

Computational Analysis of South Korean Local Government Mascot Characters Using Vision-Language Models and Semantic Embedding

Jiin Lee¹, In Kwon Park^{2*}

¹MS. Student, Department of City Planning, Graduate School of Environmental Studies,
Seoul National University, Seoul, Republic of Korea

(E-mail: sota_jiin@snu.ac.kr)

²Professor, Department of City Planning, Graduate School of Environmental Studies,
Seoul National University, Seoul, Republic of Korea

*Corresponding author: parkik@snu.ac.kr

1. Introduction and Motivation

Local government mascot characters have become a prominent instrument of place branding in South Korea. Municipalities routinely deploy these visual symbols across public communication channels, including tourism promotion, festivals, policy campaigns, and civic services. This expansion reflects a broader shift in which place branding operates as a strategic response to intensified inter-urban competition (Jensen, 2005). Yet despite their practical salience and public investment, municipal mascots remain understudied in systematic, nationwide terms. Existing scholarship has often relied on case-based qualitative interpretation, offering rich contextual readings but limited comparability across jurisdictions. As a result, the field lacks a reproducible analytical framework that can answer a central question for place branding and public administration: do mascots achieve meaningful inter-regional differentiation, or do they instead converge toward standardized designs that reflect institutional isomorphism and diffusion of shared repertoires?

25

26 This study addresses that gap by conducting a computational multimodal analysis of
27 municipal mascots in South Korea. We examine whether mascots encode distinctive local
28 identity signals that differentiate places or whether they display homogenizing tendencies
29 consistent with institutional isomorphism. The empirical focus on mascots is theoretically
30 useful because mascots sit at the intersection of symbolic governance and communication
31 efficiency. On the one hand, municipalities promote mascots as identity-bearing
32 representations of local history, nature, culture, and industry. On the other hand, mascots must
33 be legible, friendly, and widely acceptable to diverse audiences, which can incentivize the
34 adoption of familiar templates. Observing mascot designs at scale therefore provides an
35 opportunity to test the tension between distinctiveness and standardization in contemporary
36 urban governance.

37

38 **2. Research Question and Contributions**

39 The study is guided by three research questions. First, what representational components
40 structure South Korean municipal mascots, particularly with respect to entity types, symbolic
41 connections to local resources, and thematic framing? Second, do reproducible design
42 paradigms exist in the national corpus, and if so, what features characterize each paradigm?
43 Third, are these paradigms geographically patterned by administrative region, suggesting
44 regional differentiation, or do they diffuse nationwide in a manner consistent with institutional
45 isomorphism?

46

The contributions are both methodological and substantive. Methodologically, we propose a scalable multimodal pipeline that integrates vision–language model extraction of structured attributes with semantic embeddings, similarity networks, community detection, and representative clustering. Substantively, we provide nationwide empirical evidence on whether municipal mascot branding in South Korea is oriented more toward differentiation or toward symbolic homogenization. We also identify which local resources municipalities most frequently foreground as their branding “selling points,” thereby clarifying how local identity is translated into standardized brand language.

3. Data and Sample Construction

We compiled a nationwide dataset of South Korean municipal mascots using official visual materials and accompanying textual descriptions. From an initial collection of 229 mascots, we retained 207 cases for analysis after excluding 22 cases with insufficient visual quality, missing official descriptions, or incomplete multimodal information. Each observation includes a mascot image and a textual description used by the municipality to explain the character’s identity, symbolism, and local references. This multimodal structure is crucial because symbolism is often communicated through the joint interpretation of visual cues and narrative framing. A single visual motif may remain ambiguous without the text, while the text may amplify or redirect the meaning of the visual representation.

4. Multimodal Attribute Extraction Using a Vision–Language Model

To render mascot representations comparable at scale, we employed OpenAI’s GPT-4o vision–language model (VLM) to extract structured attributes from mascot images and their associated textual descriptions. The extraction schema targeted attributes that are central to place branding analysis, including entity type (animal, human-form, object, or mixed), anthropomorphism and character framing, symbolic connections to local resources (such as agricultural specialties, historical figures, and geographic landmarks), and thematic elements (such as nature, culture, and future orientation). By converting multimodal content into structured variables, the VLM enables national-scale comparative analysis that would be costly and time-intensive to conduct purely through manual coding.

To validate the reliability of VLM-based extraction, a subset of cases was independently coded by human annotators using the same schema. The validation yielded substantial inter-rater agreement, supporting the use of VLM-assisted structured coding as a replicable approach for multimodal place branding research. This step is important for establishing credibility and interpretability, particularly because automated extraction must be assessed against human judgment to ensure that inferred attributes align with plausible readings of the image-text pair.

5. Semantic Embeddings and Similarity Network Construction

Following attribute extraction, we generated 3,072-dimensional semantic embeddings using the text-embedding-3-large model. These embeddings provide continuous representations in which semantically similar descriptions are located near one another in high-dimensional space. We computed pairwise cosine similarities between all 207 mascot embeddings to quantify

semantic proximity. Rather than constructing a fully connected similarity network, which would be dense and potentially noisy, we applied k-nearest neighbor sparsification with k set to 8. This procedure retains, for each mascot, edges connecting it only to its eight most similar counterparts, thereby preserving the strongest similarity relations while filtering weaker links. The resulting sparsified network captures the local neighborhood structure of similarity and provides a tractable graph representation for community detection.

6. Multi-Stage Clustering Strategy

We implemented a multi-stage clustering strategy that combines network-based community detection with partition-based clustering to leverage complementary strengths. In the first stage, we applied the Louvain algorithm for community detection to the kNN similarity network (Blondel et al., 2008). Louvain iteratively reassigns nodes to communities to maximize modularity, producing a partition in which within-community connectivity is higher than expected under a random baseline. This step identifies natural community structure without requiring a priori specification of the number of clusters. Applied to the mascot similarity network, the Louvain algorithm yielded four communities, and the modularity score indicated meaningful structure above chance. We adopted this four-community solution as a data-driven target cluster count for subsequent analysis.

In the second stage, we performed k-medoids clustering with k set to 4 using the Partitioning Around Medoids (PAM) algorithm and cosine distance (Kaufman and Rousseeuw, 1990). We preferred k-medoids over k-means because medoids are actual data points rather than abstract

centroids. This property is especially valuable for interpretability because each cluster can be represented by a real mascot that functions as an exemplar for qualitative reading and communication to non-technical audiences. The PAM algorithm iteratively optimizes cluster assignments and medoid selection to minimize within-cluster dissimilarity and converges on a stable partition in which each mascot is assigned to exactly one cluster.

7. Validation and Robustness Checks

We evaluated the clustering solution using multiple complementary metrics. First, intra-cluster average cosine similarity measured semantic coherence within each cluster. The four clusters exhibited high coherence, with intra-cluster similarities ranging from 0.826 to 0.843. Second, we computed the silhouette coefficient to assess cluster separation. The average silhouette score was 0.103, suggesting modest separation and indicating that mascot designs occupy a continuous semantic space with overlapping boundaries rather than forming sharply discrete categories. Third, we conducted chi-square analysis to test whether cluster assignment is statistically independent of administrative region. This test assesses whether design paradigms vary geographically, consistent with differentiation, or instead diffuse nationwide, consistent with institutional isomorphism.

8. Empirical Findings

The analysis reveals several consistent patterns in how municipalities represent place identity through mascots. With respect to representational composition, mixed entity types that combine multiple symbolic elements, along with human-form representations, dominate the

corpus and account for 57 percent of all designs. This suggests that municipalities frequently employ composite designs capable of encoding multiple local references within a single character, which may enhance narrative flexibility and recognizability.

Symbolic resources concentrate on tangible regional assets. The most frequent symbolic connections were local agricultural specialties (21.3 percent), historical figures (18.8 percent), and geographic landmarks (17.4 percent). This distribution indicates that municipal branding through mascots privileges assets that are concrete and communicable to broad audiences, reinforcing an interpretation of mascots as translation devices that convert local resources into legible brand cues. Thematic framing similarly shows patterned emphasis. Nature-oriented themes comprised 34.3 percent of cases, followed by culture-oriented themes at 28.0 percent and future-oriented themes at 26.1 percent. Taken together, these results suggest that municipalities balance authenticity claims grounded in nature and culture with forward-looking narratives that align with modernization and aspiration.

The clustering results identified four design paradigms with high internal coherence, supporting the presence of reproducible patterns in mascot representation. However, the modest silhouette score implies that these paradigms should be interpreted as overlapping repertoires rather than sharply bounded categories. This finding is consistent with an environment in which municipalities draw from shared symbolic and stylistic conventions, blending and recombining familiar elements rather than inventing entirely unique forms.

Finally, the chi-square analysis found no statistically significant association between administrative region and cluster assignment. This suggests that the observed design paradigms are not strongly localized by geography. Instead, mascot conventions appear to diffuse nationwide, transcending administrative boundaries and implying a degree of institutional isomorphism in symbolic practices.

9. Discussion: Differentiation and Isomorphism in Symbolic Governance

These findings indicate that South Korean municipal mascots operate simultaneously as tools of place differentiation and as vehicles of symbolic homogenization. On one side, municipalities clearly attempt to anchor mascots in place-specific assets such as specialties, landmarks, and historical figures. On the other side, the continuous semantic space, overlapping cluster boundaries, and lack of geographic association point to a shared design repertoire that circulates across the country. Such diffusion can arise through mimetic adoption of perceived successful templates, shared vendor and designer ecosystems, and legitimacy-seeking behavior in public communication. In this context, standardized designs may reflect rational adaptation to audience expectations and communication constraints, but they also risk weakening the differentiation that place branding is meant to achieve.

The results therefore complicate a simplistic expectation that branding efforts naturally generate distinction. Instead, they suggest a dual process in which municipalities pursue distinctiveness through localized symbolic references while simultaneously converging in

stylistic form and representational logic. This duality is central to understanding symbolic governance in competitive yet institutionally interconnected environments.

10. Limitations and Future Work

The study has several limitations that motivate future research. First, while VLM-based extraction was validated on a subset, broader robustness checks would further strengthen generalization, including expanded human coding, prompt sensitivity testing, and comparisons across alternative models. Second, our analysis focuses on symbolic content and semantic similarity rather than downstream outcomes. Future work should link mascot paradigms to performance indicators such as social media engagement, tourism campaign reach, public sentiment, or policy communication effectiveness. Third, geographic region may be too coarse to capture meaningful institutional variation. Incorporating municipal fiscal capacity, tourism dependency, industrial structure, governance networks, and procurement practices may reveal mechanisms that better explain diffusion and convergence. Finally, cross-national comparisons would clarify whether the observed balance of differentiation and isomorphism is specific to South Korea or characteristic of municipal branding more broadly.

11. Conclusion

This study provides a nationwide, reproducible computational analysis of South Korean municipal mascots as multimodal place branding artifacts. By integrating VLM-based structured attribute extraction with semantic embeddings, kNN similarity networks, Louvain community detection, and k-medoids clustering, we demonstrate a scalable approach for

199 analyzing symbolic governance at scale. Substantively, we show that municipalities frequently
200 foreground tangible regional assets and recurring thematic frames, while simultaneously
201 exhibiting diffusion-driven convergence toward standardized design repertoires. The findings
202 provide empirical evidence that place branding practices can encode local identity yet also
203 reproduce institutional isomorphism, raising questions about the effectiveness of mascot-based
204 branding as a mechanism for durable regional distinction. Methodologically, the study
205 illustrates how multimodal AI tools can extend place branding research beyond small-n
206 qualitative analysis toward systematic, comparative, and replicable empirical inquiry.