an identified face.
frames[*].start	string (s.fff)	required	Time of the frame, in seconds.
frames[*].objects	list	required	List of bounding boxes in the frame containing identified faces.
frames[*].objects[*]	object	required	A bounding box in the frame containing an identified face.
frames[*].objects[*].name	string	required	The name of the face within the bounding box.
frames[*].objects[*].score	object	optional	A confidence or relevance score for the face.
frames[*].objects[*].score.type	string (confidence relevance)	required	The type of score, confidence or relevance.
frames[*].objects[*].score.value	number	required	The score value, typically a number in the range of 0-1.

frames[*].objects[*].vertices	object	optional	The top left (xmin, ymin) and bottom right (xmax, ymax) relative bounding coordinates.
frames[*].objects[*].vertices.xmin	number	required	The top left x coordinate.
frames[*].objects[*].vertices.ymin	number	required	The top left y coordinate.
frames[*].objects[*].vertices.xmax	number	required	The bottom right x coordinate.
frames[*].objects[*].vertices.ymax	number	required	The bottom right y coordinate.

JSON Schema

Schema

```
{
  "$schema": "http://json-schema.org/schema#",
  "type": "object",
  "title": "Facial recognition Schema",
  "required": [
    "media",
    "frames"
  ],
  "properties": {
    "media": {
      "type": "object",
      "title": "Media",
      "description": "Wrapper for metadata about the source media file.",
      "required": [
        "filename",
        "duration"
      ],
      "properties": {
        "filename": {
          "type": "string",
          "title": "Filename",
          "description": "Filename of the source file.",
          "default": "",
          "examples": [
            "myfile.wav"
          ]
        },
        "duration": {
          "type": "string",
          "title": "Duration",
          "description": "Duration of the source file audio.",
          "default": "",
          "examples": [
            "25.888"
          ]
        }
      },
      "frameRate": {
        "type": "number",
        "title": "Frame rate",
        "description": "The frame rate of the video, in FPS.",
        "default": 0,
        "examples": [
          29.97
        ]
      },
      "numFrames": {
        "type": "integer",
        "title": "Number of frames",
        "description": "The number of frames in the video.",
        "default": 0,
        "examples": [
          1547
        ]
      },
      "resolution": {
```

```

        "type": "object",
        "title": "Resolution",
        "description": "Resolution of the video.",
        "required": [
            "height",
            "width"
        ],
        "properties": {
            "height": {
                "type": "integer",
                "title": "Height",
                "description": "Height of the frame, in pixels.",
                "default": 0
            },
            "width": {
                "type": "integer",
                "title": "Width",
                "description": "Width of the frame, in pixels.",
                "default": 0
            }
        }
    },
    },
    "frames": {
        "type": "array",
        "title": "Frames",
        "description": "List of frames containing identified faces.",
        "items": {
            "type": "object",
            "required": [
                "start",
                "objects"
            ],
            "properties": {
                "start": {
                    "type": "string",
                    "title": "Start",
                    "description": "Time of the frame, in seconds.",
                    "default": "",
                    "examples": [
                        "23.594"
                    ]
                },
                "objects": {
                    "type": "array",
                    "title": "Bounding boxes",
                    "description": "List of bounding boxes in the frame containing
identified faces.",
                    "items": {
                        "type": "object",
                        "required": [
                            "name"
                        ],
                        "properties": {
                            "name": {
                                "type": "string",
                                "title": "Text",
                                "description": "The name of the identified face
within the bounding box.",
                                "default": ""
                            }
                        },
                        "score": {
                            "type": "object",
                            "title": "Score",
                            "description": "A confidence or relevance score for the
entity.",
                            "required": [
                                "type",
                                "scoreValue"
                            ],

```

```

        "properties": {
          "type": {
            "type": "string",
            "title": "Type",
            "description": "The type of score, confidence or
relevance.",
            "enum": [
              "confidence",
              "relevance"
            ]
          },
          "scoreValue": {
            "type": "number",
            "title": "Score value",
            "description": "The score value, typically a float
in the range of 0-1.",
            "default": 0,
            "examples": [0.437197]
          }
        },
        "vertices": {
          "type": "object",
          "title": "Vertices",
          "description": "The top left (xmin, ymin) and
bottom right (xmax, ymax) relative bounding coordinates.",
          "required": [
            "xmin",
            "ymin",
            "xmax",
            "ymax"
          ],
          "properties": {
            "xmin": {
              "type": "number",
              "title": "Xmin",
              "description": "The top left x
coordinate.",
              "default": 0
            },
            "ymin": {
              "type": "number",
              "title": "Ymin",
              "description": "The top left y
coordinate.",
              "default": 0
            },
            "xmax": {
              "type": "number",
              "title": "Xmax",
              "description": "The bottom right
x coordinate.",
              "default": 0
            },
            "ymax": {
              "type": "number",
              "title": "Ymax",
              "description": "The bottom right
y coordinate.",
              "default": 0
            }
          }
        }
      }
    }
  }
}

```

Sample output

Sample Output

```
{
  "media": {
    "filename": "myfile.mov",
    "duration": "8334.335",
    "frameRate": 30.000,
    "numFrames": 1547,
    "resolution": {
      "width": 654,
      "height": 486
    }
  },
  "frames": [
    {
      "start": "625.024",
      "objects": [
        {
          "name": "Herman B. Wells",
          "score": {
            "type": "confidence",
            "scoreValue": 0.9903119
          },
          "vertices": {
            "xmin": 219,
            "ymin": 21,
            "xmax": 340,
            "ymax": 53
          }
        }
      ]
    }
  ]
}
```

Recommended tool(s)

Python face_recognition

Official documentation: [Library documentation](#) | [Custom code](#)

Language: Python

Description: OpenCV-based face recognition library.

Cost: Free (open source)

Social impact: We retain full control of use of the images/face data.

Notes: Tests run on Charlie Nelms and Herman B Wells images/videos.

Installation & requirements

Install via pip (face_recognition).

Requires opencv-python

Parameters

Input formats

For training: Images labelled with person's name (currently via file path, but this should perhaps change-- discussion to have with dev)

For identifying: A model trained on the relevant people

Example Usage

See [Colab notebook](#).

Example Output

List of timestamps where face was found

Custom FR Tool Output
00:02:28
00:02:30
00:02:39
00:03:15
00:03:18
00:03:26
00:03:27
00:03:28
00:03:31
00:03:42

Evaluation summary

Precision, recall, and F1 scores for ground truth testing of five videos are in the project [Google Drive](#).