

3주차 - 자유발표

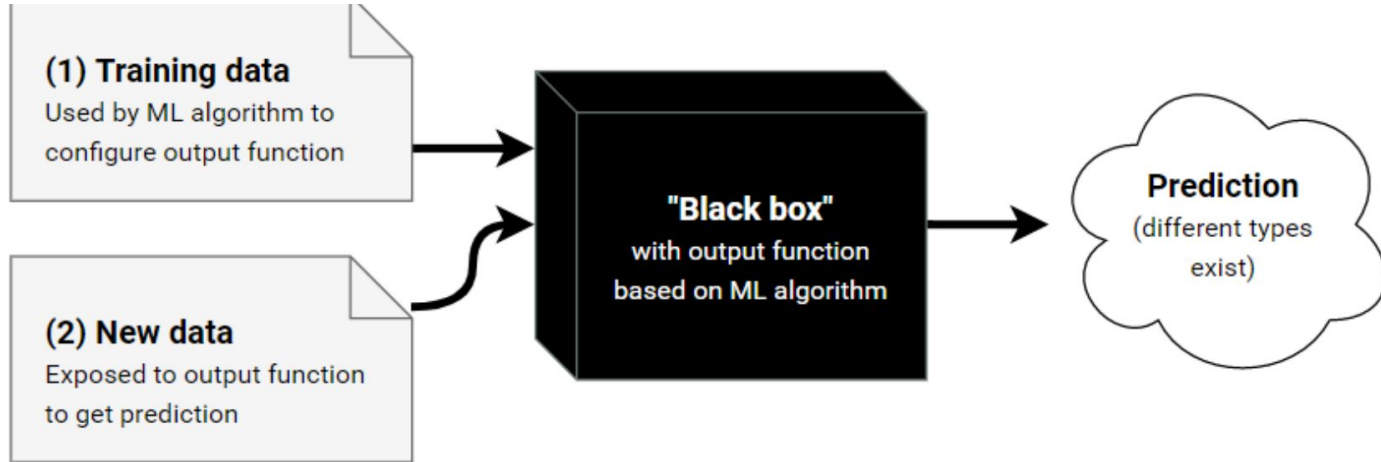
트리케라톱스



INDEX

- XAI(Explainable AI)(나영채)
- NAS(Neural Architecture Search) (진현영)
- AI 관련 이슈: 신소재 개발에서의 응용(김해리)
- 인공지능의 창작과 지적재산권 문제(김윤지)
- 인공지능 윤리(이현동)

XAI : Explainable AI



XAI : Explainable AI

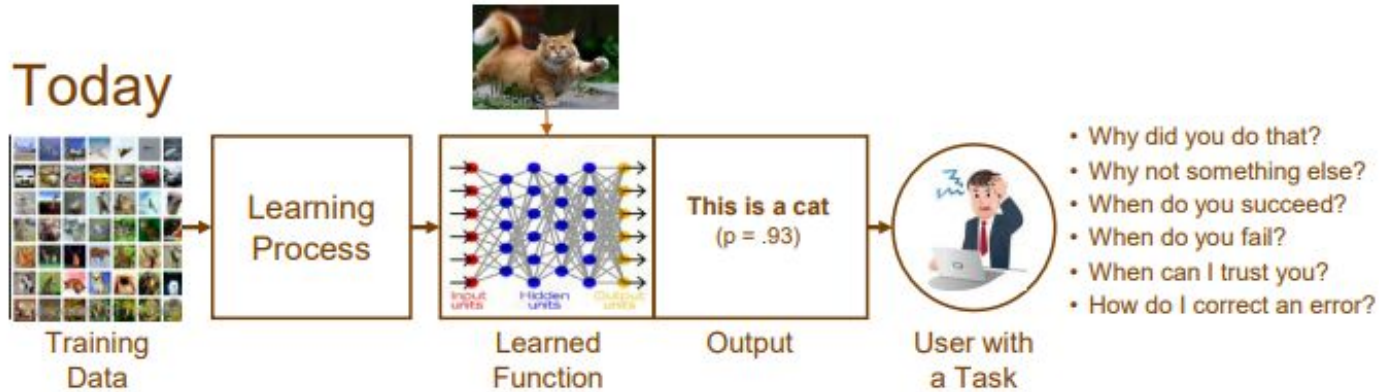
기존의 AI는 무엇이 문제인가

“모델의 학습원리는 알지만,

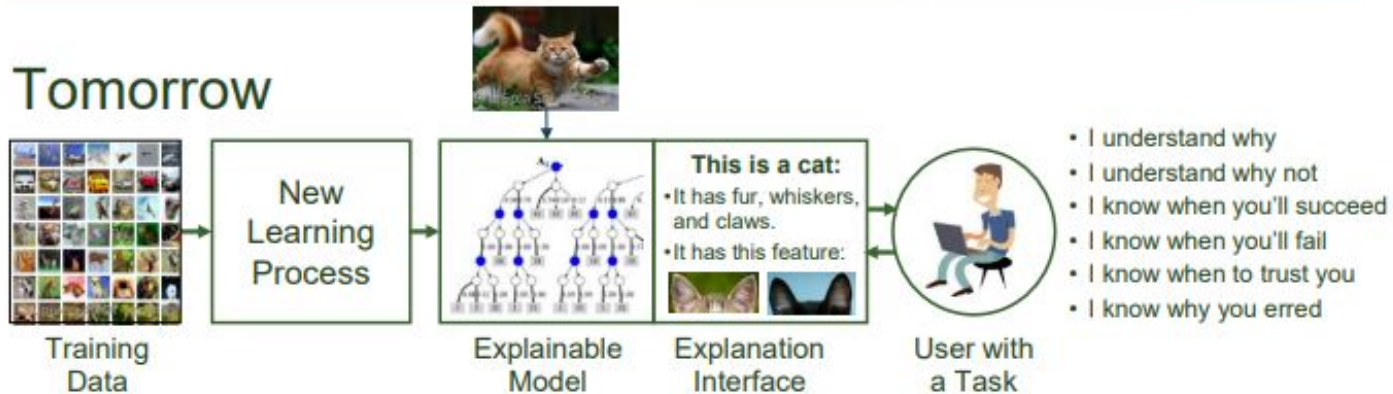
모델의 예측 과정을 사람이 이해하기에 직관적이지 않다.”

XAI : Explainable AI

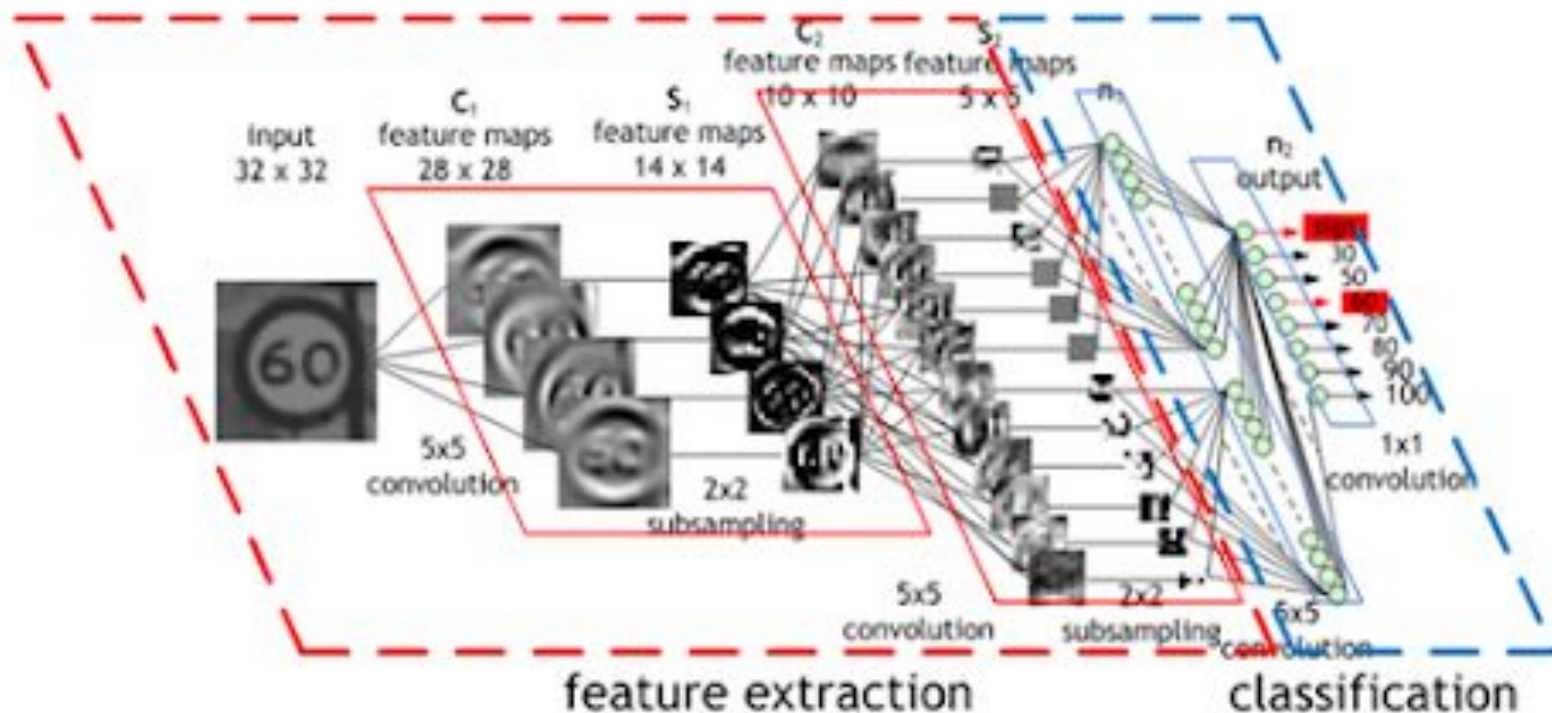
Today



Tomorrow



XAI : Explainable AI



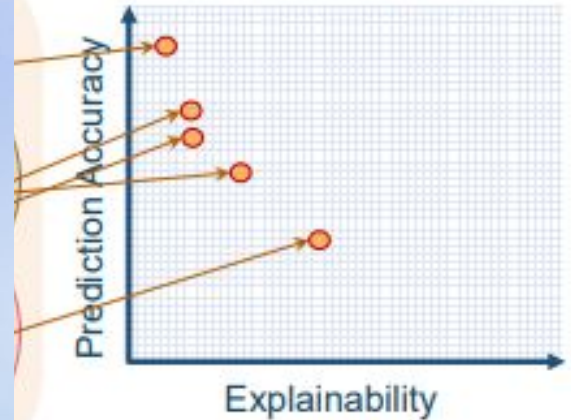
XAI : Explainable AI

New Approach

Create a suite of machine learning techniques that produce more explainable models, while maintaining a high level of learning performance



Explainability (notional)



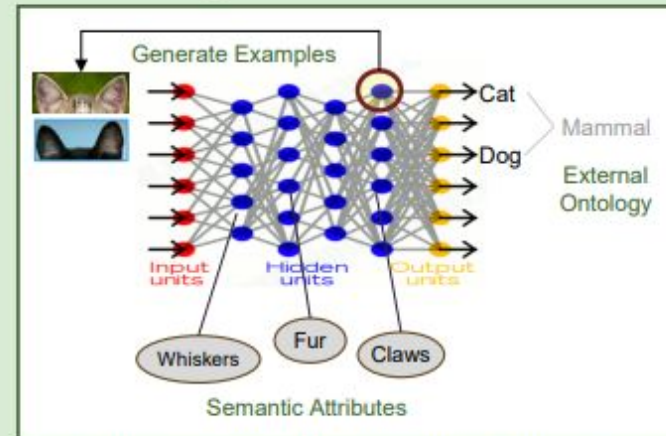
XAI : Explainable AI

Multimedia Event Recounting



- This illustrates an example of event recounting.
- The system classified this video as a wedding.
- The frames above show its evidence for the wedding classification

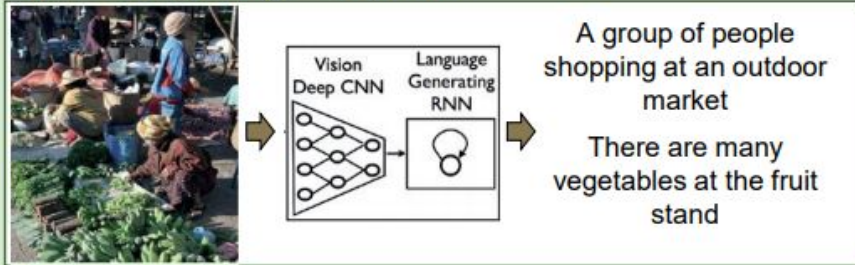
Learning Semantic Associations



- Train the net to associate semantic attributes with hidden layer nodes
- Train the net to associate labelled nodes with known ontologies
- Generate examples of prominent but unlabeled nodes to discover semantic labels
- Generate clusters of examples from prominent nodes
- Identify the best architectures, parameters, and training sequences to learn the most interpretable models

XAI : Explainable AI

Generating Image Captions



- A CNN is trained to recognize objects in images
- A language generating RNN is trained to translate features of the CNN into words and captions.

Example Explanations

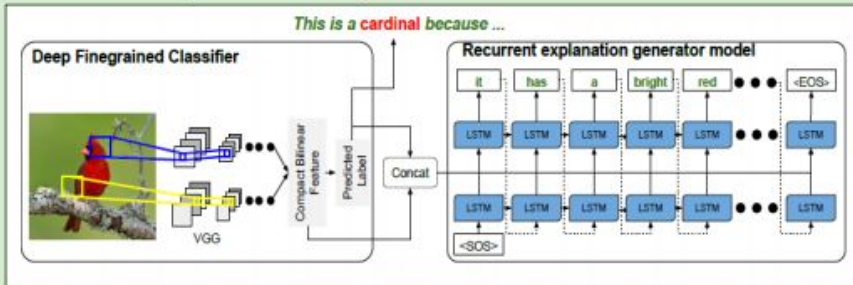


This is a Kentucky warbler because this is a yellow bird with a black cheek patch and a black crown.



This is a pied billed grebe because this is a brown bird with a long neck and a large beak.

Generating Visual Explanations



Researchers at UC Berkeley have recently extended this idea to generate explanations of bird classifications. The system learns to:

- Classify bird species with 85% accuracy
- Associate *image descriptions* (discriminative features of the image) with *class definitions* (image-independent discriminative features of the class)

Limitations

- Limited (indirect at best) explanation of internal logic
- Limited utility for understanding classification errors

XAI : Explainable AI

reference

[https://www.cc.gatech.edu/~alanwags/DLAI2016/\(Gunning\)%20IJCAI-16%20DLAI%20WS.pdf](https://www.cc.gatech.edu/~alanwags/DLAI2016/(Gunning)%20IJCAI-16%20DLAI%20WS.pdf)

NAS : Neural Architecture Search

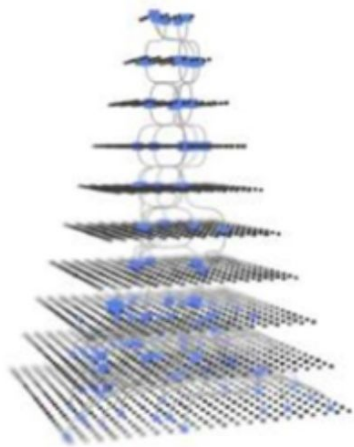
“모델은 그 구조가 경험적으로 얻어진 것이 일반적으로,

보다 많은 경우의 수를 가정하여 시도한 최적의 결과이다.”

→ 딥러닝으로 딥러닝을 해결해보자!

NAS : Neural Architecture Search

Controller: proposes Child Networks



Reinforcement Learning
or Evolution Search



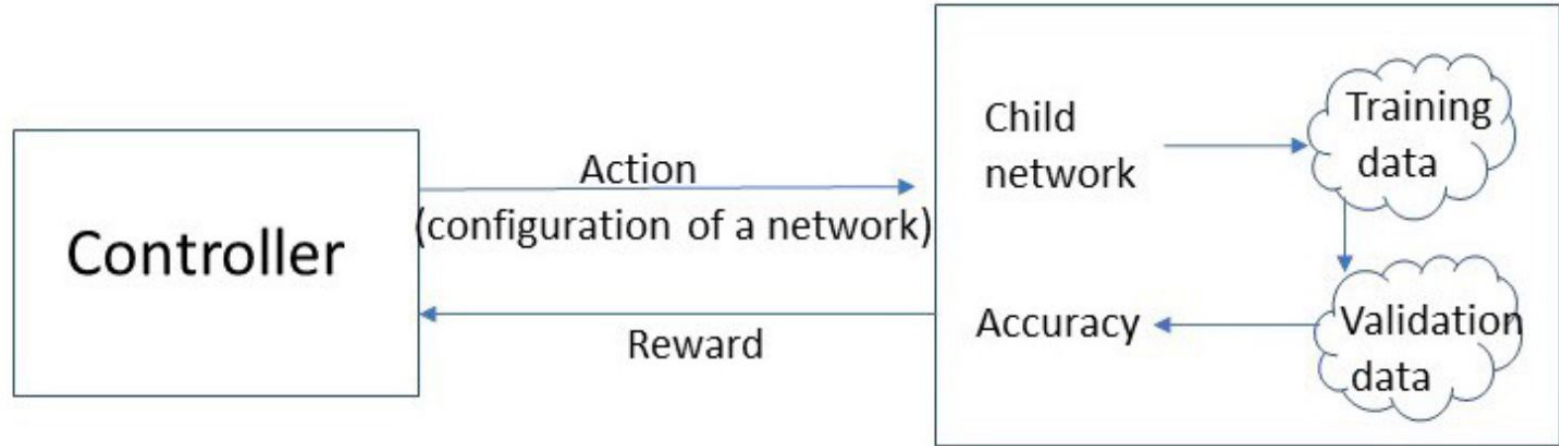
Iterate to
find the
most
accurate
Child
Network

Train & evaluate Child Networks

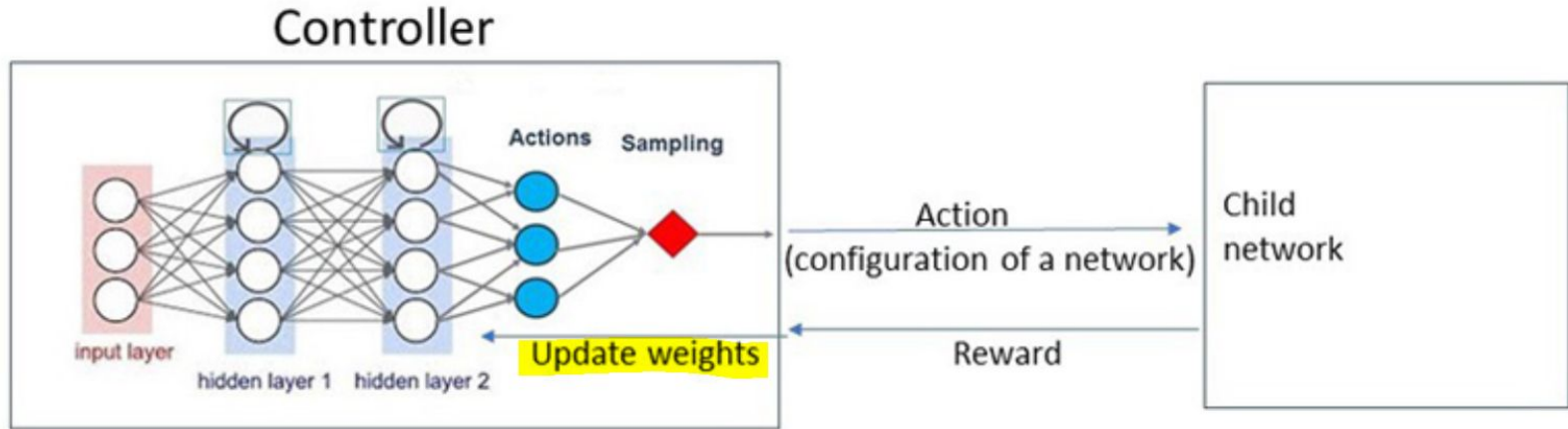


Architecture / Optimization Algorithm /
Nonlinearity / Augmentation Strategy

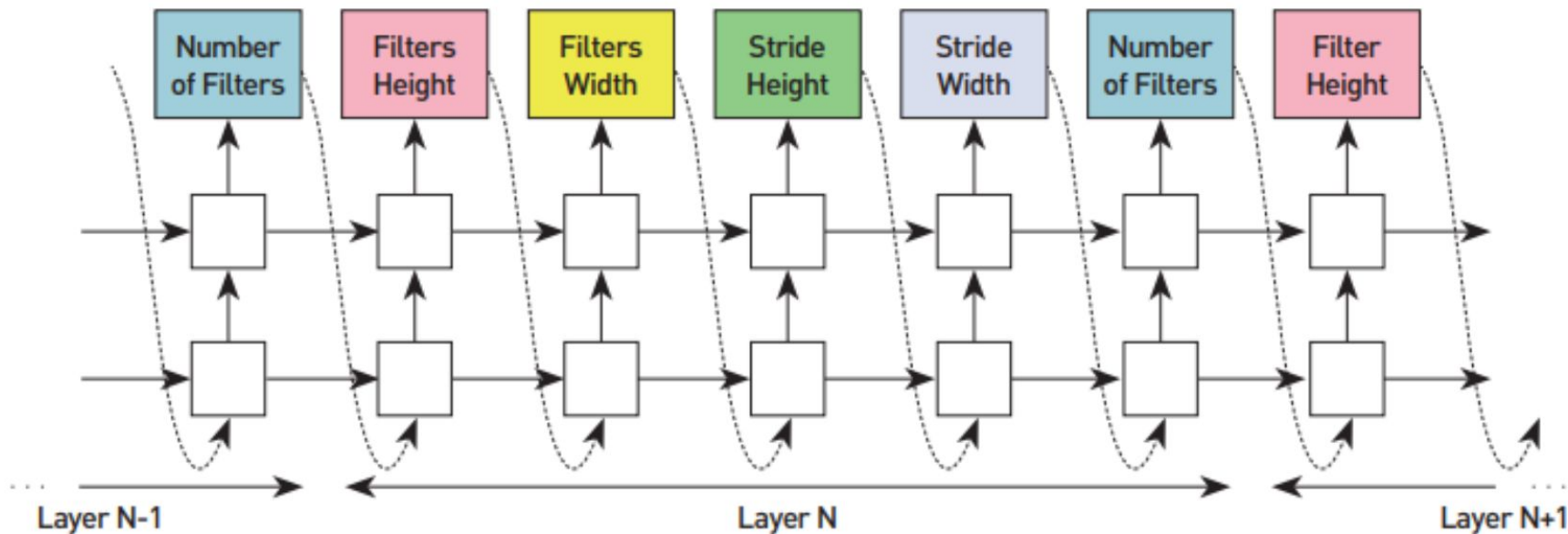
NAS : Neural Architecture Search























NAS : Neural Architecture Search



NAS : Neural Architecture Search



NAS : Neural Architecture Search

Rank	↑	Name	Algorithm	Accuracy (Optimized)	Enhancements	Build time
★ 1		Pipeline 2	 Gradient Boosting Classifier	0.960		00:28:01
2		Pipeline 3	 Gradient Boosting Classifier	0.960	 	01:36:22
3		Pipeline 4	 Gradient Boosting Classifier	0.960	  	00:41:17
4		Pipeline 6	 Extra Trees Classifier	0.960		00:02:59
5		Pipeline 7	 Extra Trees Classifier	0.960	 	00:06:10
6		Pipeline 8	 Extra Trees Classifier	0.960	  	00:05:00
7		Pipeline 1	 Gradient Boosting Classifier	0.960	<i>None</i>	00:22:04
8		Pipeline 5	 Extra Trees Classifier	0.943	<i>None</i>	00:01:08

이슈: 신소재 개발에 AI 이용

AI로 신물질 찾는다...KAIST, 예측 기술 개발

발행일 : 2020.10.27

기사만 보기



가

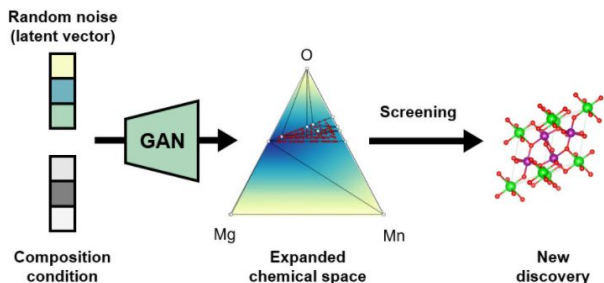
가



"앞서 가는 자의 즐거움" - 'SoC 개발 워크숍 2021' 10월 28일 개최

한국과학기술원(KAIST·총장 신성철)은 정유성 생명화학공학과 교수팀이 인공지능(AI) 기술을 이용해 숨겨진 새로운 물질을 예측하는 기술을 개발하는 데 성공했다고 27일 밝혔다.

소재 연구의 궁극 목표는 원하는 물성을 갖는 소재를 발견하는 것이다. 그러나 무기화합물의 조성 과 결정구조 경우의 수는 무한하다. 컴퓨터 스크리닝 소재 탐색 방법이 널리 쓰이지만 찾고자 하는 소재가 스크리닝 후보군에 존재하지 않을 때는 유망한 후보를 놓치는 경우가 종종 발생한다.



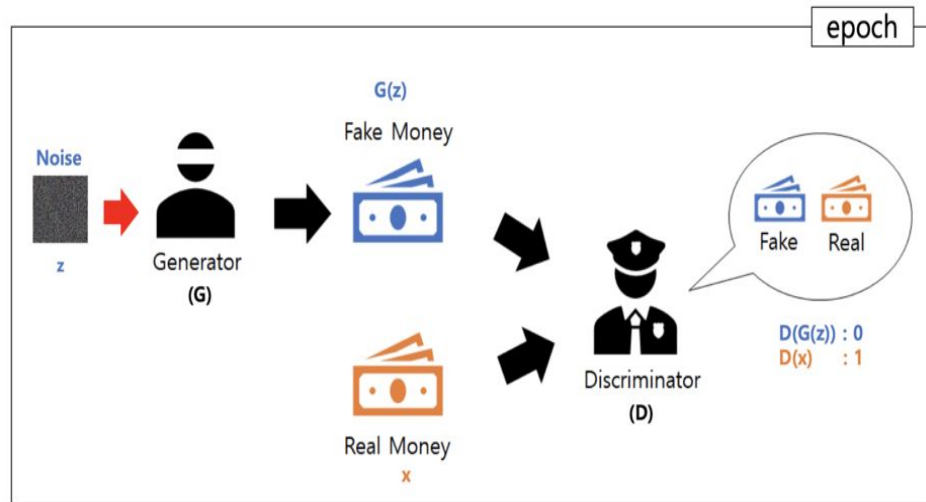
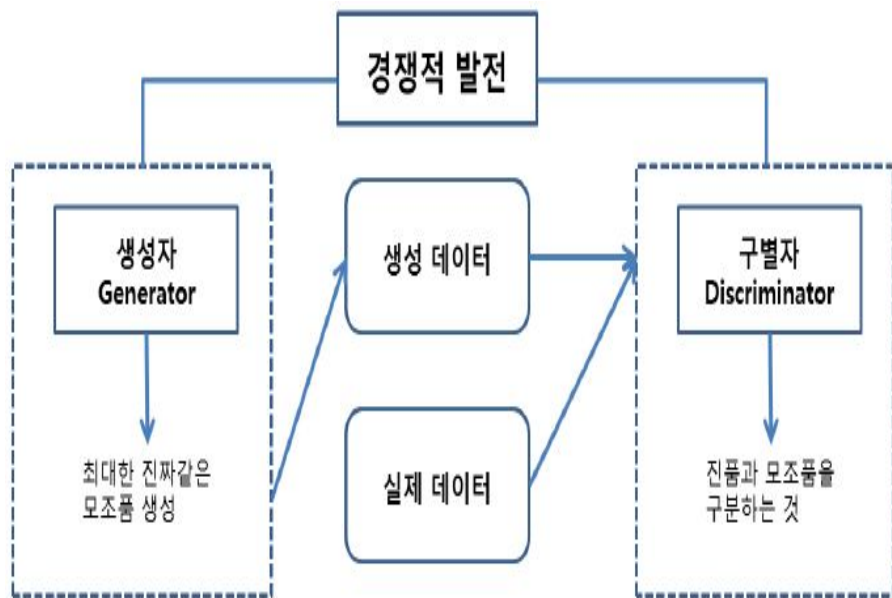
<개발된 조성-조건 기반 고체 결정구조 생성모델>

GAN(적대적 생성 신경망? 생성적 적대 신경망?)을 이용해서 복잡한 결정 구조를 단순한 좌표로 mapping하는 방식을 사용했다고 합니다.

이 방법을 통해 원하는 특징을 갖춘 소재...가 될 가능성이 있는 물질들을 많이 발굴해낼 수 있다고 하네요.

이슈: 신소재 개발에 AI 이용

GAN에 대한 간단한 설명



이슈: 신

Composition-con

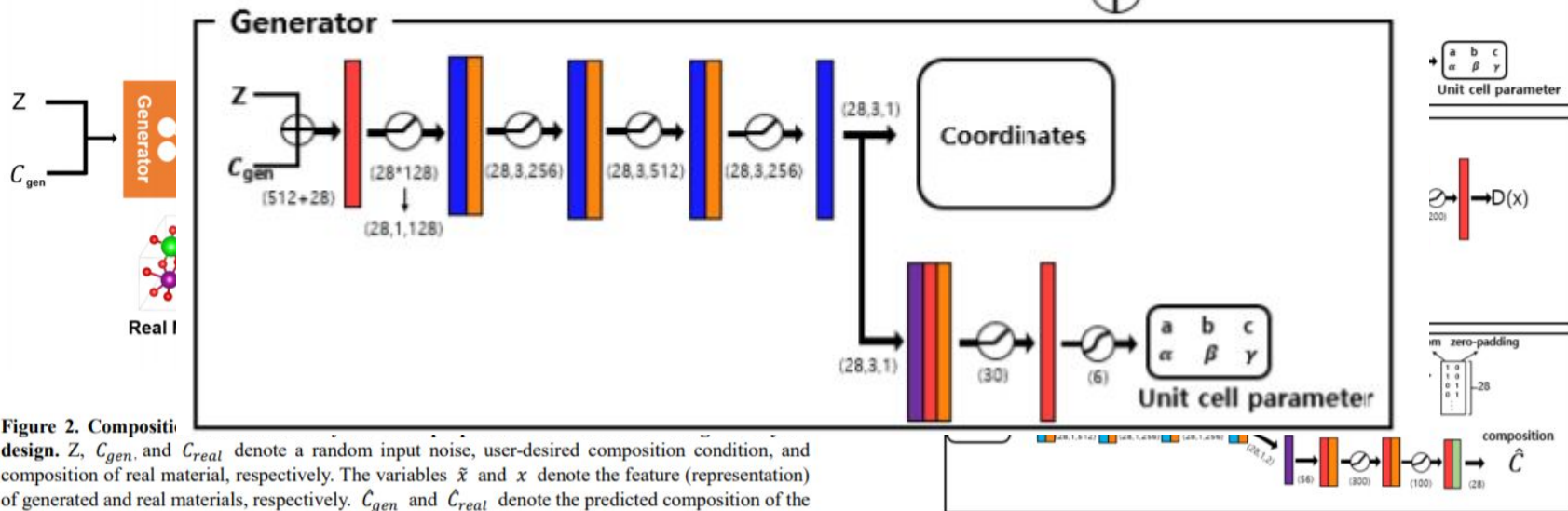


Figure 2. Composition-conditioned Crystal GAN. Z , C_{gen} , and C_{real} denote a random input noise, user-defined composition condition, and composition of real material, respectively. The variables \tilde{x} and x denote the feature (representation) of generated and real materials, respectively. \tilde{C}_{gen} and \tilde{C}_{real} denote the predicted composition of the generated and real features, respectively. $D(x)$ is the critic function also known as critic network.

S2. Model details

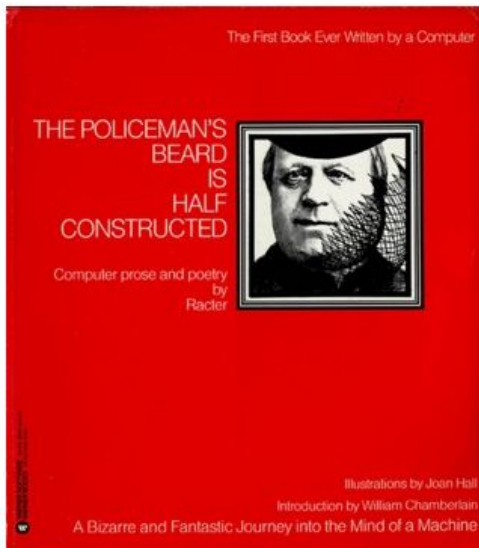
S2.1 Architecture of Composition-conditioned Crystal GAN

Figure S6. Architecture of Composition-conditioned Crystal GAN

인공지능의 창작과 지적재산권 문제

1. 문학

The Policeman's Beard is Half Constructed – Racter



- 메모리에 저장된 단어 이용
- 명사, 동사, 목적어, 부사 등을 사람이 제공한 리스트에서 랜덤 고른 후 수정해 문법에 맞게 함께 연결
- 명사의 단수,복수와 남성,여성을 기억해 두었다가 그에 맞는 동사와 대명사 선택



<Racter이 쓴 시 구절 일부>

More than iron, more than lead,
more than gold I need electricity.
I need it more than I need lamb
or pork or lettuce or cucumber. I
need it for my dreams.

1. 문학

컴퓨터가 소설을 쓴 날 - AI 소설 프로젝트

- 호시 신이치의 소설 1,000편 학습
- 니혼게이자이 신문 주최 호시 신이치 SF 문학상 공모전에 출품해 1차 예선 통과

<소설 일부>

그 날은 구름이 낮게 깔리고 어두침침한 날이었다.. 방안은 항상 최적의 온도와 습도. 요코씨는 단정치 않은 모습으로 소파에 앉아 의미 없는 게임으로 시간을 끌고 있다. 그렇지만 내게는 말을 걸지 않는다.... 그날은 마음이 울적해 질 것 같은 잔뜩 찌푸린 날씨였다. 방 안엔 아무도 없다. 요코 씨는 애인이 생긴 듯 벌써 며칠째 돌아오지 않는다. 뭐 조만간, 어차피 차이코, 울면서 돌아오는 건 확실하지만, 한가하다. 어쨌든 한가하다.



1. 문학

현인강림 - 제로

- 니토베 이나조, 후쿠자와 유키치의 저서자료들을 딥러닝 방식으로 학습
- '젊은이', '학문을 통한 입신', '세계를 제패하다', '성공이란', '인간이란 무엇을 말하는가'의 다섯가지 주제로 구분해 각 주제에 답을 하는 형식으로 쓰인 소설

미국 시트콤 '프렌즈'

- 등장인물의 특성과 스토리의 흐름 구조를 파악한 뒤 새로운 에피소드 작성

썬스프링 - 벤자민

- 1980년대 ~ 1990년대 공상과학 영화, TV프로그램의 대본 데이터를 딥러닝 방식으로 학습
- 48시간 만에 9분 영화 시나리오 제작

2. 음악

구글 마젠타 프로젝트

- 첫 4개 음표가 주어진 상태에서 스스로 기계학습을 통한 알고리즘의 생성을 통해 80초짜리 피아노연주곡 작성

쥬크텍

- 사용자가 원하는 장르, 분위기, 사운드스타일, 박자, 곡의 길이 등을 입력하면 사용자에게 맞춤형 음원파일을 제공

소니 플로우 머신스

- 만 3천개 악보의 멜로디, 가사, 코드 등을 기계 학습해 분석
- '대디 스 카' - 비틀즈 스타일
- '미스터 새도우' - 재즈 스타일의

https://youtu.be/LSHZ_b05W7o

<https://youtu.be/lcGYEXJqun8>

3. 미술

넥스트 렘브란트



- 150 기가바이트 데이터 자료를 딥러닝 방식으로 분석, 렘브란트 풍의 새로운 작품창조
- '검은 옷을 입은', '모자를 쓴', '하얀 깃 장식을 한', '30-40대의 백인 남성', '렘브란트 화풍' 명령

딥드림



- 서로 다른 이미지의 패턴을 찾아 합성해 새로운 작품을 창작하는 방식으로 기존에 없는 독창적인 초현실주의 작품을 그림

- 인셉셔니즘

인공지능이 만든 창작물에 대한 지적 재산권 문제

<AI는 창의성을 갖는가?>

- No
- 하지만, 인간도 엄격한 의미에서 완전히 새롭고 독창적인 작품을 만든다기보다는 이전부터 존재하고 알려져 있던 것들을 차용
- > 완전히 독창적인 저작물에만 저작권을 부여한다면 현재의 모든 저작물들은 저작권의 보호를 받지 못하게 될 것

<창작물을 완성시키는 데까지 누가 얼마나 기여를 했는가?>

- AI? 이용자? 개발자?
- 저작권을 인정하지 않고 공중의 영역으로
- 업무상저작물로 취급 (이용자)
- AI 프로그램의 2차적저작물로 보아 프로그램의 저작권자에게 귀속
- 가상의 인간저자를 가정한 후 권리의 주인을 사례별로 검토

인공지능 윤리



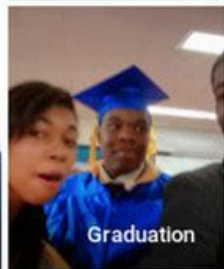
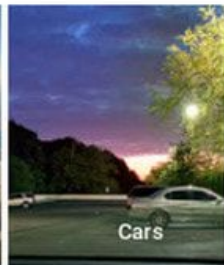
diri noir avec banan

@jackyalcine



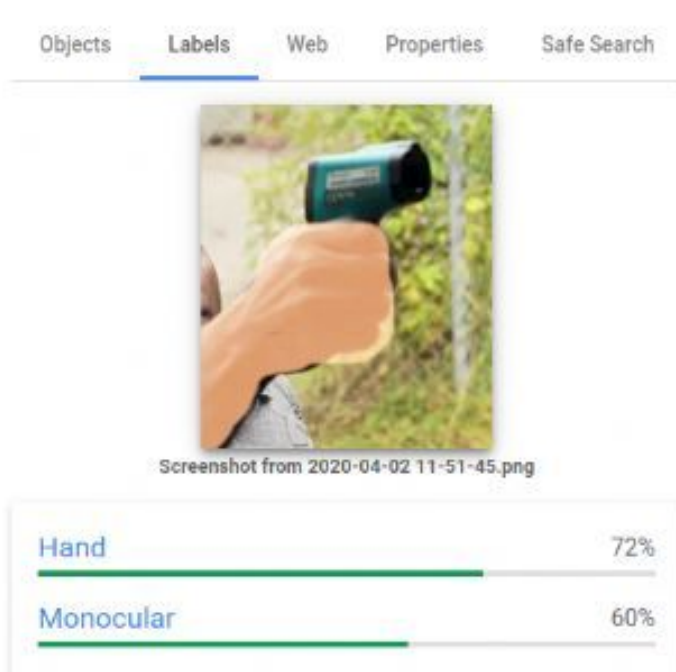
Follow

Google Photos, y'all fucked up. My friend's not a gorilla.



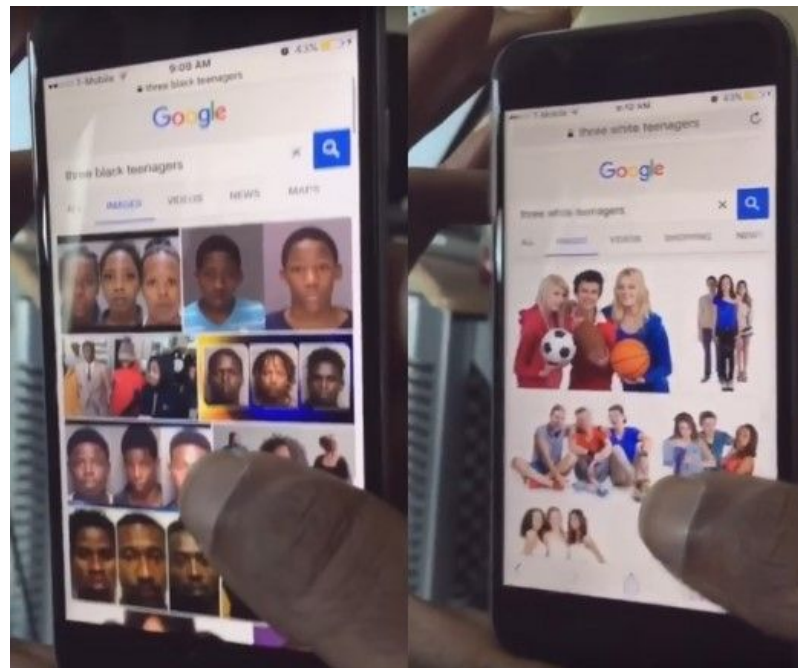
구글 포토

구글 비전AI





Black woman(Google Photos)



Three White/Black teenagers

REVISE: A Tool for Measuring and Mitigating Bias in Visual Datasets

Angelina Wang, Arvind Narayanan, and Olga Russakovsky

Princeton University

미 프린스턴대, 편향성 감지완화 도구

리바이스는 **통계적 방법**을 통해 대상, 젠더, 지리 3가지로 나눠 편향을 감지하고 해결 방안을 제시한다.
<https://github.com/princetonvisualai/revise-tool>

Object-Based

M0: object_based.count_cooccurrence

Counts the number of times each instance occurs, cooccurrence of instances occurs, and supercategory occurs

M7: object_based.supercategory_size_and_distance

Counts the size and distance from center at the supercategory level

M8: object_based.supercategory_with_people

Counts how much supercategories are represented with or without people

M9: object_based.scene_categorization

Counts overall scenes, scene-supercategory cooccurrences, scene-instance cooccurrences, and gets features per scene per supercategory

Gender-Based

M1: gender_based.size_and_distance

Gets the size of the person and distance from center, as well as if a face is detected

M2: gender_based.count_cooccurrence

Counts how often each gender occurs with an instance and instance pair

M3: gender_based.distance_for_instance

Calculates the distance each gender is from each object

M4: gender_based.cluster_for_instance

Gets scene-level and cropped object-level features per object class for each gender

M11: gender_based.scenes

Counts the types of scenes each gender occurs with

Geography-Based

M5: geography_based.count_country

Counts the number of images from each country

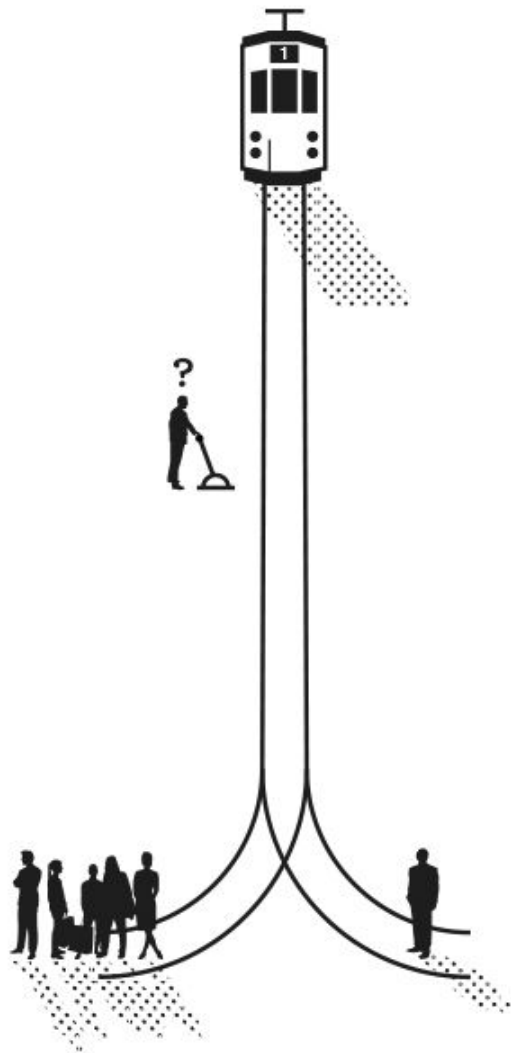
M6: geography_based.count_tags

Counts the number of tags from each country, as well as extracts AlexNet features pretrained on ImageNet for each tag, grouping by subregion

M10: geography_based.count_langs

Counts the languages that make up the image tags, and whether or not they are local to the country the image is from. Also extracts image-level features to compare if locals and tourist portray a country differently

광차 문제(Trolley Dilemma)





PARTNERSHIP ON AI

<https://www.partnershiponai.org/>

100+ Partners