한국소음진동공학회 AI혁신위원회 주관
2022년 인공지능 강습회

기계이상진단을 위한 인공지능 학습 기법

제 1강 이상진단기법 소개

한국과학기술원 전기및전자공학부

최정우
jwoo@kaist.ac.kr

KAIST EE

이상진단이란?

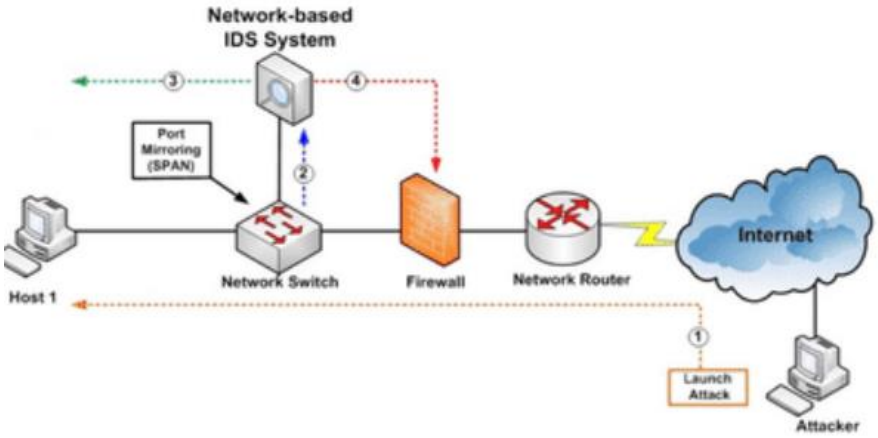
Introduction


Normal data
Outlier

이상진단 응용분야

- Security & Surveillance

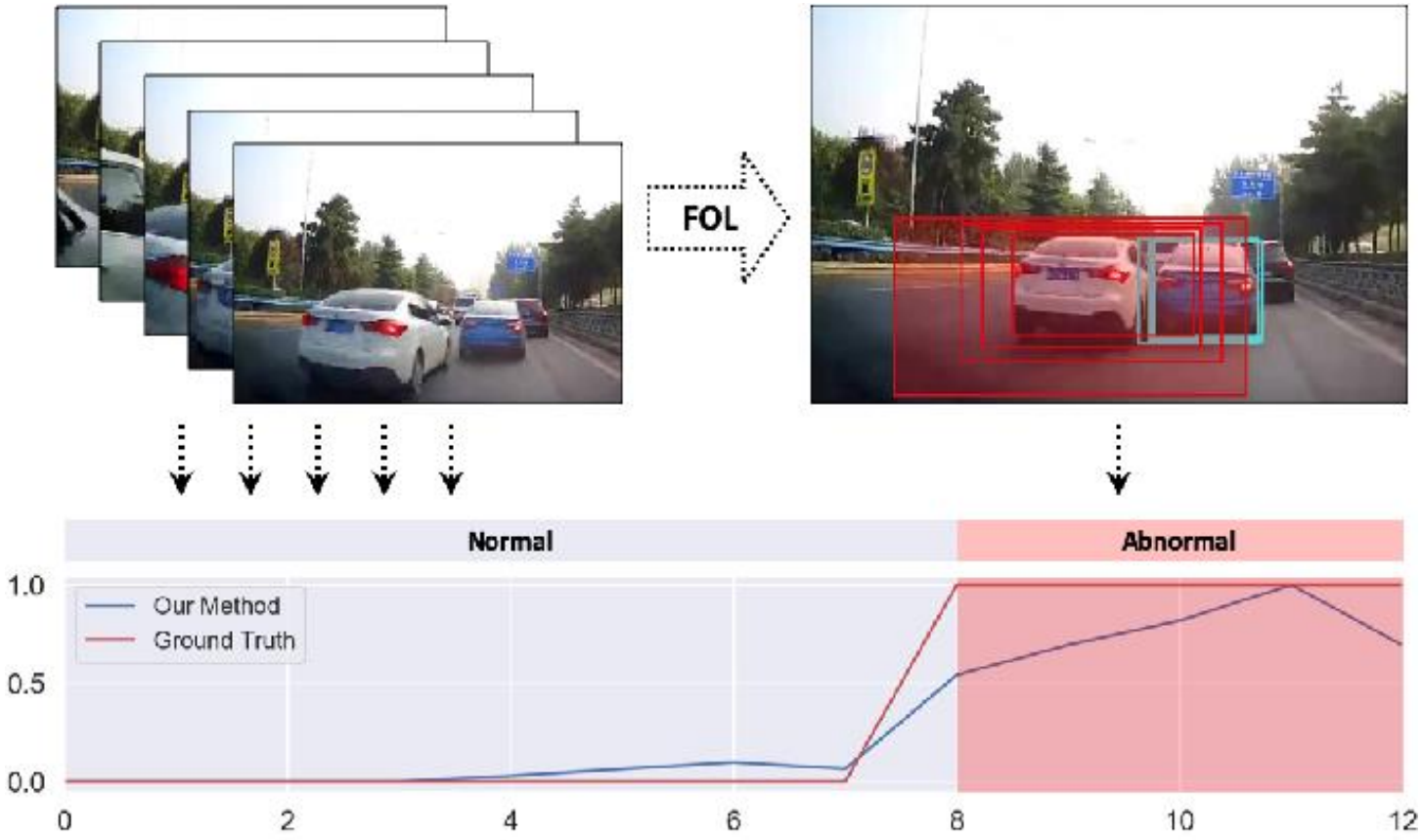
Illegal traffic flow detection



Cyber-network intrusion detection

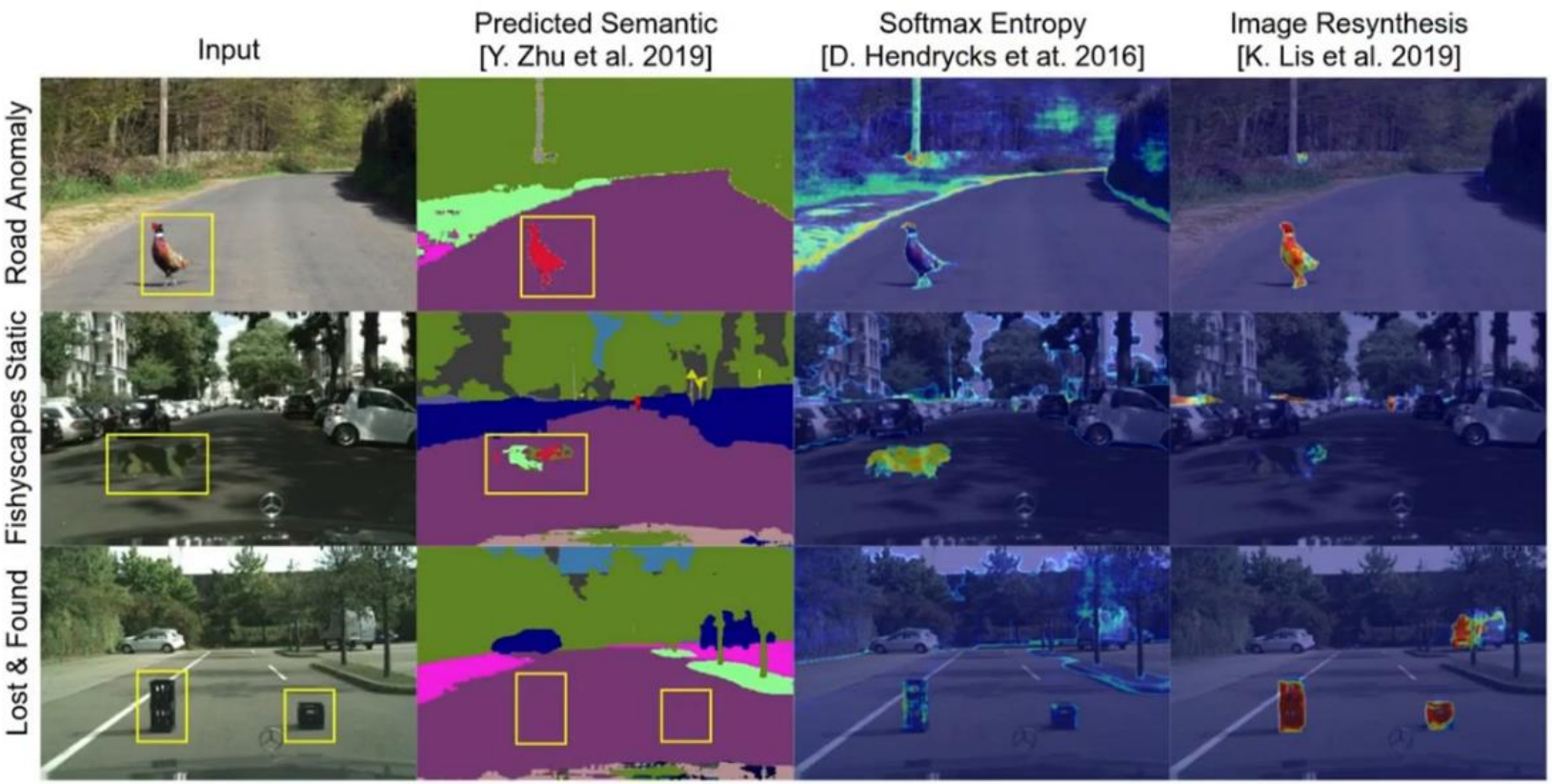
이상진단 응용분야

- Accident detection in self-driving cars



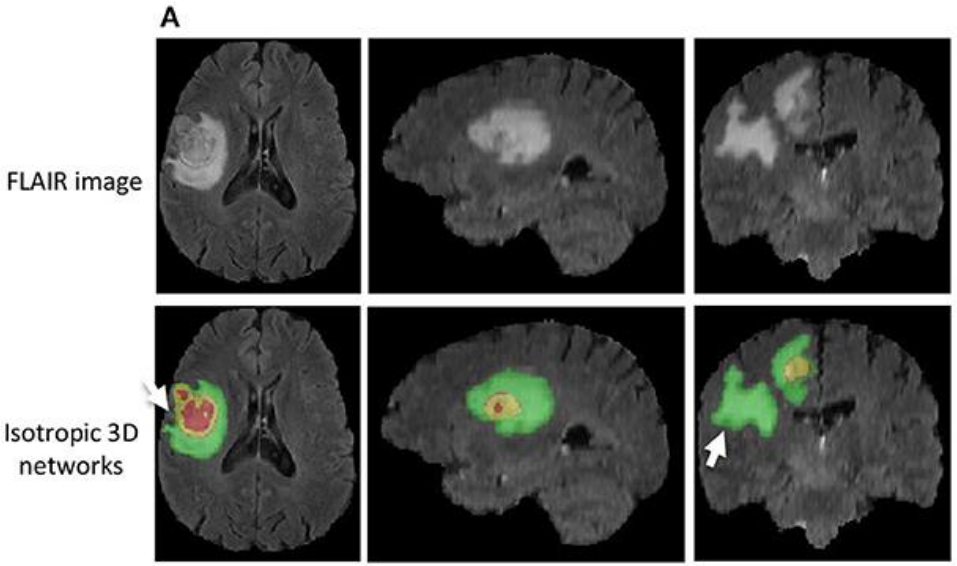
이상진단 응용분야

- Anomalous object detection for self-driving cars

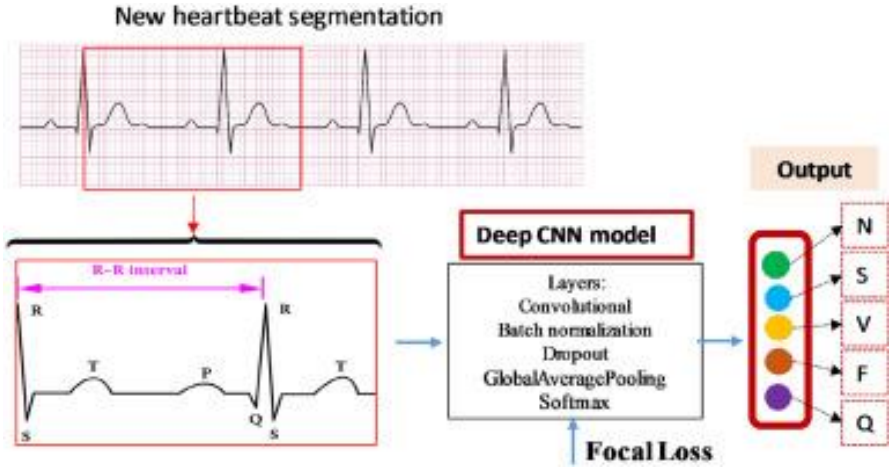


이상진단 응용분야

- Biomedical applications
  - Health condition monitoring & diagnosis



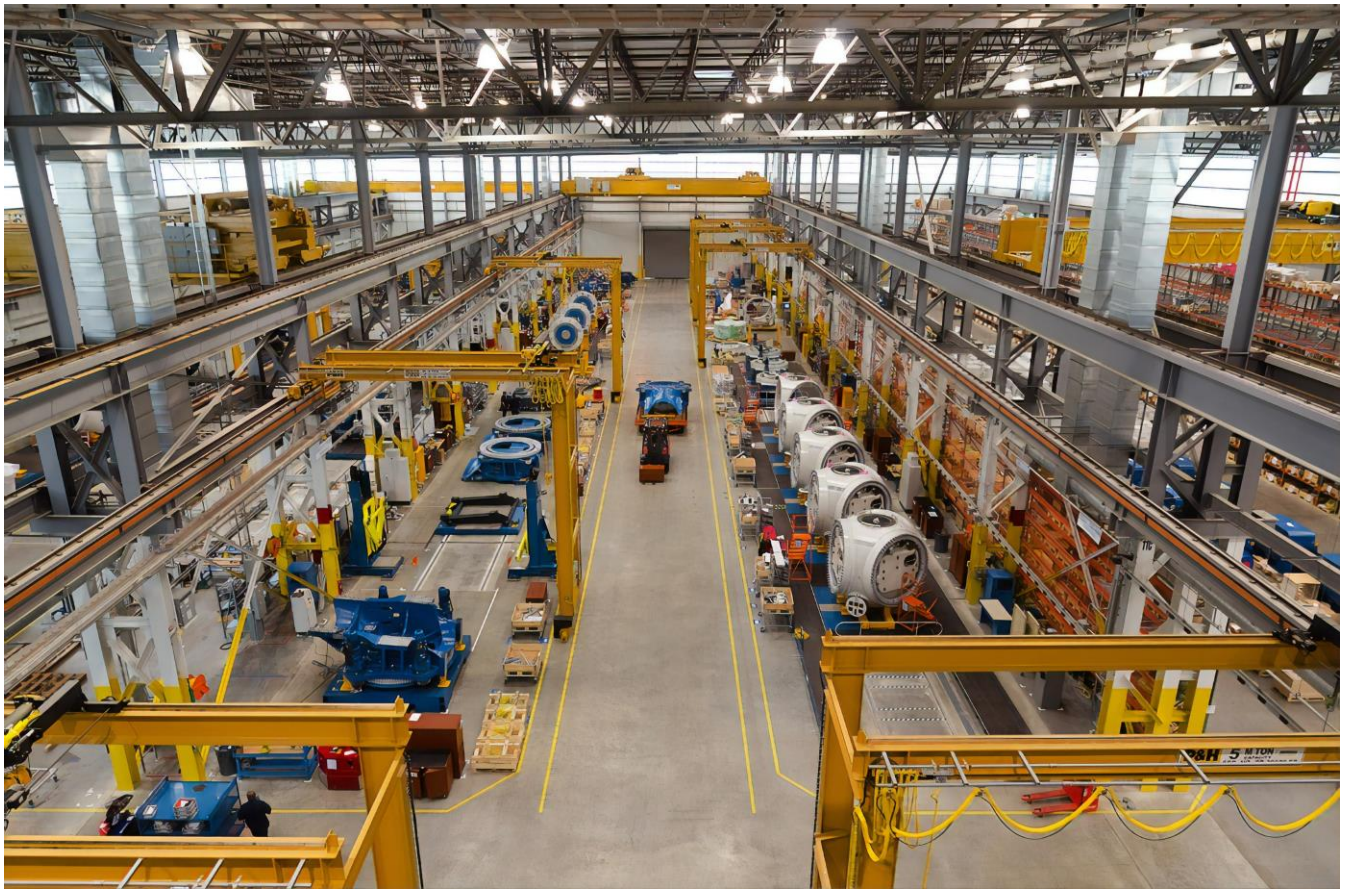
Brain tumor segmentation https://doi.org/10.3389/fncom.2019.00056



heartbeat signal anomaly detection & classification
https://doi.org/10.3389/fncom.2019.00056

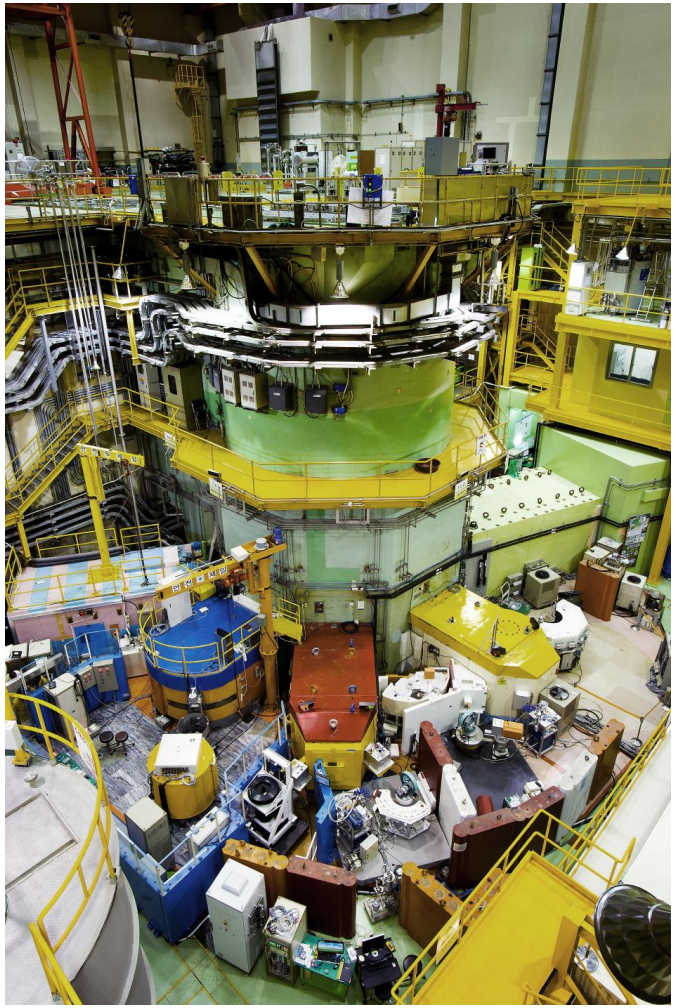
이상진단 응용분야

- Machinery fault detection
  - Finding faulty products in the production line



이상진단 응용분야

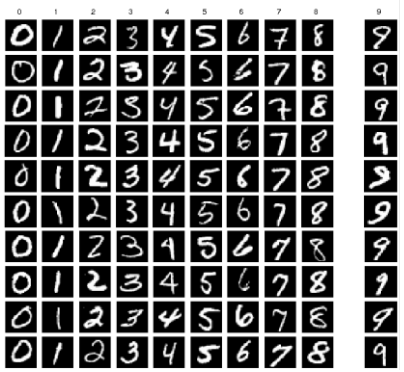
- Machinery condition diagnosis
  - In-situ diagnosis of working conditions
  - Early alarm of malfunctioning
  - Nuclear power plant
    - Vibration & Noise signals
  - Home appliances
    - Acceleration signals, RPMs



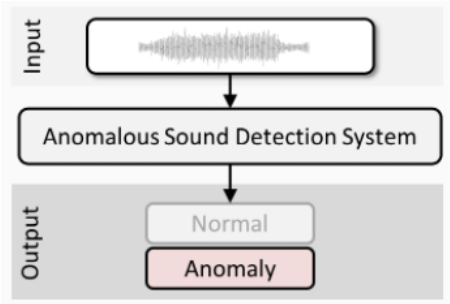
오늘의 이상진단 학습

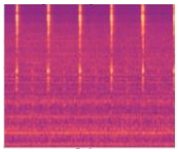
- Finding an imposter !

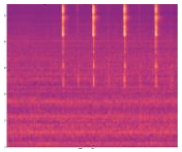
MNIST



http://dcase.community/challenge2020/task-unsupervised-detection-of-anomalous-sounds



Normal 

Anomaly 

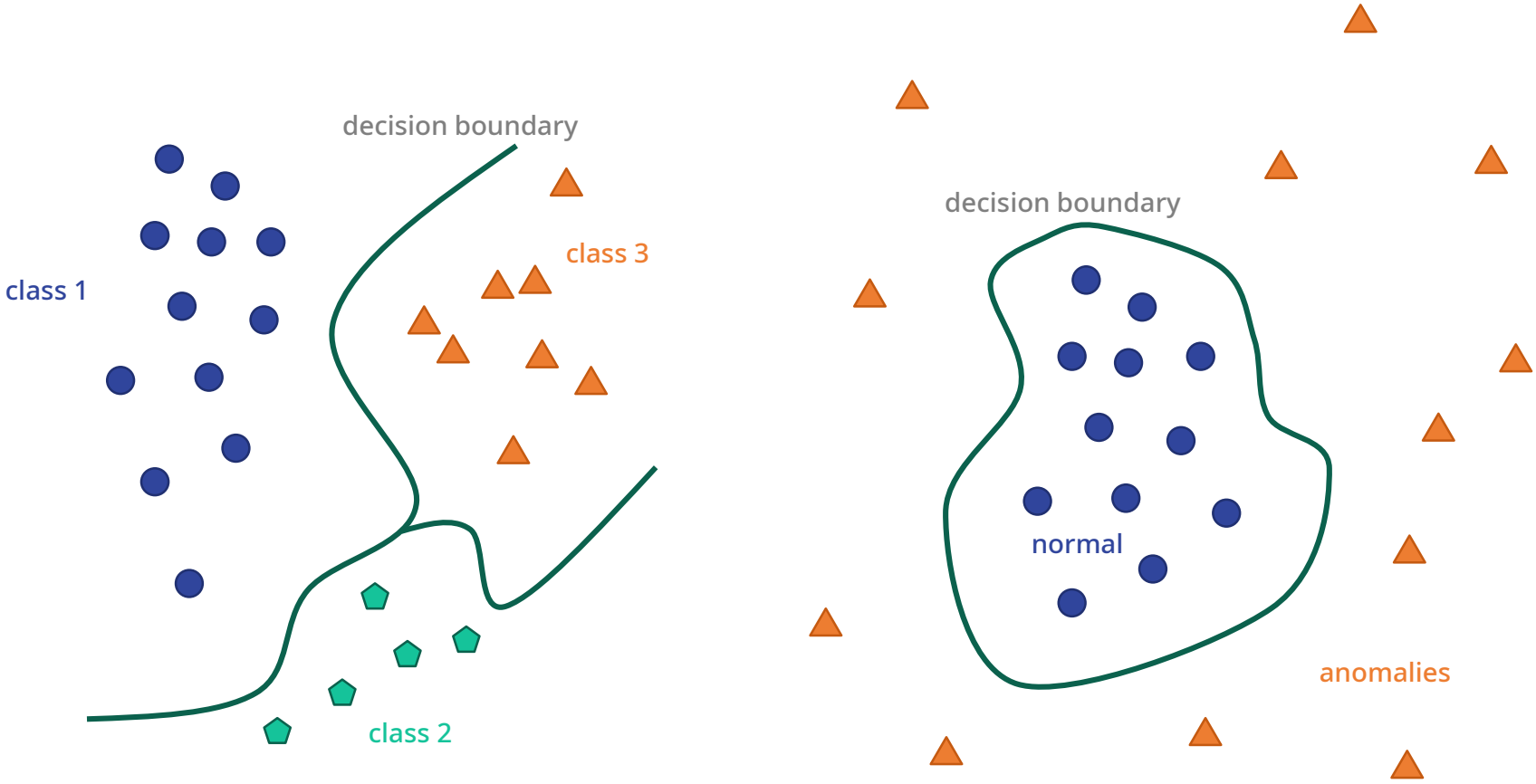
이상 진단의 어려움

- Scarcity of anomalies
  - The number of anomalous samples is relatively small at the training stage
  - DNNs tend to “overfit” when supervised learning is applied to a small number of anomalies

- Types of anomalies are diverse
  - Causes of anomalies are countless
  - We cannot measure all possible anomalies at the training stages

분류기법(classification)으로서의 이상진단

- One-class classification problem

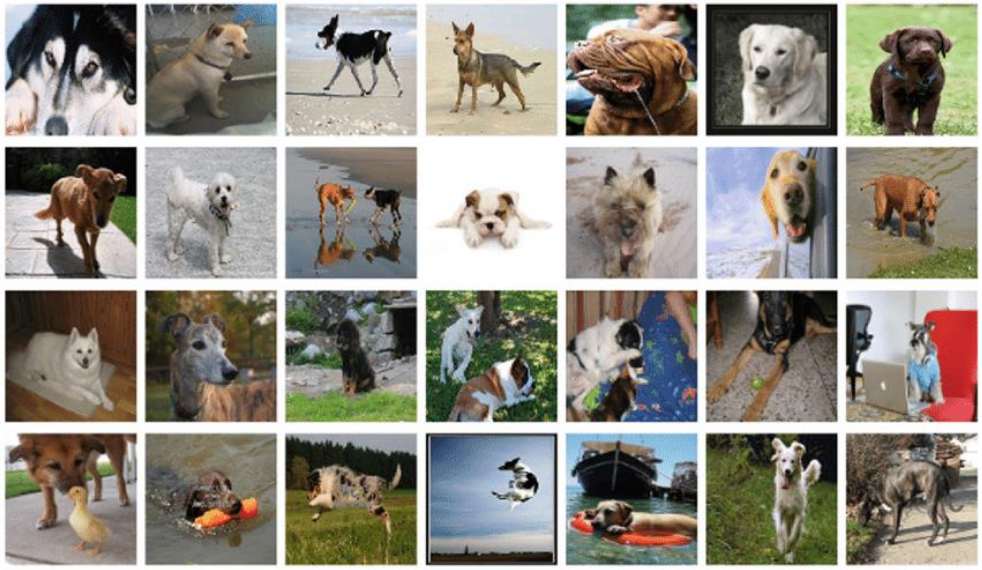


이상 진단을 위한 심층 신경망 훈련법

- Training only with normal data
  - DNNs only capture the “features” of normal data
  - Some intra-class can be provided for the normal data

- Test & Validate with normal + anomaly data

Training data

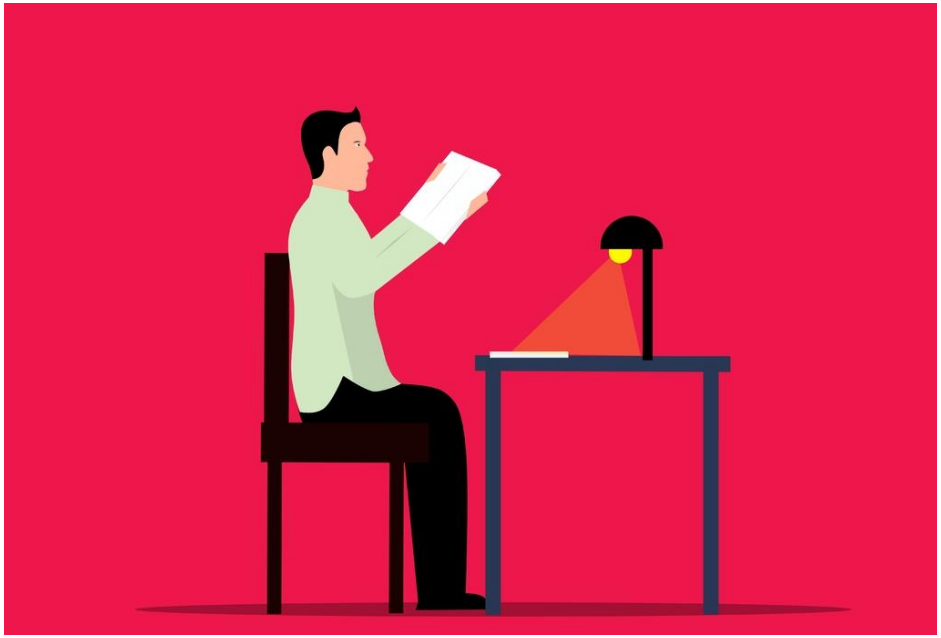


Test data

자기지도 학습 Self-supervised learning

- Supervised learning

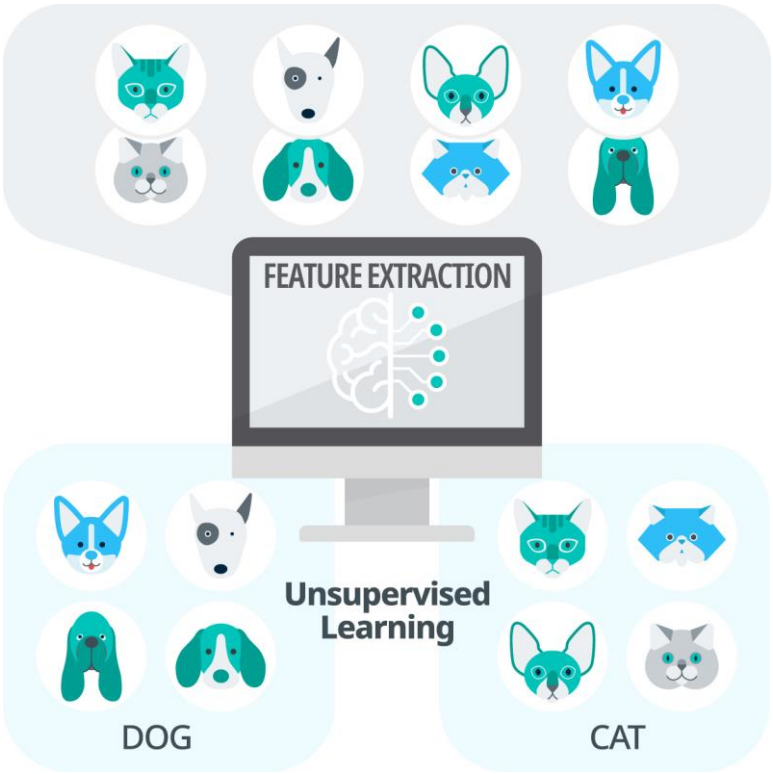
- Self-supervised learning



자기지도 학습 Self-supervised learning

- Supervised learning
  - Train with labeled data

- Self-supervised learning
  - One type of unsupervised learning
  - No label in the training data



images from http://terms.tta.or.kr/

A dark-matter-of-intelligence (Facebook AI)

https://ai.facebook.com/blog/self-supervised-learning-the-dark-matter-of-intelligence/

- If AI systems can glean a deeper, more nuanced understanding of reality **beyond what's specified in the training data set**, they'll be more useful and ultimately bring AI closer to human-level intelligence.

- As babies, we learn how the world works largely by observation. We form **generalized predictive models** about objects in the world by learning concepts such as object permanence and gravity.
  Later in life, we observe the world, act on it, observe again, and **build hypotheses to explain** how our actions change our environment by trial and error.

- A working hypothesis is that generalized knowledge about the world, or **common sense**, forms the bulk of biological intelligence in both humans and animals.

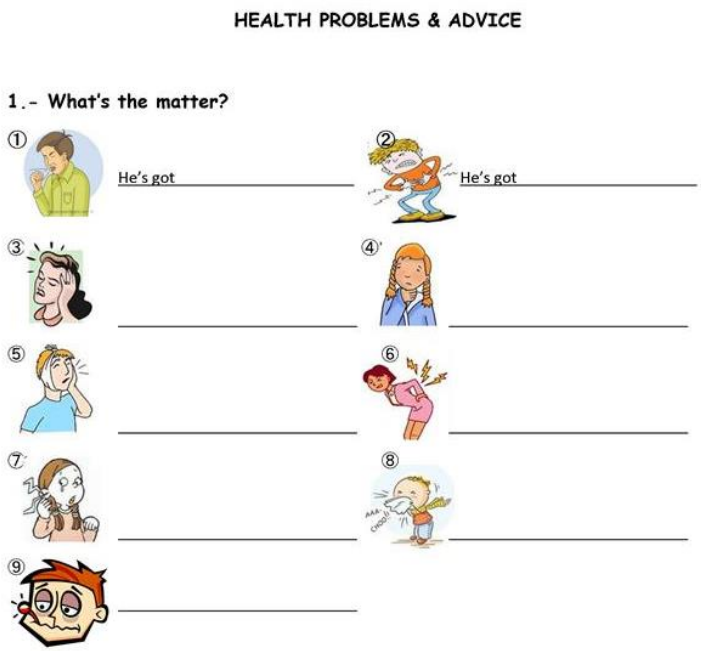
Self-supervised learning

자기지도 학습법

예습 Pretext task

- How a DNN model can learn without labels?
  - Through proxy tasks that can be accomplished by self-study

Pretext task



Downstream task

34. Precision and determinacy are a necessary requirement for all meaningful scientific debate, and progress in the sciences is, to a large extent, the ongoing process of achieving ever greater precision. But historical representation puts a premium on a proliferation of representations, hence not on the refinement of one representation but on the production of an ever more varied set of representations. Historical insight is not a matter of a continuous "narrowing down" of previous options, not of an approximation of the truth, but, on the contrary, is an "explosion" of possible points of view. It therefore aims at the unmasking of previous illusions of determinacy and precision by the production of new and alternative representations, rather than at achieving truth by a careful analysis of what was right and wrong in those previous representations. And from this perspective, the development of historical insight may indeed be regarded by the outsider as a process of creating ever more confusion, a continuous questioning of _____, rather than, as in the sciences, an ever greater approximation to the truth. [3점]

* proliferation: 증식

① criteria for evaluating historical representations
② certainty and precision seemingly achieved already
③ possibilities of alternative interpretations of an event
④ coexistence of multiple viewpoints in historical writing
⑤ correctness and reliability of historical evidence collected

가능한 예습 방법들 (pretext tasks)

1. Natural language processing (NLP) model (BERT, GPT)

- Language translation model
- Pretraining with a large corpus of text data
- DNN model learns language structure without any label

Training data

It is important to impedance match the output of an amplifier to a load. A matched amplifier/load combination will result in maximum power being delivered to the load. In addition, the impedance matching network may be designed as a low pass filter, thus improving the spectral output of the amplifier and reducing the possibility of harmonic output. Finally, if the load is matched to the source, minimum reflected power occurs, thereby reducing the power dissipation of the final device and the associated output components.

The maximum power transfer theorem states that maximum power is transferred to the load if the impedance of the load is equal to the output impedance of the source. While proven in a circuits course for DC sources and resistive loads, the theorem also hold true for AC circuits and loads with reactive components. If the source impedance is purely resistive then the matching network and load resistance combined must have the same equivalent resistance as the source resistance. In the case where the output impedance of the source is complex, the matching network and load combined must have an impedance that is the complex conjugate of the source impedance for maximum power transfer to occur.

The design of the matching network for impedance matching often requires a capacitor in parallel with the load and an inductor in series with the output of the source. This combination is analogous to an "L" network, which is a prototypical example of a low pass filter. Since the reactance of the inductor increases linearly with frequency and the reactance of the capacitor is inversely proportional to frequency, the net effect of the series L - shunt C combination is a voltage-divider type circuit whose output decreases as frequency increases. This circuit configuration will reduce the harmonic output of the amplifier.

The power not delivered to the load must be dissipated somewhere else in the circuit. In the case of a mismatched amplifier, this reflected power is dissipated in the output transistor, where it is converted to heat. The heat must be removed from the circuit, usually by attaching the transistor to a heat sink. As the transistor heats up, its gain increases, thereby creating more reflected power and more heat. If the heat sink cannot dissipate this heat, the internal temperature of the transistor will rise and lead to an effect known as thermal runaway. The result of thermal runaway is usually a burned out transistor.

Thus, impedance matching at the output of an amplifier circuit is necessary to ensure maximum power to the load while simultaneously reducing the power dissipated by the final device due to the mismatch. The matching circuit, itself, may act as a filter to reduce the harmonic output of the amplifier. Impedance matching of the amplifier to the load is an important part of electronic circuit design.

언어 학습 (빈칸 채우기, 순서 맞히기)

A. Unveiling masked words

Q) Fill in the blank

Training data: the man went to the store, he bought a gallon of milk.

Input: the man went to the [_____], he bought a [_____] of milk.

B. Identifying sentence orders

Q) Which sentence comes first?

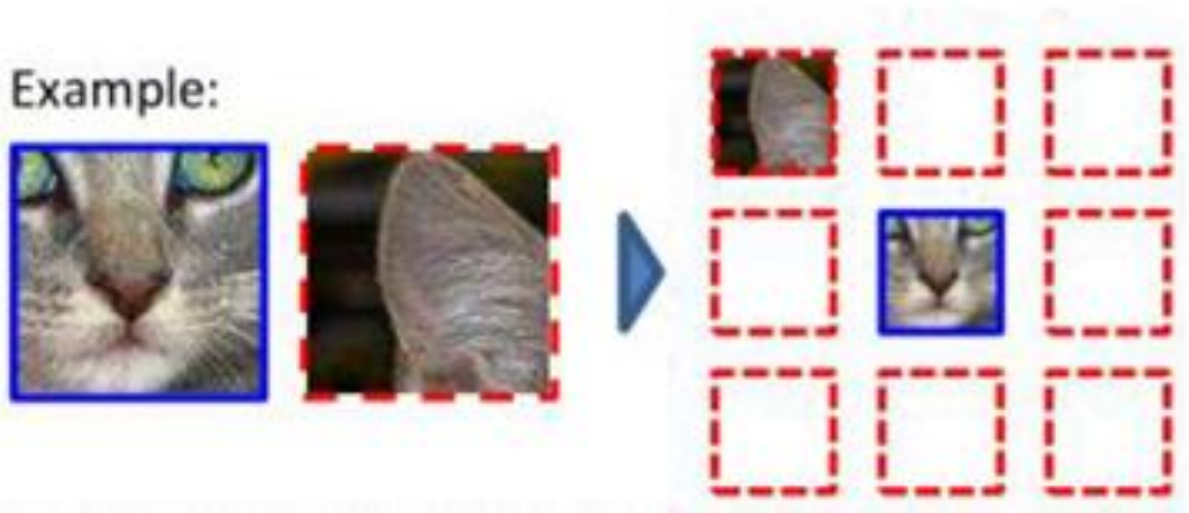
Sentence A: the man went to the store.

Sentence B: he bought a gallon of milk.

이미지 학습 (상대 위치 맞히기)

- Making a model understanding the context of images

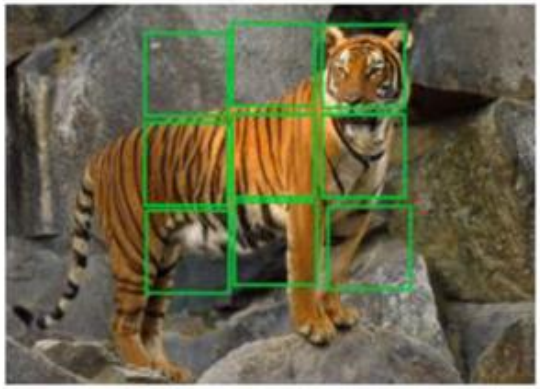
- Finding relative patch location (ICCV 2015)
  - Given the center patch image and second patch,
    a DNN model is trained to answer where the second patch should be positioned



이미지 학습 (순서 맞추기)

- How can we make a model that understands the context of images?

- Jigsaw Puzzle (ECCV 2016)
  - Scrambles the order of a sample image patches
  - DNN model is designed to find one correct order among 100 possible permutations



Sample image

Extract 9 patches

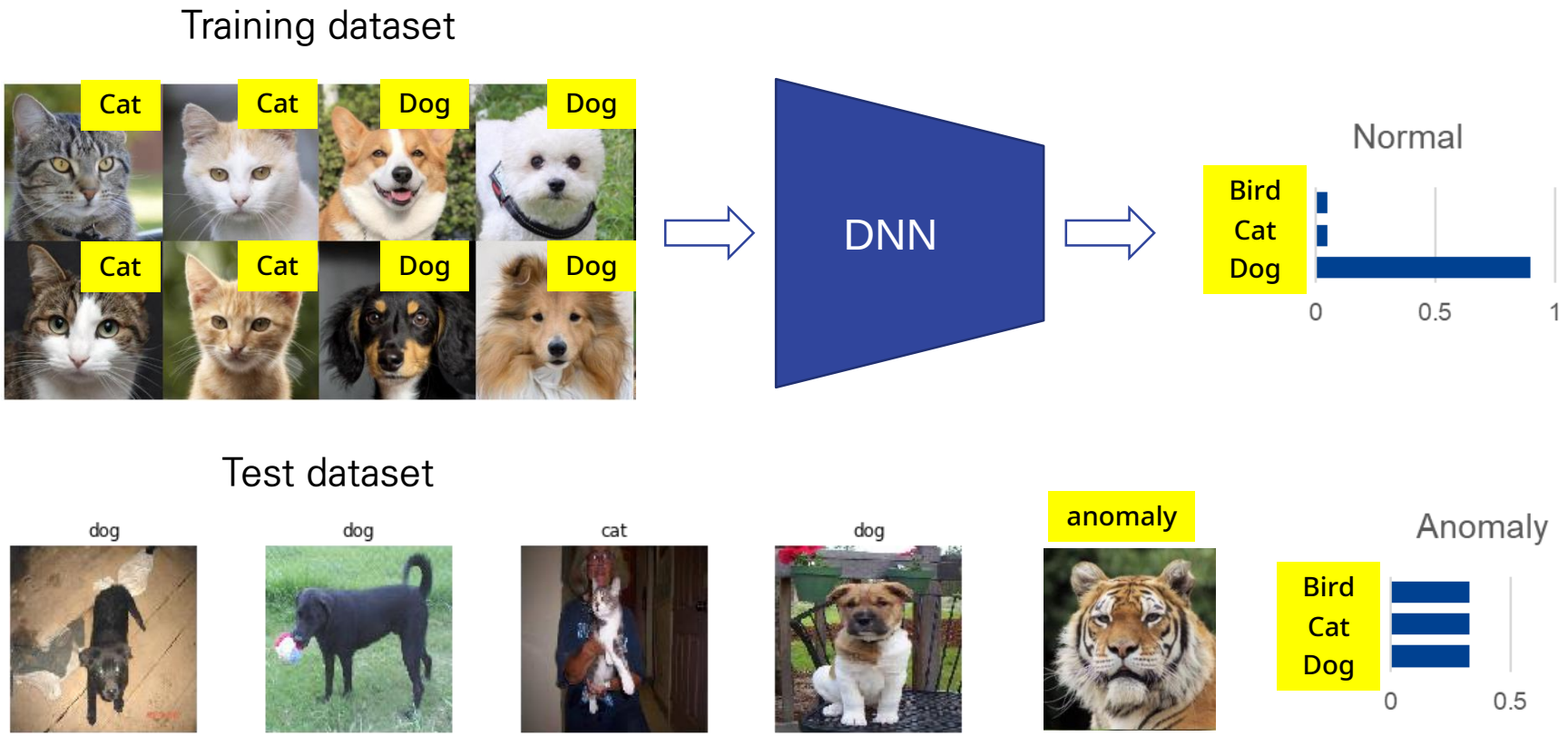
Index (0~99)
61

Permutation
9, 5, 8, 3, 2, 4, 7, 1, 6

Permutate 9 patches

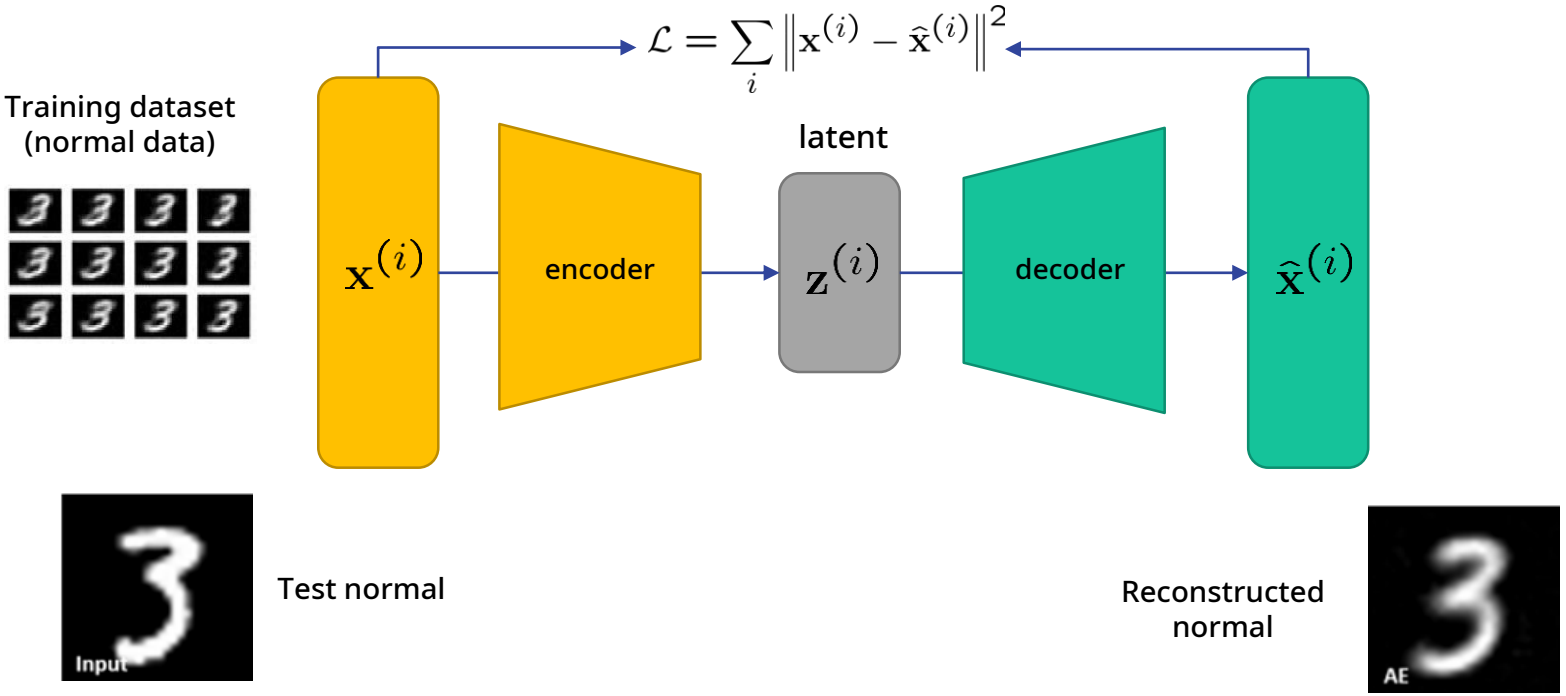
분류 학습 (classification task)

- When train data has internal labels
  - DNN learns important features of normal data from the classification task

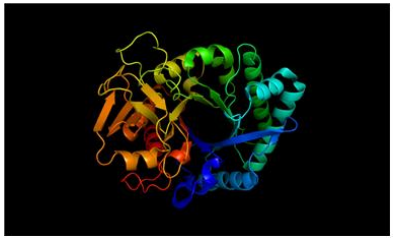


재현 학습 (reconstruction task)

- Autoencoder example



요약



- 자기지도학습 Self-supervised learning
  - Runs pretext tasks to learn the context of training (normal) data
  - Requires no prior information on anomalous data

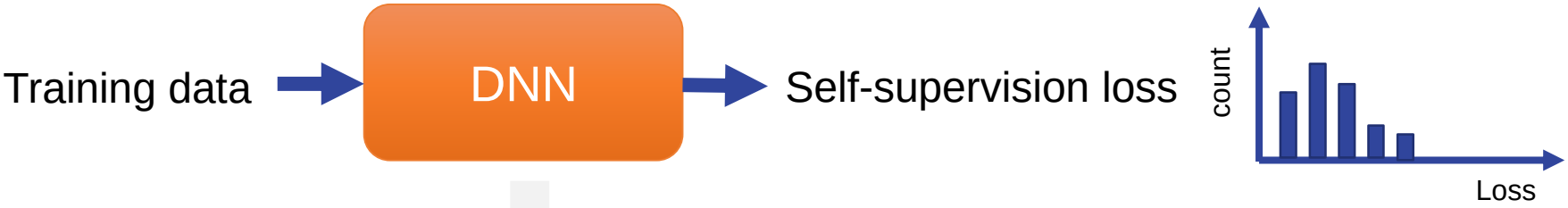
- Creativity matters!
  - Connectivity: spatial relation (relative/absolute position), temporal relation (order)
  - Low dimensional key features: autoencoder + reconstruction
  - Prior information on normal data
    - Classification: using internal labels
    - Augmentation within the boundary of normal data
      (rotation, flip, color change, clip: human's prior knowledge on normal data)
  - Devise your own way of self-study to understand the normal data better

학습된 모델의 이상 진단 수행 방법

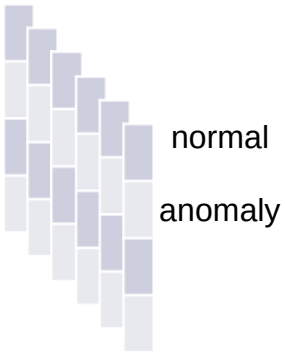
Anomaly detection using pre-trained model

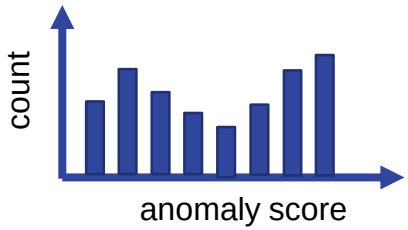
이상 진단 과정

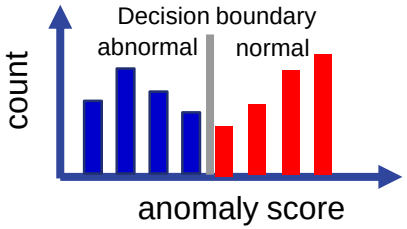
- Pretraining using pretext task



- Downstream task (anomaly detection)

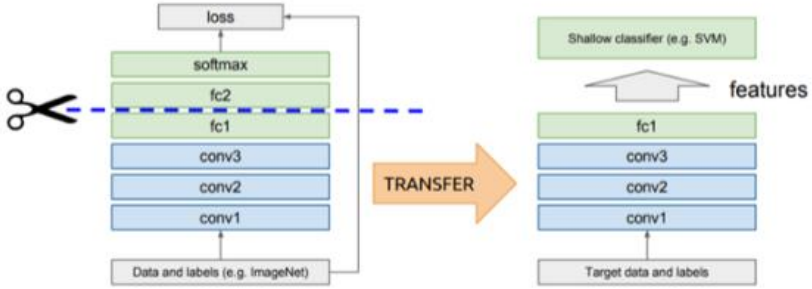






본 게임 Downstream task

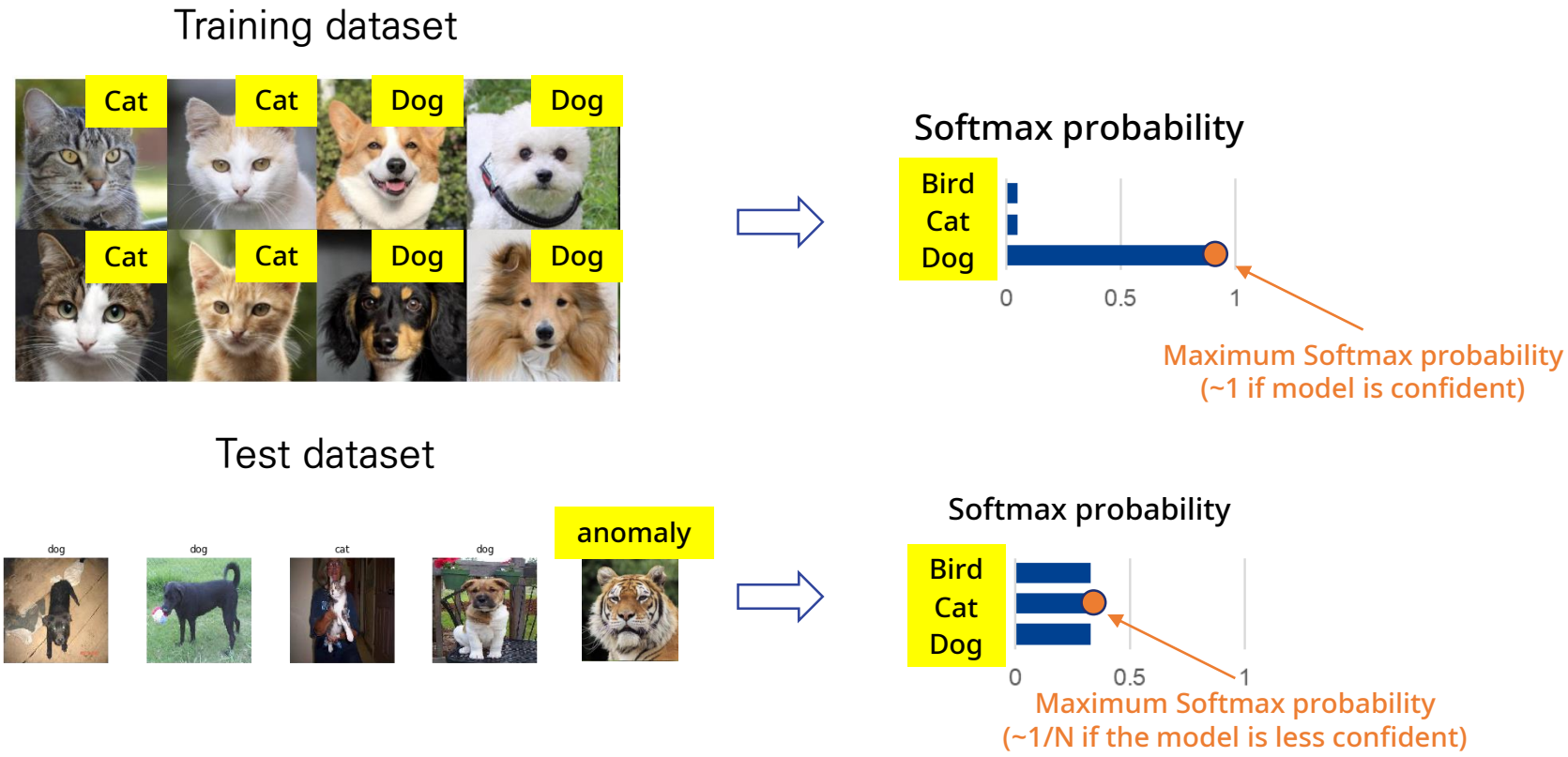
- Choice of downstream task
  - Usually, downstream task for anomaly detection is the same as the pretext task
  - For different tasks, various network architectures can be used



- Strategy
  - The model trained by normal data cannot solve the pretext task well for the anomaly data
  - The loss function evaluated for anomalous data will be positioned out of the normal data distribution

본 게임 Downstream task

- Example: classification task



예: 재현학습모델 Reconstruction task

- Autoencoder example