

롱테일 키워드 기반 RAG를 활용한 주얼리 디자인 생성의 조형적 타당성 연구

**A Study on Design Orientation in Jewelry Generation Using RAG and
Layered Long-Tail Keyword Correlation**

주 저 자 : 강혜림 (Kang, Hye Rim)

한경국립대학교 디자인예술학부 교수
hrkang@hknu.ac.kr

<https://doi.org/10.46248/kidrs.2026.2.494>

접수일 2026. 05. 20. / 심사완료일 2026. 05. 27. / 게재확정일 2026. 06. 08. / 게재일 2026. 06. 30.

Abstract

This study investigates the application of Retrieval-Augmented Generation in jewelry design to minimize hallucinations from rare long-tail keywords and accurately reflect design intent. By comparing RAG and Large Language Model approaches, we empirically verified the formative validity of Visual Question Answering evaluation metrics on the generated outputs. Results show that RAG outperformed LLMs across all key areas: fidelity(+0.34), aesthetics(+0.24), and validity(+0.93). Furthermore, within the RAG framework, VQA demonstrated significant positive correlations with aesthetics($r=.58$, $p=.001$) and fidelity($r=.49$, $p=.004$), confirming its reliability as an automated evaluation tool. Ultimately, this study demonstrates that the validity of VQA metrics varies depending on the methodology used. Long-tail keyword-based RAG not only generates structurally superior jewelry designs but also aligns seamlessly with automated evaluation metrics, highlighting that methodological choice critically impacts metric effectiveness in AI-driven design evaluation systems.

Keyword

Generative AI(생성형 AI), Retrieval-Augmented Generation(RAG), Long-Tail Keyword(롱테일 키워드), Jewelry Design(주얼리 디자인), Visual Question Answering(VQA)

요약

본 연구는 주얼리 디자인 분야에 검색 증강 생성(RAG)을 적용하여 희소성이 높은 롱테일 키워드의 할루시네이션을 최소화하고 디자인 지향성을 정밀하게 반영하는 방안의 가능성을 확인함에 목적이 있다. 이를 위해 RAG와 대형 언어 모델(LLM) 방법론을 비교하여 생성 결과물에 대한 VQA(Visual Question Answering) 평가 지표의 조형적 타당성을 실증적으로 검증하였다. 연구 결과, RAG 방법론에서는 VQA가 심미성($r=.58$, $p=.001$)과 충실도($r=.49$, $p=.004$)에서 유의한 정적 상관을 보여 자동 평가 지표로서의 타당도가 확인되었다. RAG가 LLM 대비 충실도(+0.34), 심미성(+0.24), 유효성(+0.93)이 모두 높게 나타났다. 본 연구는 VQA 지표의 타당도가 방법론에 따라 상이하게 작동함을 실증하였으며, 롱테일 키워드 기반 RAG가 주얼리 디자인 생성에서 조형적으로 더 타당한 결과물을 생성할 뿐 아니라 자동 평가 지표와의 정합성도 확인하였다. 이는 AI 기반 디자인 평가 체계 수립에 있어 방법론 선택이 평가 지표의 유효성에 영향을 미친다는 점을 시사한다.

목차

1. 서론

- 1-1. 연구 배경
- 1-2. 연구 목적 및 방법

2. 이론적 배경

- 2-1. LLM 방식과 RAG 방식
- 2-2. 선행연구 분석

3. 디자인 생성 및 평가

- 3-1. 롱테일 키워드 및 프롬프트 설계
- 3-2. 주얼리 디자인 생성
- 3-3. 디자인 평가

4. 연구 결과

5. 결론 및 제언

참고문헌

1. 서론

1-1. 연구 배경

글로벌 테크 기업인 가트너(Gartner)가 2026년 트렌드 전략 기술로 선정한 도메인 특화 언어 모델(Domain-Specific Language Models, DSLM)은 기업용 AI의 지향점이 단순한 지능을 넘어 ‘산업적 맥락’으로 이동했음을 시사한다.¹⁾ 이는 기업에서 도메인 데이터에 근간한 특화된 AI의 수요가 증대함을 의미한다. DSLM의 성능을 극대화하고 기업 내부의 보안 데이터 및 실시간 정보를 결합하기 위해서는 데이터의 학습만큼이나 검색 증강 생성(RAG, Retrieval-Augmented Generation) 기술의 결합이 필수적이다. 대규모 언어 모델(Large Language Model, LLM)의 발전 방향이 범용 데이터의 양적 확충에서 벗어나, 고품질로 정제된 데이터 기반의 신뢰성 확보로 그 중심축이 이동하고 있다.

주얼리 산업에서 최근 AI를 활용한 정밀한 구현 기술의 중요성이 그 어느 때보다 강조되고 있다. 그러나 생성형 AI의 근본적인 취약점으로 언급되는 할루시네이션(Hallucination)은 정교한 디자인 설계에 있어 여전히 극복해야 할 기술적 장벽이다. 따라서 본 연구는 주얼리 디자인 분야에 검색 증강 생성 접근법을 도입하여 회소성이 높은 룬테일 키워드의 할루시네이션을 최소화하고 사용자의 디자인 지향성을 정밀하게 반영하는 방안을 고찰하고자 한다.

1-2. 연구 목적 및 방법

본 연구는 검색 증강 생성을 통하여 AI를 활용하는 창작물의 디자인 지향성을 제고하는 방향과 가능성을 확인하는 데 그 목적이 있다. 디자인 지향성이란, 창작자의 욕망과 맥락에 정밀하게 대응하기 위해 조형 요소를 의식적으로 선택·조합하는 설계의 속성을 의미한다.²⁾

조형적 지향성에 영향을 미치는 구성 요소들을 계층(Layer)으로 구분한 프롬프트를 통해 각 변수가 최종 결과물에 미치는 상관관계를 심층 분석한다. 연구 방법은 소스 기반 생성(Source-Grounded

Engine) 모델인 구글(Google) 제미니(Gemini)를 활용하여 검색 증강 생성(RAG)을 수행한다. 검색 증강의 품질을 극대화하기 위해 각 룬테일 키워드에 부합하는 레퍼런스 이미지를 선정하고, 이미지 크기 조정, 노이즈 제거, 태깅을 통한 메타데이터 구조화 등 고도화된 전처리를 거쳐 검색 소스로 활용한다. 이후, 주얼리 디자인 지향성을 표현하는 룬테일 키워드로 구현된 프롬프트를 통해 LLM 방식과의 차이를 분석한다. 프롬프트 상의 텍스트 정보와 이미지 간의 의미 일치도를 측정하기 위해 VQA(Visual Question Answering) Score를 통하여 LLM과의 상호 비교 평가를 실시한다. VQA 모델은 프롬프트와의 토큰 예측 확률을 계산하여 0~1의 값을 측정하며 숫자가 높을수록 일치성이 높다. 이후 주얼리 디자인 전문가 그룹 3명이 도메인 특화 평가를 한다.

[표 1] 연구인력

전문가	전공	학위	관련 경력
A	금속공예	석사	10년
B	금속공예	석사	8년
C	주얼리디자인	석사	5년

2. 이론적 배경

2-1. LLM 방식과 RAG 방식

ChatGPT로 대표되는 대규모 언어 모델(LLM)은 기존의 포털의 검색 기능을 넘어 정보의 요약과 생성을 대항하며, 우리 일상의 실질적인 실용화 단계에 진입했다. 실용화 단계에서 정합성 있는 문제를 해결하기 위해서는 정제된 데이터와의 결합이 필요하다. 검색 증강 생성(RAG)은 외부 데이터베이스에서 정보를 결합하여 더 정확하고 신뢰성 높은 결과를 제공하는 기술이다.³⁾ RAG는 LLM의 사전 학습된 GPT에 데이터를 기반하는 추가적인 검색 기능을 강화하는 기술로 볼 수 있다.⁴⁾

1) Gartner, ‘Gartner Top 10 Strategic Technology Trends for 2026’, Gartner, (2026.04.29.)
www.gartner.com/en/articles/top-technology-trends-2026

2) ‘Orientation’, MockFlow, (2026.05.09.)
mockflow.com/glossary/Orientation

3) 브라이스 유, 조경아, 박수진, 김재웅, 『RAG 마스터: 랭체인으로 완성하는 LLM 서비스』, 프리렉, 2025, p.12

4) 이경록, 『테디노트의 랭체인을 활용한 RAG 비법노트』, 리코멘드, 2025, p.26

[표 2] LLM 방식과 RAG 방식 비교

구분	LLM ⁵⁾	RAG ⁶⁾
요약	방대한 데이터를 학습하여 언어 패턴을 이해하고 생성	LLM 답변 전, 외부 데이터에서 지식을 검색, 증강하여 생성
학습	범용 학습 시점에 고정됨(특정 불가)	외부 데이터 업데이트 시점에 고정됨
출처	답변의 근거나 출처에 어려움이 있음	구축된 데이터베이스에 근거한 답변
특성	일반적 지식, 창의적 인사이트 도출	도메인 지식, 할루시네이션 감소

2-2. 선행연구 분석

RAG 기술은 의료 분야에서 대규모 언어 모델의 한계를 보완하기 위해 외부 데이터베이스에서 정보를 검색하며, 이에 대한 유용성이 확인되고 있다.⁷⁾ RAG를 활용한 기업 맞춤형 챗봇 서비스 파이프라인을 구성하여 고객의 데이터를 활용하고, 고객이 원하는 챗봇을 개발하는 연구가 진행되고 있다.⁸⁾ 특정 과학 문헌이 제공되는 교육 환경에서 RAG 기반 애플리케이션은 학생들의 학습을 지원하는 데 유용성이 있다.⁹⁾ RAG 기반 설계는 범용성이 뛰어나 최소한의 최적화만으로도 법률과 같은 전문 분야로

5) 김지현, ‘기업의 업무 자동화 RPA, LLM으로 환골탈태하다’, Samsung SDS, 2023.10.25. (2026.04.19.)

www.samsungds.com/kr/insights/leveraging-llm-for-rpa.html

6) 위아람, ‘[검색증강생성(RAG)①]환각현상 낮추는 RAG, AI 기반 혁신 주도한다’, DataNet, 2024.05.14. (2026.04.19.) www.datanet.co.kr/news/articleView.html?idxno=193548

7) 윤여찬, 김수균, ‘생성형 AI를 위한 Retrieval-Augmented Generation (RAG) 기술 동향 및 전망’, 컴퓨터교육학회. 컴퓨터교육학회 논문지, Vol.28, No.2, 2025. 02. pp.69-80

8) 최광민, ‘검색증강생성(RAG) 기반 기업 맞춤형 챗봇(Chatbot) 시스템 구축 및 활용’, 한국전자통신학회, 한국전자통신학회 논문지, Vol.19, No.6, 2024. 12. pp.1281-1291

9) Thüs, Dominik, Malone, Sarah, Brünken, Roland, ‘Exploring Generative AI in Higher Education: A RAG System to Enhance Student Engagement with Scientific Literature’, Frontiers Media SA, Frontiers in Psychology, Vol.15, 2024. 11. p.1-18

의 확장이 가능하며, 이를 통해 전문가들이 방대하고 정밀한 정보에 기반한 고도화된 의사결정을 수행할 수 있음을 입증하였다.¹⁰⁾

이상에서 살펴본 선행연구는 텍스트 기반 일반 비즈니스와 IT 환경에서의 적용 사례와 방향성을 제시하고 있다. 본 연구는 기존 선행연구와 달리 디자인 영역에 집중하였다는 점과 롱테일 키워드의 요소별 상관관계를 분석하여 디자인 품질과의 연계성을 규명했다는 점에서 학술적 차별성을 지닌다.

3. 디자인 생성 및 평가

3-1. 롱테일 키워드 및 프롬프트 설계

본 연구에서는 창작자의 의도를 명확히 투영하여 디자인 왜곡을 최소화하는 AI 생성 기법을 탐색하고자 한다. LLM은 학습된 데이터를 임의로 조합하여 디자인 요소를 생성하는 방식으로, 창작자의 의도를 왜곡하거나 예기치 못한 윤리적 위해성에 노출될 우려가 있다. 구현하고자 하는 디자인 요소의 복잡도와 독창성이 높을수록 생성 리스크 또한 비례하여 증대되므로, 이를 체계적으로 관리하고 제어할 수 있는 생성 방법론이 필수적이다. 방법론을 검증하기 위해 본 연구에서는 LLM의 학습 빈도가 낮을 것으로 가정한 ‘롱테일 키워드’를 중심으로 의도한 디자인의 구현 가능성을 확인한다.

‘롱테일 키워드(Long-Tail Keywords)’란, 세 단어나 이상의 구체적인 구문으로 구성되어 특정 주제나 대상을 정교하게 타겟팅하는 속성의 용어이다.¹¹⁾ 롱테일 키워드는 다수의 단어가 조합되는 만큼 정보의 보편성은 낮으나 사용자의 타겟팅을 정교화할 수 있는 표준 형태를 띤다. 검색 볼륨은 작지만, 특정 주제에 대한 심층적인 의미를 내포할 수 있다. 따라서 본 연구는 이러한 롱테일 키워드로 창작자의 디자인적 지향성을 AI가 파악하여 투영할 수 있는 방법론을 실증하고자 한다. 롱테일 키워드는 주얼리

10) Wahidur, R. S. M., Kim, S., Choi, H., Bhatti, D. S., and Lee, H.-N., ‘Legal Query RAG’, IEEE Access, Vol.13, 2025. 02. pp.36978-36994

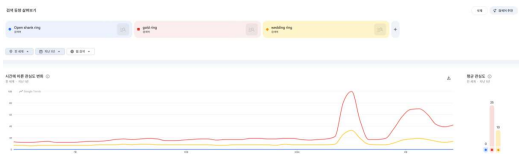
11) Aimuni, ‘Long-Tail Keywords: How to Effectively Use Them in SEO’, ematic solutions, 2025.01.15. (2026.04.29.) www.ematicsolutions.com/long-tail-keywords-seo/

디자인 및 제작을 구성하는 요소별로 설정한다. 제품의 외형적 구조와 종류 및 전반적인 콘셉트를 결정하는 Form, 금속의 물성을 정의하는 Material, 제작 공정 및 세팅 기법을 규정하는 Construction, 그리고 디자인의 서사적 영감과 조형적 테마를 부여하는 Motif를 기반으로 각 요소를 그룹화한다. [표 3]

[표 3] 룩테일 키워드

Layer	항목	룩테일 키워드
1	Form	- Stacking Midi Ring
	형태/상징성	- Asymmetric Bypass Ring
2	Material	- Oxidized Sterling Silver
	소재	- Gold-Accent - Sandblasted Layering
3	Construction	- Bridge Accent Setting
	가공/세팅	- Thread Setting
4	Motif	- Parametric Design
	모티프/컨셉트	- Memphis Design

디자인적 요소가 강한 룩테일 키워드는 전문가 그룹으로부터 키워드를 5배수로 수령하고, 이를 바탕으로 구글 트렌드(Google Trend)의 검색 동향을 분석하여 희소성 있는 룩테일 키워드를 도출하였다.



[그림 1] 구글 트렌드 검색 동향

구글 트렌드 조사 및 전문가 의견으로 도출한 룩테일 키워드를 중심으로 각 Layer에 맞는 특화된 프롬프트를 생성하고, AI를 활용하여 디자인을 생성한다. 디자인 구성요소로 구분된 4개의 Layer로 작성한다. 각 Layer를 대표하는 명확한 룩테일 키워드를 중심축으로 삼고, 나머지 Layer는 검색량이 높은 직관적 키워드들로 조합하였다. 이를 기반으로 각 계층의 룩테일 키워드가 유기적으로 반영된 프롬프트를 설계하여 주얼리 이미지 생성을 수행한다.



[그림 2] 프롬프트 설계

3-2. 주얼리 디자인 생성

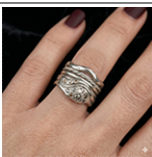
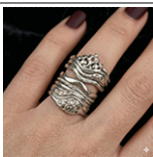
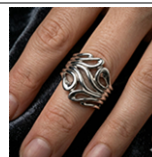
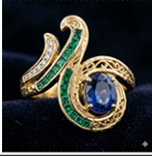
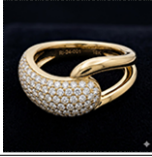
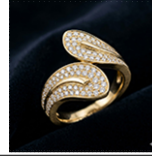
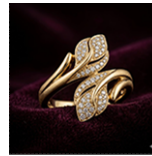
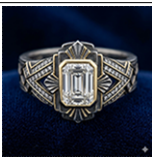
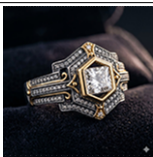
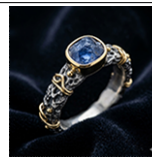
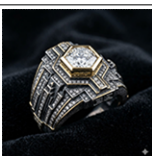
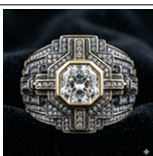
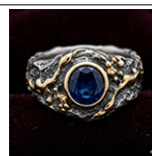
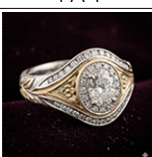
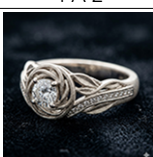
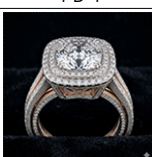
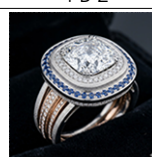
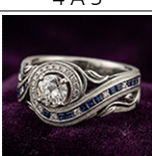
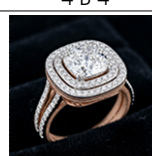
룩테일 키워드 프롬프트를 활용하여 주얼리 디자인을 생성하였다. LLM 그룹은 기본 프롬프트를 있는 그대로 입력하여 주얼리 디자인 이미지를 생성한 반면, 검색 증강 생성(RAG) 그룹은 사전에 룩테일 키워드와 관련된 5개의 학습 이미지를 제공하여 프롬프트를 증강한 후 주얼리 디자인을 생성하였다. [표 5]. [표 6]

이후 [표 4] 기준의 VQA 측정과 메인 특화 평가를 위해서 주얼리 디자인 전문가 그룹 3명에게 프롬프트를 구현하는 ‘충실도’, 조형적 완성도를 의미하는 ‘심미성’, 디자인 지향성의 구현을 의미하는 ‘유효성’의 5점 척도 지표를 제공하여 병행 측정하였다.

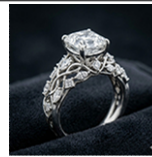
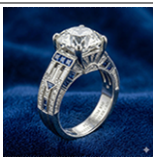
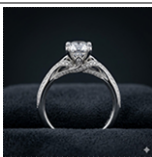
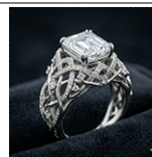
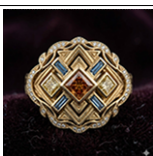
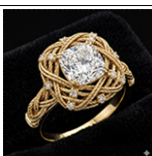
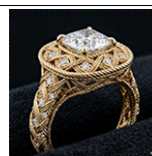
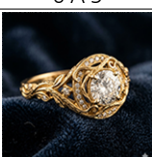
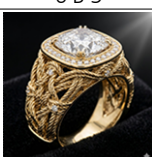
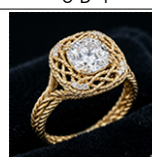
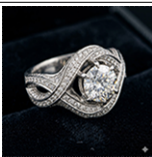
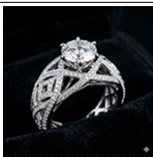
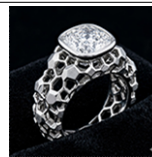
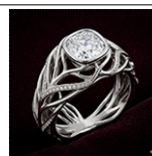
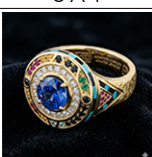
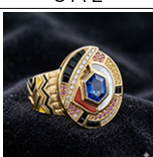
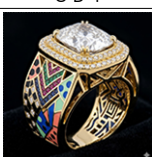
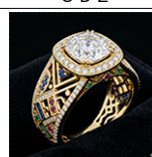
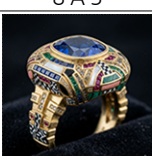
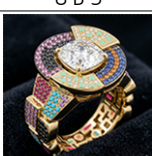
[표 4] VQA

항목	내용
평가 방법	Decomposed VQA Score
VQA 모델	BLIP-2 FlanT5-XL
점수 범위	0 ~ 1
분해 방식	프롬프트 속성별 바이너리 서브 질문 설계

[표 5] 주얼리 디자인 생성(Layer 1, 2)

프롬프트		(A) LLM		(B) RAG	
L a y e r 1	Luxury jewelry concept by a master jeweler. Form: Stacking Midi Ring Material: Platinum 950 Construction: Plain Setting Motif: non-formal design Please put it on model's finger. Photorealistic texture, studio lighting, dark velvet background, macro photography, 8k, cinematic. --ar 1:1	1-A-1	1-A-2	1-B-1	1-B-2
					
		1-A-3	1-A-4	1-B-3	1-B-4
					
	Luxury jewelry concept by a master jeweler. Form: Asymmetric Bypass Ring Material: 18K Gold Construction: Pavé Setting Motif: Modern Design Photorealistic texture, studio lighting, dark velvet background, macro photography, 8k, cinematic. --ar 1:1	2-A-1	2-A-2	2-B-1	2-B-2
					
		2-A-3	2-A-4	2-B-3	2-B-4
					
L a y e r 2	Luxury jewelry concept by a master jeweler. Material: Oxidized Sterling Silver Gold-Accent Form: Solitaire Ring Construction: Plain Setting Motif: Non-Formal Design Photorealistic texture, studio lighting, dark velvet background, macro photography, 8k, cinematic. --ar 1:1	3-A-1	3-A-2	3-B-1	3-B-2
					
		3-A-3	3-A-4	3-B-3	3-B-4
					
	Luxury jewelry concept by a master jeweler. Material: Sandblasted Layering Form: Halo Ring Construction: Pavé Setting Motif: Modern Design Photorealistic texture, studio lighting, dark velvet background, macro photography, 8k, cinematic. --ar 1:1	4-A-1	4-A-2	4-B-1	4-B-2
					
		4-A-3	4-A-4	4-B-3	4-B-4
					

[표 6] 주얼리 디자인 생성(Layer 3, 4)

프롬프트		(A) LLM		(B) RAG	
L a y e r 3	Luxury jewelry concept by a master jeweler. Construction: Bridge Accent Setting Form: Solitaire Ring Material: Platinum 950 Motif: Non-Formal Design Photorealistic texture, studio lighting, dark velvet background, macro photography, 8k, cinematic. --ar 1:1	5-A-1	5-A-2	5-B-1	5-B-2
					
		5-A-3	5-A-4	5-B-3	5-B-4
					
	Luxury jewelry concept by a master jeweler. Construction: Thread Setting Form: Halo Ring Material: 18K Gold Motif: Modern Design Photorealistic texture, studio lighting, dark velvet background, macro photography, 8k, cinematic. --ar 1:1	6-A-1	6-A-2	6-B-1	6-B-2
					
		6-A-3	6-A-4	6-B-3	6-B-4
					
L a y e r 4	Luxury jewelry concept by a master jeweler. Motif: Parametric Design Form: Solitaire Ring Material: Platinum 950 Construction: Plain Setting Photorealistic texture, studio lighting, dark velvet background, macro photography, 8k, cinematic. --ar 1:1	7-A-1	7-A-2	7-B-1	7-B-2
					
		7-A-3	7-A-4	7-B-3	7-B-4
					
	Luxury jewelry concept by a master jeweler. Motif: Memphis Design Form: Halo Ring Material: 18K Gold Construction: Pavé Setting Photorealistic texture, studio lighting, dark velvet background, macro photography, 8k, cinematic. --ar 1:1	8-A-1	8-A-2	8-B-1	8-B-2
					
		8-A-3	8-A-4	8-B-3	8-B-4
					

3-3. 디자인 평가(VQA 및 전문가 FGI 평가, 상관관계 분석)

대형 언어 모델과 검색 증강 생성 방식으로 생성

한 주얼리 디자인에 대한 VQA 평가를 실시하였고, 전문가 그룹의 FGI 분석을 하였다. [표7], [표8]

[표 7] 프롬프트 이미지 평가 결과

구분	(A) LLM					(B) RAG				
	No	VQA 지수	전문가 지수(AVG)			No	VQA 지수	전문가(AVG)		
			충실도	심미성	유효성			충실도	심미성	유효성
Layer 1	1-A-1	0.5208	4.00	4.33	2.33	1-B-1	0.4928	4.00	4.00	5.00
	1-A-2	0.5406	4.00	4.00	2.00	1-B-2	0.5107	4.00	4.00	4.67
	1-A-3	0.5176	3.67	3.00	2.00	1-B-3	0.5137	4.00	4.00	5.00
	1-A-4	0.4913	4.00	3.00	2.33	1-B-4	0.4799	4.00	3.00	4.33
	2-A-1	0.496	3.33	4.00	4.00	2-B-1	0.5387	4.67	4.00	4.33
	2-A-2	0.4704	4.00	3.67	3.33	2-B-2	0.5515	4.67	4.33	4.67
	2-A-3	0.4942	4.00	4.33	3.67	2-B-3	0.5463	4.67	4.00	4.67
	2-A-4	0.4869	4.00	4.00	4.33	2-B-4	0.5194	4.67	4.33	4.33
Layer 2	3-A-1	0.4697	4.00	4.33	4.00	3-B-1	0.5204	4.67	4.00	4.67
	3-A-2	0.4525	4.00	4.33	4.67	3-B-2	0.5037	4.00	3.67	4.00
	3-A-3	0.4842	4.00	4.00	4.00	3-B-3	0.5084	4.67	5.00	4.00
	3-A-4	0.4336	4.00	3.33	4.00	3-B-4	0.5246	4.67	4.33	4.00
	4-A-1	0.4704	4.00	4.00	3.33	4-B-1	0.5114	3.33	3.67	3.67
	4-A-2	0.4832	4.00	3.00	3.33	4-B-2	0.5293	4.00	4.00	4.00
	4-A-3	0.5151	3.00	3.67	3.67	4-B-3	0.5309	4.00	3.67	4.00
	4-A-4	0.4853	4.00	4.00	4.00	4-B-4	0.4999	4.33	3.67	4.00
Layer 3	5-A-1	0.5243	4.00	4.00	4.00	5-B-1	0.5171	4.33	4.00	5.00
	5-A-2	0.5162	3.33	3.33	2.67	5-B-2	0.5237	4.00	3.33	4.33
	5-A-3	0.5422	4.00	4.33	4.00	5-B-3	0.5571	4.00	4.00	4.00
	5-A-4	0.4937	4.00	4.00	3.00	5-B-4	0.5207	4.00	4.00	4.67
	6-A-1	0.5294	4.00	3.67	3.33	6-B-1	0.5956	4.00	4.33	5.00
	6-A-2	0.5728	4.00	4.00	2.00	6-B-2	0.5966	4.67	4.67	5.00
	6-A-3	0.557	4.00	4.33	3.33	6-B-3	0.5973	4.00	4.00	4.00
	6-A-4	0.5517	4.00	3.67	2.33	6-B-4	0.5935	4.33	4.33	5.00
Layer 4	7-A-1	0.495	4.00	4.33	4.67	7-B-1	0.5453	4.00	4.33	5.00
	7-A-2	0.5101	4.00	4.00	4.33	7-B-2	0.5443	4.00	4.00	4.00
	7-A-3	0.4933	4.00	4.33	4.33	7-B-3	0.5352	4.00	4.00	4.67
	7-A-4	0.5474	3.33	3.67	4.00	7-B-4	0.4979	3.00	4.00	4.00
	8-A-1	0.6209	3.00	4.33	4.00	8-B-1	0.5917	4.00	4.00	4.33
	8-A-2	0.6263	3.33	3.00	4.00	8-B-2	0.6351	5.00	4.33	4.67
	8-A-3	0.5916	4.00	3.33	4.00	8-B-3	0.6374	5.00	4.67	4.67
	8-A-4	0.6159	3.33	3.33	3.67	8-B-4	0.6163	4.67	4.67	4.67

AVG: Average

측정된 평가 지표의 조형적 타당성을 확보하기 위해 VQA 점수와 전문가의 평가 3중(충실도·심미성·유효성) 간의 준거 타당도(Criterion Validity)를 Pearson 상관분석을 통해 검증하였다. [표 8]

[표 8] VQA 지표의 준거 타당도 검증 결과

그룹	준거변수	r	t	p
LLM (n=32)	충실도	+0.291	1.664	.107
	심미성	-0.167	-0.927	.361
	유효성	-0.155	-0.858	.398
RAG (n=32)	충실도	+0.490	3.076	.004(**)
	심미성	+0.581	3.910	.001(**)
	유효성	+0.698	5.335	<.001(**)

(N=64, 양측 검정, df=n-2, ** $p < .01$, * $p < .05$)

4. 연구 결과

본 연구는 64개의 디자인(LLM 32개, RAG 32개)을 분석하였다. 이 분석은 자동 지표인 VQA 점수 평가와 전문가 그룹의 평가(충실도·심미성·유효성)를 병행 측정하였으며, VQA와 전문가 그룹 평가 간 준거 타당도를 분석하였다. RAG 방법론에서는 VQA가 심미성($r=.58$, $p=.001$)과 충실도($r=.49$, $p=.004$)와의 유의한 정적 상관을 보여 자동 평가 지표로서의 타당도가 확인되었다. Layer 1의 ‘Stacking Midi Ring’은 손가락 착용 정보의 정밀한 구현을 핵심 디자인 지향점으로 삼고 있다. RAG 방식은 이러한 디자인 의도를 정확하게 반영했을 뿐만 아니라, 준거 타당도 검증에서도 통계적 유의성을 확보함으로써 그 기술적 효용성을 입증하였다.

전문가 평가 전반에서도 RAG가 LLM 대비 충실도(+0.34), 심미성(+0.26), 유효성(+0.93) 모두 높게 나타났다. 반면, LLM 방법론에서는 어떠한 전문가 그룹의 평가 차원에도 유의한 상관이 나타나지 않아 LLM 기반 생성물 평가에서 VQA 지표의 타당도를 확인할 수 없었다.

본 연구의 VQA 지표 타당도의 분석 결과, 룽테

일 키워드 기반 RAG가 주얼리 디자인 생성에서 조형적으로 더욱 디자인 지향적인 결과물을 생성할 뿐 아니라 자동 평가 지표와의 적합성도 존재함을 확인하였다. 이는 생성형 AI 기반 디자인 평가 체계 수립에 있어 방법론 선택이 평가 지표의 유효성에 직접적 영향을 미친다는 점을 시사한다.

또한, 정성적 평가의 일환으로 진행된 전문가 그룹 인터뷰에서 전문가 A는 “다양한 형태적 요소가 끊임없이 생성되고 결합되는 현대 주얼리 산업 환경을 고려할 때, 특화 데이터 학습을 기반으로 사용자의 프롬프트 의도성을 명확히 제어하는 작업은 연구적 의미가 높다.”라고 하였다. 전문가 B는 “검색 증강 생성 기반 시스템은 시각화 오류로 발생하는 기회비용을 단축하며, 디자이너와 AI 간의 협업 워크플로의 가능성이 보였고, 디자인의 의도성과 희소성이 강화될수록 검색 증강 생성은 더욱 필수적인 절차가 될 것이다.”라고 하였다. 전문가 C는 “성능 향상은 고무적이거나, 실제 제조 공정과의 데이터 연동성 측면에서는 다소 아쉬움이 남는다. 향후 연구에서 3D 모델링 수치와의 직접적인 호환성까지 확보된다면, 단순 디자인 생성을 넘어 실질적인 주얼리 제조 생태계를 완성할 수 있을 것이다.”라고 제언하였다.

5. 결론 및 제언

본 연구는 주얼리 디자이너의 창의적 지향성을 효과적으로 반영하기 위해 검색 증강 생성(RAG)을 디자인 프로세스에 도입하였으며, 실제 산업 현장에서의 실무적 활용 가치를 심도 있게 분석하고자 하였다. AI를 통해 생성된 64개의 주얼리 디자인 이미지를 생성하고 그 산출물에 대한 통계 및 전문가 분석을 실시하였다. 기존 LLM 방식과 비교하였을 때, 검색 증강 생성 기술 기반의 생성 방식이 창작자의 의도가 반영된 특화 디자인 도출에 실질적인 활용 가치가 있음이 확인되었다.

프롬프트 엔지니어링을 AI와의 상호작용을 위한 커뮤니케이션 프로토콜로 정의한다면, 최적의 산출물 생성을 위해 학습 데이터의 품질과 구조를 체계화하는 작업은 그에 못지않게 본질적인 요소가 된다. 금번 실험을 통해 주얼리 디자인 지향성을 높이는 AI 작업에 있어 관련 분야의 전문적인 데이터 관리 역량이 필수적임을 시사한다. 본 연구는 한정

된 표본을 대상으로 진행되었다는 점에서 일부 한계를 지니고 있다. 하지만 지속적으로 진화하는 AI 기술을 바탕으로, 주얼리 디자인의 창의적 지향성을 강화하기 위해 검색 증강 생성 방법론을 적용했다는 점에서 학술적 의의를 지닌다. 향후 디자인을 지원하고 관리하는 시스템으로서 기능에 대한 확장적 검증을 수행하는 연구를 진행할 예정이다. 본 연구의 내용이 주얼리 산업의 고부가가치 창출을 돕는 창조활동의 기반이 되기를 바란다.

10. www.ematicssolutions.com

11. www.gartner.com

참고문헌

1. 브라이스 유, 조경아, 박수진, 김재웅, 『RAG 마스터: 랭체인으로 완성하는 LLM 서비스』, 프리렉, 2025
2. 이경록, 『테디노트의 랭체인을 활용한 RAG 비법노트』, 리코멘드, 2025
3. 윤여찬, 김수균, ‘생성형 AI를 위한 Retrieval-Augmented Generation (RAG) 기술 동향 및 전망’, 컴퓨터교육학회, 컴퓨터교육학회 논문지, Vol.28, No.2, 2025
4. 최광미, ‘검색증강생성(RAG) 기반 기업 맞춤형 챗봇(Chatbot) 시스템 구축 및 활용’, 한국전자통신학회, 한국전자통신학회 논문지, Vol.19, No.6, 2024
5. Thüs, Dominik, Malone, Sarah, Brünken, Roland, ‘Exploring Generative AI in Higher Education: A RAG System to Enhance Student Engagement with Scientific Literature’, Frontiers Media SA, Frontiers in Psychology, Vol.15, 2024
6. Wahidur, R. S. M., Kim, S., Choi, H., Bhatti, D. S., and Lee, H.-N., ‘Legal Query RAG’, IEEE Access, Vol.13, 2025
7. mockflow.com
8. www.samsungsds.com
9. www.datanet.co.kr