CVD Discriminability: Worst-Case Palette Separation Under Simulated Color Vision Deficiency, for Evaluating Accessibility Checkers

Satyajeeth Suresh Kannan
Innovation Academy · Mary's Lab
github.com/JeetuSK0808/maryslab

Datasheet · August 2026 · Mary's Lab dataset DS-03 · CC-BY-4.0

ABSTRACT

A datasheet for the `cvd-discriminability` dataset: 4320 rows, each one categorical palette evaluated for one observer model — its colours, its minimum pairwise separation in ΔE_{OK} as that observer sees it, the identity of the worst pair, and a thresholded convenience label. Observers span normal trichromacy and protan/deutan/tritan deficiency at graded severity; palettes come from three provenances, including colours of real measured materials. Under normal vision 70.3% of palettes clear the 0.05 threshold; under full deuteranopia 31% do. The labels are the output of a physiological simulation model, not human judgements — this datasheet says so wherever it matters, because the dataset's intended use is evaluating accessibility checkers against a stated model, and mistaking that for human ground truth would overclaim it.

Keywords: dataset; accessibility; color vision deficiency; palettes; data visualization; wcag

1. Motivation

Palette accessibility checkers disagree with each other, and there is no labelled corpus to evaluate them against — most check colour pairs against a fixed contrast rule and are graded, in practice, against other checkers. This dataset supplies a defined reference: the actual worst-case pairwise separation under a named physiological model, at graded severities, over palettes that include the kinds designers actually ship. A checker can then be scored on how well its verdicts track a stated model, which is a weaker but honest substitute for human data that does not yet exist at scale.

2. Composition

4320 rows: 720 palettes (3–10 colours) × six observer conditions. Train 3552, validation 360, test 408, grouped so all six observer rows for one palette stay on one side.

Table 1. Share of rows whose minimum pairwise separation clears 0.05 ΔE_OK, by observer condition. The gradient down this column is the point of the dataset.

observer	rows	clears 0.05
normal	720	70.3%
protan 0.5	720	53.5%
protan 1.0	720	48.1%
deutan 0.5	720	49.2%
deutan 1.0	720	31%
tritan 1.0	720	33.6%

Palette provenances: `random` (uniform sRGB), `hue_wheel` (evenly spaced hues at fixed lightness and chroma — the naive "accessible" construction, and a common real mistake, since fixed lightness removes the channel a dichromat retains), and `material` (colours of measured USGS reflectances under D65 — colours a physical surface can actually produce). Under full deuteranopia the hue-wheel palettes pass at 15%, against 52.5% for random and 25.4% for material.

3. Collection and labeling

Every label is computed: the palette is passed through the Machado, Oliveira and Fernandes (2009) simulation at the stated type and severity, differences are taken in Oklab, and the minimum over pairs is recorded with the pair that attains it. The continuous separation is the real label. The boolean at 0.05 is a convenience whose threshold was chosen by measuring the label distribution rather than assumed — at stricter thresholds the boolean degenerates and a classifier scores by always answering "no" — and the card publishes percentiles so it can be re-thresholded.

4. Uses

Checker evaluation is the intended use: run an accessibility checker over `palette_hex` and compare its verdicts to the model-derived labels, per observer and per severity. Regression (predict the separation) and classification (predict the boolean) are well-posed learning tasks; a model that does well is imitating the Machado-plus-Oklab computation, which is useful as a fast surrogate and must not be described as predicting human perception.

5. Distribution, license, and maintenance

JSONL splits plus a dataset card; CC-BY-4.0. Material palette colours derive from public-domain USGS data (cite Kokaly et al. 2017). Seeded and reproducible from `datasets/build/`; engine 0.4.1, snapshot usgs-splib07a-1. Corrections arrive as a new version, never a silent overwrite.

6. Limitations

This section states what the result does not establish. It is written to be read before the abstract is quoted.

- **Model output, not human data.** The Machado model is fitted to population averages; real colour vision deficiency varies between individuals, and anomalous trichromacy is a continuum one severity parameter represents coarsely. No participant judged any of these palettes (TR-14 records the same gap for the lab's stimulus designer, and it is closed by a psychophysical study or not at all).
- **No viewing conditions.** Size, background, adaptation, display calibration and ambient light all affect whether a separation is usable, and none appear here.
- **Pairwise colour separation only.** A passing palette is not an accessible design: nothing here measures whether colour is the sole channel carrying information, which is the failure that actually excludes people.
- **The 0.05 threshold is an engineering choice.** The continuous column is the honest label; re-threshold as your use requires.

7. Acknowledgements and disclosure

The instruments described here were implemented in Mary's Lab, an open-source computational color science project. Software design, implementation, and the preparation of this report were carried out with substantial assistance from a large language model (Anthropic Claude), under the author's direction and review. All numerical results were produced by executing the released code; none were generated by a language model. The author is responsible for the claims made here.

References

1. Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., and Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, 64(12), 86–92.
2. Machado, G. M., Oliveira, M. M., and Fernandes, L. A. F. (2009). A physiologically-based model for simulation of color vision deficiency. *IEEE Transactions on Visualization and Computer Graphics*, 15(6), 1291–1298.
3. Ottosson, B. (2020). *A perceptual color space for image processing (Oklab)*. bottosson.github.io.
4. Kokaly, R. F., et al. (2017). *USGS Spectral Library Version 7*. U.S. Geological Survey Data Series 1035. doi:10.3133/ds1035.
5. Kannan, S. S. (2026). *Bracketing the Best Possible Palette: Covering Bounds on Categorical Discriminability Under Dichromatic Vision*. Mary's Lab technical report TR-03.
6. Kannan, S. S. (2026). *Computing Ishihara-Style Stimuli From Spectra*. Mary's Lab technical report TR-14.

Generator, statistics, and this datasheet's renderer: `github.com/JeetuSK0808/maryslab` under `datasets/build/`. The build is seeded — the same commit reproduces the same dataset byte for byte — and every number above resolves from the generator's own stats.json. Datasheet structure follows Gebru et al. (2021). Prepared with AI assistance (see Acknowledgements).