

sub**vision**

vision-driven subtitle detection

User Manual

1. Overview

SubVision is a macOS app designed to extract burned-in (hardcoded) subtitles from video using OCR (Optical Character Recognition). Once extracted, you can review, correct, and export the result as an SRT file.

The application provides a complete workflow for subtitle extraction, refinement, and validation, with real-time feedback and timeline consistency to ensure that what you see in the interface matches the final exported result.

What SubVision Is (and Isn't)

Best use case

- Videos where subtitles are rendered into the image and cannot be turned off.

Not the best tool for

- Videos that already contain embedded subtitle tracks (SRT / CC / WebVTT).
OCR can still work, but extracting the original track is usually cleaner.

SubVision supports both automatic OCR extraction and manual subtitle creation, allowing hybrid workflows directly on the timeline.

2. Supported Media & Requirements

Supported Formats

SubVision can open most video files supported by macOS playback frameworks:

- .mp4, .m4v, .mov, .qt — natively supported via AVFoundation
- .mkv — supported when internal codecs are compatible with macOS native playback

Local files only: SubVision works with video files stored locally on your Mac or connected drives. Streaming content and DRM-protected videos are not supported.

File Validation

SubVision validates files before loading them. Unsupported file types, incompatible codecs, or files without a valid video track are rejected with a warning.

If a file fails validation, the current session — including loaded video, OCR results, and subtitle list — remains unchanged.

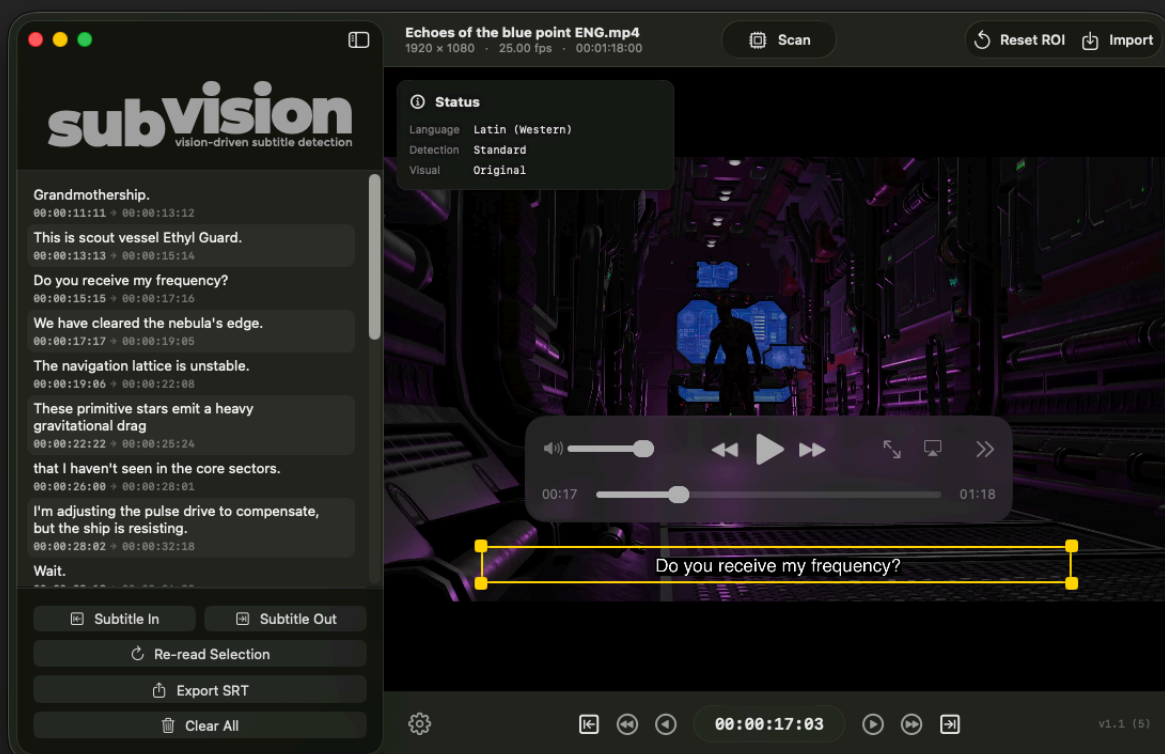
Technical Requirements

- Resolution: SubVision works better with source videos with a minimum resolution of 720p (HD) for optimal results.
- Performance: SubVision runs OCR in batch mode by scanning frames. Faster or more accurate settings increase scan time.
- Privacy: All OCR processing is performed locally on your device. No media content or subtitle data is collected or transmitted.
- Dictionary correction: Disabled by default to prevent unintended alterations of proper nouns or technical terminology.

3. Interface Overview

The SubVision window is organized into four main areas:

- Video Player: Playback reference during setup and review.
- ROI Overlay (rectangle on video): The Region of Interest — defines where OCR searches for text.
- Subtitles Sidebar: List of detected subtitles; select items to review, edit, and refine timing. The sidebar also provides visual feedback for timing adjustments and OCR updates.
- Settings Panel: OCR tuning and image enhancement (contrast, grayscale, invert, threshold, etc.).



Drag & Drop Interaction

When dragging a video file over the SubVision window, a highlighted drop area appears to indicate where the file can be released. This visual feedback confirms that the app is ready to receive the file.

Guided Quick Start

The Drop Zone features a three-step visual guide to ensure the optimal workflow is followed:

1. Import Video — drag and drop or use File → Open Video...
2. Define ROI — position the Region of Interest rectangle over the subtitle area
3. Start Scan — press the Scan button to begin OCR extraction

4. Recommended Workflow (End-to-End)

Step 1 — Load a Video

Load a file via:

- File → Open Video...
- Drag & Drop a video file into the app window

After loading, the app displays basic media information: duration, resolution, and frame rate.

Step 2 — Set the ROI (Region of Interest)

This is the most important step for speed and accuracy.

A poorly placed ROI is the most common cause of bad OCR output.

- Resize and position the ROI rectangle to tightly include **only the subtitle area** (typically bottom center).
- Avoid including logos, watermarks, UI overlays, or other text.
- Click **Reset ROI** to restore the default region.

Step 3 — (Optional) Set a Scan Range (Mark In / Mark Out)

By default, the scan runs on the full video. To limit the scan to a specific segment:

- Move the playhead to the start point and set Mark In
- Move the playhead to the end point and set Mark Out

Use this to quickly test settings on a small excerpt before scanning the entire file.

Step 4 — Tune Settings (OCR + Visual Enhancement)

Open Settings / Preferences and adjust based on your footage.

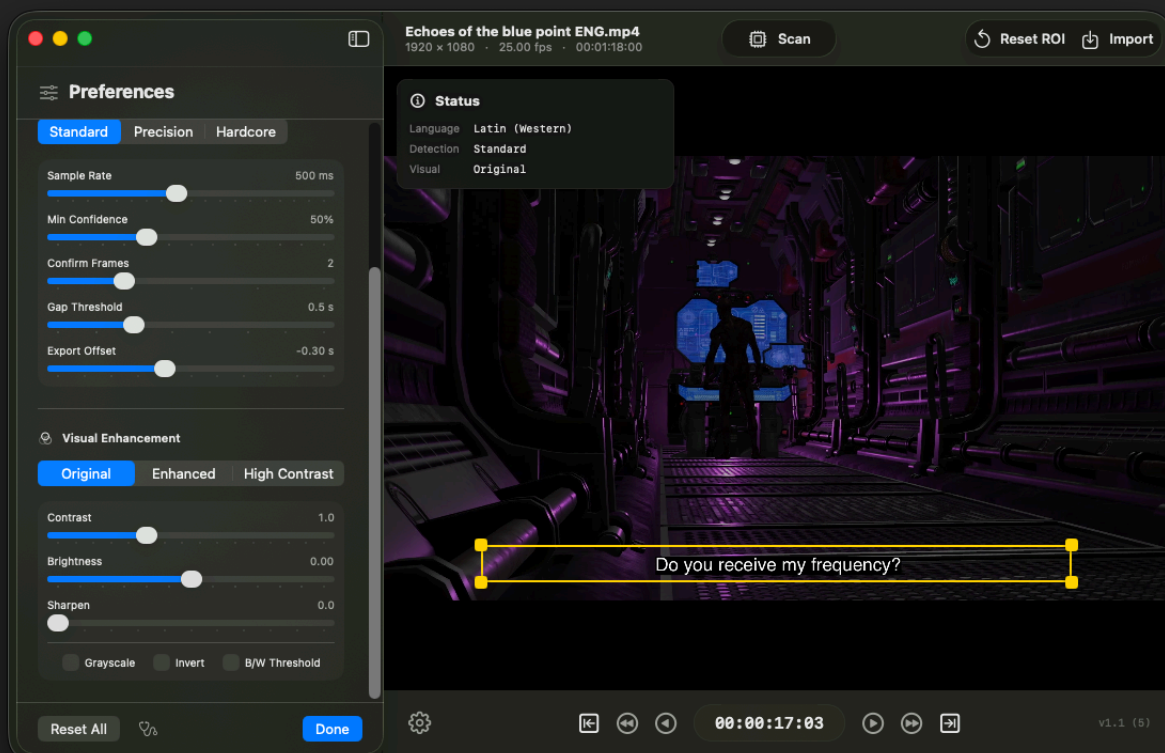
OCR Tuning (Speed vs. Accuracy)

- Sample Rate / Extraction Interval: Smaller intervals catch fast subtitle changes but increase processing time.
- Minimum Confidence: Higher values reduce nonsense characters but may miss faint text.
- Confirm Frames / Confirmation Threshold: Requires the same text to appear across multiple frames before it becomes a subtitle. Reduces flicker/false positives, but may miss very brief subtitles.
- Gap Threshold: Determines when two detections should become separate subtitle entries.
- Export Offset: Applies a global timing shift to the exported file — useful when everything is consistently early or late.

Visual Enhancement (What OCR “Sees”)

Use the following settings to improve subtitle readability while simplifying the background:

- Contrast / Brightness
- Sharpen
- Grayscale
- Invert
- Black / White Threshold



Step 5 — Scan

OCR accuracy depends on the source. Results vary depending on subtitle language, font style, size, outlines/glow, motion blur, compression, and overall contrast. If recognition is inconsistent, tighten the ROI and try grayscale/invert/threshold adjustments before rescanning.

Start the scan from the main toolbar. SubVision scans the video frame-by-frame and populates the subtitle list once complete. A real-time progress bar provides continuous feedback on the scanning process.

During a scan:

- You cannot play, pause, scrub, or change the ROI while a scan is running.
- You can cancel the scan at any time.

Step 6 — Review, Fix, and Refine

After scanning, use the sidebar to validate text and timing.

A) Edit Subtitle Text

- Select a subtitle and enter edit mode (click the edit icon or press Return).
- Confirm edits with Return / Enter.

B) Fix Timing with Subtitle In / Subtitle Out

- Select a subtitle.
- Move the playhead to the desired start frame → press I (Subtitle In).
- Move the playhead to the desired end frame → press O (Subtitle Out).

These controls operate directly on the timeline and are always available, regardless of which UI element is focused. This allows you to navigate the video freely and adjust timing without changing context.

If no subtitle is selected, pressing I or O creates a new subtitle at the current playhead position. The new subtitle can then be edited manually.

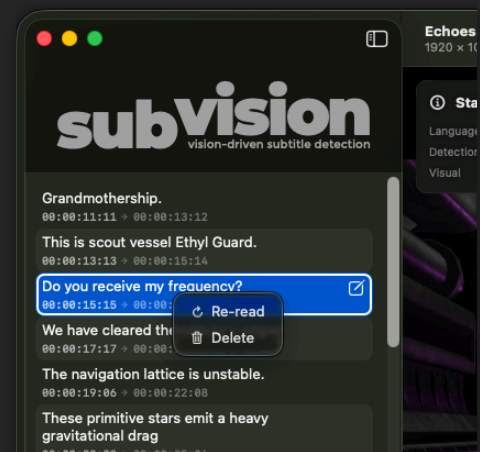
This behavior allows you to create subtitles from scratch or correct missing OCR detections directly during playback.

Context Menu (Quick Actions)

Right-clicking on a subtitle entry opens a context menu with quick access to common actions:

- Re-read — Re-run OCR on the selected subtitle
- Delete — Remove the selected subtitle from the list

These actions operate on the currently selected subtitle and do not affect the rest of the timeline.



Notes:

- If a timing adjustment would create an overlap with an adjacent subtitle, SubVision automatically adjusts the timeline to maintain consistency.
- Timing changes are reflected immediately in the interface. The subtitle list always matches the final exported SRT output.

C) Re-read Selection (Fix One Problematic Subtitle)

If a single subtitle is wrong due to complex background or poor readability:

- Select the problematic subtitle.
- Adjust Settings (contrast / invert / threshold / confidence).
- Click Re-read Selection to rescan only that subtitle's time window and update the result.

This is faster than rescanning the entire video.

Pro Tip for Accuracy

If you encounter a segment with low-contrast subtitles:

- Set the In/Out points for that segment.
- Open Preferences and enable Invert or Contrast Boost.
- Click Rescan Range to update only that section with the enhanced filters.

Step 7 — Export as SRT

When satisfied, click Export SRT, choose a destination folder, and save the file.

The exported SRT file reflects exactly the timing shown in the interface, with all overlap corrections already applied.

5. Scanning Modes & Precision Recovery

SubVision offers three distinct levels of text analysis, allowing you to move from a full-video automated scan to the surgical correction of a single subtitle.

1. Standard Scan

The primary mode for initial subtitle acquisition.

How it works

- Full Scan: If no In/Out points are defined, SubVision analyzes the entire video from start to finish.
- Range Scan: If In/Out points are set, only the specified portion of the timeline is processed.
- Running a Standard Scan clears previous results and generates a new subtitle list.

When to use

Use this upon first importing a video, or when you wish to restart the entire analysis with new global preferences.

2. Rescan Range (Merge Mode)

A targeted scan designed to preserve your existing work. The Scan button automatically updates to Rescan Range when the app detects existing subtitles and the user has defined a new In/Out interval.

How it works

- The engine analyzes only the video segment between the defined In and Out points.
- Existing subtitles within that specific interval are removed and replaced.
- New results are seamlessly merged into the existing list, leaving all subtitles before and after the range untouched.

When to use

Ideal for fixing a specific scene where OCR performed poorly (e.g., white text on a bright background). You can adjust parameters — such as increasing contrast or refining the ROI — specifically for those 30 seconds without re-processing the entire video.

3. Re-read Selection (Surgical OCR)

The most precise correction tool, available by selecting a specific subtitle row and clicking "Re-read Selection" in the sidebar.

If the text is updated, the affected subtitle is briefly highlighted to make the change immediately visible.

How it works

- The engine focuses exclusively on the single selected subtitle entry.
- It performs a new OCR analysis using the current Preferences and Visual Enhancements only for that brief timestamp (calculated as Start Time -0.2s to End Time +0.2s).
- The selected subtitle text is instantly updated with the new result.

When to use

Use this to perfect a single "difficult" subtitle that OCR misread. It allows you to experiment with different filters on the fly for a single frame without affecting the rest of the timeline.

6. Language Recognition & Script Mode

SubVision can automatically adapt to different languages and scripts, allowing flexible workflows without requiring strict manual configuration.

To maximize extraction accuracy, SubVision allows you to configure the OCR engine based on the character set (Script) present in the video. This setting,

found in the Preferences sidebar as Script Mode, instructs the app which alphabets and language dictionaries to prioritize.

By selecting a specific script, the engine significantly reduces “noise” and false positives by ignoring irrelevant characters and background graphical elements.

Available Options

Latin / Western (Default)

- Description: Optimized for Western languages using the Latin alphabet.
- Included languages: English (US), Italian, French, German, Spanish, and Portuguese.
- Extended Support: It also supports many other European languages using Latin characters with diacritics (accents, tildes, cedillas, etc.), such as Dutch or Scandinavian languages.
- Best use: Recommended setting for the majority of European and American content. Active by default on first launch.

Cyrillic

- Description: Specifically tuned for the Cyrillic alphabet.
- Included languages: Russian and Ukrainian.
- Best use: Use exclusively for East Slavic language content.

Asian (CJK)

- Description: Optimized for Chinese (Simplified), Japanese, and Korean logograms.
- Best use: Essential for correctly identifying East Asian subtitles.

Arabic

- Description: Optimized for Arabic script and grammar.
- Best use: Specifically designed for Arabic language content.

Universal (Auto Detection)

- Description: A restriction-free mode that enables global automatic detection.
- Behavior: The engine does not limit its search to a specific alphabet, allowing the system to identify the language frame-by-frame.
- Best use: Recommended when the language is unknown, mixed, or when working with multi-language content. This mode enables automatic detection without requiring manual configuration.
- Note: While this mode allows maximum flexibility, restricting the script to a specific alphabet may still improve accuracy in controlled scenarios.

7. System Diagnostic Tool

SubVision includes a diagnostic utility designed to provide general system and session information.

- Access: Click the Stethoscope icon in the Preferences sidebar.

The diagnostic tool provides a summary of the current session, including:

- File status: Confirmation that the video file is accessible and properly loaded
- OCR session data: Information about the most recent scan
- System information: Details about the OCR environment in use

This tool is intended as an informational reference and is not required for normal operation of the app.

8. Keyboard Shortcuts

Focus rules apply to playback and editing shortcuts. Some shortcuts may behave differently depending on which UI element currently has keyboard focus.

Playback Shortcuts

When the video player has focus:

- Space — Play / Pause
- J / K / L — Rewind / Pause / Fast-forward (when player has focus)

Focus note: Space may not work if keyboard focus is on the subtitle list or a text field. Click the video player area to restore playback focus.

Global In-Window Shortcuts

- ← / → — Step backward / forward by 1 frame
- I — Subtitle In
Set the start time of the selected subtitle. If no subtitle is selected, a new subtitle is created at the current playhead position.
- O — Subtitle Out
Set the end time of the selected subtitle. If no subtitle is selected, a new subtitle is created and can be edited manually.

Sidebar / Editing Shortcuts

- **Return / Enter** — Start or confirm editing (when the subtitle list or text field is active)

Subtitle In / Subtitle Out are designed for fast timeline-based editing. They can be used while navigating playback, and if no subtitle is selected, they can also create a new subtitle entry for manual editing.

9. Known Limits & Common Pitfalls

- OCR is probabilistic. Expect occasional errors on: low-resolution footage, motion blur, heavy compression, glow effects, stylized fonts, and low-contrast subtitles.
- ROI placement is decisive. A bad ROI produces bad output quickly.
- Spacebar depends on focus. If play/pause doesn't respond, click the video player area first.
- Subtitle timing is protected against invalid overlaps. If a timing adjustment would conflict with an adjacent subtitle, SubVision automatically refines the result to maintain timeline consistency.
- Re-read Selection is localized. It improves only the selected subtitle entry without affecting the rest of the timeline. After re-reading, the subtitle remains selected and the app indicates whether the text was updated or unchanged.
- Global offset is often the correct fix. If every subtitle is uniformly early or late, prefer Export Offset over manual shifting.
- Subtitle In / Subtitle Out and Mark In / Mark Out serve different purposes. Subtitle In / Subtitle Out adjust individual subtitle timing, while Mark In / Mark Out define the scan range for OCR analysis.
- Manual subtitle creation is supported. If OCR misses a subtitle or the source is unclear, you can create a new subtitle directly on the timeline and edit its text manually.

10. Troubleshooting

“No Subtitles Found”

- ROI is in the wrong position or too small → reposition/resize and retry
- Minimum Confidence is too high → lower it
- Confirm Frames is too strict → reduce it
- Sample Rate is too slow → decrease the interval to catch short subtitles
- Subtitles are too faint → increase contrast, try Invert or Threshold

“Text Is Garbage / Random Characters”

- Tighten ROI (exclude logos and background text)
- Increase Minimum Confidence
- Increase Confirm Frames
- Apply grayscale, invert, and threshold to simplify the image

“Space Doesn’t Play / Pause”

- Click the video player to restore focus, then press Space again

11. App Store & Privacy

Sandboxing & Security

SubVision operates within the macOS sandbox environment to ensure maximum system security. The app only accesses video files explicitly opened or dropped by the user and does not collect or transmit any media content or extracted subtitle data.

Licensing

Your purchase is tied to your Apple ID. You can restore or install SubVision on any Mac authorized with your account via the Mac App Store Purchased tab.

12. Customer Support

For technical support or to suggest new features for future updates, visit marcobucci.eu or contact via the support link on the App Store page.

About this Guide

This guide may be updated independently from the app version. When requesting support, include both the Manual Version and the App Version (Build).

SubVision is an independent tool and is not affiliated with or endorsed by Apple Inc.