

SNAP-TESTER

On-Device Visual QA for Windows UIs

Team: PIXEL PERFECT

Problem Statement

Windows applications must look and function correctly across a vast range of display environments—1080p up to 4K, 100%–200% DPI, dark and high-contrast themes, multiple languages, and varied window sizes. Conventional CI pipelines catch crashes and functional bugs but *cannot* detect pixel-level visual regressions: unreadable text, poor contrast, overlapping UI elements, or tiny controls that fail accessibility guidelines. Manual review of dozens of variants is slow, error-prone, and unscalable, while cloud-based visual-test services introduce latency, expose proprietary UI captures, and rely on external infrastructure.

Proposed Solution

Snap-Tester is a turnkey, on-device CI plug-in for the Snapdragon X Elite Copilot+ PC that automates *pixel-perfect visual and accessibility QA* without any cloud dependency.

1. Automated Screenshot Capture

Leverage existing UI-automation (PowerShell UIA, WinAppDriver, FlaUI) to navigate key screens under each resolution/DPI/theme, saving snapshots in a standardized folder.

2. AI-Powered Audit Pipeline

- **Text Detection (BYO Tiny-OCR ONNX)** A compact OCR model (team-supplied ONNX, compiled via `qai_hub compile`) extracts every UI text region and bounding box in tens of milliseconds on the Hexagon NPU.
- **Control Detection (BYO MobileNet-SSD ONNX)** A lightweight SSD, fine-tuned on UI controls (buttons, fields, icons), also compiled via AI Hub, locates interactive elements in tens of milliseconds.

- **Rule-Based Checks (CPU)** Computes WCAG contrast ratios (minimum 4.5:1) and enforces control dimensions (48×48 dp) directly on pixel data.

3. Actionable Guidance

Integrate the **AnythingLLM 1B-parameter QNN** (pre-optimized on AI Hub) to translate numeric failures (e.g. “contrast = 2.8:1 on ‘OK’”) into concise, prioritized fix suggestions (“Use white text for 7:1 contrast”) in under 100 ms.

4. CI Integration & Reporting

Runs as a native CI step on one Copilot+ PC; all screenshots and AI inferences remain local. Generates a JSON/HTML report and flags builds with critical regressions—providing comprehensive feedback in under one minute for 50–100 screens.

Edge-AI & AI Hub Integration

- **On-Device Inference:** All OCR, control detection, and LLM runs execute on the Hexagon NPU, showcasing real-time edge-AI performance without cloud latency.
 - **AI Hub Models:** Utilizes the AnythingLLM 1B ONNX directly from AI Hub; uses AI Hub’s “Bring Your Own Model” flow to compile our OCR and SSD networks into Hexagon QNN format.
 - **Unified Runtime:** Inference is driven through AI Engine Direct or ONNX Runtime’s QNN Execution Provider, highlighting the complete Snapdragon X edge-AI stack.
-

Use-Case Scenarios

- *Automotive Infotainment:* Automatically validate day/night and multilingual screens for contrast and layout faults before production.
- *Enterprise Kiosks:* Ensure high-DPI and accessibility theme variants of self-service software remain readable and tappable—without sending images to the cloud.
- *Productivity Suites:* Guard custom skins and dynamic layouts across diverse monitor setups, catching UI regressions early and reducing support tickets.

Impact

Snap-Tester empowers development teams to *shift left* their visual and accessibility testing—catching every pixel-level issue on the same device end-users run and delivering consistent, private QA feedback that scales across all display environments.