

# 머신러닝 기반 아파트 주동형상 자동 판별 모형 개발 및 적용 - 주동형상에 따른 아파트 개발 특성분석을 중심으로 -

한상욱<sup>1</sup> · 서정석<sup>1</sup> · Sri Utami Purwaningati<sup>2</sup> · 오지훈<sup>3</sup> · 김정섭<sup>4\*</sup>

## Application and development of a machine learning based model for identification of apartment building types

### - Analysis of apartment site characteristics based on main building shape -

Sanguk HAN<sup>1</sup> · Jungseok SEO<sup>1</sup> · Sri Utami Purwaningati<sup>2</sup> ·  
Jihun OH<sup>3</sup> · Jeongseob KIM<sup>4\*</sup>

#### 요 약

본 연구의 목적은 GIS와 머신러닝 알고리즘을 활용하여 아파트 단지의 주동형상을 자동으로 판별해주는 모형을 개발하고, 이를 주동형상과 단지특성 간의 관계 분석에 적용하는 것이다. 지리정보데이터를 사용하여 아파트단지별 주동 데이터베이스를 구축하고 랜덤포레스트 알고리즘을 활용하여 단지 내 개별동을 형태에 따라 판상형, 탑상형, 혼합형으로 분류하였다. 또한, 아파트단지별 주동형상별 비중과 개발밀도, 층수 등 단지특성 정보간의 관계를 분석하여 부동산 분야 지리정보 응용 가능성을 제안하였다. 본 연구는 인공지능 기반 건축물 유형 분류와 관련한 기초연구로서 다

2023년 05월 10일 접수 Received on May 10, 2023 / 2023년 05월 21일 수정 Revised on May 21, 2023 / 2023년 05월 30일 심사완료 Accepted on May 30, 2023

\* 이 논문은 한국국토정보공사 공간정보연구원 산학협력R&D사업의 지원을 받아 수행된 연구임.(과제명:노후계획도시 재정비를 위한 디지털트윈 기반 정비사업 시뮬레이션 및 평가체계 개발, 과제번호:2022-503) / This (research ect.) was supported by industry-academic Cooperation R&D program funded by LX Spatial Informaion Research Institute(LXSIRI, Republic of Korea) [Project Name: Simulation and evaluation framework for redevelopment of old planned cities using digital twin model / Project Number: 2022-503]

1 울산과학기술원 도시환경공학과 박사과정 / Ulsan National Institute of Science and Technology, Department of Urban and Environmental Engineering, Graduate student

2 울산과학기술원 도시환경공학과 석·박통합과정 / Ulsan National Institute of Science and Technology, Department of Urban and Environmental Engineering, Graduate student

3 울산과학기술원 도시환경공학과 박사후연구원 / Ulsan National Institute of Science and Technology, Department of Urban and Environmental Engineering, Postdoctoral Associate

4 울산과학기술원 도시환경공학과 부교수 / Ulsan National Institute of Science and Technology, Department of Urban and Environmental Engineering, Associate Professor

\* Corresponding Author E-mail: jskim14@unist.ac.kr

양한 공간분석 및 부동산 분석에 활용될 것으로 예상된다.

**주요어** : 아파트, 주동형상, 지리정보시스템, 랜덤포레스트, 머신러닝

## ABSTRACT

This study aims to develop a model that can automatically identify the rooftop shape of apartment buildings using GIS and machine learning algorithms, and apply it to analyze the relationship between rooftop shape and characteristics of apartment complexes. A database of rooftop data for each building in an apartment complex was constructed using geospatial data, and individual buildings within each complex were classified into flat type, tower type, and mixed types using the random forest algorithm. In addition, the relationship between the proportion of rooftop shapes, development density, height, and other characteristics of apartment complexes was analyzed to propose the potential application of geospatial information in the real estate field. This study is expected to serve as a basic research on AI-based building type classification and to be utilized in various spatial and real estate analyses.

**KEYWORDS** : Apartment, Main Building Shape, GIS, Random Forest, Machine Learning.

## 서 론

1962년 마포아파트를 시작으로 본격적으로 공급된 단지형 아파트는 현재 우리나라 주택의 대표적인 유형이 되었다. 통계청에서 실시한 '2021년 인구주택총조사'에 따르면 전체 주택호수 중 아파트는 63.5%를 차지하여 국내 주택시장에서 아파트가 차지하는 비중이 매우 높다. 아파트의 형태는 건축물 자체의 시각적인 특성뿐만 아니라 건축물의 실내 공간 특성을 결정하는 중요한 요소 중 하나이다. 특히, 아파트의 주동형상은 주택의 외관과 평면구성에 큰 영향을 미치며 아파트 단지의 밀도, 층수, 성능 등과 밀접하게 관련되어 있다.

1990년대까지 아파트 단지 배치계획은 보통 판상형 주동을 남향이 확보되도록 배치하는 것이 일반적이었다(Son and Kang, 2001). 1990년대 후반, 2000년대 초반부터는 개발 밀도의 상승과 기술력의 발전에 따라 탑상형 주동을 기초로 하는 초고층 아파트나 주상복합 개발이 본격화되었다(Lee and Lee, 2014). 고층 고밀

개발에서는 밀도를 높이며 건축법상 이격거리를 만족하기 위해 동서로 긴 판상형과는 달리 정사각형, X자형, Y자형과 같은 탑상형 아파트가 나타났다. 최근에는 판상형과 탑상형의 특징을 혼합한 혼합형 아파트도 등장하면서 주동형상이 더욱 다양해졌다.

판상형 주동은 건축물의 외관뿐만 아니라 주동 내 개별호수의 내부 평면이 유사하지만 탑상형 주동은 세대 위치에 따라 평면이 다르고 창을 갖는 벽면의 수나 방위도 각각 달라진다. 즉, 주동형상은 건축물의 외관과 주동 내에 구성되는 각 내부 평면에 모두 영향을 미쳐 아파트 단지의 주요한 개발 특성중 하나라고 볼 수 있다. 또한, 실제로 아파트 주동형상별로 주택시장에서의 선호도가 다르기 때문에, 동일 단지 내에서도 주동형상 차이에 따른 각 세대의 가격 차이가 나타나기도 한다. 아파트 주동형상이 아파트의 평면과 층수, 밀도 및 가격과 관련이 높지만 주동형상에 대한 체계적인 연구는 부족하다. 아파트 주동형상의 중요성에도 불구하고 주동형상을 데이터베이스화하여 아파트단지 특성분석이나 공간분석에 활용한 연구는 제한적이었는

데, 이는 일반적으로 부동산 정보 제공 서비스 업체들이 건축물대장에서 제공되는 속성정보에 한정하여 데이터베이스를 구축하였기 때문이다. 공간정보를 활용하여 아파트 주동형상을 자동으로 판별하여 판상형, 탑상형 등으로 자동분류할 수 있다면 향후 아파트 단지설계 특성과 주택시장에서의 선호도, 주택가격 간의 영향 관계를 분석하는 기초자료로서 활용도가 매우 높을 것이다.

따라서, 본 연구에서는 지리정보를 활용하여 서울시 내 아파트 단지를 대상으로 주동형상을 자동으로 판별해주는 모형을 구축하고, 이를 분석에 적용하여 아파트 단지의 개발특성과 주동형상이 어떠한 관련이 있는지 비교해보고자 한다. 이를 통해 아파트 단지의 개발과 설계에서 주동형상이 가지는 중요성을 인식하고, 공간분석과 부동산분석을 위한 기초 자료를 제공하고자 한다.

### 선행연구 및 이론적 고찰

아파트의 배치계획과 주동형상과 관련한 연구들을 고찰한 결과, 기존 연구들은 주로 외관의 시각적 모습에 따라 주동형상을 구분하고, 이후 세부적인 분석을 진행하였다. 최무현(Choi, 1997)은 공동주택단지의 형태 구성을 주거동의 물리적 형태에 따라 5가지를 제시한 바 있다.

이를 건축년도, 용도지역별로 분류하고 각 형태별 특징을 비교하였다. 一자형과 ㄴ자형의 주거동으로 구성된 단지일수록 낮은 건폐율, 중고층 구조, 공공공급비율이 높고, ㄷ자형, ㄹ자형의 주거동이 주를 이룰수록 높은 건폐율, 중저층의 구조, 민간공급비율이 높다는 것을 확인하였다. 김영배(Kim, 2006)는 공동주택 배치기법의 변화과정을 고찰하고 주거동 형태 4가지와 배치형태 5가지로 분류하였다. 수도권 400세대 이상의 20개 단지 사례를 상세히 분석하고 시대별 형태 변화를 일본의 경우와 비교하여 분석하였다.

김형진 외(Kim *et al.*, 2010)는 선행연구 배치패턴을 재정리하여 주거동 형태 2가지, 주동 배치 10가지로 분류하였다. 이후 2007년 열린 행정중심복합도시 설계공모안을 바탕으로 당시의 아파트 단지 설계 경향을 파악하였다. 경관이 양호한 쪽으로 탑상형을 배치하고, 도로변에 판상형을 배치하는 형태가 주를 이루었음을 파악하였다. 이지은과 김지민(Lee and Kim, 2015)은 주동형상 및 배치패턴을 조사하고, 거주자 설문조사를 통해 거주민의 아파트 형태 선호도를 살펴보았다. 이를 통해 판상형 주동을 선호하고, 그 중에서도 ㄴ자 형태의 주동이 가장 높은 만족도를 주는 것을 파악하였다. 고승일과 이정재(Ko and Lee, 2022)는 Open CV를 활용하여 공동주택의 주거동 형태를 13가지

TABLE 1. Literature review related to apartment types

Author (Year)	Study area	Classification form	Analysis content
Choi (1997)	Daegu, 606 public housing complexes	—, ㄴ, ㄷ, ㄹ, Tower	Characteristics of apartment types by year of construction and area of use
Kim (2006)	Metropolitan area, 20 apartment complexes	—, ㄷ, Y, Grid	Trend of change by year and comparative analysis with Japan
Kim et al. (2010)	design proposal, 10 apartment complexes	Plate, Tower	Identify the latest apartment design trends
Lee and Kim (2015)	Daegu, 10 apartment complexes	ㄴ, Y, Plate	Resident preference survey by apartment type
Ko and Lee (2022)	More than 7 buildings, 500 households, A total of 4900 apartment complexes	ㄴ, ㄷ, T, Y, Rectangle, Obtuse-angle, Butterfly, Wing, Polygon, Staggered-squares, Circle, Curve, Special type	Main shape of apartment discrimination method and supply change by year

TABLE 2. Examples of types by main building shape of apartment

	Plate-Type		Tower-Type		Mixed-Type
—		Y		Wing	
L		□		Butterfly	
ㄷ		Polygon		Complex	

\*<https://www.google.com/maps/@/data=!3m1!1e3?hl=ko> (Google map)

로 분류하였고, 각 유형의 연도별 공급변화를 살펴보았다. 이를 통해 직사각형 형태의 판상형 아파트 공급량은 줄어들고, 탑상형 및 복합형 형태의 아파트 공급량이 늘어나고 있음을 파악하였다(표 1 참고).

주동형태는 연구자에 따라 2가지부터 13가지 까지 다양하게 분류되고 있지만, 최근에는 일반적으로 세가지 형태로 구분할 수 있다. 판상형은 복도식 아파트에서 복도를 따라 세대가 길게 분포되어 있거나, 계단식 아파트에서 코어(계단실 및 엘리베이터)를 중심으로 양측으로 세대를 배치한 유닛이 반복되어 일자형으로 배치된 형태이다. 기존 연구에서 분류한 —자 형태가 여기에 포함되며, —자가 조합된 형태(‘L’ 또는 ‘ㄷ’)를 가지는 경우도 있다. 탑상형은 코어(엘리베이터 홀)를 통해 세대로 이어지기 때문

에 중앙으로 밀집된 형태를 보이고 있고, 일반적으로 3세대 이상이 하나의 코어를 공유한다. 선형연구에서 분류한 Y자, □자, 다각형 등이 여기에 포함된다. 혼합형은 판상형과 타워형이 혼합된 형태로 볼 수 있으며 선형연구에서는 날개형, 나비형, 복합형 등으로 분류하고 있다(표 2).

아파트 주동형상이 단지의 밀도와 배치, 주거 환경, 호별 평면을 결정하는 중요한 요소임에도 불구하고, 아파트 주동형상의 분류와 그 영향에 대한 체계적인 분석 연구는 부족하다. 대부분의 연구들이 일부 아파트 단지 사례에 기초하여 주동형상을 분류하거나 설계 변천 과정에 초점을 두고 있어, 전체 주택시장에서의 특성으로 일반화하기 어렵다. 따라서 본 연구에서는 지리정보를 활용하여 주동형상 분류를 자동으로 분류하

는 모형을 개발하고, 이를 바탕으로 아파트 주동형상의 특성과 설계 요소와의 관계를 고찰함으로써 향후 공간정보 기반의 주택 연구의 기초 정보를 제공하고자 한다.

## 분석 방법

### 1. 데이터

대부분의 아파트는 건물평면의 윤곽선과 높이, 이 두 개의 비교적 단순한 정보만으로 건물의 형상을 표현하는 것이 가능하다. 그 중에서도 주동형상은 평면정보만으로도 파악할 수 있다(Yoo and Han, 2005). 따라서, 본 연구에서는 아파트단지별 주동 데이터베이스 구축을 위하여 국토지리정보원에서 제공하는 수치지형도 v2.0(1:5000)의 건축물의 평면 도형 정보를 사용하였다. 또한, 아파트 단지 단위로 주동형상 특성을 분석하기 위하여 ‘아파트 단지’ DB를 국토교통부 건축행정시스템 세움터에서 제공하는 ‘건축물대장 표제부’와 환경부에서 제공하는 ‘토지피복지도’를 활용하여 구축하였다.

분석 대상은 서울특별시 아파트이며 2019년에서 2021년 3년간 매매 거래가 1건 이상 보고된 아파트 단지를 대상으로 하였다. 또한, 1997년 IMF 경제위기 전후로 브랜드 아파트가 등장하고 아파트 단지배치계획 패턴이 크게 변화되었으므로 1997년 이후 준공된 아파트 단지를 대상으로 한정하여 분석을 수행하였다.

아파트 단지 내 개별 동의 주동형상 도형정보는 수치지형도v2.0(1:5000)상의 ‘건물’ 소분류 ‘아파트’와 더불어 재개발·재건축 등 정비사업이 진행중인 아파트 단지의 이전 주동형상 도형 정보를 포함하기 위해 ‘공사중건물’을 추가로 활용하였고, 주상복합아파트를 포함하기 위해 ‘주택외건물’로 구분되어 있는 항목들을 추출하여 사용하였다.

‘건축물대장 표제부’에는 각 아파트별 건축물 정보, 세대수, 주소 등의 정보가 포함되어 있으며, ‘토지피복지도’에는 토지 경계, 주소 등의 정보가 포함되어 있다. 「주택법」 및 「건축법」상 주택건설사업으로 조성된 아파트를 대상으로 주동형상 자동판별모형을 개발함에 따라

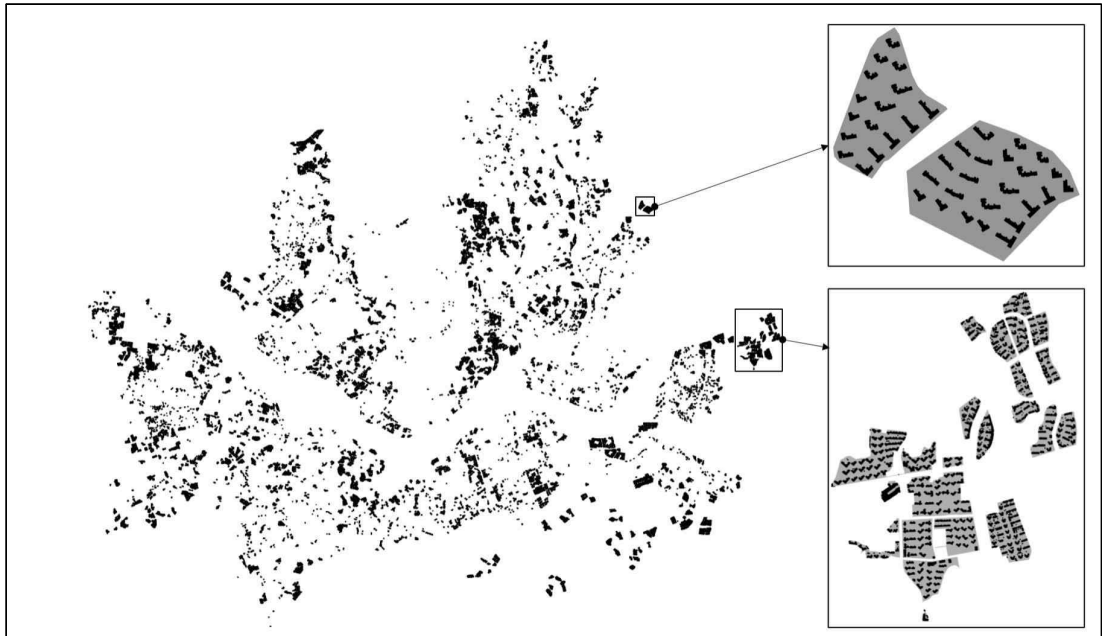


FIGURE 1. Apartment complexes and building geographic information data

‘건축물대장 표제부’ 정보를 기준으로 20세대 이상이고 최고층수가 5층 이상인 공동주택을 대상으로 아파트의 속성정보(용적률, 건폐율, 최고층수, 건축년도 등) 데이터베이스를 구축하였다. 수치지형도 상 주택의 도형 정보, 아파트 속성 정보 데이터베이스 및 ‘도지피복지도’ 의 아파트 단지 경계 정보를 매칭하여 주동형상 판별의 초 자료로 사용하였다(그림 1).


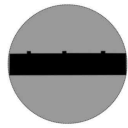

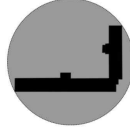











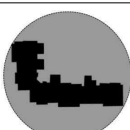

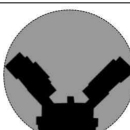
2. 주동형상 자동판별 모형

본 연구에서 주동형상은 선행연구에서 검토한 바와 같이, 판상형, 탑상형, 혼합형의 3가지로 구분한다. 판상형은 일반적으로 건축물의 폭이 길이와 비교하여 좁게 설계되어 있다. 이와 대조적으로 타워형은 그 형태가 판상형과 비교하여 다양하지만, 중앙의 엘리베이터 홀을 중심으

로 각 세대가 설계되기 때문에 판상형에 비해 폭과 길이의 차이가 적고 건축물의 중심이 명확한 편이다. 혼합형은 판상형과 탑상형의 특징이 혼합되어 있다. 본 연구에서는 각 유형의 형태적 특징들을 바탕으로 주동형상 판별모형을 개발하였다.

자동 주동형상 판별모형에서는 아파트 개별 주동에 대하여 Bounding Box와 Bounding Circle을 적용한 후 도형 형태특성 속성정보를 구축하고 머신러닝 알고리즘을 적용하여 판상형, 탑상형, 혼합형 등 주동형상을 분류한다. 도형의 경계를 잇는 Bounding Box와 Bounding Circle을 만드는 기능은 보통 사용 GIS 프로그램의 기본 분석도구로 제시된다. 하지만 이를 아파트 주동형상 분석에 활용한 선행연구는 없으며, 본 연구에서 창의적으로 적용한 부분이다. 우선, ArcGIS의 Minimum Bounding Geometry

TABLE 3. Example of Output Bounding

	Type 1		Form Type 2		Type 3		
	Box	Circle	Box	Circle	Box	Circle	
Plate	Drawing						
	RLB	0.226		0.651		0.725	
	RAB	0.988		0.264		0.343	
	RAC	0.150		0.192		0.239	
Tower	Drawing						
	RLB	0.557		0.991		0.925	
	RAB	0.827		0.520		0.639	
	RAC	0.438		0.572		0.683	
Mixed	Drawing						
	RLB	0.574		0.633		0.628	
	RAB	0.383		0.409		0.374	
	RAC	0.301		0.375		0.322	

기능을 사용하여 각 건축물 Shape의 Output Bounding을 생성한다. 이 기능은 각 Shape의 가장 작은 경계도형을 만들어주는 도구로써, 정사각형, 직사각형, 원, 다각형 등 다양한 형태로 생성할 수 있다. 이 중 직사각형과 원을 사용하여 2가지 Output Bounding을 실행하면 넓이, 직사각형의 각 변의 길이, 원의 반지름 등의 속성 정보가 생성된다. 각각의 생성된 직사각형을 Bounding Box, 원을 Bounding Circle이라고 하였다. 이후, Bounding Box에서 긴 변과 짧은 변의 길이 비율(Ratio of Lengths in Box: RLB), 건축물 Shape과 Box의 면적 비율(Ratio of Area in Box: RAB), 건축물 Shape과 Bounding Circle의 면적 비율(Ratio of Area in Circle: RAC)라고 명명하였다. 각각의 RLB, RAB, RAC 값들은 표 3과 같이 각 주동형상 별로 명확한 차이를 보인다. 각 주동 유형 별로 가장 빈번하게 나타나는 상위 3개 유형을 Type 1, 2, 3로 명명하였다(예: 판상형 Type 1, 탑상형 Type 2 등).

판상형의 경우, 모든 건축물이 Type 1, 2, 3으로 나뉘질 만큼 단조로운 형태를 지니고 있다. Type 1의 경우 RLB 값이 매우 낮고, RAB 값이 극단적으로 높는데, 이는 직사각형과 거의 일치하는 형태이기 때문이다. 판상형 Type 1은 탑상형의 Type 1과 유사하게 높은 RAB값을 가지지만 탑상형보다 복도식 또는 계단식의 연속으로 길게 배치되어 있는 판상형의 특성상 RAC 값은 낮게 형성된다. 보다 복잡한 판상형 형태인 Type 2와 3는 꺾이는 부분의 각도에 따라 RLB가 다르게 나타날 수 있지만, RAB와 RAC 값을 동시에 고려하여 타워형이나 혼합형과는 구분될 수 있다. 두 유형 모두 Type 1과 마찬가지로 건축물의 폭에 비하여 전체 길이가 길다는 판상형의 특징 때문에 RAB와 RAC 값이 다른 유형에 비하여 상대적으로 낮게 나타난다. 탑상형은 중앙 코어를 중심으로 2호, 3호, 4호 등 다양한 수의 세대를 배치할 수 있어, 배치 구조에 따라 전체 건축물의 형태가 달라진다. 하지만 중앙의 엘리베이터 홀을 통해 각 세대로 이어지므로 중앙에 집중된 형태를 보인다.

이 때문에 Type 1, 2, 3 모두 RAB와 RAC가 높게 나타나는 것이 특징이며 그 값은 판상형 Type 1 과 나머지 주동형상의 중간에 위치한다. 혼합형은 타워형을 중심으로 판상형이 혼합되어 있는 형태이기 때문에 표 1에서 예시로 든 Type 1, 2, 3 이외에도 다양한 형태로 나타나며 RLB, RAB, RAC 값도 다양하다.

주동형상의 분류를 위해 본 연구에서는 의사결정나무 학습법(decision tree learning)을 사용하였다. 머신러닝 기반의 예측 모델 중 일반적으로 많이 사용되는 의사결정나무 학습법은 복수의 입력 변수를 이용해 목표 변수값을 예측하는 모델을 생성한다. 의사결정나무의 나무 구조에서, 각 내부 노드들은 하나의 입력 변수에, 자녀 노드들로 이어지는 가지들은 입력 변수의 가능한 값에 대응된다. 종단 노드는 각 입력 변수들이 시작 노드로부터 종단 노드로 이어지는 경로에 해당되는 값들을 가질때의 목표 변수 값에 해당된다.

일반적으로 의사결정나무를 이용해 학습하는 경우, 학습데이터 또는 데이터 품질에 따라 예측성능의 변동 폭이 크다는 단점을 가지고 있다. 특히 학습데이터에 따라 생성되는 결정 나무가 무작위성에 의해 달라지기 때문에 결정 나무를 일반화하여 사용하기에 어려움이 따른다. 이로 인해 의사결정나무는 학습데이터에서는 높은 일치율을 보일 수 있지만, 데이터가 변형된 평가데이터에서는 일치율이 크게 감소되는 과적합(overfitting) 현상이 쉽게 발생한다. 본 연구에서는 의사결정나무의 과적합 단점을 극복하고 더 나은 일반화 성능을 갖기 위해 랜덤포레스트(random forest) 기법을 적용하였다. 의사결정나무를 이용하는 기술 중 하나인 랜덤포레스트는 입력된 데이터를 임의의 수만큼 추출하여 의사결정나무 형태의 예측기를 만든 후, 생성된 모형들을 앙상블(ensemble) 기법을 통해 결합하여 최종모형을 만드는 방법이다(Heo *et al.*, 2018). 랜덤포레스트는 랜덤성에 의해 서로 조금씩 다른 특성을 가지는 결정나무들을 여러 개를 만들고 평균화된 결과값을 출력하는 기능을 수행한다. 이를 통해 하나의 의사결정나무를 이

TABLE 4. Random forest parameters

Name	Value
n_estimators	50
max_features	7
min_samples_leaf	1
min_samples_split	2

용했을 때보다 과적합 현상을 감소시키는 효과를 얻을 수 있다.

본 연구에 사용된 데이터는 서울시 내 아파트 단지 내 건축물 15,471개를 대상으로 한다. 이 중, 절반에 해당하는 7,736개 건축물을 무작위로 선정하여 직접 주동형상을 분류하여 학습값(관상형, 탑상형, 혼합형)을 생성하였다. 이 중 절반인 3,868개를 모형 구축을 위한 학습데이터로, 나머지 3,868개는 모형의 성능을 평가하기 위한 데이터로 사용한다. 이를 통해, 성능이 검증된 주동형상 판별 모형을 나머지 학습되지 않은 데이터에 적용하여 전체 건축물에 대한 주동형상 분류를 수행하였다.

개별 건축물 도형의 주동형상 분류 결과를 아파트 단지 단위로 집계하여 단지별로 각 주동형상 유형의 비중 정보를 구축하였다. 또한, 아파트 단지의 주동형상 특성과 건축년도, 밀도, 최고층수 등 개발 특성간의 교차분석을 수행하여 주동형상과 아파트 단지 특성 간의 관계를 고찰한다.

## 분석 결과

분석 결과는 1) 주동형상 자동판별 모형의 성능, 2) 아파트단지 단위 주동형상 특성 분석, 3) 주동형상과 아파트단지 개발특성간의 관계 분석으로 구분하여 제시한다.

### 1. 주동형상 자동판별 모형의 성능

본 연구에서는 다수의 의사결정나무를 이용해 정확도와 안정성이 높은 랜덤포레스트를 이용해 주동형상 판별을 위한 머신러닝 모델을 개발하였다. 랜덤포레스트의 예측 성능에 가장 큰 영향을 미치는 의사결정나무 수를 선정하기 위해

해당 매개변수를 변화시켜가며 예측 정확도를 관찰하였다. 정확도는 식 1로 정의된 결정계수를 이용해 평가하였다.

$$R^2 = \frac{1 - \sum (y_i - \tilde{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (1)$$

결정계수는 추정된 모델이 주어진 자료에 적합한 정도를 재는 통계적 척도이다. 결정계수의 범위는 0에서 1까지이며 모형의 예측 정확도가 높을수록 결정계수는 1에 가까워진다.

분석 결과, 의사결정나무의 수가 50개 이상일 경우 평가데이터에서 정확도 향상이 거의 없는 것으로 나타났기 때문에, 분류기 수(n\_estimators)는 50개로 선정해 모형을 개발하였다. 최대로 선택할 특성의 수(max\_features)는 일반적으로 특성 개수의 제곱근을 사용하므로 7개로 설정하였다. 최소 샘플 가지 수(min\_samples\_leaf)는 노드를 분할하기 위해 필요한 최소한의 샘플 데이터 수를 지정하는 것을 의미한다. 이 값을 작게 설정하면 일반화 능력은 높아지지만, 훈련 데이터에 대해 과도하게 적합될 가능성이 높아지며, 새로운 데이터에 대한 성능이 떨어지는 과적합 모델을 생성할 수 있으므로 평가 데이터를 통해 얻어진 정확도와 비교하여 설정할 필요가 있다. 표 4에서와 같이 학습 데이터와 평가 데이터에서 모형의 정확도 차이가 미미하므로 일반화 능력을 높이기 위해 1로 설정하였다. 최소 샘플 분할 수(min\_samples\_split)는 노드를 분할하기 위해 필요한 최소한의 샘플 데이터 수를 의미하며 이 값보다 적은 샘플이 있는 노드는 더 이상 분할하지 않기 때문에 일반적으로 2로 설정하고, 모형의 복잡도를 줄이기 위해 10까지 설정하는 경우도 있다. 따



TABLE 5. Comparison of model performance (Accuracy)

	Training Data	Test Data
Decision tree	0.9982	0.9453
XGBoost	0.9959	0.9873
LightGBM	0.9964	0.9878
Random forest	0.9971	0.9878

라서 최소 샘플 분할 수는 일반적 기준인 2로 설정하였다.

랜덤포레스트의 학습데이터와 테스트데이터의 비율은 전체 데이터의 개수를 보고 판단한다. 학습데이터가 너무 적으면 알고리즘이 효과적으로 학습하기에 충분치 않을 수 있다. 반면, 검증 데이터가 너무 적으면 모형의 예측/분류 성능의 정확도의 신뢰도가 낮아질 수 있다. 학습데이터와 테스트데이터는 일반적으로 6:4의 비율로 나누며, 전체 데이터의 개수가 적을수록 7:3, 8:2 등 학습 데이터의 수를 늘린다. 본 연구에서는 전체 데이터가 7,736개로 충분히 많은 숫자라 판단되고, 실제로 7:3부터 5:5의 비율로 학습데이터와 테스트데이터를 나눌 때 정확도 차이가 거의 발생하지 않았다. 따라서, 본 연구에서는 주동형상 자동판별 모형의 신뢰도를 보여주기 위해 절반인 3,868개를 모형 구축을 위한 학습데이터로, 나머지 3,868개는 모형의 성능을 평가하기 위한 데이터로 사용하고 그 결과를 제시하였다.

본 연구에서 개발한 주동형상 자동판별 모형은 랜덤포레스트 알고리즘을 적용한 결과이다. 하지만, Bagging 방법에 기초한 랜덤포레스트 외에도 최근에는 Boosting 방법에 기초한 다양한 머신러닝 알고리즘이 적용되고 있다. 본 연구에서는 의사결정나무 대비 랜덤포레스트를 이용한 정확도 개선 및 과적합 문제 해결을 확인하기 위해 표 5와 같이 학습데이터와 평가데이

터에 대한 결정계수 값을 측정하였고, Boosting 방법에 기초한 대표적인 머신러닝 알고리즘인 XGBoost 및 LightGBM 알고리즘 적용 결과와 모델 성능을 비교하였다. 의사결정나무를 사용할 경우 학습데이터에서는 일치율이 높은 것으로 나타나지만, 데이터가 변형된 평가데이터에서는 일치율이 감소하는 과적합 현상이 나타났다. 반면 랜덤포레스트를 사용할 경우 평가데이터에서도 높은 일치율이 나타나 과적합 현상이 줄어든 것을 알 수 있다. XGBoost와 LightGBM는 랜덤 포레스트와 비교하여 같거나 소폭 낮은 정확도를 보인다. 최종적으로 가장 학습 속도가 빠르고 사용이 간편한 랜덤포레스트를 이용하였으며, 평가데이터에 대한 결정계수는 0.9878로 나타나 높은 예측력을 보이고 있다.

## 2. 아파트단지 단위 주동형상 특성 분석

학습된 주동형상과 모형으로 추정된 주동형상 유형을 종합하여 서울시 내 아파트 건축물 15,471개의 주동형상을 구분하였다. 파악된 주동형상은 아파트 단지별로 집계해서 가장 많은 형태의 주동형상을 아파트단지의 대표 주동형상으로 구분하였다. 대표 주동형상이 2개인 경우는 해당 유형을 각각 0.5개, 3가지 유형이 모두 같은 비율일 경우 각각 0.33개로 계산하였다. 분석에 사용된 아파트 단지의 개수는 3,734개이다.

‘개별 건축물 단위’에서는 탑상형이 9,205

TABLE 6. Main building shape analysis result

	Plate-Type Number (%)	Tower-Type Number (%)	Mixed-Type Number (%)	Total Number (%)
Building	4,928 (31.85)	9,205 (59.5)	1,338 (8.65)	15,471 (100)
Apartment Complex	871 (23.33)	2,674 (71.61)	189 (5.06)	3,734 (100)

개로 전체의 약 60%를 차지하고, 판상형이 4,928개로 31.85%, 혼합형이 1,338개 8.65%로 나타났다. ‘아파트 단지의 대표 주동형상’ 별 집계 결과에서는 탑상형이 2,674개 단지로 71.61%, 판상형이 871개 단지 23.33%, 혼합형이 189개 단지 5.06%로 탑상형의 비중이 크게 늘어난 것을 알 수 있다. 이는 분석에 사용된 아파트 단지가 소규모 단지인 경우가 많기 때문으로 판단된다. 300세대 이하 소규모 아파트 단지는 총 2,521개로 전체 아파트 단지의 67.5%를 차지하고, 대부분 3개 동 이하로 구성되어 있다. 이러한 소규모 아파트 단지는 탑상형인 경우가 많기 때문에 ‘아파트 단지의 대표 주동형상’ 결과에서는 탑상형의 비율이 높게 나타난다(표 6).

### 3. 주동형상과 아파트단지 개발특성과의 관계 분석

앞서 도출한 주동형상은 20세대 이상 주택건설단지를 대상으로 함에 따라 소규모 아파트나 주상복합을 모두 포함한다. 소규모 주택건설단지의 경우, 소규모 필지에서 개발된 경우가 많고 주민편의시설 등을 갖추지 못한 경우가 많아 일반적으로 대중에게 인식되는 아파트단지와 다르다. 따라서 주동형상과 아파트단지 개발특성과의 관계 분석은 소규모 주택건설단지를 제외

하고 300세대 이상 아파트, 100세대 이상 주상복합과 도시형생활주택을 대상으로 한정하여 분석을 수행하였다. 분석에 사용된 아파트 단지의 개수는 1,231개이고, 건물의 총 개수는 7,233개이다.

분석결과 탑상형이 주를 이루는 단지는 582개, 전체의 47.24%로 가장 많은 비중을 차지했으나 소규모 아파트 단지까지 포함했던 이전 결과보다는 비율이 상당부분 줄어들었음을 알 수 있다. 이와 반대로 판상형과 혼합형이 주를 이루는 단지는 각각 468개, 181개로 전체의 38.02%, 14.70%를 차지해 모든 아파트를 대상으로 한 결과에 비해 소폭 증가하였다(표 7).

이들의 주동형상과 단지 특성(용적률, 건폐율, 최고층수, 건축년도)과의 관계를 분석하였다. 각 주동형상별로 단지 특성과 교차분석을 진행하였으며, 분석을 위해 각 단지 특성 변수를 특정 구간으로 나누어 해당 구간에서 주동형상이 차지하는 비율을 대상으로 하였다.

건축년도 별 주동형상 분포를 살펴보면, 2002년 이전 건축된 아파트 단지는 판상형의 비중이 57.77%로 가장 높았다. 이후 2002년부터 2016년까지는 판상형이 감소하고 탑상형의 비율이 계속 늘었으며 혼합형은 초기엔 감소했다가 다시 늘어나는 추세를 보이고 있다. 2016년 이후 건축된 아파트 단지는 혼합형의 비율이 크게 증가해 32.26%를 차지하고 있으며, 상대

TABLE 7. Main building shape of large-scale apartment analysis result

	Plate-Type Number (%)	Tower-Type Number (%)	Mixed-Type Number (%)	Total Number (%)
Building	2,562 (35.42)	3,486 (48.19)	1,185 (16.39)	7,233 (100)
Apartment Complex	468 (38.02)	582 (47.28)	181 (14.70)	1,231 (100)

TABLE 8. Result of main building shape distribution by construction year

	Year of construction				
	Before 2002	2002~2006	2007~2011	2012~2016	After 2016
Plate (%)	57.77	40.90	27.85	17.62	20.37
Tower (%)	30.02	50.86	56.74	60.63	47.36
Mixed (%)	12.21	8.24	15.41	21.76	32.26
Number of apartment complex	293	311	236	211	180

TABLE 9. Main building shape distribution result by floor area ratio

	Floor area ratio					
	<200%	200~250%	250~300%	300~350%	350~400%	>400%
Plate (%)	41.87	34.65	42.19	62.31	43.22	15.34
Tower (%)	37.91	38.67	37.71	26.87	44.96	79.11
Mixed (%)	20.21	26.69	20.10	10.83	11.82	5.55
Number of apartment complex	122	290	286	110	109	314

TABLE 10. Main building shape distribution result by building coverage ratio

	Building coverage ratio				
	<20%	20~30%	30~40%	40~50%	>50%
Plate (%)	39.33	44.67	32.31	18.53	15.74
Tower (%)	42.57	33.54	53.42	76.56	78.09
Mixed (%)	18.10	21.79	14.27	4.91	6.17
Number of apartment complex	307	543	65	72	244

적으로 탑상형의 비율은 감소하였다(표 8).

용적률 별 주동형상의 분포를 살펴보면, 용적률 300% 이하의 아파트 단지에서는 두드러진 변화가 보이지 않는다. 탑상형은 세 구간 모두 일정한 비율로 유지되고, 판상형과 혼합형은 200~250% 구간만 소폭 변동이 있다. 두드러진 변화는 용적률 300% 이상에서부터 나타난다. 용적률 300~350% 구간에서는 판상형의 비중이 20% 이상 급격하게 상승하고, 탑상형과 혼합형의 비율이 각각 약 10%씩 감소한다. 이후 350% 이상 구간에서는 탑상형의 비중이 계속 상승하여 350~400% 구간에서는 44.96%, 400%이상 구간에서는 79.11%를 차지하고 있다. 이와 반대로 판상형의 비중은 계속 감소하여 각각 43.22%, 15.34%의 비율을 보인다(표 9).

건폐율과 주동형상 관계에서는 건폐율 20~30% 구간을 기점으로 큰 변화를 보인다. 판상형과 혼합형은 건폐율 20~30% 구간까지 상승하여 각각 44.67%, 21.79%이고, 탑상형은 반대로 감소하여 33.54%의 비율을 보인다. 이후로는 판상형과 혼합형은 감소하고, 탑상형은 계속 증가하는 현상이 나타난다(표 10).

최고층수 별 주동형상 분포를 살펴보면, 10층 이하에서는 판상형의 비중이 50.62%로 가장 높고, 탑상형이 43.18%, 혼합형이 6.2%의 분포를 보인다. 이후 11~30층의 경우 87%의 아파트 단지가 이 구간에 속해있으며, 주동형상 분포도 전체 아파트 단지과 흡사한 비율을 보인다. 이후로 층수가 높아질수록 판상형은 급격히 줄어들고 대부분이 탑상형으로 구성되는 현상이 나타난다. 41층 이상의 아파트 단지에서는 판상

TABLE 11. Main building shape distribution result by highest floors

	Highest floors				
	~10th	11th~20th	21st~30th	31st~40th	41st~
Plate (%)	50.62	36.96	38.65	17.68	0
Tower (%)	43.18	47.08	44.20	60.46	95.56
Mixed (%)	6.20	15.96	17.15	21.86	4.44
Number of apartment complex	30	600	468	103	30

형은 없으며 95.56%의 아파트 단지가 탑상형으로 구성되어 있다(표 11).

종합적으로 살펴보면 2002년까지 우리나라 아파트의 대표적인 형태는 판상형이라고 볼 수 있으며, 용적률 300~350%, 건폐율 20~30%, 최고층수 10층 이하의 저층 고밀 아파트 단지의 주된 유형이 판상형이다. 이후 2003년 일반 주거지역 중 세분화가 시행되면서 각 용도지역별로 용적률과 층수를 달리 적용하면서 탑상형의 비중이 늘어났다. 특히 고층·고밀의 주상복합이 주로 탑상형으로 지어졌다는 것을 알 수 있다. 2006년 이후로는 혼합형의 비중이 계속 늘어나는데, 용적률 200~300%의 제3종 일반 주거지역에 비교적 최근에 지어진 아파트 일수록 혼합형의 비중이 늘어난다.

## 결 론

본 연구에서는 서울특별시 내 아파트 단지의 주동형상을 자동으로 판별하는 모형을 개발해 실제 분석에 적용해보았다. 각각의 건축물을 판상형, 탑상형, 혼합형 3가지로 구분하는 모형을 GIS와 랜덤포레스트 머신러닝을 기반으로 개발하였으며, 해당 모형은 높은 정확도를 보였다. 주동형상 자동판별모형을 서울시 아파트단지에 적용하여 아파트 개발 특성과 비교 분석한 결과, 개발시기, 용적률, 건폐율, 최고층수 등에 따라 대표 주동형상의 비중이 변하는 점을 확인할 수 있었다. 본 연구에서는 아파트 주동형상 자동판별 모형 적용 결과를 바탕으로 탐색적으로 주동형상과 단지특성 간의 관계를 제시하였다. 후속연구에서는 도시형태 및 도시설계 이론에 기초하여 아파트 단지설계 특성을 규명하는 연구로 확대될 필요가 있다.

본 연구는 GIS와 머신러닝을 활용하여 건축물 도형정보를 유형화한 기초연구로서 의의가 있다. 본 연구에서 개발한 주동형상 자동판별 모형은 다른 GIS로 도형정보가 제공된다면 아파트 뿐 아니라 다양한 건축물에도 적용이 될 수 있다. 건축물 도형의 형상과 형태는 도시공간을 구성하는 핵심요소가 된다. 따라서, 다양한

용도의 건축물에 대한 유형 분류 연구가 진행되어 관련 정보 제공이 가능하다면 다양한 도시형태와 도시공간분석 연구로 확장될 수 있을 것이다. 또한, 본 연구에서 개발한 모형에서 생산된 주동형상 정보는 아파트 단지에 대한 다양한 속성정보를 종합적으로 제공하는 부동산 정보 서비스와 연계하여 단지배치계획의 핵심정보로서 활용성이 높다. 본 연구에서는 주동형상을 대표적인 3가지 유형으로 한정하여 분석하였으나, 필요에 따라 각 유형을 세분화하는 모형 개발을 후속 연구로 진행할 수 있겠다. **KAGIS**

## REFERENCES

- Choi, M.H. 1997. A study on the characteristics of the physical form and the spatial layout of multi-family housing in Taegu., *Journal of the Architectural Institute of Korea Structure & Construction* 13(2): 117-124 (최무현, 1998. 공동주택단지 주거동의 형태 및 배치형태별 계획특성의 연구. 대한건축학회논문집 13(2):117-124).
- Heo, S.Y., J.Y. Kim, and T.H. Moon, 2018. Predicting crime risky area using machine learning. *Journal of the Korean Association of Geographic Information Studies* 21(4):64-80 (허선영, 김주영, 문태현, 2018. 머신러닝기반 범죄발생 위험지역 예측. 한국지리정보학회지 21(4):64-80).
- Kim, H.J., Y.S. Kim, S.J. Kim, J.G. Jo, and C.K. Park, 2010. An analysis on the layout patterns and their trends in the apartment housing design competitions of Sejong city. *Journal of the Korea Academia-Industrial cooperation Society* 11(4): 1381-1390 (김형진, 김영석, 김상진, 조중근, 박찬규, 2010. 행정중심도시 공동주택 설계공모안의 배치 계획 및 경향 분석. 한국산학기술학회논문집 11(4):1381-1390).
- Kim, Y.B. 2005. An architectural study on

- changes of the apartment housing layout techniques : focused on the cases capital area during 1970~2000. Thesis, Univ. of Hanyang. Seoul. Korea. 85pp (김영배, 2005. 공동주택배치기법의 변천에 관한 계획적 연구. 한양대학교 대학원 석사학위논문. 85쪽).
- Ko, S.I. and J.J. Lee. 2022. A study of the building types and arrangements of apartment building complex using image analysis. Journal of the Architectural Institute of Korea Structure & Construction 42(2):340-343 (고승일, 이정재. 2022. 이미지 분석을 이용한 국내 공동주택의 주동형태 및 배치 분석연구. 대한건축학회논문집 42(2):340-343).
- Lee, J.E. and J.M. Kim. 2015. A study on the preference of residents about apartment building space and arrangement. Journal of the Architectural Institute of Korea Structure & Construction 31(7):39-47 (이지은, 김지민. 2015. 아파트 주동형태 및 단지계획에 대한 거주자 선호도 분석. 대한건축학회논문집 31(7):39-47)
- Lee, J.E. and G.U. Lee. 2014. Comparison of daylighting and dnergy performance according to apartment floor shapes. Journal of the Architectural Institute of Korea Structure & Construction 30(3):113-120 (이지은, 이강엽. 2014. 아파트의 평면 형태에 따른 채광 및 에너지성능 비교. 대한건축학회논문집 30(3):113-120).
- Population Census Division, Census Management Bureau. 2022, The results of the 2022 and 2021 population and housing census. Statistical Office (조사관리국 인구총조사과, 2022, 2021년 인구주택총조사 결과. 통계청)
- Son, S.G. and G.H. Kang. 2001. A study on the change in the characteristic of outdoor planning of Korean public apartment project. Journal of the Korean Housing Association 12(4):1-8 (손세관, 강경호. 2001. 우리나라 공영 아파트단지의 외부공간의 특성 변천에 관한 연구. 한국주거학회논문집 12(4):1-8).
- Yoo, B.H. and S.H. Han. 2005. Image-based modeling of urban buildings using aerial photographs and digital maps. Journal of the Korean Association of Geographic Information Studies 8(1):49-62 (유병현, 한순홍. 2005. 항공사진과 수치 지도를 이용한 도시 건물의 이미지 기반 모델링. 한국지리정보학회지 8(1):49-62). [KAGIS](#)