IMAGIONARY : Jurnal Seni, Desain dan Multimedia
Vol. 4, No. 1, October 2025
P-ISSN: 2808-4144, E-ISSN: 2808-442X

AI Thumbnails on YouTube for MSME Brands: A Study on GETI Media SME Live

Edy Chandra
Program Studi Desain Komunikasi Visual-FSRD, Universitas Tarumanagara, Jakarta, Indonesia
edyc@fsrd.untar.ac.id

History:
Submit : August 25, 2025
Review : September 30, 2025
Publish : October 31, 2025

Keywords:
desain komunikasi visual; visual branding; YouTube; AI thumbnails; CTR; AVD; UMKM

ABSTRACT

This study evaluates the effectiveness of AI-generated or AI-assisted thumbnails in enhancing content performance and reinforcing visual branding on the educational YouTube channel "GETI Media – UKM Live." The evaluation focuses on two key YouTube analytics metrics: Impressions Click-Through Rate (CTR), which serves as an indicator of visual appeal and its ability to drive clicks, and Average View Duration (AVD), used as a proxy for viewer engagement and retention. The research design employs an A/B experimental approach involving a minimum of 12 newly published videos over a three-week period. Each video is presented with two thumbnail conditions: Condition A (manually designed by a human designer) and Condition B (AI-generated or assisted using documented prompts). Controls are applied to variables such as video title, description, release time, tags, and duration to ensure consistency across conditions. In addition, a visual content analysis of the thumbnails is conducted to map elements that may moderate performance outcomes. These elements include: presence and expression of faces, clarity of logos or wordmarks, color contrast, information density, presence of visual call-to-action (CTA), brand identity consistency, composition, and gaze direction. Data analysis involves paired t-tests or Wilcoxon signed-rank tests to compare CTR and AVD across conditions, and logistic or linear regression models to assess the influence of AI-generated thumbnails and visual moderators, while controlling for covariates such as topic, video duration, impressions, and release timing. The study aims to provide empirical evidence and practical guidance for MSME (UMKM) practitioners and educational content creators in optimizing thumbnails while maintaining ethical standards and brand identity alignment. The findings are also contextualized within YouTube's "Test & Compare Thumbnails" feature and the potential for cross-lingual visual localization.

1. Introduction

YouTube is a strategic medium for education and marketing among Indonesian UMKM (micro, small, and medium enterprises). The micro-canvas of the thumbnail acts as the key visual that triggers click decisions before exposure to content. This article addresses three questions: (1) do AI-generated/assisted thumbnails (hereafter, AI thumbnails) raise CTR relative to manual baselines; (2) what is their impact on AVD; and (3) which visual elements moderate these effects (face/expression, contrast, logo/text legibility, information density, identity fit).

Empirical work shows that sentiment/emotion in thumbnails relates to views/clicks (Cui, 2024); (Li et al., 2022). Psychovisual evidence indicates that faces and text systematically attract gaze (Cerf et al., 2009), while thumbnail visual-expression elements relate to viewer attitude/satisfaction(Lee, 2023). In visual branding, consistent logo/color/type lock-ups strengthen recognition (Henderson & Cote, 1998); in Indonesia, sustained identity practice correlates with UMKM performance (Saifulloh & Fadillah, 2021);(Indrajaya & Redintan, 2024);(Christine & Lestari, 2025). IMAGIONARY studies highlight logo complexity → brand awareness, Gestalt-based rebranding for a BUMN, and a local rebranding case (Kurniadi & Harnoko, 2025);(Telaumbanua & Darmawan, 2025);(Chandra & Mutiara, 2022)

Why this is an ideal experimental setting. First, at the platform level, YouTube offers Test & Compare Thumbnails, enabling controlled A/B tests at video level or—if unavailable—a scheduled thumbnail swap. Second, at the production level, GETI Media – UKM Live uploads consistently on education-for-UMKM topics with sufficient early impressions, allowing repeated A/B pairs within 48-hour and seven-day windows. Third, at the audience/branding level, returning viewers of education channels encounter repeated brand cues (color/logo/type), enabling a joint test of identity consistency and CTR optimization ("YouTube Is Testing a New Feature to Help Videos Travel around the World," 2025).

Contributions. We provide (a) an applied A/B framework tailored to an education-for-UMKM channel; (b) a replicable thumbnail coding matrix; and (c) models linking AI condition, visual moderators, and performance outcomes.

2. Method

The research method describes: the approach, scope or object, operational definition in each variable/description of research focus, place, population and sample/informant, main source and tool, technique of data collection, and technique of data analysis.

2.1. Design and Site

We conduct A/B experiments on the channel GETI Media – UKM Live for three weeks (Week-1 calibration; Weeks-2–3 testing), targeting ≥12 new uploads (ideally 12–18) to obtain ≥6 A/B pairs. Where available, we use Test & Compare Thumbnails; otherwise, we perform 24-hour thumbnail swaps.

2.2. Unit of Analisis and Sample

The unit is a video with two thumbnail conditions: A (manual/designer) and B (AI assisted/generated). Inclusion criteria: new uploads within the experiment window; no paid promotion; comparable duration per pair.

The unit is a video with two thumbnail conditions: A (manual/designer) and B (AI assisted/generated). Inclusion criteria: new uploads within the experiment window; no paid promotion; comparable duration per pair.

2.3. Procedure

Prompt families. Prepare ≥3 prompt families varying composition (face close-up vs product hero), style (flat vs photorealistic), and ≤3-word hooks. Log prompt text, versions, and seeds. Thumbnail production. For each video produce two variants:

a. A (Manual): by the channel designer, following identity guidelines (Henderson & Cote, 1998)
b. B (AI): via generator/assistive design with minor post-edits (crop, logo overlay, color balance) to preserve identity.

Controls. Keep titles, descriptions, tags, upload time, and duration constant across the pair—following online controlled-experiments principles (Kohavi et al., 2020);(Kohavi, 2023). A/B implementation. Use Test & Compare or 24-hour swaps with randomized order to mitigate carryover.

Observation windows. Collect metrics at 48 hours (primary) and 7 days (robustness) in line with reporting recommendations (Lakens, 2013).

2.4. Intstrument and Metric

CTR: percentage of impressions that turn into views—indicator of thumbnail+title pull. AVD: average viewing duration—indicator of retention. Impressions and Views-48h control for exposure and early interest. Watch Time, End-screen click rate, Subscribers gained/video complement brand/engagement outcomes. Data are sourced from YouTube Analytics (owner access) and exported to CSV.

2.5. Thumbnail Content Coding

Coding is applied to A and B thumbnails using the following scheme (Cerf et al., 2009);(Lee, 2023): (1) Face (0/1) & expression (−1/0/+1); (2) Logo/wordmark (0/1) & legibility (0–2); (3) Main words (0=0–3; 1=>3); (4) Global contrast (0 low/med; 1 high); (5) Visual CTA (0/1); (6) Identity fit (0–2); (7) Density (1–3); (8) Composition (0/1); (9) Gaze to camera (0/1). Inter-coder reliability is assessed with Cohen's kappa.

2.6. Staitiscal Analysis

1. Paired tests (t or Wilcoxon) for CTR and AVD (A vs B per video), reporting effect sizes (Hedges, 1981) and 95% CIs(Lakens, 2013).
2. Regression models: linear/GLM for AVD; quasi-binomial/logistic for CTR with covariates (topic, duration, day/time, impressions).
3. Moderation: interactions AI × (face, contrast, legibility, etc.).
4. Robustness: 7-day window; leverage checks; Holm correction for multiplicity.
5. Experiment reliability: design and execution follow trustworthy online experiments guidelines (Kohavi et al., 2020);(Kohavi, 2023).

2.7. Ethic

Use licensed assets and written consent for faces; avoid unconsented face-swap or misleading clickbait. Fully document prompts and post-edits for audit.

Results and Discussion

Note. Numbers currently shown are simulated for format preview. Replace with GETI Analytics values

3.1.Sample Description

Table 1. Sample Characteristics and A?B Pairs

Video ID	Release Date	Duration (min)	Topic	Condition A (Manual) – Brief	Condition B (AI) – Brief	Impressions 48h	Views 48h
V01	2025-04-02	18.4	IG Optimization	Half-body host + 5-word text, blue BG	Face-centric + 2-word hook, yellow BG	12,540	1,036
V02	2025-04-03	22.1	Food Packaging	Product flat lay + 6-word text	Product hero + 3-word headline	10,118	789
V03	2025-04-05	15.6	Business Legal	Text slide + small logo	Host pointing + CTA arrow	9,430	702
V04	2025-04-06	12.9	Price Psychology	Diagram + small text	Infographic-lite number "3X"	11,207	905
V05	2025-04-08	28.3	Live Coaching	Live screenshot + overlay	Red LIVE tag + 2-word headline	13,991	1,112
V06	2025-04-09	17.7	Phone Product Photo	4-photo collage	Single packshot + clean BG	8,764	624

3.2.CTR and AVD (48 hours)

Table 2. CTR (%) and AVD (min) – A/B per pair (48h)

Video ID	CTR A	CTR B	Δ CTR (B−A)	AVD A	AVD B	Δ AVD (B−A)
V01	6.6	9.1	+2.5	5.9	6.2	+0.3
V02	6.4	8.2	+1.8	6.8	6.5	−0.3
V03	6.9	7.6	+0.7	5.1	5.4	+0.3
V04	6.1	8.1	+2.0	7.2	7.1	−0.1
V05	6.7	8.0	+1.3	9.4	9.0	−0.4
V06	6.2	7.1	+0.9	5.6	6.0	+0.4
Mean	6.5	8.0	+1.5	6.7	6.7	0.0

Table 3. Paired Difference Tests (48h)

Variable	n	Mean A	Mean B	Mean Diff	t / Z	p-value	95% CI
CTR (%)	6	6.5	8.0	+1.50	5.12	0.004	[0.81, 2.19]
AVD (min)	6	6.67	6.70	+0.03	0.18	0.866	[−0.35, 0.40]

3.3. Moderation by Visual Elements

Table 4. Visual-Coding Summary (A and B)

Code	Face	Expr	Logo	Logo Legibility	Words≤3	Contrast	CTA	Identity	Density	Composition	Gaze
V01-A	1	0	1	1	0	0	0	1	3	0	1
V01-B	1	+1	1	2	1	1	1	2	2	1	1
V02-A	0	0	1	1	0	0	0	1	3	0	0
V02-B	0	0	1	2	1	1	0	2	1	1	0
V03-A	0	0	1	1	0	0	0	1	3	0	0
V03-B	1	+1	1	2	1	1	1	2	2	1	1
V04-A	0	0	1	1	0	0	0	2	2	0	0
V04-B	0	0	1	2	1	1	0	2	1	1	0
V05-A	0	0	1	1	0	0	0	1	3	0	0
V05-B	1	+1	1	2	1	1	1	2	2	1	1
V06-A	0	0	1	1	0	0	0	1	3	0	0
V06-B	0	0	1	2	1	1	0	2	1	1	0

Table 5. Regression Model (CTR as outcome, 48h) *(simulated)*

Predictor	Coef (β)	SE	z/t	p	Note
Intercept	5.80	0.42	13.8	<0.001	manual baseline
AI condition	+1.10	0.28	3.9	0.001	main AI effect
Face	+0.65	0.25	2.6	0.018	social cue
High contrast	+0.48	0.22	2.2	0.041	legibility
Logo legibility	+0.30	0.14	2.2	0.048	brand cue
Words ≤3	+0.27	0.15	1.8	0.094	positive trend
Visual density	−0.33	0.13	−2.5	0.024	crowding
AI×Face	+0.42	0.18	2.3	0.036	key moderator
AI×Contrast	+0.31	0.16	1.9	0.082	borderline

Table 6. Regression Model (AVD as outcome, 48h) *(simulated)*

Predictor	Coef (β)	SE	t	p	Note
Intercept	6.40	0.38	16.9	<0.001	manual baseline
AI condition	+0.04	0.11	0.4	0.701	small effect
Face	+0.22	0.10	2.2	0.047	minor retention
High contrast	+0.06	0.09	0.7	0.502	ns
Logo legibility	+0.05	0.06	0.8	0.446	ns
Words ≤3	+0.03	0.07	0.4	0.686	ns
Visual density	−0.18	0.07	−2.6	0.028	drop-off
AI×Identity fit	+0.12	0.05	2.4	0.036	brand fit matters

3.4. Robustness (7-day window)

Table 7. CTR and AVD – 7 days (simulated)

Video ID	CTR A (7d)	CTR B (7d)	Δ	AVD A (7d)	AVD B (7d)	Δ
V01	6.3	8.7	+2.4	6.1	6.2	+0.1
V02	6.2	7.9	+1.7	7.0	6.6	−0.4
V03	6.7	7.4	+0.7	5.3	5.5	+0.2
V04	5.9	7.8	+1.9	7.3	7.2	−0.1
V05	6.5	7.7	+1.2	9.6	9.1	−0.5
V06	6.1	6.9	+0.8	5.7	6.0	+0.3
Mean	6.3	7.7	+1.4	6.8	6.8	0.0

Table 8. Paired Difference Tests — 7-day Window [simulated]

Variable	n	Mean A	Mean B	Mean Diff (B−A)	t / Z	p-value	95% CI
CTR (%)	6	6.30	7.70	+1.40	4.86	0.005	[0.73, 2.07]
AVD (min)	6	6.80	6.80	+0.00	0.02	0.984	[−0.38, 0.39]

Discussion

Three threads structure the interpretation. First, AI→CTR. If the lift is significant, AI's faster visual iteration contributes to pre-view pull, aligning with evidence on thumbnail–performance links (Cui, 2024; Li et al., 2022) and online controlled-experiments practices (Kohavi et al., 2020). Second, AVD. AVD is often content-sensitive; AI thumbnails influence it indirectly via expectation calibration. Over-promise can raise clicks yet hurt retention; fit between promise and delivery is essential (YouTube Help, n.d.). Third, moderators. Expressive faces strengthen AI effects on CTR through social-cue salience (Cerf et al., 2009). High contrast supports legibility on the micro-canvas and relates to satisfaction (Lee, 2023). Logo/text legibility integrates branding and CTA; this is consistent with design guidelines (Henderson & Cote, 1998) and IMAGIONARY findings on logo complexity and brand awareness (Kurniadi & Harnoko, 2025). In the UMKM context, consistent identity links to better performance (Saifulloh & Fadillah, 2021; Indrajaya & Redintan, 2024; Christine & Lestari, 2025). Practical implications include AI for ideation with human curation, concise brand guidelines for thumbnails, a steady A/B cadence, disciplined metric monitoring (CTR/AVD at 48h), and ethical guardrails—aligned with platform practice and OTT trends (YouTube Help, n.d.; Business Insider, 2025).

Generalizability from a single education-for-UMKM channel; topical variation independent of thumbnails; potential carryover in swaps; seasonality/trends; limited sample size for complex moderation. Future work: cross-channel replication, thumbnail language localization, and lightweight brand-lift surveys to complement CTR/AVD.

4. Conclusion

We propose an applied experimental framework to quantify AI thumbnail effectiveness on CTR and AVD for an Indonesian UMKM education channel. By combining platform metrics with visual-content coding and moderation tests, we derive evidence-based design rules that balance click intent with retention while safeguarding brand identity and ethics.

Acknowledgments

We thank GETI Media – UKM Live and UMKM partners for their collaboration.

References

Cerf, M., Frady, E. P., & Koch, C. (2009). Faces and text attract gaze independent of the task. *Journal of Vision*, *9*(12), 10. https://doi.org/10.1167/9.12.10

Chandra, E., & Mutiara, M. W. (2022). New Ancol Logo Design, Brings the Meaning of "Happiness" or "Disappointment" for Indonesian People. *IMAGIONARY*, *1*(1), 10–15.

Christine, J., & Lestari, R. (2025). The effect of branding and digital marketing on UMKM sales performance. *VICIDI: Visual Communication Design Journal*, *6*(1). https://journal.uc.ac.id/index.php/vicidi/article/view/5243

Cui, G. (2024). Predicting video views through a sentiment analysis of YouTube thumbnail images. *Journal of Business Research*. https://www.sciencedirect.com/science/article/abs/pii/S0148296324003539

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://www.jstor.org/stable/1164588

Henderson, P. W., & Cote, J. A. (1998). Guidelines for Selecting or Modifying Logos. *Journal of Marketing*, *62*(2), 14–30.

Indrajaya, G., & Redintan, M. J. (2024). Visual identity as a visual representation media in branding Bakso Atok. *IMAGIONARY*, *3*(1). https://journal.paramadina.ac.id/index.php/imagionary/article/view/922

Kohavi, R. (2023). Online controlled experiments and A/B tests. In *Encyclopedia of Machine Learning and Data Science*. https://exp-platform.com/Documents/2023-03-11EncyclopeiaMLDSABTestingFinal.pdf

Kohavi, R., Tang, D., & Xu, Y. (2020). *Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing*. Cambridge University Press. https://pmc.ncbi.nlm.nih.gov/articles/PMC7007661/

Kurniadi, T., & Harnoko, I. (2025). Effectiveness of logo complexity on the success of company brand awareness. *IMAGIONARY*, *3*(2), 40–47. https://journal.paramadina.ac.id/index.php/imagionary/article/view/931

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00863

Lee, S. (2023). Focusing on video thumbnails: Visual expression elements and user satisfaction. *Journal of Web Engineering*, *22*(8). https://journals.riverpublishers.com/index.php/JWE/article/view/18891

Li, Y., Kim, H.-J., Do, B., & Choi, J. (2022). The effect of emotion in thumbnails and titles of video clips on online audiences' viewing behavior and pre-roll advertising effectiveness. *Journal of Business Research*, *151*, 232–243. https://doi.org/10.1016/j.jbusres.2022.06.051

Saifulloh, M., & Fadillah, A. (2021). Branding product of UMKM in the digital era. *DIANMAS: Jurnal Pengabdian Kepada Masyarakat*, *10*(2), 166–176. https://journal1.moestopo.ac.id/index.php/dianmas/article/view/1498

Telaumbanua, J. P. J., & Darmawan, A. J. (2025). Efforts to rejuvenate the visual identity of state-owned enterprises: Barata Indonesia based on Gestalt. *IMAGIONARY*, *3*(2), 66–76. https://doi.org/10.51353/jim.v3i2.1075

YouTube is testing a new feature to help videos travel around the world. (2025). *Business Insider*. https://www.businessinsider.com/youtube-testing-new-thumbnail-feature-to-help-videos-travel-globally-2025-6