Framework

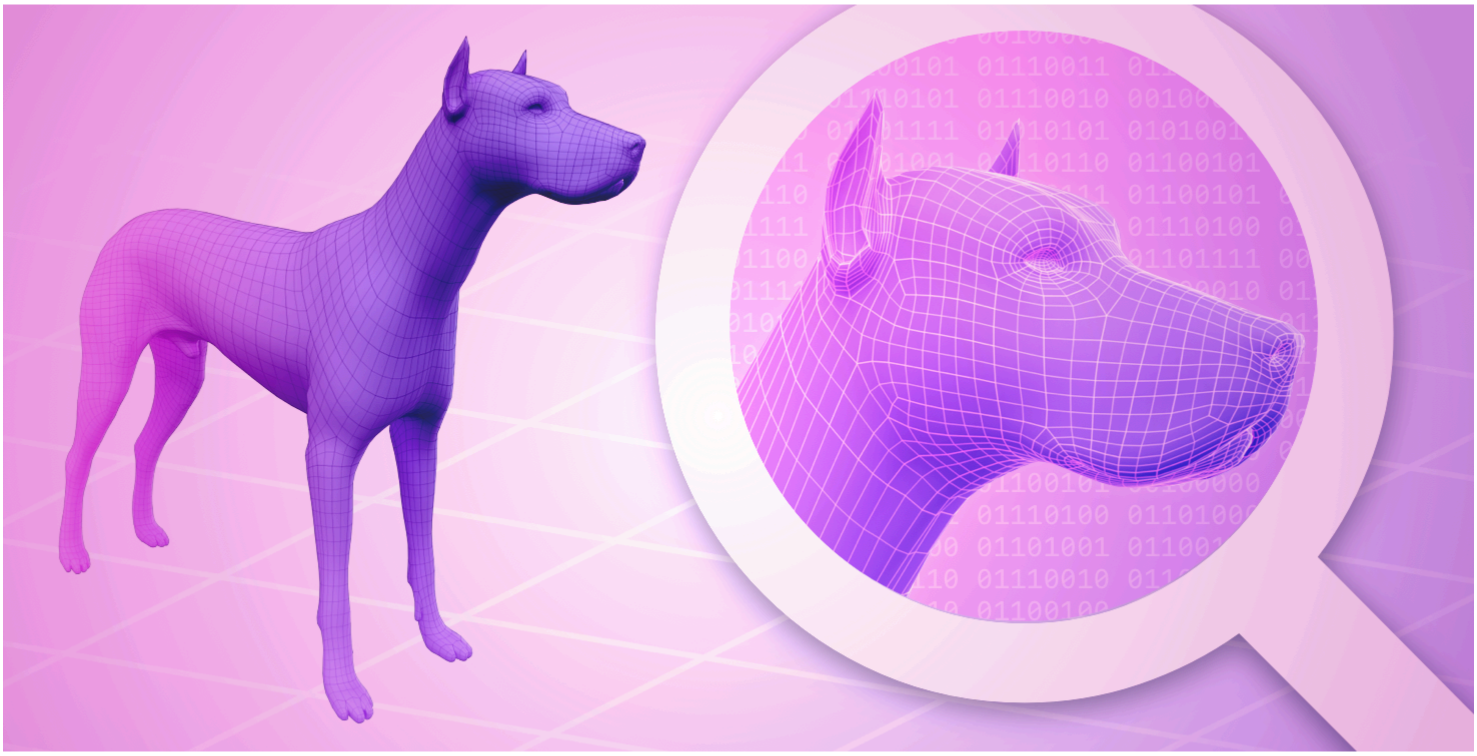
Vision

Apply computer vision algorithms to perform a variety of tasks on input images and videos.

Overview

The Vision framework combines machine learning technologies and Swift’s concurrency features to perform computer vision tasks in your app. Use the Vision framework to analyze images for a variety of purposes:

- Tracking human and animal body poses or the trajectory of an object
- Recognizing text in 18 different languages
- Detecting faces and face landmarks, such as eyes, nose, and mouth
- Performing hand tracking to enable new device interactions
- Calculating an aesthetics score to determine how memorable a photo is



To begin using the framework, you create a request for the type of analysis you want to do. Each request conforms to the `VisionRequest` protocol. You then perform the request to get an observation object — or an array of observations — with the analysis details for the request. There are more than 25 requests available to choose from. Vision also allows the use of custom Core ML models for tasks like classification or object detection.

> Note
>
> Starting in iOS 18.0, the Vision framework provides a new Swift-only API. See Original Objective-C and Swift API to view the original API.

Topics

Still-image analysis

{} Classifying images for categorization and search

Analyze and label images using a Vision classification request.

`struct ClassifyImageRequest`

A request to classify an image.

`protocol ImageProcessingRequest`

A type for image-analysis requests that focus on a specific part of an image.

`class ImageRequestHandler`

An object that processes one or more image-analysis requests pertaining to a single image.

`protocol VisionRequest`

A type for image-analysis requests.

`protocol VisionObservation`

A type for objects produced by image-analysis requests.

`struct DetectLensSmudgeRequest`

A request that detects a smudge on a lens from an image or video frame capture.

`struct SmudgeObservation`

An observation that provides an overall score of the presence of a smudge in an image or video frame capture.

Image sequence analysis

`class GeneratePersonSegmentationRequest`

A request that produces a matte image for a person it finds in the input image.

`struct GeneratePersonInstanceMaskRequest`

A request that produces a mask of individual people it finds in the input image.

`struct DetectDocumentSegmentationRequest`

A request that detects rectangular regions that contain text in the input image.

`protocol StatefulRequest`

The protocol for a type that builds evidence of a condition over time.

Image aesthetics analysis

{} Generating high-quality thumbnails from videos

Identify the most visually pleasing frames in a video by using the image-aesthetics scores request.

`struct CalculateImageAestheticsScoresRequest`

A request that analyzes an image for aesthetically pleasing attributes.

Saliency analysis

`struct GenerateAttentionBasedSaliencyImageRequest`

An object that produces a heat map that identifies the parts of an image most likely to draw attention.

`struct GenerateObjectnessBasedSaliencyImageRequest`

A request that generates a heat map that identifies the parts of an image most likely to represent objects.

Object tracking

`class TrackObjectRequest`

An image-analysis request that tracks the movement of a previously identified object across multiple images or video frames.

`class TrackRectangleRequest`

An image-analysis request that tracks movement of a previously identified rectangular object across multiple images or video frames.

Face and body detection

{} Analyzing a selfie and visualizing its content

Calculate face-capture quality and visualize facial features for a collection of images using the Vision framework.

`struct DetectFaceRectanglesRequest`

A request that finds faces within an image.

`struct DetectFaceLandmarksRequest`

An image-analysis request that finds facial features like eyes and mouth in an image.

`struct DetectFaceCaptureQualityRequest`

A request that produces a floating-point number that represents the capture quality of a face in a photo.

`struct DetectHumanRectanglesRequest`

A request that finds rectangular regions that contain people in an image.

Body and hand pose detection

`struct DetectHumanBodyPoseRequest`

A request that detects a human body pose.

`struct DetectHumanHandPoseRequest`

A request that detects a human hand pose.

`protocol PoseProviding`

An observation that provides a collection of joints that make up a pose.

`enum Chirality`

The hand sidedness of a pose.

`struct Joint`

A pose joint represented as a normalized point in an image, along with a label and a confidence value.

3D body pose detection

`class DetectHumanBodyPose3DRequest`

A request that detects points on human bodies in 3D space, relative to the camera.

`struct Joint3D`

An object that represents a body pose joint in 3D space.

Text detection

{} Recognizing tables within a document

Scan a document containing a contact table and extract the content within the table in a formatted way.

{} Locating and displaying recognized text

Perform text recognition on a photo using the Vision framework’s text-recognition request.

`struct RecognizeDocumentsRequest`

An image-analysis request to scan an image of a document and provide information about its structure.

`struct DocumentObservation`

Information about the sections of content that an image-analysis request detects in a document.

`struct DetectTextRectanglesRequest`

An image-analysis request that finds regions of visible text in an image.

`struct RecognizeTextRequest`

An image-analysis request that recognizes text in an image.

Barcode detection

`struct DetectBarcodesRequest`

A request that detects barcodes in an image.

Trajectory, contour, and horizon detection

`class DetectTrajectoriesRequest`

A request that detects the trajectories of shapes moving along a parabolic path.

`struct DetectContoursRequest`

A request that detects the contours of the edges of an image.

`struct DetectHorizonRequest`

An image-analysis request that determines the horizon angle in an image.

Animal detection

`struct DetectAnimalBodyPoseRequest`

A request that detects an animal body pose.

`struct RecognizeAnimalsRequest`

A request that recognizes animals in an image.

Optical flow and rectangle detection

`class TrackOpticalFlowRequest`

A request that determines the direction change of vectors for each pixel from a previous to current image.

`struct DetectRectanglesRequest`

An image-analysis request that finds projected rectangular regions in an image.

Image alignment

`class TrackTranslationalImageRegistrationRequest`

An image-analysis request that you track over time to determine the affine transform necessary to align the content of two images.

`class TrackHomographicImageRegistrationRequest`

An image-analysis request that you track over time to determine the perspective warp matrix necessary to align the content of two images.

`protocol TargetedRequest`

A type for analyzing two images together.

Image feature print and background removal

`struct GenerateImageFeaturePrintRequest`

An image-based request to generate feature prints from an image.

`struct GenerateForegroundInstanceMaskRequest`

A request that generates an instance mask of noticeable objects to separate from the background.

Machine learning image analysis

`struct CoreMLRequest`

An image-analysis request that uses a Core ML model to process images.

`struct CoreMLFeatureValueObservation`

An object that represents a collection of key-value information that a Core ML image-analysis request produces.

`struct ClassificationObservation`

An object that represents classification information that an image-analysis request produces.

`struct PixelBufferObservation`

An object that represents an image that an image-analysis request produces.

Image locations and regions

`struct NormalizedPoint`

A point in a 2D coordinate system.

`struct NormalizedRect`

The location and dimensions of a rectangle.

`typealias NormalizedRegion`

A polygon composed of normalized points.

`struct NormalizedCircle`

The center point and radius of a 2D circle.

`protocol BoundingBoxProviding`

A protocol for objects that have a bounding box.

`protocol BoundingRegionProviding`

A protocol for objects that have a defined boundary in an image.

`protocol QuadrilateralProviding`

A protocol for objects that have a bounding quadrilateral.

`enum CoordinateOrigin`

The origin of a coordinate system relative to an image.

Request Handlers

`class ImageRequestHandler`

An object that processes one or more image-analysis requests pertaining to a single image.

`class TargetedImageRequestHandler`

An object that performs image-analysis requests on two images.

Utilities

`enum ComputeStage`

Types that represent the compute stage.

`class VideoProcessor`

An object that performs offline analysis of video content.

Errors

`enum VisionError`

The errors that the framework produces.

Legacy API

- Original Objective-C and Swift API