심층 영상 군집화

Deep Image Clustering: Algorithms, Applications, and Emerging Trends

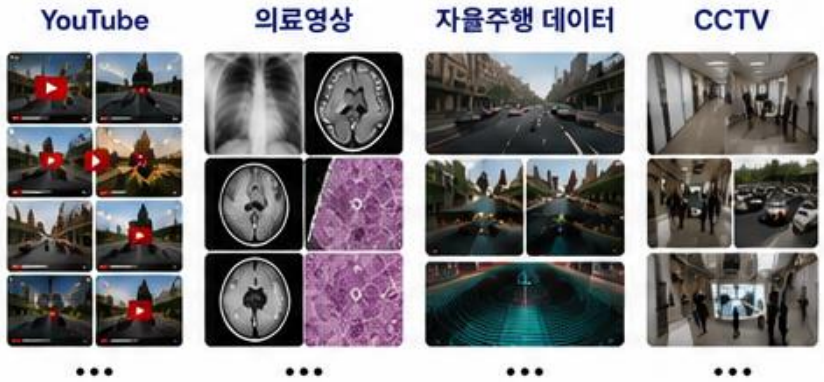
"인공지능이 라벨 없이 이미지를 스스로 분류할 수 있을까?"

왜 이미지 자동 분류가 필요한가?

Why Is Automatic Image Classification Needed?

- **다양한 분야에서 이미지 데이터가 폭발적으로 증가하면서 분석해야 할 영상의 규모가 빠르게 확대**
  Image data are growing explosively across diverse domains, rapidly increasing the scale of visual data to be analyzed.
- **대규모 이미지에 정답(label)을 부여하는 수작업 라벨링은 높은 시간과 비용을 요구**
  Manual annotation of large-scale image data requires substantial time and cost.
- **YouTube · 의료영상 · 자율주행 · CCTV처럼 지속적으로 생성되는 데이터를 사람이 일일이 분류하기에는 현실적인 한계**
  Human experts cannot manually classify the continuously growing images from YouTube, medical imaging, autonomous driving, and CCTV at scale.

❶ 다양한 분야에서 이미지 데이터 폭증
Explosive growth of image data



지속적으로 생성되는 방대한 이미지 데이터
Massive and continuously generated image data

❷ 사람의 수작업 라벨링
Manual labeling by humans

수많은 이미지를 일일이 확인하고 분류
Time-consuming and labor-intensive process

❸ 라벨이 붙은 데이터는 극히 일부
Only a small portion is labeled

라벨 부족으로 인한 활용의 한계
Severe shortage of labels limits utilization

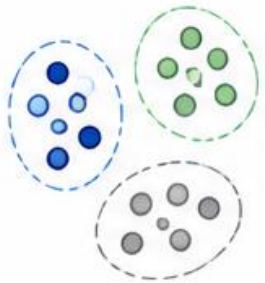
발표 개요

Overview · Presentation Roadmap

01

군집화란?

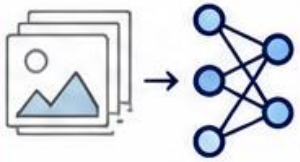
Clustering



02

심층 영상 군집화 개념

Deep Image Clustering



03

대표 알고리즘

Representative Algorithms

- K-means
- Hierarchical Clustering
- DBSCAN
- ...

04

실제 응용 사례

Real-world Applications

05

최신 연구 동향

Recent Research Trends

06

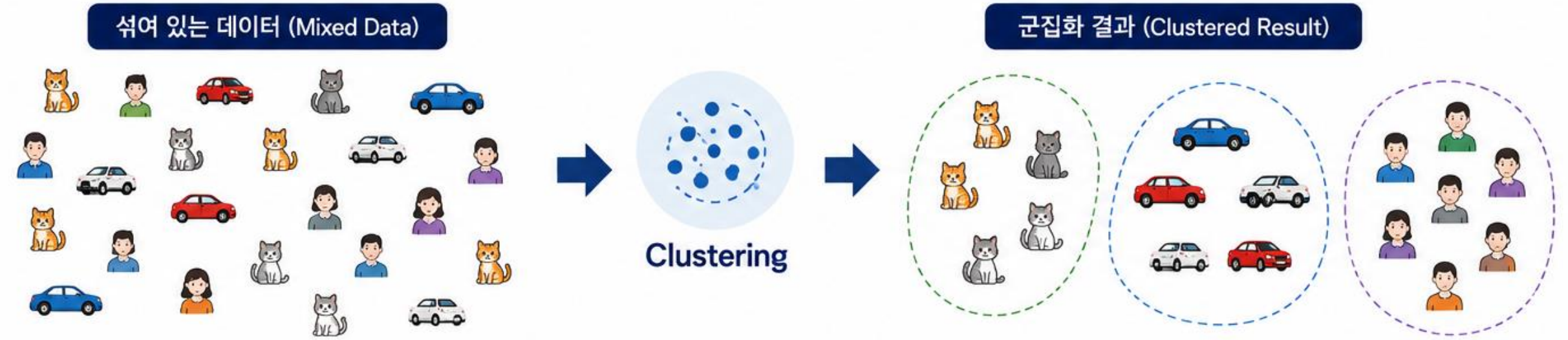
미래 과제

Future Challenges

군집화(Clustering)란 무엇인가?

What is Clustering?

- **유사한 데이터를 자동으로 그룹화하여 데이터 내부의 구조를 발견**
  Clustering automatically groups similar data to discover underlying structure.
- **고양이·자동차·사람처럼 서로 유사한 특성을 가진 데이터가 각각 하나의 군집을 형성**
  Data with similar characteristics, such as cats, cars, and people, form separate clusters.
- **사전에 정답(label)을 제공하지 않아도 데이터 간 유사성을 기반으로 그룹을 구성**
  Clusters are formed from data similarity without predefined labels.



전통적인 군집화 방법

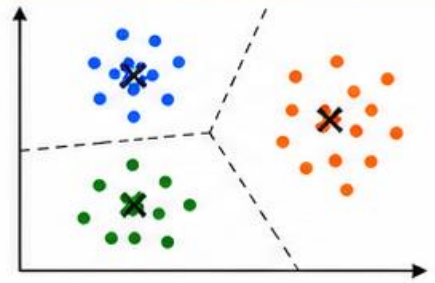
Related Work · Classical Clustering

- **대표 전통 군집화 기법: K-means, Hierarchical Clustering, DBSCAN**
  Representative classical clustering methods: K-means, hierarchical clustering, and DBSCAN.
- **기법별 원리: 중심점 반복 할당 / 병합 기반 계층 형성 / 밀도 기반 군집 및 이상점 분리**
  Operating principles: centroid-based iterative assignment / merging-based hierarchy / density-based clustering with outlier separation
- **이미지 데이터 한계: 고차원성, 의미 부족, 비선형 구조, 특징 학습 부재로 복잡 시각 패턴 분리 한계**
  Limitations for image data: high dimensionality, lack of semantics, nonlinear structure, and no feature learning, leading to limited separation of complex visual patterns

대표 전통 기법과 이미지 데이터에서의 한계

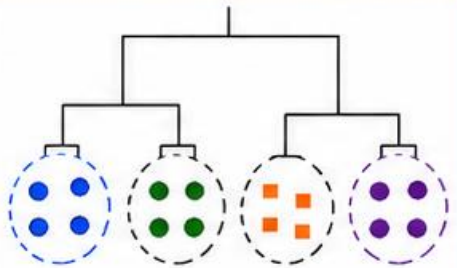
Representative classical methods and limitations for image data

K-means



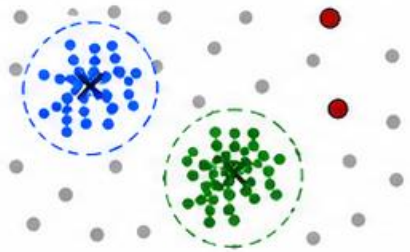
중심점 반복 할당
Centroid-based iterative assignment

Hierarchical



병합 기반 계층 형성
Merging-based hierarchy

DBSCAN



밀도 기반 군집 및 이상점 분리
Density-based clustering
with outlier separation

이미지 데이터에서의 한계

Limitations for image data

- **고차원성**
  High dimensionality
- **의미 부족**
  Lack of semantics
- **비선형 구조**
  Nonlinear structure
- **특징 학습 부재**
  No feature learning

이미지는 숫자 배열이다

Introduction · Images as Numerical Arrays

- **컴퓨터는 이미지를 직접 이해하지 못하고 각 픽셀의 숫자 값 배열로 입력받는다**
  A computer does not understand an image directly; it receives an array of pixel values as input.
- **이미지 한 장은 수천~수만 개의 숫자로 표현되어 고차원 공간에 놓이게 된다**
  A single image is represented by thousands to tens of thousands of numbers, placing it in a high-dimensional space.
- **원시 픽셀은 의미 정보를 거의 담지 못하므로 군집화를 위해서는 특징 표현 학습이 필요하다**
  Raw pixels contain little semantic meaning, so feature representation learning is needed for clustering.

이미지가 숫자로 표현되는 과정

From visual content to numerical representation

1. 사람이 보는 이미지
Visual meaning

고양이
Cat

2. 컴퓨터가 받는 입력
Pixel matrix

픽셀 행렬 (H × W × 3)
Pixel matrix (H × W × 3)

3. 숫자 벡터
Flattened values

R	34	29	31	28	37	···
G	118	112	109	121	116	···
B	207	199	191	204	210	···

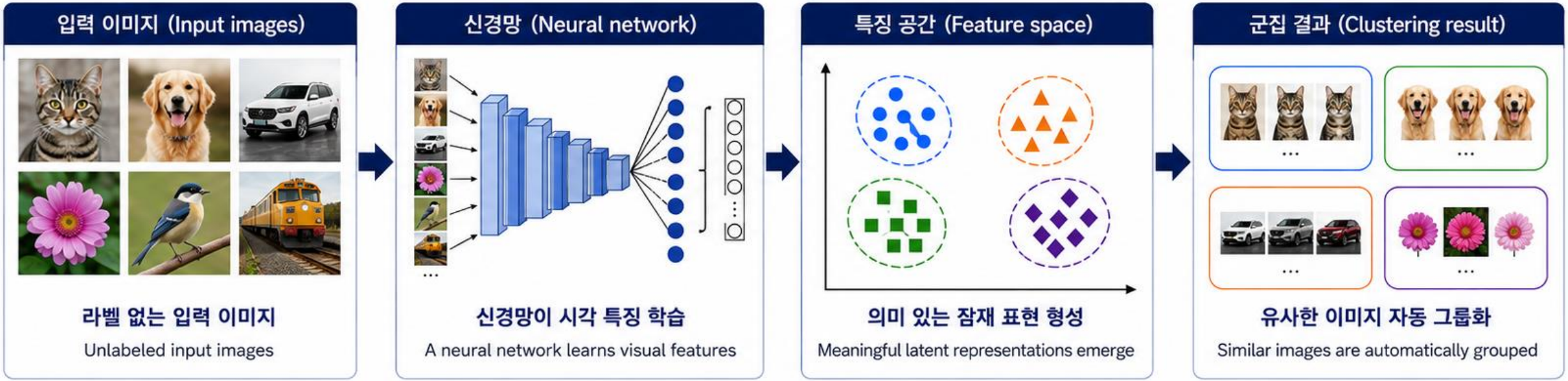
작은 RGB 이미지 예시 · 100 × 100 × 3
Example RGB image · 100 × 100 × 3

30,000개 숫자
30,000 numbers

심층 영상 군집화란?

Introduction · What Is Deep Image Clustering?

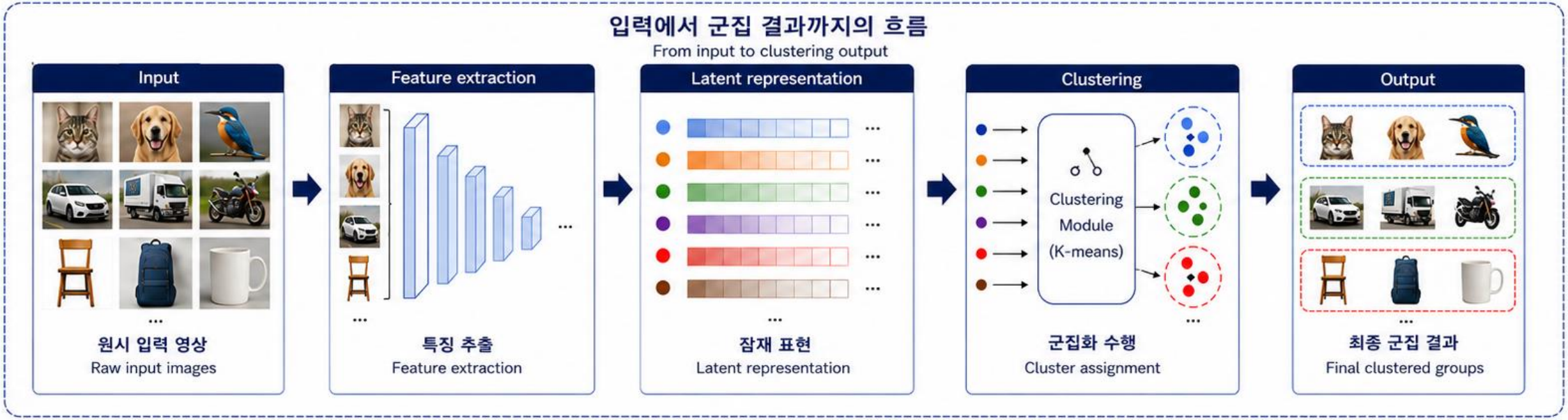
- **심층 영상 군집화는 신경망으로 유용한 시각 표현을 학습한 뒤 유사한 이미지를 자동으로 그룹화하는 방법**
  Deep image clustering learns useful visual representations with neural networks and then automatically groups similar images.
- **원시 픽셀 대신 특징 공간을 학습하여 클래스 구조를 더 잘 반영하는 잠재 표현을 형성**
  Instead of raw pixels, it learns a feature space that forms latent representations reflecting classss structure more effectively.
- **특징 학습과 군집화를 결합함으로써 전통 기법보다 더 의미 있는 그룹 형성이 가능**
  By combining representation learning and clustering, it produces more meaningful groups than classical methods.



심층 영상 군집화 기본 구조

Method Overview · Basic Pipeline

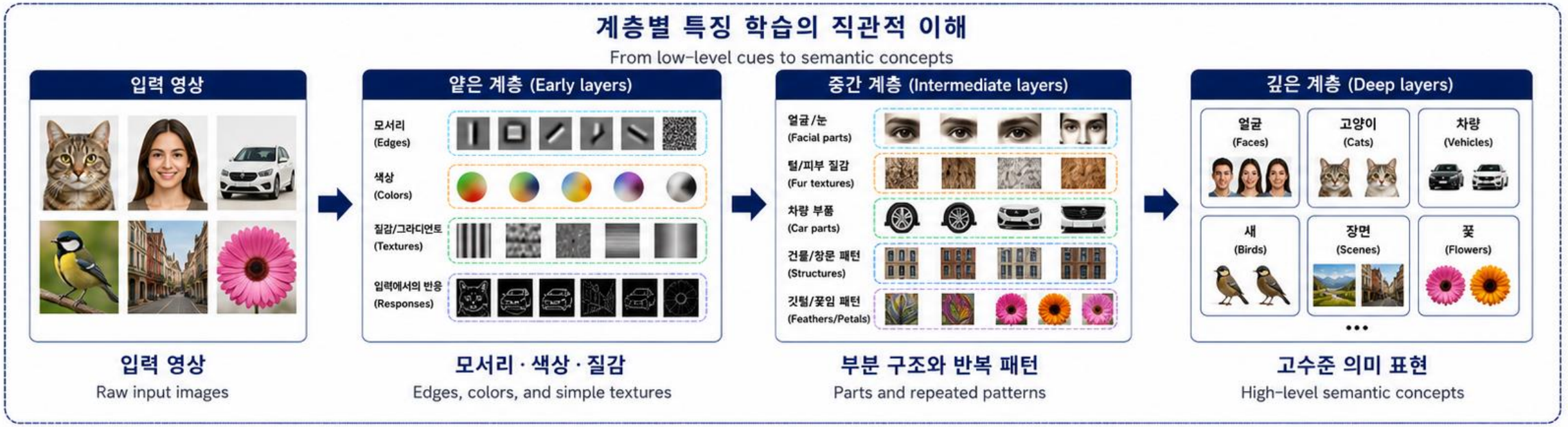
- **입력 영상은 인코더를 통해 군집화에 적합한 특징 벡터로 변환**
  Input images are transformed into feature vectors through an encoder suitable for clustering.
- **잠재 표현은 고차원 원시 픽셀보다 더 압축되고 의미적인 정보 구조를 담음**
  Latent representations contain more compact and semantic structure than high-dimensional raw pixels.
- **학습된 특징 공간에서 군집화 모듈이 유사 샘플을 자동으로 묶어 최종 그룹을 형성**
  In the learned feature space, a clustering module automatically groups similar samples to form final clusters.



신경망은 무엇을 배우는가?

Method Overview · Hierarchical Feature Learning

- **신경망은 얕은 계층에서 모서리 · 색상 · 간단한 질감과 같은 저수준 특징을 먼저 학습**
  Early layers learn low-level cues such as edges, colors, and simple textures.
- **중간 계층에서는 반복 패턴과 부분 구조를 조합해 텍스처 · 부품 수준의 표현을 형성**
  Intermediate layers combine repeated patterns into texture-level and part-level representations.
- **깊은 계층으로 갈수록 얼굴 · 차량 · 장면과 같은 고수준 의미 개념을 구분하는 표현을 학습**
  Deeper layers capture high-level semantic concepts such as faces, vehicles, and scenes.



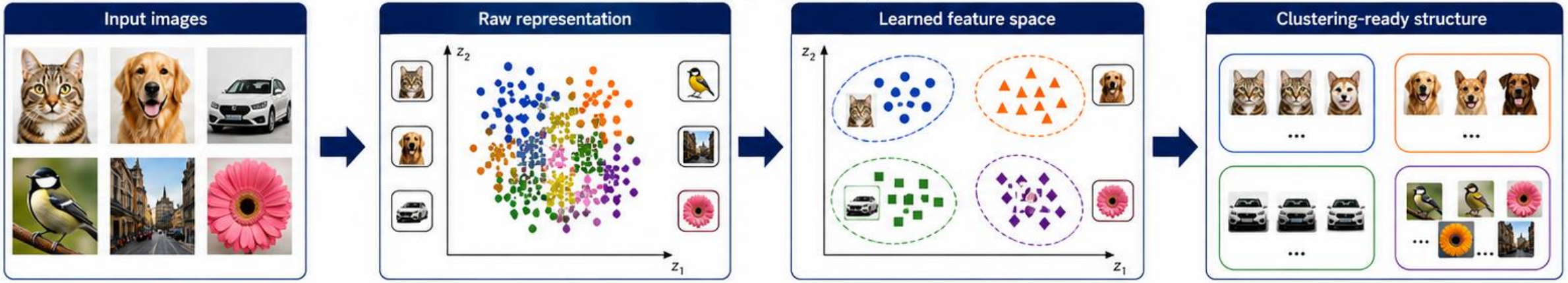
특징 공간은 어떻게 형성되는가?

Method Overview · Feature Space Formation

- **특징 공간은 입력 영상을 군집화에 유리한 잠재 벡터로 변환해 데이터의 의미 구조를 표현하는 공간**
  A feature space represents the semantic structure of data by transforming input images into latent vectors suitable for clustering.
- **좋은 특징 공간에서는 유사한 이미지는 가깝게 모이고 서로 다른 이미지는 멀어져 클래스 구조가 더 뚜렷해짐**
  In a good feature space, similar images move closer together while dissimilar images move farther apart, making class structure more distinct.
- **표현 학습이 진행될수록 군집 간 분리가 향상되어 이후 군집화 단계가 더 안정적이고 해석 가능해짐**
  As representation learning progresses, separation between groups improves, making downstream clustering more stable and interpretable.

특징 공간의 직관적 이해

From mixed representations to separable clusters



라벨 없는 입력 영상
Unlabeled input images

원시 표현에서는 클래스가 뒤섞임
Classes are entangled in raw representation

학습된 특징 공간에서 유사 샘플이 가까워짐
Similar samples become closer in the learned feature space

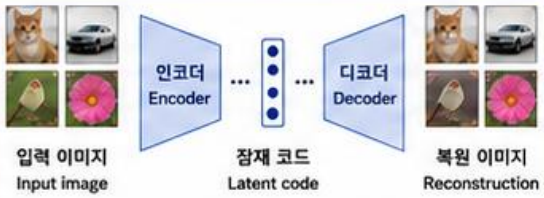
군집화에 적합한 구조 형성
A cluster-friendly structure is formed

심층 영상 군집화 방법 분류

Method Overview · Taxonomy of Deep Image Clustering Methods

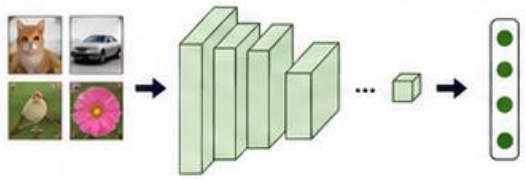
- **심층 영상 군집화 방법은 특징을 학습하는 방식과 군집화를 결합하는 전략에 따라 여러 계열로 구분될 수 있음**
  Deep image clustering methods can be categorized according to how they learn representations and how they integrate clustering.
- **대표적으로 오토인코더 기반, CNN 기반, 대조학습 기반, Transformer 기반 접근이 널리 사용됨**
  Representative families include autoencoder-based, CNN-based, contrastive-learning-based, and Transformer-based approaches.
- **각 계열은 표현 학습 방식, 장점, 계산 특성이 다르며 데이터와 목적에 따라 적합성이 달라짐**
  Each family differs in representation learning style, strengths, and computational characteristics, making suitability task-dependent.

오토인코더 기반
Autoencoder-based



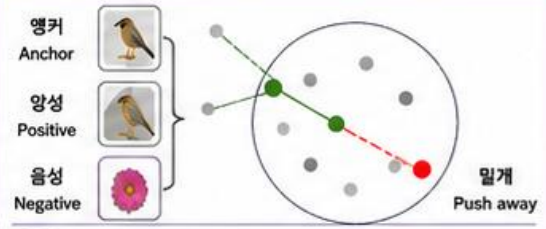
- **복원 기반 표현 학습**
  Reconstruction-driven representation learning
- **구현 용이**
  Easy to implement
- **복잡한 시각 구조에는 한계**
  Limited on highly complex visual structure

CNN 기반
CNN-based



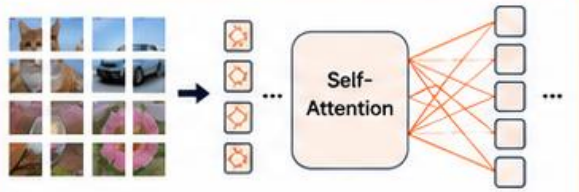
- **공간 정보 학습**
  Learns spatial structure
- **이미지 처리에 강점**
  Strong for image analysis
- **국소 패턴 중심**
  Focused on local patterns

대조학습 기반
Contrastive-learning-based



- **유사 샘플을 가까이**
  Pulls similar samples together
- **다른 샘플은 멀게**
  Pushes dissimilar samples apart
- **강한 표현 학습**
  Produces strong representations

Transformer 기반
Transformer-based



- **전역 관계 학습**
  Captures global relationships
- **장거리 의존성 표현**
  Models long-range dependencies
- **대규모 데이터에 적합**
  Suitable for large-scale data

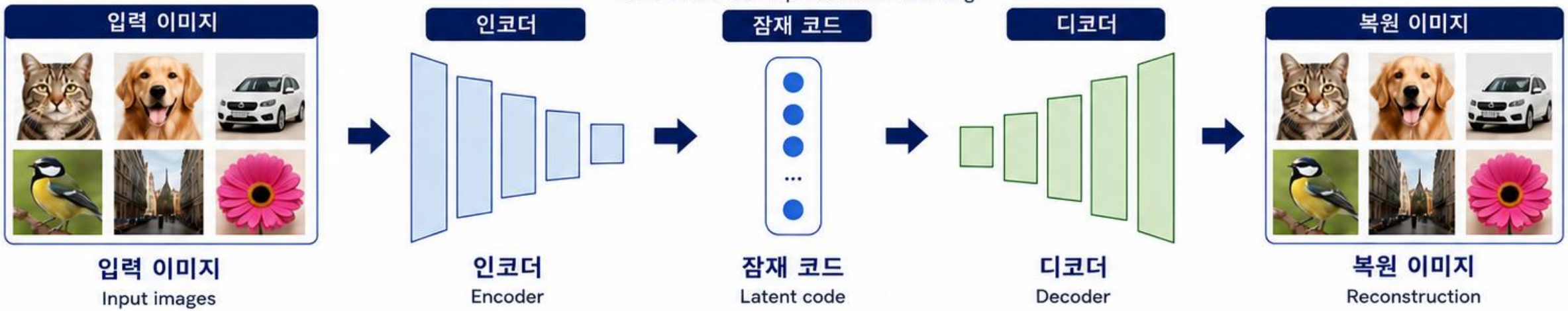
오토인코더 기반 접근

Method Overview · Autoencoder-based Approach

- **오토인코더는 입력 영상을 압축한 뒤 다시 복원하도록 학습하며 중요한 잠재 표현을 형성**
  Autoencoders learn to compress an input image and reconstruct it, forming informative latent representations.
- **인코더는 고차원 이미지를 저차원 코드로 변환하고 디코더는 이를 이용해 원본 영상을 재구성**
  The encoder maps a high-dimensional image into a low-dimensional code, and the decoder reconstructs the original image from it.
- **병목 구조에서 얻은 잠재 코드는 군집화에 활용될 수 있는 핵심 특징 표현으로 사용 가능**
  The bottleneck code can be used as a compact feature representation for downstream clustering.

오토인코더의 기본 구조

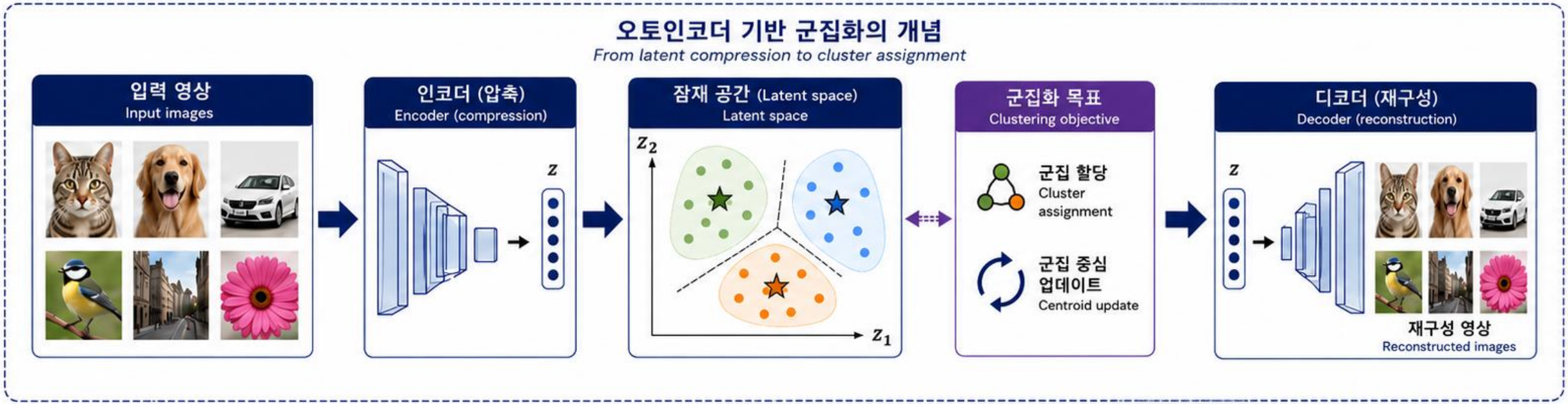
Basic bottleneck representation learning



오토인코더 + 군집화

Method Overview · Autoencoder with Clustering

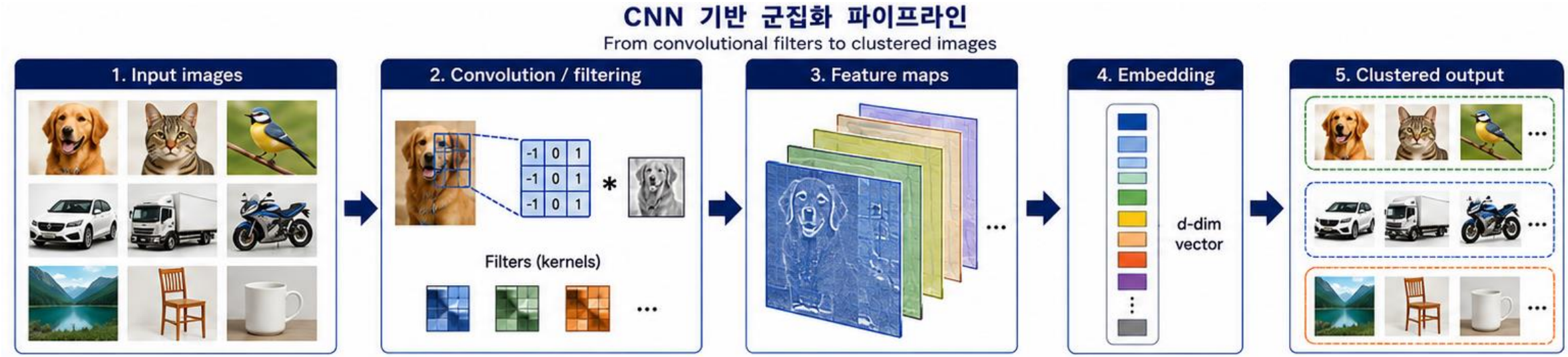
- **오토인코더 기반 군집화는 잠재 표현 학습과 군집 할당을 결합해 더 구조화된 특징 공간을 형성**
  Autoencoder-based clustering combines latent representation learning with cluster assignment to form a more structured feature space.
- **대표 아이디어는 특징을 압축하면서 군집 중심을 함께 최적화해 유사 샘플을 같은 그룹으로 유도하는 것**
  A representative idea is to compress features while jointly optimizing cluster centers so that similar samples are guided into the same group.
- **구현은 비교적 단순하지만 복잡한 시각 데이터에서는 재구성 목표만으로 분리가 충분하지 않을 수 있음**
  The approach is relatively simple to implement, but reconstruction alone may not provide enough separation for complex visual data.



CNN 기반 심층 영상 군집화

Method Overview · CNN-based Deep Image Clustering

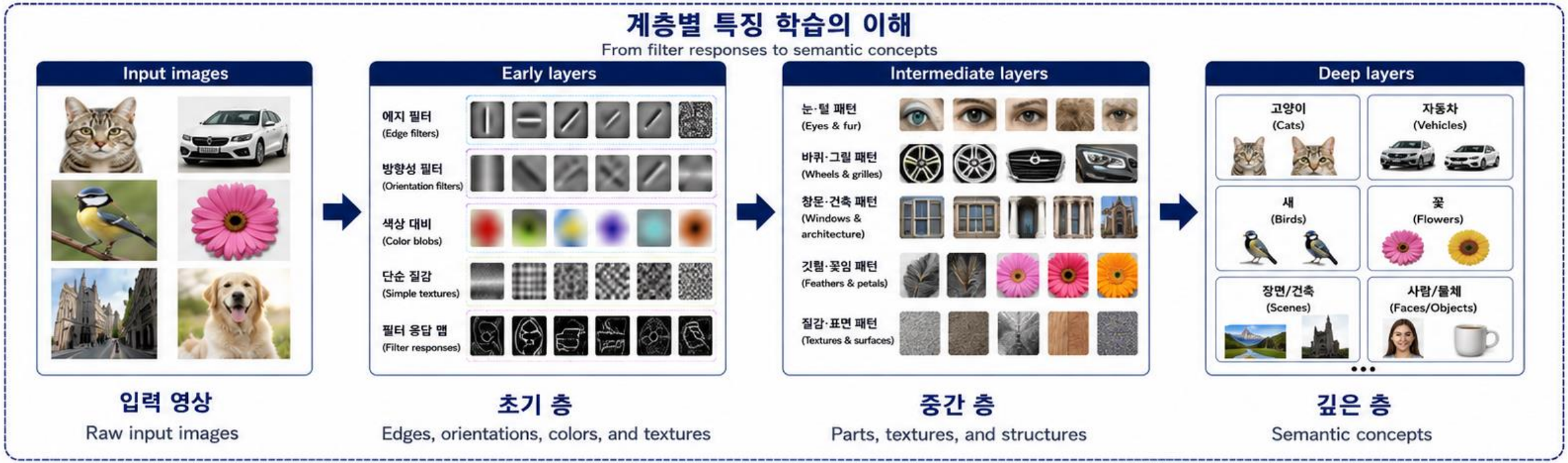
- **CNN은 합성곱 필터를 통해 국소 패턴과 공간 구조를 학습해 이미지 표현을 효과적으로 추출**
  CNNs extract image representations effectively by learning local patterns and spatial structure through convolutional filters.
- **초기 층에서는 모서리와 질감, 깊은 층에서는 부품과 객체 수준의 정보를 점차 포착**
  Early layers capture edges and textures, while deeper layers progressively encode parts and object-level information.
- **이러한 표현은 군집화 단계에서 유사 이미지가 더 잘 모이도록 도와 다양한 시각 데이터에 강점을 보임**
  These representations help similar images gather more effectively in the clustering stage, making CNNs strong for many visual data types.



CNN은 무엇을 학습하는가?

Method Overview · What Does a CNN Learn?

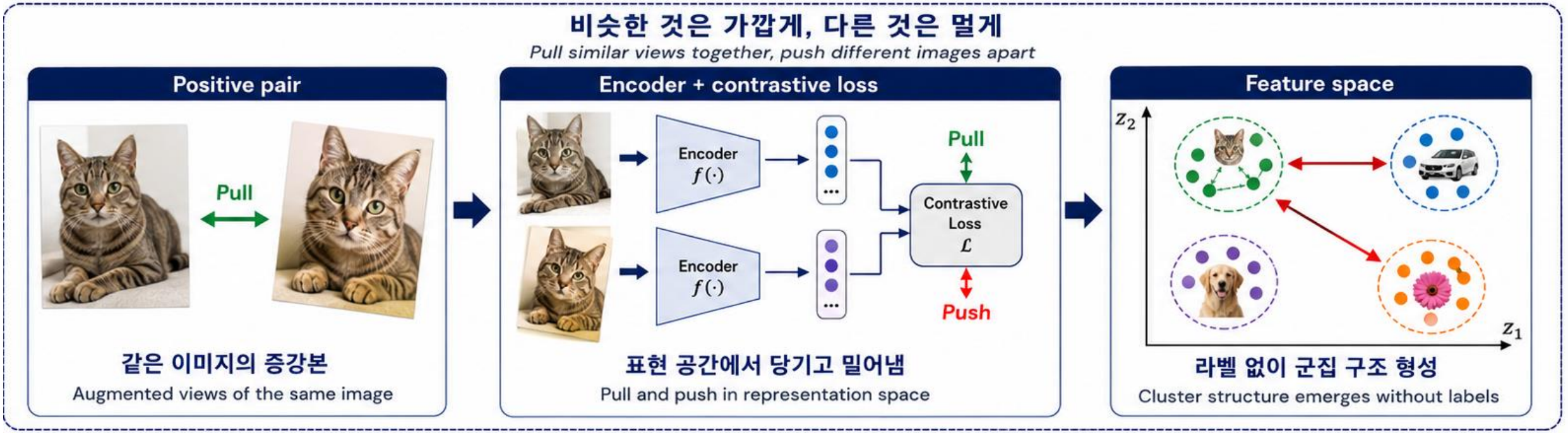
- **초기 합성곱 층은 모서리, 방향성, 색상 대비와 같은 저수준 시각 단서를 먼저 포착**
  Early convolutional layers first capture low-level visual cues such as edges, orientations, and color contrast.
- **중간 층에서는 반복 패턴과 부분 구조가 조합되어 질감, 부품, 형태 수준의 표현이 형성**
  Intermediate layers combine repeated patterns and partial structures to form texture-, part-, and shape-level representations.
- **깊은 층으로 갈수록 고양이, 자동차, 새, 꽃과 같은 의미적 객체 개념을 더 분명히 구분**
  As layers deepen, semantic object concepts such as cats, cars, birds, and flowers become more clearly separated.



대조학습

Algorithm · Contrastive Learning

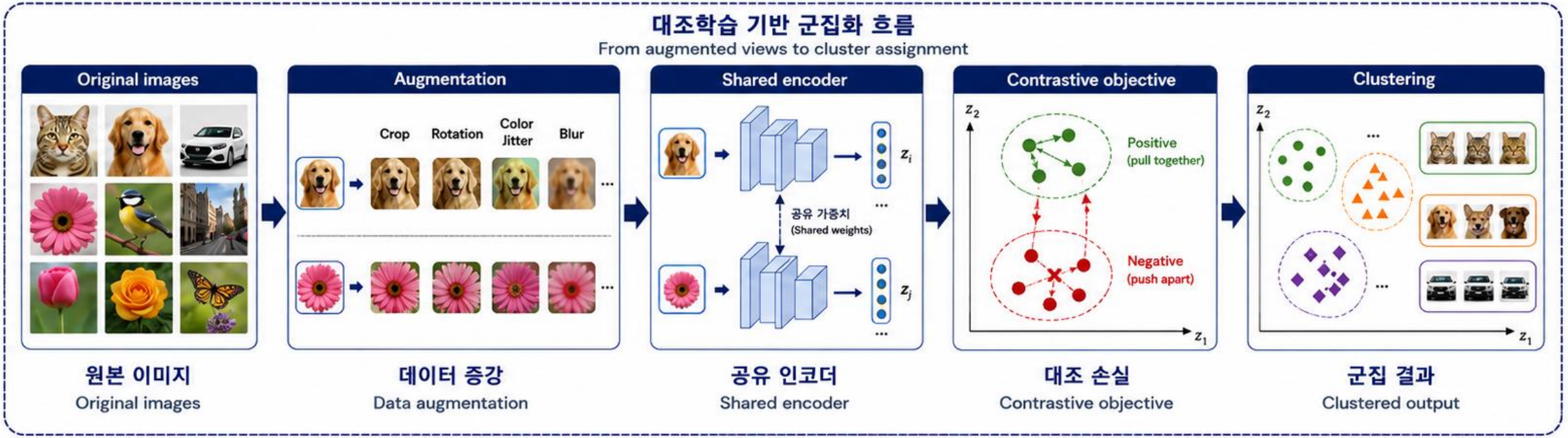
- **대조학습은 같은 이미지의 서로 다른 증강본은 가깝게, 다른 이미지는 멀게 배치하도록 표현을 학습**
  Contrastive learning pulls augmented views of the same image together and pushes different images apart in feature space.
- **정답 라벨 없이 positive pair와 negative pair를 구성해 의미 있는 표현을 스스로 학습**
  It learns meaningful representations without labels by constructing positive and negative pairs.
- **학습된 특징 공간에서는 유사한 샘플이 자연스럽게 모여 이후 군집화에 유리한 구조가 형성**
  The learned feature space naturally groups similar samples, forming a structure favorable for clustering.



대조학습 기반 심층 영상 군집화

Algorithm · Contrastive Learning for Clustering

- **원본 이미지를 여러 방식으로 증강해 동일 샘플의 다양한 관측을 positive pair로 구성**
  Multiple augmentations of the same image are treated as positive pairs.
- **공유 인코더는 증강에 강인한 특징을 학습하고 군집화에 적합한 잠재 표현을 형성**
  A shared encoder learns augmentation-invariant features and forms clustering-friendly latent representations.
- **표현 학습 후 특징 공간에서 유사 샘플을 자동으로 묶어 라벨 없는 군집 결과를 도출**
  After representation learning, similar samples are automatically grouped in the feature space without labels.



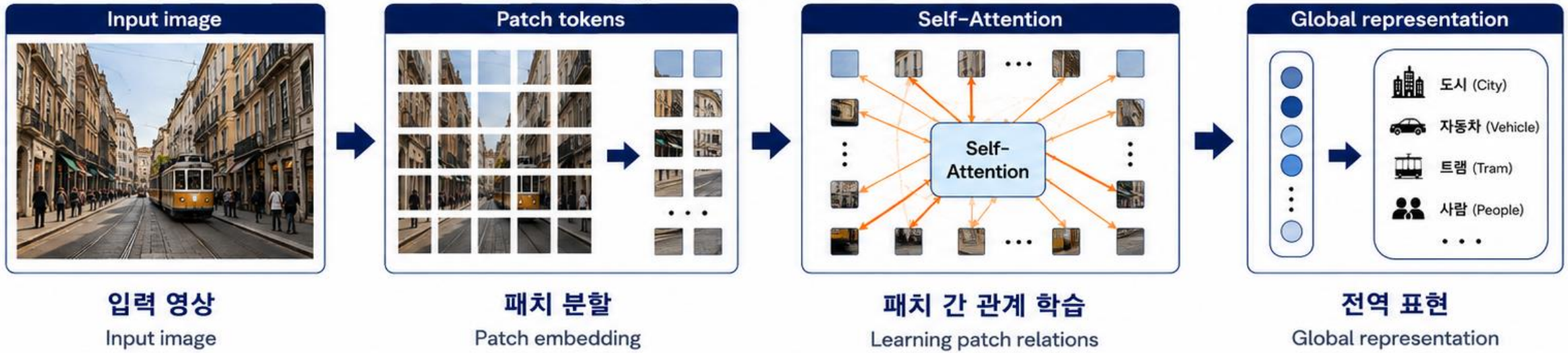
Transformer란?

Method Overview · Vision Transformer and Attention

- **Transformer는 이미지를 작은 패치로 나눈 뒤 각 패치 간 관계를 Attention으로 학습**
  A Transformer divides an image into patches and learns relationships among patches through attention.
- **CNN이 주로 가까운 국소 영역을 처리하는 반면 Transformer는 영상 전체의 전역 관계를 동시에 고려**
  While CNNs mainly process local neighborhoods, Transformers consider global relationships across the whole image.
- **전역 문맥을 반영한 표현은 복잡한 장면과 장거리 의존성이 종요한 데이터에서 유용**
  Representations with global context are useful for complex scenes and long-range dependencies.

이미지 패치와 Self-Attention 구조

Image patches and self-attention structure



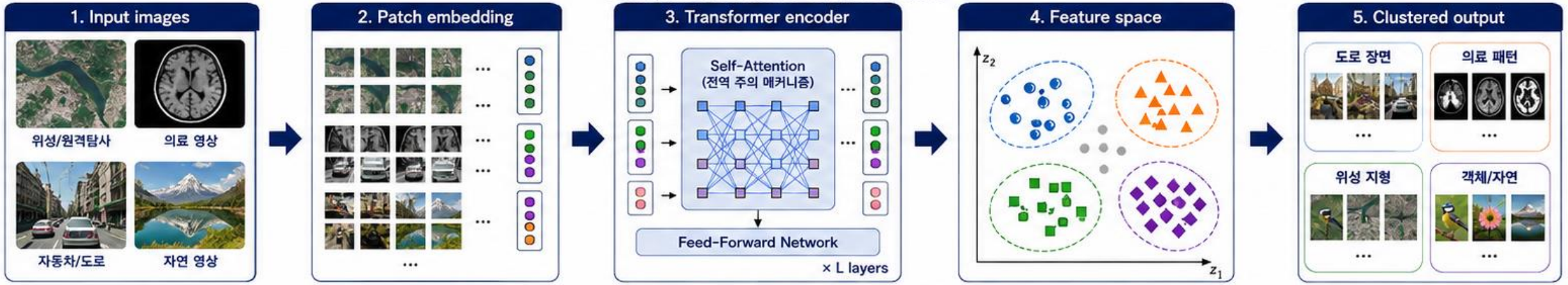
Transformer 기반 심층 영상 군집화

Method Overview · Transformer-based Deep Image Clustering

- **Transformer 기반 접근은 패치 간 장거리 관계를 학습해 복잡한 이미지의 전역 구조를 표현**
  Transformer-based approaches learn long-range relationships among patches to represent global structure in complex images.
- **대규모 데이터에서 풍부한 표현을 학습할 수 있어 다양한 이미지 유형의 군집화에 활용 가능**
  They can learn rich representations from large-scale data and support clustering across diverse image types.
- **강력한 표현 능력과 함께 높은 계산 비용과 데이터 요구량이 중요한 고려사항으로 남음**
  High representation capacity is accompanied by important considerations such as computational cost and data demand.

Transformer 기반 군집화 파이프라인

From patch attention to semantic clusters



입력 영상 Input images

패치 토큰화 Patch embedding

전역 관계 학습 Global relation learning

전역 표현 형성 Global representations

의미별 군집 Semantic clusters

대표 알고리즘 비교

Summary · Method Comparison

- **심층 영상 군집화 방법은 표현 학습 방식에 따라 장점과 한계가 뚜렷하게 달라짐**
  Deep image clustering methods differ clearly in strengths and limitations depending on how representations are learned.
- **오토인코더와 CNN은 비교적 직관적인 구조를 제공하고, 대조학습과 Transformer는 더 강한 표현 학습에 초점**
  Autoencoders and CNNs offer relatively intuitive structures, while contrastive learning and Transformers focus on stronger representations.
- **실제 적용에서는 데이터 규모, 계산 비용, 해석 가능성, 목표 작업을 함께 고려해 방법을 선택해야 함**
  In practice, method selection should consider data scale, computational cost, interpretability, and the target task.

방법 / *Method*	대표 모델 / *Models*	장점 / *Strengths*	단점 / *Limitations*	계산 비용·성능 *Cost & Performance*
오토인코더 *Autoencoder*	DEC · DCN	구현 단순 · 압축 표현 *Simple implementation · Compact representation*	복잡한 데이터에서 분리 한계 *Limited separation for complex data*	비용: 낮음 성능: ★★☆☆☆
CNN	DeepCluster	공간 구조 학습 · 이미지 처리 강점 *Learns spatial structure · Strong for image analysis*	국소 패턴 중심 *Focused on local patterns*	비용: 중간 성능: ★★★☆☆
대조학습 *Contrastive*	SimCLR · SwAV	라벨 불필요 · 강한 표현 학습 *Label-free · Strong representation learning*	증강 설계에 민감 *Sensitive to augmentation design*	비용: 중간-높음 성능: ★★★★☆
Transformer	DINO · ViT	전역 관계 학습 · 대규모 데이터 적합 *Global relation learning · Suitable for large-scale data*	계산량 큼 · 데이터 요구량 높음 *High compute and data demand*	비용: 매우 높음 성능: ★★★★★

어디에 사용되는가?

Applications · Where It Is Used

- **심층 영상 군집화는 라벨이 부족한 실제 영상 환경에서 유사 패턴 탐색과 자동 그룹화에 활용**
  Deep image clustering is useful for discovering similar patterns and automatically grouping images in real-world settings with limited labels.
- **제조, 의료, 원격탐사, 자율주행, 지능형 감시 등 다양한 분야에서 데이터 구조 파악과 이상 탐지를 지원**
  It supports structure discovery and anomaly screening across manufacturing, medical imaging, remote sensing, autonomous driving, and intelligent surveillance.
- **입력 영상은 다르지만 공통적으로 무라벨 데이터에서 의미 있는 그룹을 형성하는 것이 핵심 목표**
  Although the input images differ by domain, the shared goal is to form meaningful groups from unlabeled data.

대표 응용 분야

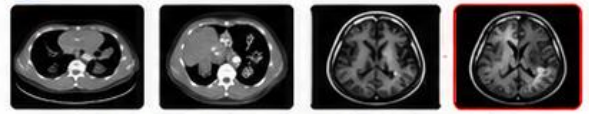
Representative application areas

심층 영상 군집화
Deep Image Clustering

산업 검사 / *Industrial Inspection*

정상 제품과 결함 제품 자동 그룹화 및 이상 검출
Automatic grouping of normal and defective products

의료 영상 / *Medical imaging*



정상/비정상 패턴 군집을 통한 이상 병변 탐지
Abnormality detection via normal/abnormal pattern clustering

원격탐사 / *Remote sensing*

토지피복 유형 자동 분류 및 변화 지역 탐지
Land cover classification and change area detection

자율주행 / *Autonomous driving*

도로 상황 및 객체 그룹화로 이상 상황 탐지 지원
Support anomaly detection via clustering of road scenes and objects

지능형 감시 / *Intelligent surveillance*

정상 행동 / Normal behavior　　이상 행동 / Anomalous behavior

일상 패턴 학습 후 비정상 행동 자동 식별
Learning routine patterns and identifying abnormal behavior

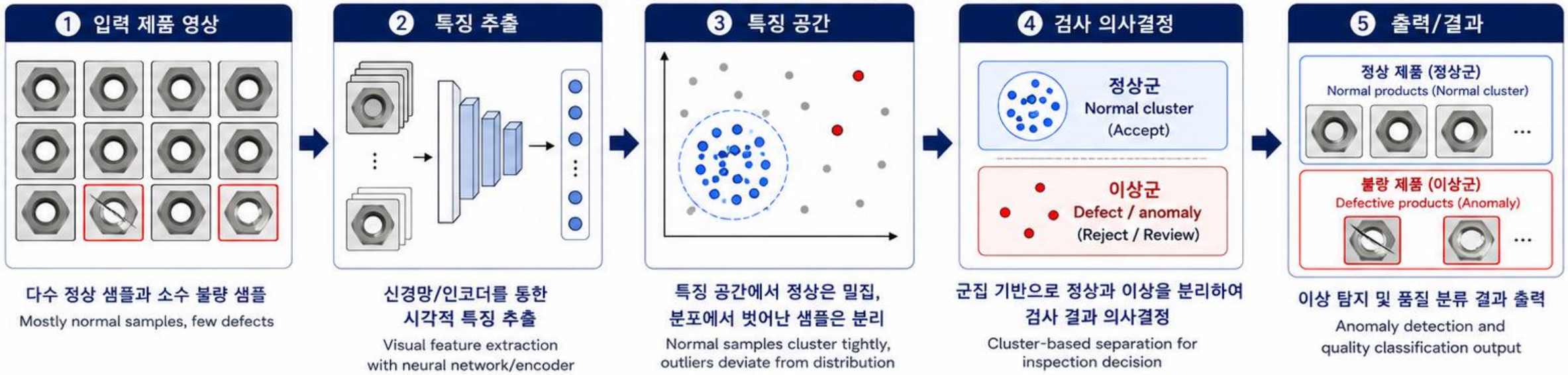
산업 품질 검사

Applications · Industrial Inspection

- **산업 검사에서는 대부분의 제품 영상이 정상 샘플이며 불량 사례는 상대적으로 적어 라벨 확보가 어려움**
  In industrial inspection, most product images are normal samples, while defective cases are scarce and costly to label.
- **심층 영상 군집화는 정상 패턴을 자동으로 묶고 분포에서 벗어난 샘플을 이상군으로 분리하는 데 유용**
  Deep image clustering is useful for grouping normal patterns and separating samples that deviate from the distribution as anomalies.
- **반도체·제조 공정과 같은 품질 검사 환경에서 시각적 검사 자동화와 이상 탐지에 활용 가능**
  It can support visual inspection automation and anomaly detection in quality-control settings such as semiconductors and manufacturing processes.

산업 검사에서의 군집화 흐름

From product images to anomaly-aware grouping



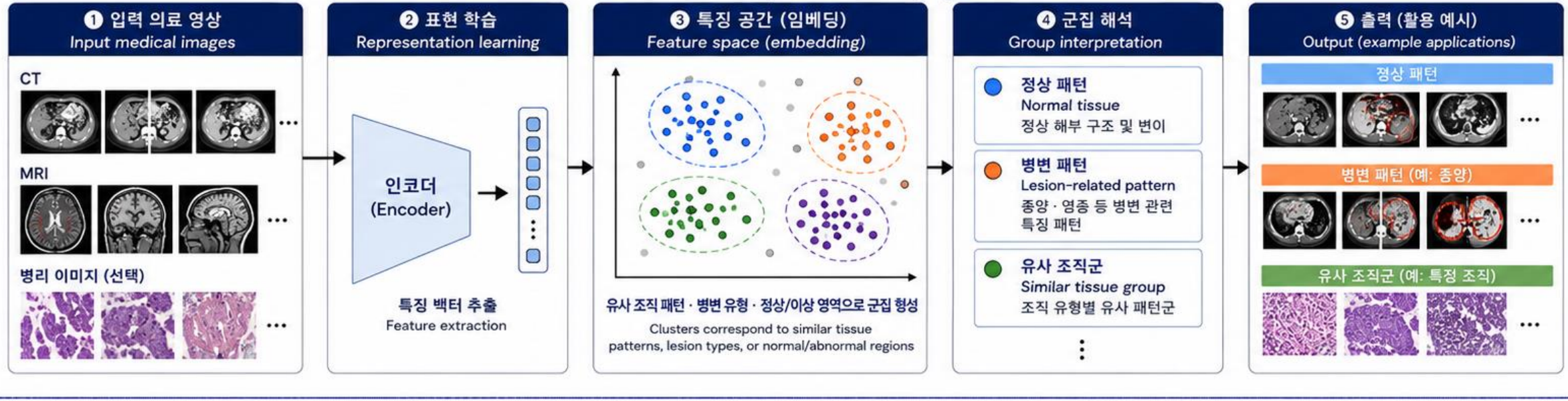
의료 영상 응용

Applications · Medical Imaging

- **의료영상은 전문가 라벨링 비용이 높고 희귀 질환처럼 충분한 정답 데이터를 확보하기 어려운 경우가 많음**
  Medical imaging often involves expensive expert labeling, and sufficient labeled data are difficult to obtain, especially for rare diseases.
- **심층 영상 군집화는 CT · MRI 등에서 유사 조직 패턴을 자동으로 발견해 패턴 탐색과 이상 영역 탐지에 도움을 줄 수 있음**
  Deep image clustering can automatically discover similar tissue patterns in CT and MRI, supporting pattern discovery and abnormal-region screening.
- **암 영상 분류, 조직 패턴 탐색, 이상 탐지와 같은 의료 영상 분석 작업에 활용 가능**
  It can be applied to tasks such as cancer image categorization, tissue-pattern exploration, and anomaly detection in medical images.

의료 영상에서의 군집화 개념

Pattern discovery in medical imaging



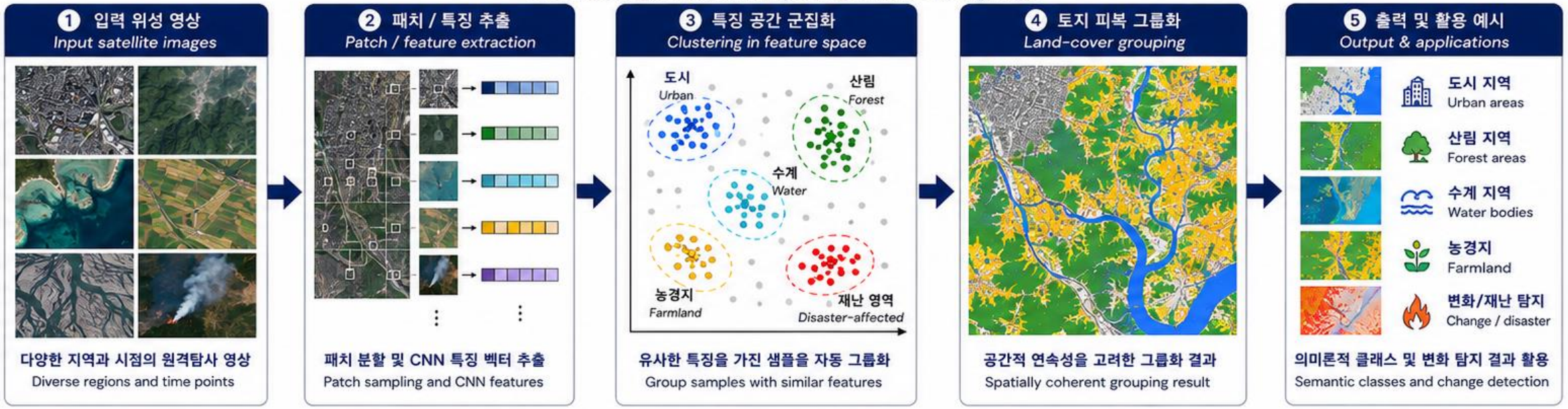
위성 영상 분석

Applications · Remote Sensing

- **위성 영상은 넓은 지역과 대량의 무라벨 데이터를 포함해 수작업 분류와 해석에 많은 시간과 비용이 필요**
  Satellite imagery covers broad regions and large volumes of unlabeled data, making manual classification and interpretation time-consuming and expensive.
- **심층 영상 군집화는 유사한 토지 피복과 지형 패턴을 자동으로 그룹화해 토지 이용 분석과 변화 탐지에 유용**
  Deep image clustering automatically groups similar land-cover and terrain patterns, which is useful for land-use analysis and change detection.
- **도시·산림·수계·농경지 구분뿐 아니라 산림 변화, 홍수, 화재 등 재난 탐지 보조에도 활용 가능**
  It can support the separation of urban, forest, water, and farmland regions, as well as forest-change, flood, and fire-related screening.

원격탐사 영상 군집화 흐름

From earth observation images to semantic land groups



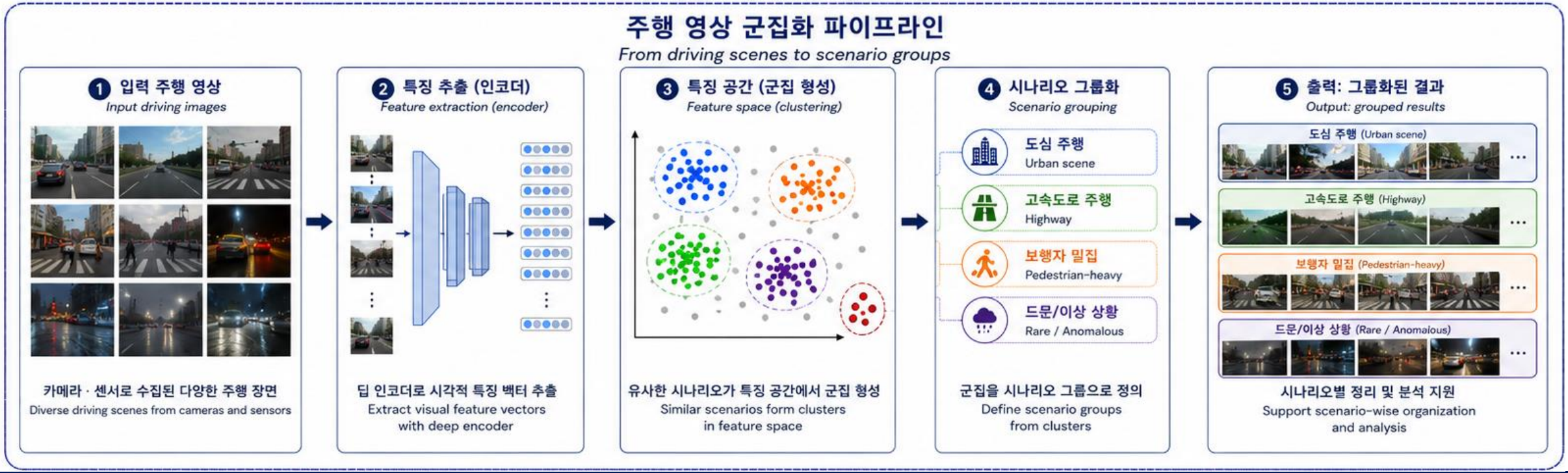
자율주행 데이터 분석

Applications · Autonomous Driving

- **자율주행 시스템은 카메라와 센서로부터 방대한 무라벨 주행 영상을 생성해 모든 장면을 수작업으로 정리하기 어려움**
  Autonomous driving systems generate massive unlabeled driving data from cameras and sensors, making manual organization of all scenes impractical.
- **심층 영상 군집화는 유사한 도로 장면과 객체 패턴을 그룹화하고 드문 상황이나 이상 장면을 찾는 데 도움을 줄 수 있음**
  Deep image clustering can group similar road scenes and object patterns while helping identify rare situations or anomalous scenes.
- **도로 상황 분류, 객체 패턴 탐색, 이상 상황 탐지 등 자율주행 데이터 정리와 분석에 활용 가능**
  It can support autonomous-driving data organization and analysis for road-scene categorization, object-pattern discovery, and anomaly detection.

주행 영상 군집화 파이프라인

From driving scenes to scenario groups

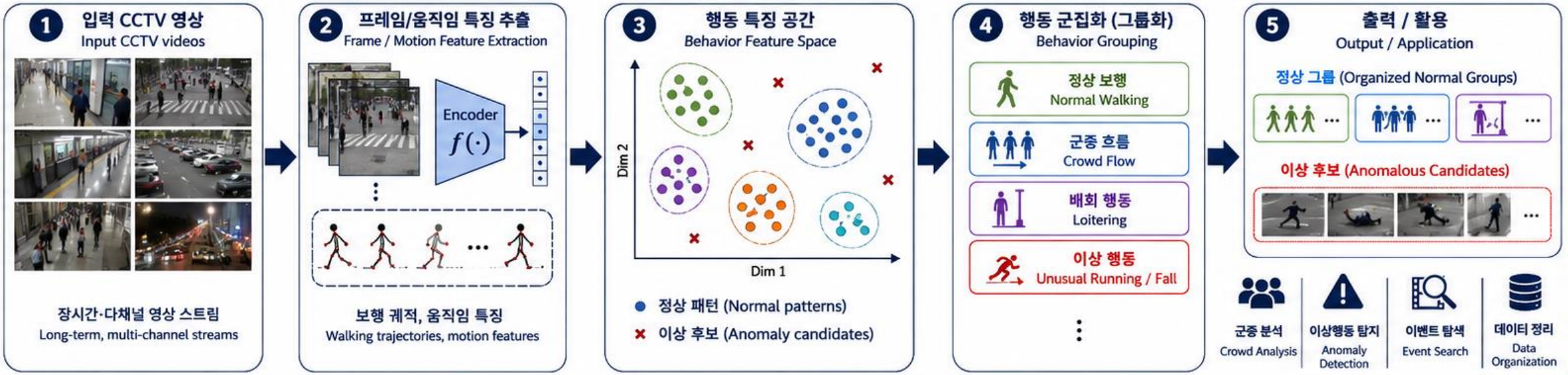


지능형 영상 감시

Applications · Intelligent Surveillance

- **감시 영상은 장시간 · 다채널로 축적되어 사람이 모든 장면을 직접 확인하기 어려움**
  Surveillance videos accumulate over long periods and multiple channels, making manual review difficult.
- **심층 영상 군집화는 정상 행동 패턴을 자동으로 묶고 분포에서 벗어난 행동을 이상 후보로 분리**
  Deep image clustering can group routine behavior patterns and separate outlying behaviors as anomaly candidates.
- **군중 행동 분석, 이상행동 탐지, 이벤트 탐색 등 지능형 감시 시스템의 데이터 정리에 활용 가능**
  It can support crowd behavior analysis, abnormal-event detection, and video data organization.

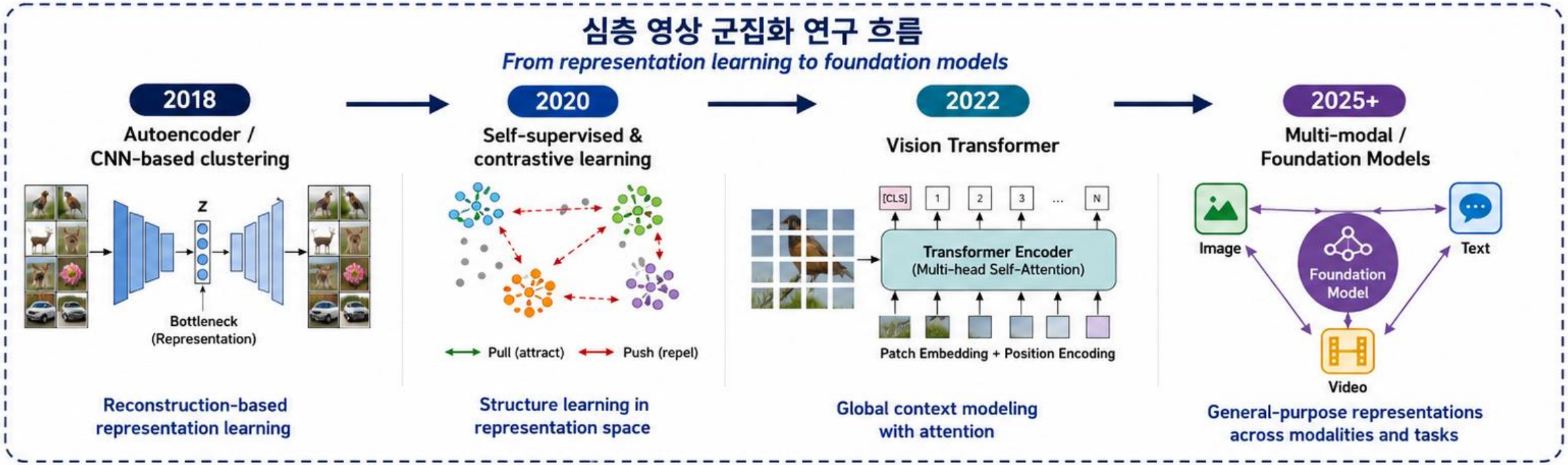
지능형 감시에서의 군집화 흐름
From CCTV videos to behavior groups



최신 연구 동향

Research Trends · Emerging Directions

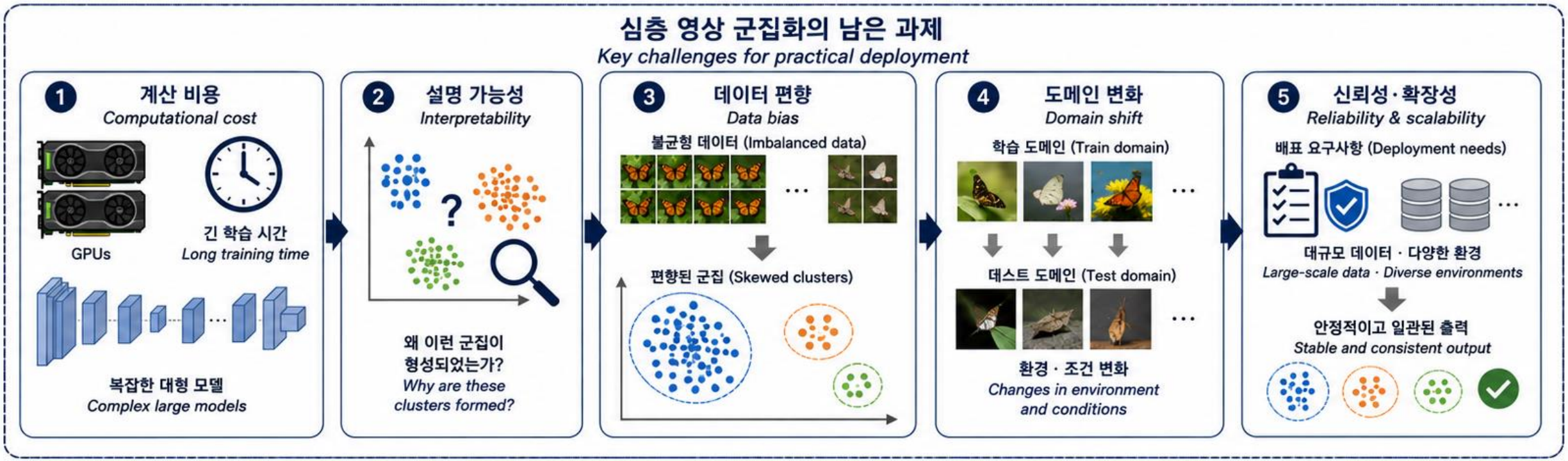
- **최근 연구는 더 적은 라벨, 더 강한 표현, 더 큰 모델을 활용하는 방향으로 발전**
  Recent research is moving toward fewer labels, stronger representations, and larger models.
- **Self-supervised Learning, Vision Transformer, Multi-modal Learning, Foundation Models가 주요 흐름**
  Self-supervised learning, Vision Transformers, multi-modal learning, and foundation models are major trends.
- **군집화는 단일 알고리즘을 넘어 범용 표현을 활용한 데이터 구조 발견 문제로 확장**
  Clustering is expanding from a single algorithmic task to discovering data structure through general-purpose representations.



아직 해결해야 할 문제

Challenges · Open Issues

- **높은 계산 비용과 대규모 학습 자원 요구는 실제 적용의 주요 부담**
  High computational cost and large training requirements remain practical barriers.
- **왜 특정 군집이 형성되었는지 설명하기 어려워 해석 가능성과 신뢰성 확보가 중요**
  It is often difficult to explain why clusters are formed, making interpretability and reliability important.
- **데이터 편향, 도메인 변화, 확장성 문제는 다양한 환경에서 안정적 적용을 제한**
  Data bias, domain shift, and scalability can limit robust use across diverse environments.



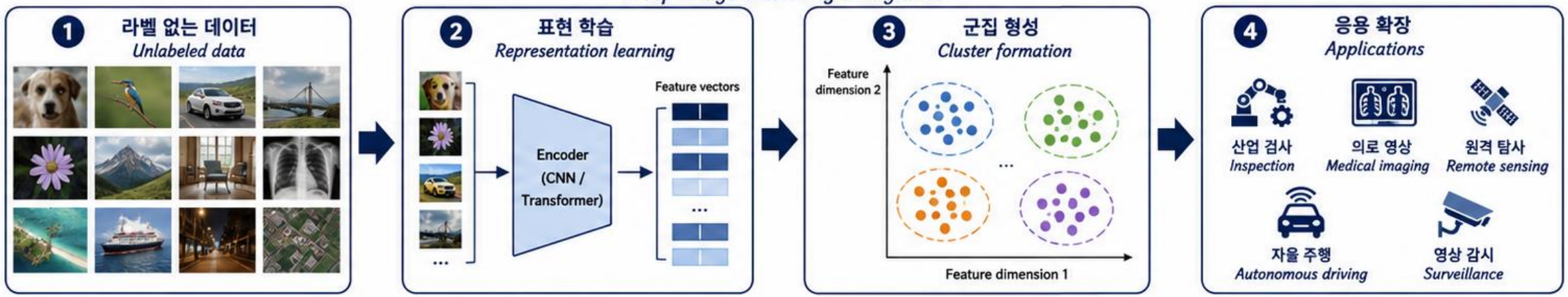
오늘의 핵심 메시지

Summary · Key Takeaways

- **라벨이 부족한 영상 데이터에서도 의미 있는 구조를 발견할 수 있음**
  Deep image clustering can discover meaningful structure even when labels are limited.
- **핵심은 원시 픽셀이 아니라 군집화에 적합한 표현 공간을 학습하는 것**
  The key is learning representation spaces suitable for clustering, rather than using raw pixels directly.
- **대조학습과 Transformer, Foundation Models로 확장되며 다양한 실제 응용으로 연결**
  Recent trends connect contrastive learning, Transformers, foundation models, and practical applications.

심층 영상 군집화 한눈에 보기

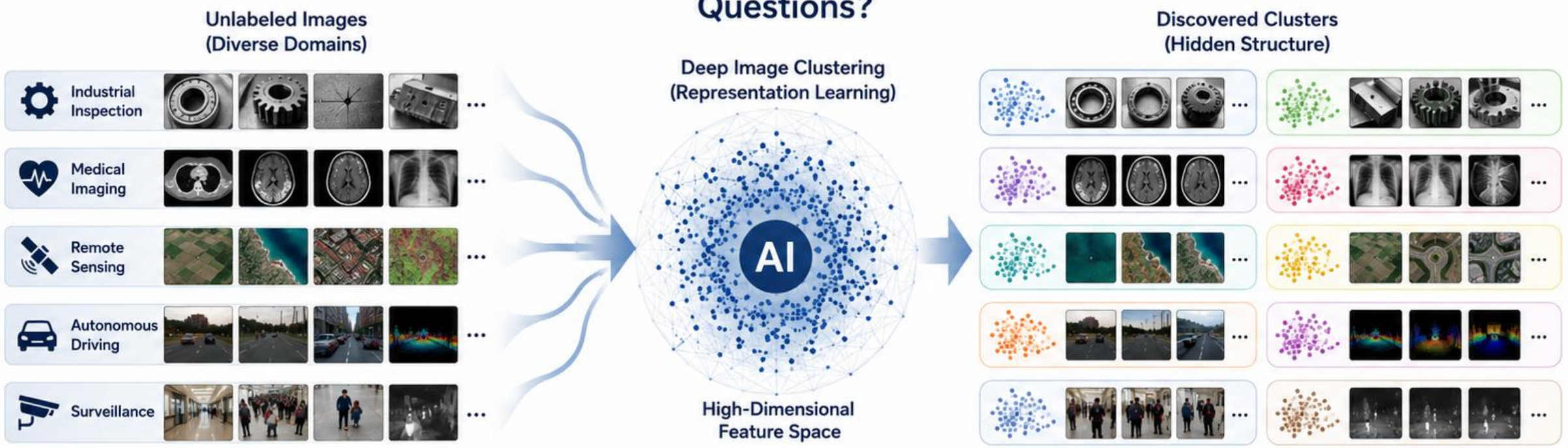
Deep image clustering at a glance



- **라벨 불필요** — *No labels required*
- **특징 자동 학습** — *Automatic feature learning*
- **다양한 응용** — *Diverse applications*
- **미래 발전 가능성** — *Future scalability*

감사합니다

Questions?



심층 영상 군집화는 데이터 속 숨겨진 의미를 발견하는 핵심 기술로 발전하고 있습니다.

Deep image clustering is evolving into a key technology for discovering hidden meaning in visual data.