

생성형 AI를 이용한 서예 유튜브 콘텐츠 내 내재적 정보 추출 및 조회수에 미치는 영향에 관한 이질성 분석 연구

곽영식*¹, 허해초*², 이두희** , 곽윤식***

The Heterogeneous Effects of Intrinsic Information from Calligraphy YouTube Content on View Counts via Generative AI

Young Sik Kwak*¹, Hai Chao Xu*², Du Hee Lee**, and Yoon Sik Kwak***

요 약

웹 크롤링으로 수집 가능한 유튜브 영상 외적 정보 외에도 생성형 AI 추론 기능을 활용해 콘텐츠 내재적 특성을 추출하고 해당 요인이 조회수에 미치는 영향을 파악하고자 본 연구를 시작하였으며, NotebookLM을 이용해 서예 관련 유튜브 영상 237개에서 총 42개의 외재적 및 내재적 특성을 추출하고 조회수에 미치는 영향을 분석하였다. 분석 결과 전체 표본에서는 외재적 정보 2개와 내재적 정보 7개가 조회수에 유의한 영향을 보였고, 하위집단 분석에서는 무배경음악 환경의 대자 행서 콘텐츠와 남성 작가가 전서 외 서체를 장시간·큰 글씨로 제작한 영상의 반응이 뚜렷하게 나타났다. 본 연구는 생성형 AI를 통한 영상 콘텐츠 내재 특성의 체계적 추출 가능성을 실증하였으며, 서예 크리에이터의 조회수 상승을 위한 콘텐츠 전략 수립에 실무적 시사점을 제공한다.

Abstract

This study aims to extract intrinsic content characteristics using generative AI reasoning beyond YouTube video external data collected via web crawling, and explore their effects on view counts. With NotebookLM, 42 external and intrinsic features were extracted from 237 calligraphy-related YouTube videos for impact analysis. The results showed two external and seven intrinsic factors significantly influenced view counts in the overall sample. Subgroup analysis indicated outstanding audience responses to large running script videos without background music, as well as long-duration large-font videos of scripts other than seal script created by male creators under the same condition. This study empirically proves generative AI can systematically extract intrinsic features of online video content, and provides practical implications for calligraphy creators to formulate content strategies for higher view counts.

Keywords

generative AI, big data analytics, intrinsic information, calligraphy, mixture regression model

* 경상국립대학교 경영학과(*² 교신저자)
 - ORCID¹: <https://orcid.org/0000-0002-5993-6984>
 - ORCID²: <https://orcid.org/0009-0008-3960-9338>
 ** 원광대학교 미술대학 외래교수
 - ORCID: <https://orcid.org/0009-0002-7884-1929>
 *** 국립한국교통대학교 컴퓨터공학과 교수
 - ORCID: <https://orcid.org/0000-0002-7407-207X>

· Received: Feb. 20, 2026, Revised: Mar. 08, 2026, Accepted: Mar. 11, 2026
 · Corresponding Author: Hai Chao Xu
 Dept. of Business administration, Gyeongsang National University
 501 Jinju-daero, Jinju-si, Gyeongsangnam-do, 52828, Korea
 Tel.: +82-55-772-3454 Email: haidao hao@hotmail.com

I. 서 론

한국의 유튜브 사용자는 2025년에 월간활성 사용자수(MAU, Monthly Active Users)가 4,813만명으로 전체 인구 도달률이 94.7%에 달한다[1][2]. 크리에이터는 이런 수요행태를 기반으로 수익을 창출할 목적으로 활동한다. 크리에이터의 가장 큰 관심 중의 하나는 어떤 특징을 가진 콘텐츠를 유튜브에 업로드해야만 조회수나 좋아요 수를 높일 수 있을 것인가 하는 것이다.

이 콘텐츠 특징은 수집되는 정보의 성질에 따라 동영상길이나 조회수, 좋아요 수 같은 양적 특징(Quantitative features) 또는 외재적 정보(Extrinsic information)와 해쉬태그의 유인성, 등장인물의 매력성, 콘텐츠 내용 구성과 같은 질적 특징(Qualitative features) 또는 내재적 정보(Intrinsic information)로 나눌 수 있다. 동영상에 나타나는 외재적 정보는 web crawling으로 수집이 용이하다. 반면에 동영상 내부의 맥락, 시각적 요소, 연출방식, 콘텐츠 내 관찰되는 형태, 내용, 특성 등 직접적인 데이터 추출이 어려운 정보는 심층적 정보(Deep information), 비정형 정보(Unstructured information), 내재적 정보라고 정의 내릴 수 있는데, 이 정보는 웹 크롤링으로 쉽게 추출할 수 없다. 그래서 지금까지는 전문가인 인간이 개입하여 내재적 정보를 수집하여 왔다.

기존 빅데이터 연구는 주로 웹크롤링이나 유튜브 측이 제공하는 정보 중 표면적 정보인 외재적 정보를 이용한 경우가 많았다[3]-[5]. 하지만 표면적 정보로는 유튜브 조회수를 설명하는데 제한적이다. 따라서 최근 활발히 개발되어 출시되고 있는 생성형 AI(Generative AI)를 활용하여 유튜브 동영상 콘텐츠의 양적 물적 특징 중 심층적 정보인 내재적 정보를 추출한 시도가 있었다[6]. 하지만, 이 기존 연구는 유튜브 동영상 조회수에 영향을 미친 정도를 분석하면서 전체표본수준의 연구를 진행하였다. 이런 전체 데이터를 하나의 회귀식으로 분석하는 경우에는 데이터 내에 숨겨진 여러 개의 하위 집단(Latent classes)간 이질적인(Heterogeneous) 유튜브 동영상 선호행태를 반영할 수 없다.

이 연구의 목적은 유튜브 예술 동영상과 같이 빅

데이터 내에서 단기간에 내재적 정보를 추출하는 일련의 과정을 실증적으로 실행하고, 이렇게 추출된 내재적 정보를 하위집단 겸 세분시장 수준 분석에 의해 소비자의 선호이질성을 확보한 상태에서 조회수에 영향을 미치는 유튜브 콘텐츠 특성을 확인하는데 있다. 이 연구 목적을 달성하기 위해 여러 산업 중 비정형적이고 내재적 속성의 심층정보가 상대적으로 많이 동영상에 담겨있을 문화산업 중 서예분야 크리에이터의 유튜브 동영상을 대상으로 연구를 진행하였다. 이를 통해 실무적으로는 빅데이터 형태의 유튜브 동영상 내부의 심층적 정보를 추출하는 일련의 과정을 학술적으로 소개한다는 공헌점을 가지며, 서예 유튜브 크리에이터에게 조회수를 증가시킬 수 있는 콘텐츠 특징의 전술적 가이드라인을 제공하게 된다. 학술적으로는 연구분석대상측면에서 문화산업이자 예술분야인 서예분야에서 거의 이루어지지 않는 온라인관련 연구결과를 축적한다는 공헌점과 연구분석단위 측면에서 내재적 정보가 조회수와 같은 성과변수에 미치는 영향여부를 전체표본데이터수준과 세분시장수준에서 확인할 수 있는 공헌점을 갖게 될 것이다.

II. 연구 방법

2.1 연구대상의 선정

소셜 미디어 플랫폼(Social media platform)은 카톡, 페이스북, 유튜브, 틱톡, 인스타그램, X 등 다양하며, 동시에 플랫폼마다 동영상 특징이 상이하다. 따라서 이 연구에서는 가장 많은 사용자와 많은 학술적 연구가 진행된 유튜브를 연구대상으로 삼았다.

한편, 유튜브 API에서는 크리에이터가 영상을 업로드할 때 크리에이터가 직접 지정하는 표준 동영상 카테고리(Standard video categories)가 있는데 이 카테고리별로 외재적이며 표면적 정보와 내재적이고 심층적 정보가 매우 상이할 것으로 예상된다. 따라서 여러 카테고리 중 특정 카테고리나 산업으로 연구 분석대상을 제한할 필요성이 대두되었다.

또한 심층적 정보가 많이 들어 있을 개연성이 높은 카테고리나 산업을 선택할 필요성이 있었다. 이

에 한국정부가 분류하는 문화산업 12개 분류와 유튜브의 12개 표준 동영상 카테고리도 동시에 발견되는 예술분야에서 연구대상을 선택하였다. 특히 예술분야 중 서예는 한자문화권의 15억명이 사용하지만, 산업크기에 비해 유튜브 콘텐츠가 활성화되지 않았고, 이 예술분야의 작가들의 소득이 상대적으로 다른 예술에 비해 낮다고 알려졌다[7]. 이에 이 분야 예술가들에게 수익측면에서 본 연구의 연구결과가 실질적인 도움이 되기 위해 서예 유튜브 동영상을 연구대상으로 삼았다.

2.2 데이터수집

이 연구에서는 유튜브에 전통서예 작품에 대한 동영상을 2026년 2월 현재 537개 업로드 한 ‘국당 조성주의 하나 둘 셋’에서 데이터를 수집하였다. 데이터 수집 기간은 2023년 11월부터 2026년 1월까지로 이 기간 동안 업로드된 전체 공개 영상을 확보하였다. 정리 과정에서는 중복 영상이나 시험적 성격의 영상을 제거하였으며, 최종적으로 237편의 롱폼 영상으로 표본을 구성하였다. 연구에서 분석해야 할 동영상 상영길이는 약 93.6시간이다. 이렇게 특정 채널을 선택한 이유는 같은 예술분야 동일한 장

르이어도 채널별 특성에 따라 내재적 외재적 특징이 다를 것이라는 현상을 반영하기 위함이다. 또한 여러 채널을 선택했을 때 창작자의 발신자의 특성 변수가 조회수에 영향을 미치는 영향을 배제하기 위해서이다.

2.3 유튜브 콘텐츠 동영상 특징 정보 pool구성

이 연구에서 가장 선행해야 할 작업은 서예분야의 동영상 특징을 분류하는 양적, 질적, 표면적 심층적 기준을 선택하는 것이다. Y. Kwak et al.[7]는 이미 서예 소셜 미디어 콘텐츠 특징분류기준 풀(Pool)을 구성하여 제시한 바 있다.

이들은 전문가를 대상으로 Delphi법에 의해 5대 대분류 요인 안에 49개 유튜브 서예 특징분류 기준을 정리하였다. 본 연구에서는 표 1과 같이 Y. Kwak et al.[7]에서 제시된 변수 중 썸네일 가시성, 썸네일 내 제목유인력, 썸네일 스타일 일관성 등 측정가능성이 없는 변수를 제외하고 총 42개 동영상 특징을 생형 AI에게 질문하는 것으로 적용하였다. 표 1과 같이 대상에 대한 정보를 외재적 내재적 정보로 분류하는 것은 J. Jacoby et al.[8]이래로 전통적인 방법이다.

표 1. 본 연구에서 생형형 AI에게 분석의뢰된 양적, 질적, 표면적, 심층적 정보 풀

Table 1. Information pool of quantitative, qualitative, surface, and intrinsic content features extracted through generative AI in this study

Category	quantitative information	qualitative information
Surface information/ extrinsic information	Number of views, video duration	Presence of subtitles, presence of background music, presence of sound effects
Deep information/ intrinsic information	Number of Hangul script categories, number of Hanja (Chinese character) script categories, number of seal script categories, number of calligraphy content categories, amount of calligraphy (text volume), number of calligraphy works, font size	Filming location, presence of dialogue, presence of exhibition, exhibition characteristic classification, presence of performance, theory vs. practice distinction, episode classification, presence of Chinese calligrapher, presence of Korean calligrapher, creator gender, Chinese character script classification, Hangul script classification, type of instructional format, Chinese dynasty classification, Korean dynasty classification, Hangul historical period classification, whether artist research is included, artist gender, artist life period classification, calligraphy writing posture, brush-holding method, seal script sub-classification, calligraphy target (e.g., poetry/prose), presence of calligraphy critique, content classification of the artwork, artwork composition structure classification, color classification, color mixing method, method of artwork presentation

III. 연구절차 및 연구결과

3.1 생성형 AI 선택

이 연구에서는 연구자가 제공한 정보만을 사용하여 응답하는 생성형 AI 특징을 가진 NotebookLM을 사용하였다. 이를 통해 일반적인 생성형 AI가 갖는 환각효과를 방지하였다.

3.2 정보 추출 prompt 제공

NotebookLM에게 동영상별 링크를 제공한 후, 표 1의 42개 서예유튜브 특징별로 동영상별 정보를 추출해 줄 것을 그림 1과 같이 명령했다. 그 결과, 추출된 동영상별 42개 서예동영상 콘텐츠 특징을 데이터셋으로 만들어 달라고 생성형AI에게 다시 지시하였다. 이를 통해 조회수에 영향을 미치는 독립변

수 및 종속변수의 데이터셋을 확보하였다.

3.3 추출데이터 타당성검증

생성형 AI인 NotebookLM이 산출한 동영상별 42개 특징이 올바른지를 인간인 전문가가 판단하는 절차를 거쳤다. 중국과 일본에서 서예와 전각을 전공하였으며 현역 활동 중인 한국인 작가가 실제 동영상을 시청하면서 NotebookLM이 생성한 동영상별 정보가 올바르게 추출되었는지를 확인하는 절차를 거쳤다. 전문가는 NotebookLM은 작가창작자세가 보이지 않는데 237개 동영상 중 2번을 '현완법(붓글씨를 쓸 때 팔목을 들어 바닥에 대지 않고 붓을 곧게 쥐고 쓰는 방법)'을 사용했다고 응답하는 잘못을 발견하였다. 최종 연구에서는 전문가의 수정의견이 반영된 데이터셋을 사용하였다.

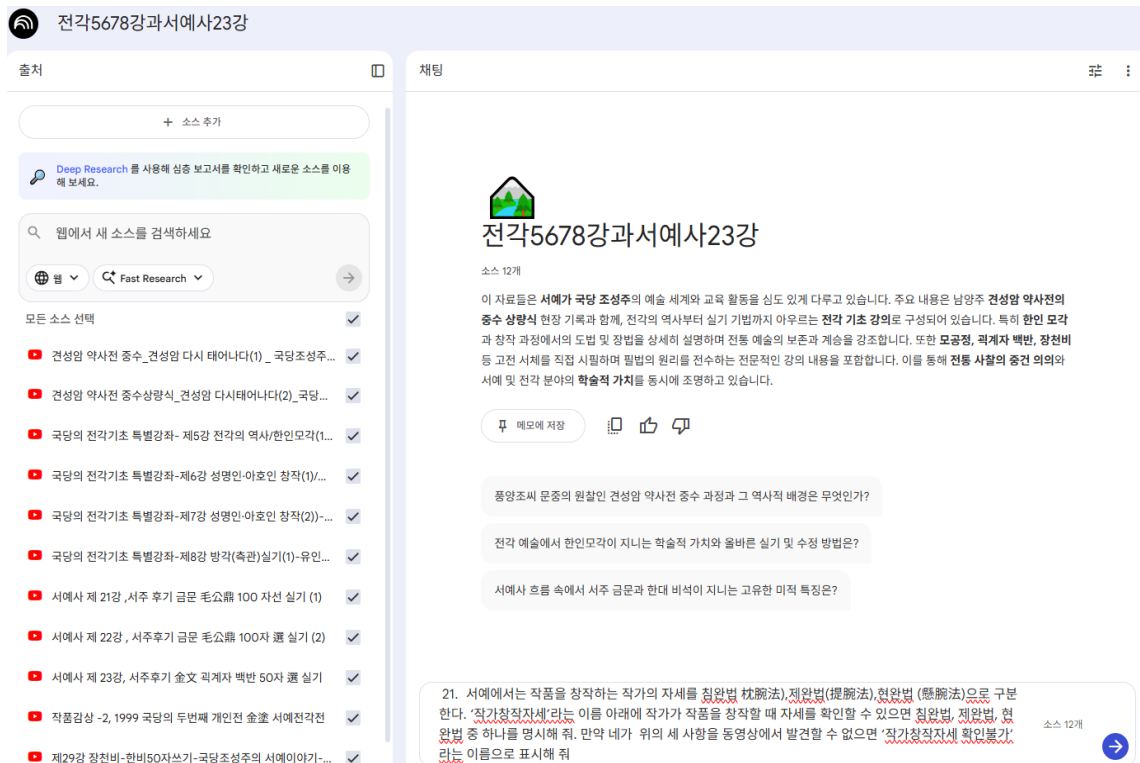


그림 1. NotebookLM에 양적 질적 표면적 심층적 정보 추출 의뢰된 사례 일부

Fig. 1. Selected cases of information extraction requests to notebookLM involving quantitative, qualitative, surface, and intrinsic content features

이 연구에 사용된 변수선정과 연구절차는 모두 Y. Kwak et al.[6][7]과 동일하게 진행하였다. 다만 연구대상 표본 수집기간이 확대되어 연구대상 동영상 수가 211개에서 237개로 증가하였다. 이를 통해 기존 논문과 이번 연구 간의 종단적 연구(Longitudinal study) 성격을 부여하고, 분석단위가 전체표본데이터수준에서 세분시장수준으로 연구단위 변경이 반영할 수 있게 되었다.

3.4 유튜브 콘텐츠 정보가 조회수에 미치는 영향을 전체표본 데이터수준에서 분석한 경우

NotebookLM이 추출한 서예 유튜브 콘텐츠 특성을 독립변수로 하고 동영상별 조회수를 로그를 취한 값을 종속변수한 다중회귀분석을 실시하였다. 이때 stepwise방식을 사용하였다. 41개 독립변수를 한꺼번에 사용하는 enter방식에 비해 유의한 변수를 확인하는데 유리하게 분산을 처리할 것으로 판단하였다. 그런데, 동영상별 조회수가 정규분포를 하지 않고 조회수가 증가함에 따라 동영상 편수가 줄어드는 오른쪽으로 치우친 분포(Right-skewed distribution)를 보인다. 대부분의 영상은 낮은 조회수를 기록하는 반면, 소수의 영상만이 매우 높은 조회수를 보이는 ‘스타콘텐츠’가 존재하기 때문이다. 따라서 조회수(View count)에 로그를 취한 값을 종속변수로 사용하였다. 변수간 다중공선성검토를 사전에 마친상태에서 분석이 실시되었다.

표 2. 로그(조회수) 기준 회귀모형 요약

Table 2. Summary of regression models in case of log(view count)

Dependent Variable	R ²	Adj-R ²	F-value	p-value
log(View count)	0.704	0.686	37.78	P<.001

회귀분석의 결정계수는 표 2에서 보듯 $R^2=0.704$, adjusted $R^2=0.686$ 으로 나타났다. 모형 전체가 통계적으로 유의하였다($F=37.78$, $p<0.001$). 내재적 정보를 포함한 독립변수 중 9개 변수가 조회수에 통계적으로 유의한 영향을 미치고 있었다. 이 중 양적 변수는 동영상 상영길이, 서예작품 수, 글자크기 등 3개이고 나머지는 질적 변수 6개였다. 외재적 정보

가 2개이고, 내재적 정보가 7개였다(표 3 참조). 예를 들어 중국 한자서체 중 행서는 조회수를 높이지만, 전서는 조회수를 낮추는 영향을 준다. 또한 동영상에 등장하는 글자크기가 매우 큰 경우에 조회수가 높았다.

표 3. 전체표본 수준 회귀분석 유의 변수 요약

Table 3. Summary of significant variables in the regression at aggregate level

Category	Variables	Items	Coefficients (B)	t_value	p_value
Video physical characteristics	q1. Video duration	Time	0.024	5.148	0.001
	q3. Presence of background music	BGM	-0.782	-5.942	0.001
Content format	q7. Instructional format	Dummy	0.707	3.698	0.001
	q22. Calligrapher's creative posture	Hyunwan	-0.936	-4.890	0.001
	q42. Artist gender	Male	0.199	2.877	0.004
Calligraphic characteristics	q27. Semi-cursive script	Semi-cursive script	0.473	4.151	0.001
	q27. Seal script	Seal script	-0.399	-3.498	0.001
	q43. Artwork size	Extra-large characters	1.378	5.713	0.001
Non-visual elements	q13. Chinese dynasty classification	china-Yuan	-1.363	-3.590	0.001
	q42. Number of artworks	no	-0.004	-3.025	0.003

결론적으로 빅데이터인 유튜브 동영상 콘텐츠를 생성형 AI로 내재적 정보까지 추출할 수 있었으며, 이렇게 추출된 동영상의 양적 질적 정보 중 일부가 조회수에 영향을 미치는 실증분석결과를 얻었다(표 4 참조). 또한 그 설명력도 69%로 매우 높았다.

표 4. 특성 체계 내 유의한 변수 요약

Table 4. Summary of significant variables in characteristics scheme

Category	Quantitative information	Qualitative information
Surface information/ extrinsic information	Video duration	Presence of background music
Deep information/ intrinsic information	Number of calligraphy works, font size	Instructional format, calligrapher's creative posture, artist gender, Chinese script type, Chinese dynasty classification

3.5 콘텐츠정보가 조회수에 미치는 영향을 하위집단인 세분시장수준에서 분석한 경우

혼합 회귀모델(Mixture regression model) 또는 잠재 계층 회귀분석(Latent class regression)은 전체 데이터를 하나의 회귀식으로 설명하는 대신, 데이터 내에 숨겨진 여러 개의 하위 집단(Latent classes)이 존재한다고 가정하고, 각 집단인 세분시장별로 독립변수가 종속변수에 미치는 영향력이 다를 수 있음을 포착한다[8][9]. 그런데, 현재 독립변수가 41개인 반면 표본수는 237개이어서 2개 집단이상으로 혼합 모델을 구성할 경우 추정해야 할 파라미터가 너무 많아진다. 따라서 학술적 관례에 따라 이전 단계적 선택법(Stepwise)에서 유의미하게 나타난 주요 변수들을 중심으로 혼합 모델을 구성하여 분석을 진행하였다. 이 분석은 LatentGold라는 프로그램을 사용하였다.

2개의 하위집단으로 혼합회귀모델을 적용하는 경우, 하위집단 1은 전체 표본의 39.7%에 해당하는 콘텐츠가 속했고, 세분시장2는 전체표본의 60.3%가 속했다(loglikelihood=-218.42, BIC=548.91). 이질성을 반영한 하위집단 수준으로 분석단위를 정교화하자 41개 독립변수 중 5개 변수만이 조회수에 영향을 미쳤다. 전체표본 수준분석에서 9개 독립변수가 조회수에 영향을 미치는 것으로 나타났던 것과 차이가 발생하고 있다.

표 5의 하위집단인 세분시장 수준분석결과를 보면 하위집단1은 배경음악이 삽입되면 조회수가 낮아지지만 서체 중 행서를 특대자로 구사하는 경우에 조회수가 증가하였다. 평균 736회의 조회수를 보

인다. 연구자는 이 세분시장을 ‘초대형 행서콘텐츠’이라 이름지었다. 하위집단2는 동영상길이가 길수록, 진행자가 남자인 경우, 큰 글씨로 쓸 때 조회수가 높았다. 하지만 배경음악이 들어가고 전서를 쓸 경우에는 오히려 조회수가 감소하였다. 평균 3,158회의 조회수를 기록하는 콘텐츠들이다. 일반적으로 전서를 대중이 어려워 하므로 연구자는 이 세분시장을 ‘대중교육형 콘텐츠’이라고 명명하였다.

표 5. 하위집단 수준 혼합회귀모형 분석 결과 요약

Table 5. Summary of the mixture regression model at disaggregate level

Category	Variables	Class 1 (Semi-cursive dominant type, 39.7% of total sample) Coefficient (B) (t)	Class 2 (Popular/General type, 60.3% of total sample) Coefficient (B) (t)
Video physical characteristics	q1. Video duration	0.009 (1.02)	0.024 (5.15)*
	q3. Presence of background music	-0.569 (-2.78)*	-0.767 (-5.94)*
Content format	q42. Artist gender (male)	0.111 (1.16)	0.218 (2.87)*
Calligraphic characteristics	q27. Semi-cursive script	0.852 (4.38)*	0.155 (1.21)
	q27. Seal script	-0.124 (-0.89)	-0.399 (-3.49)*
	q43. Artwork size (extra-large characters)	1.642 (5.11)*	0.864 (3.58)*
Dependent variable	Mean of log(view)	6.18	7.29
	Mean of views	736.6	3158.9

두 집단 간에 조회수에 차이가 있는지를 t-test하였다. 종속변수 중 로그를 취한 조회수 경우를 대상으로 한 두 하위콘텐츠 집단 간 평균차이는 $t=18.203$ 이었고, 로그를 취하지 않은 조회수를 대상

으로 한 평균차이의 t 값은 -1.780 으로 $p=0.076$ 으로 신뢰수준 90%수준에서 통계적으로 유의한 차이가 있었다. 이런 결과는 전서와 같이 어려운 글씨는 피하고 일반적인 글씨를 크고 길게 보여주는 콘텐츠를 시청하는 '대중교육형 콘텐츠(Class 2)'가 특정 서체인 행서를 선호하는 '행서선호(Class 1)' 집단에 비해 월등히 높은 조회수 수준을 기록하였다. 이 채널의 경우에는 서예 영상 콘텐츠 수요시장이 단일한 성격을 가진 시장이 아니라, 성과 구조와 소비 행태가 근본적으로 다른 두 개의 세부 시장으로 확연히 구분됨을 알 수 있다. 또한 서체선호현상이 이질적으로 존재함을 발견할 수 있다.

IV. 결 론

이 연구의 목적은 기존 빅데이터 연구에서 많이 진행된 웹 크롤링에 의한 표면적 정보 대신에 상대적으로 소홀히 연구되어 온 유튜브 동영상 내부의 심층적 정보를 생성형 AI인 NotebookLM의 추론 능력을 활용하여 추출할 수 있는지 여부를 실증분석을 통해 제시하는 것이었다. 더나가 이렇게 생성형 AI가 추출한 정보가 조회수와 같은 성과에 유의한 영향을 이질적으로 미치는 지를 점검하는데 있었다. 특히 기존 연구에 비해 서예 동영상의 연구대상을 확대하고, 연구분석단위를 전체표본에서 세분시장수준으로 확대하여 분석하였다[10][11].

이를 위해 심층정보가 상대적으로 많이 내포할 가능성이 있는 예술분야 중 15억명의 한자문화권의 문화유산인 서예를 대상으로 한 특정 채널 동영상 분석을 NotebookLM을 통해 실시하였다. 그 결과, 서예 유튜브 동영상 콘텐츠에 양적, 질적, 외재적, 내재적 정보를 42개 변수에서 추출할 수 있었으며, 이것이 인간이 추출한 것과 거의 비슷한 정보획득 결과를 갖는 것으로 나타났다. 또한 전체표본을 대상으로 한 경우에는 41개 독립변수 중 9개 독립변수가 조회수에 유의한 영향을 미치고 있었다. 또한 세분시장별 분석에서는 초대형 행서콘텐츠가 배경음악이 없는 상태에서 중국서체 중 행서를 큰 글씨로 작업하는 경우와 대중 교육형 콘텐츠로 영상길이, 배경음악여부, 전서체, 글자크기, 작가성별이 조회수

에 유의한 영향력이 있는 경우를 파악할 수 있었다. 하위집단인 세분시장단위에서는 5개의 유의한 변수가 조회수에 영향을 세분시장별로 다르게 미치고 있었다.

이를 통해 서예 크리에이터가 실무적으로 전체 고객의 조회수를 증가시키기 위해 외재적 표면적 정보 2개와 내재적 정보 7개를 활용하여 유튜브 동영상 콘텐츠를 구성해야 한다는 가이드라인을 제시하였고, 하위집단인 세분시장수준에서는 외재적 정보 2개, 내재적 정보 3개를 활용하여 콘텐츠를 별도로 제작해야 하는 것으로 나타났다.

다만, 이 연구는 1개의 서예 개인 크리에이터의 동영상 사이트만을 대상으로 했다는 한계점을 지니고 있다. 하지만 이 연구는 Y. Kwak et al.[6]과 동일한 이론적 토대와 변수 pool을 유지하면서도 연구범위와 분석수준을 확장하였다. 첫째, 표본기간을 확대하여 237개 영상으로 확장함으로써 종단적(Longitudinal) 성격을 보완하였다. 둘째, 전체표본을 하나의 동질적 시장으로 가정하지 않고 혼합회귀모형을 적용하여 잠재계층을 도출함으로써 소비자 선호의 이질성을 구조적으로 반영하였다. 그 결과 전체수준에서 9개 변수가 유의했던 것과 달리, 세분시장 수준에서는 영향변수가 재구성되어서, 초대형 행서콘텐츠와 대중교육형 콘텐츠의 조회수에 영향력이 다른 두 하위집단의 상이한 성과구조를 확인하였다. 즉, 조회수를 높이기 위해서는 배경음악을 없애고 전서 서체를 피하면서 남성작가를 동원하여 큰 글씨의 콘텐츠를 길게 만들어 업로드하는 것이 필요함을 발견하였다. 학술적으로는 서예산업연구에서 등한시 되어 온 온라인콘텐츠분석이라는 연구성과를 축적하는 성과를 얻었다. 특히, 이 연구에서 사용한 일련의 방법론이 문화산업 콘텐츠와 같이 웹크롤링에 의해 데이터표면에서 얻어지는 정보보다, 동영상 내용 중에 심층적으로 투영되어 있고 시각적으로 관찰해야한 하는 콘텐츠가 많은 산업과 분야에 확대 적용할 수 있을 것으로 판단된다.

References

- [1] Springer Blog 2025 YouTube Statistics: Global

- Overview and Key Trends, <https://www.springer.com/blog/2025-youtube-statistics>. [accessed: Dec. 18, 2025]
- [2] IGAWorks Blog. YouTube App Analysis Report, <https://www.igaworksblog.com/reports>. [accessed: Feb. 12, 2025]
- [3] B. Koh and F. Cui, "An exploration of the relation between the visual attributes of thumbnails and the view-through of videos: The case of branded video content", *Decision Support Systems*, Vol. 160, Art No. 113820, Sep. 2022. <https://doi.org/10.1016/j.dss.2022.113820>.
- [4] B. Liu, "Towards micro-video thumbnail selection via a multi-label visual-semantic embedding model", *arXiv preprint arXiv:2202.02930*, Feb. 2022. <https://arxiv.org/abs/2202.02930>.
- [5] G. Cui, Y. Chung, L. Peng, and Q. Wang, "Clicks for money: Predicting video views through a sentiment analysis of titles and thumbnails", *Journal of Business Research*, Vol. 183, Art No. 114849, Oct. 2024. <https://doi.org/10.1016/j.jbusres.2024.114849>.
- [6] Y. Nam, Y. Kwak, H. Xu, Y. Kwak, S. Park, and D. Lee, "A Study on the Extraction of Deep Information within Calligraphy YouTube Content Using Generative AI and Its Impact on View Counts", *Proc. 2026 ICobar-SMART Joint Conf.*, Binus University, Jakarta, Indonesia, Feb. 2026.
- [7] Y. Kwak, D. Lee, and H. Xu, "An exploratory study on pool extraction of calligraphy social media content feature classification criteria: Delphi analysis focused on YouTube videos", *The Journal of Asian Studies*, Vol. 28, No. 4, pp. 65-88, Dec. 2025. <http://doi.org/10.21740/jas.2025.11.30.2.65>.
- [8] J. Jacoby, J. C. Olson, and R. A. Haddock, "Price, brand name, and product quality perceptions", *Journal of Applied Psychology*, Vol. 55, No. 6, pp. 570-579, Feb. 1972. <https://doi.org/10.1037/h0032239>.
- [9] M. Wedel and W. A. Kamakura, "Market segmentation: Conceptual and methodological foundations", Kluwer Academic Publisher, 2000.
- [10] P. Hosagoudar and N. Manjunath, "A study on YouTube channel and influencer analysis", *International Journal of Advanced Research in Science, Communication and Technology*, Vol 4, No. 1, pp. 203-208. Nov. 2024. <https://doi.org/10.48175/IJARSCT-22035>.
- [11] S. Yang and K. Cho, "The science of YouTube: What factors influence user engagement with online science videos?", *PLoS One*, Vol. 17, No. 5, e0267697, May 2022. <https://doi.org/10.1371/journal.pone.0267697>.

저자소개

곽 영 식 (Young Sik Kwak)



1997년 8월 : Texas Tech University, Merchandising, (Master of Science)
1999년 2월 : 성균관대학교 경영학과(경영학박사)
2003년 3월 ~ 현재 : 경상국립대학교 경영학부 교수
관심분야 : e-biz, 가격책정, 문화마케팅

허 해 초 (Hai Chao Xu)



2016년 7월 : 중국 남산대학교 한국어학과(문학학사)
2019년 2월 : 경상국립대학교 경영학과(경영학석사)
2019년 10월 ~ 2023년 8월 : 중국 Yantai Gold College 경영학과 강사
2023년 9월 ~ 현재 : 경상국립대학교 경영학과 박사과정
관심분야 : E-commerce, 콘텐츠 마케팅, 홍보 관리

이 두 희 (Du Hee Lee)



2002년 8월 : 중국미술학원
서법과(석사)

2005년 8월 : 일본
나라교육대학대학원
미술교육 서도전공(석사)

2009년 3월 : 중부대학교
경영학과(박사)

2009년 3월 ~ 현재 : 원광대학교 미술대학 외래교수
관심분야 : 서예교육

곽 윤 식 (Yoon Sik Kwak)



1984년 6월 : 한국경희대학교
전자공학과(공학사)

1994년 3월 : 한국경희대학교
전자공학과(공학석사)

1991년 5월 ~ 현재 :
한국교통대학교 컴퓨터공학과
교수

관심분야 : 센서 네트워크, 인터넷통신