

강봉균: local plasticity rules can learn deep representations using self-supervised contrastive predictions

- 목적: 이 논문의 목적은 어떻게 뇌가 deep hierarchical representation 을 학습하는지 그 방식을 제안한 것. Method: 새로운 아이디어인 layer-wise contrastive predictive coding 을 이용하여 positive 와 negative feature 학습을 실시하게 됩니다. 기존 방식으로는 마지막 loss function 의 값을 통해 error signal 을 산출하는 방식이나, 제안한 방식은 레이어별로 loss 를 계산하여 contrastive learning 을 하는 novelty 를 가집니다. 실험: t-SNE 값을 통해 서로 특징이 뚜렷한 class 는 구분이 잘 되도록 학습이 이루어진 것을 확인할 수 있었습니다.

구정현: Self-supervised Graph-level Representation Learning with Local and Global Structure.

- 목적: 그래프 구조는 유연해서, 레이블이 달려있는 정보가 부족하기 때문에 supervised 학습이 중요함. 이를 어떻게 하면 그래프 정보를 잘 담아내어 학습을 이루어낼지에 목적이 있습니다. Method: Local instance + Global semantic structure : 이를 모두 함께 학습하는 방법. Local instance: 비슷한것들끼리는 가깝게, 아니면 멀리 떨어지는 공간이 되도록 학습. 각 그래프가 매핑된 공간내에서 로컬 인스턴스 학습은 서브그래프의 pair 와 코사인 유사도를 목적함수로 두어 계산. Global semantic structure: EM 알고리즘을 이용하여 cluster 을 생성하고 그 국소적인 center cluster 들을 이용하여 hierarchical semantic cluster 을 또 학습하는 방식. 한번에 argmax 하는 latent variable 을 글로벌하게 semantic cluster center 을 계산할 수 없기 때문에 EM 을 이용하여 진행. 실험: ZINC 데이터를 이용하여 결과 도출. global semantic structure 이용한 것이 더 잘 나옴.

김민범: CoLV: A Collaborative Latent Variable Model for Knowledge-Grounded Dialogue Generation

- 목적: knowledge-Grounded 다이얼로그를 이용하여 저녁을 뭘 먹을까 고민할 때와 같이, 같은 representation 에서 다른 답변이 나올 수 있고, 다른 knowledge 에서 같은 답변이 나올 수 있기에 이를 고려한 학습 방식을 제안. Method: representation 을 통해 추론할 때의 KL divergence 를 최소화하게 학습. 지식과 응답의 diverse inference 를 가능하게 학습. 다이얼로그 하나만 고려한 것이 아닌, context 를 고려한 knowledge 를 이용할 수 있다고 함. 실험: jointly 하게 진행했을 때 성능이 잘 나왔고, case study 에서도 같은 문맥에서 발생할 수 있는 다양한 semantic 한 문장을 생성할 수 있다고 함. novelty: End-to-end jointly learn knowledge and response representation, Reflect various relationship between knowledges and responses, Variational inference for diverse semantic responses.

김연아: InterFaceGAN Interpreting the Latent Space of GANs for Semantic Face Editing

- 목적: 이 논문은 가우시안 분포를 따르는 레이턴트 변수 Z 와 관련되어 있음. 잘 학습된 GAN의 latent space 내의 encoded된 semantics를 분석하는 것. 더불어 semantic image editing까지 보여줌. Method: 생성자에 Z 가 들어가면 X 를 생성. 이 X 를 scoring할 수 있는 semantic scoring 함수 f_s 를 정의. 여기서 SVM을 이용하여 semantic을 조정할 수 있음. 이렇게 semantic label을 이용하여 linear boundary를 조정할 수 있다고 함. 실험: 하나의 semantic을 바꿀 때 다른 semantic을 고정할 수 있는 실험도 진행.

김정현: Long-short Distance Aggregation Networks for Positive Unlabeled Graph Learning

- 목적: positive, unlabeled된 node가 있는 그래프가 주어졌을 때, binary classifier model을 리턴하는 task. 이때 binary classifier는 node가 positive인지 negative인지 분류, negative는 unlabeled되어있는 노드들중에서만 선별해야함. Method: Short-Distance Attention, Long-short Distance Attention, Unbiased PU Learning 이용하여 attention을 이용한 PU 학습이 가능하게 함. 실험: Citeseer, DBLP에서 F1 score 비교했을 때 outperform.

김정현: Revisiting the Calibration of Modern Neural Networks

- 목적: Analysis on Calibration with image classification SOTA models. Method: Calibration, Discrepancy between Model confidence and Accuracy를 의미. 즉, confidence가 8-90 퍼센트일 때, 성능은 50%밖에 안되는 상황을 말함. Expected Calibration Error: 수식 계산을 통해 calibration error 산출. Reliability Diagrams: 다이어그램을 통해 도표화하였을 때 성능과 confidence 차이를 쉽게 도표화. Reliability Diagrams: T를 점진적으로 증가하여, calibrated probability k^{-1} 을 생성. 이 calibration은 조건부 risk를 계산하는데 있어서 고려해야할 중요한 요인이 됨. 실험: SOTA 알고리즘, MLP-Mixer, ViT, BiT, ResNeXt WSL, SimCLR, Alexnet, Guo et al. 등과 위 measure 방법 3가지를 이용하여 비교 및 분석 진행. 이때 calibration이 잘 되었던 모델들은 고의적으로 distribution shift가 일어난 데이터에 대해서도 안정적으로 ECE를 많이 바뀌지 않음을 알 수 있었음. Corruption severity(학습과 테스트의 차이정도)가 클때 ECE가 크게 나타났었음. novelty: 다양한 데이터셋, 데이터셋의 impact, 학습과정등을 조절해가며 비교분석 진행.

김종완: SDE-Net: Equipping Deep Neural Networks with Uncertainty Estimates

- 목적: Equipping DNN 을 이용하여 불확실성을 estimate. DNN 은 계산된 예측값에 대해서 uncertainties 를 정량화하기 어려움. 따라서 이를 계산하는 방법으로 BNN, Non-bayesian 접근방법이 있음. 둘다 각각의 장단점이 있으며 이 둘의 한계를 잘 극복한것이 논문의 제안점. Method: ODE 는 deterministic 하여 uncertainty 를 추정하지 못했음. stochastic 한 방법(SDE)을 사용. novelty: SDE-Net 은 불확실성과 인식적 불확실성을 명시적으로 모델링하고 예측에서 불확실성의 두 가지 원인을 분리할 수 있었음. BNN(베이지안)에서와 같이 모델 사전 분포를 지정하고 사후 분포를 추론할 필요가 없으므로 구현이 효율적이고 간단. 또한, 분류 및 회귀 작업 모두에 적용 가능.

김태규: An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale

- 목적: ViT 라고, 현재까지의 비전 task 는 CNN 을 중심으로 진행하였음, 이를 Transformer 구조를 사용하여 새로운 method 를 제시하였음. 물론 이전에도 이러한 시도는 있었지만 여기서 처음 유의미한 결과를 냄. Method: Attention 은 Decoder 의 특정 time-step 의 output 이 encoder 의 모든 time-step 의 output 들 중, 어떤 output 과 가장 연관이 있는가로 이는 Seq2seq 에서 처음 제시된 개념임. RNN, LSTM 을 사용하지 않고 어텐션을 이요해서 데이터 그 자체만으로 time sequence 역할을 할 수 있게 하는 모델이 Transformer 임. 또한 source 와 target 간 관계는 기존 어텐션으로 알 수 있었지만 source 그 자체만으로 언어의 관계에 대해서는 self-attention 을 이용하여 그 관계를 계산함. ViT 는 트랜스포머와 동일하게 진행하는데, 이미지를 쪼갠 Patch 를 토큰처럼 이용하여 이미지 classification 을 풀어냄. 실험: Pre-trained 한 데이터셋의 크기가 클수록 ViT 의 정확도를 향상함을 알 수 있었고, 이는 CNN 과 다르게 convolutional inductive bias(locality, translation equivariance)가 없기 때문이며, 데이터셋의 크기를 키울수록 성능이 좋아졌음.

김희원: Contrastive Representation Distillation

- 목적: Original KD(Knowledge Distillation)**는 미리 잘 학습된 Teacher network 의 지식을 실제로 사용하고자 하는 Student network 에게 전달하는 것이 목적. Teacher network 로부터의 soft 한 output 확률 분포를 Student 가 따라가도록 학습하게 됨. 여기서 는 KD 를 임베딩 공간에서 진행하겠다는 것이고, contrastive objective function 을 이용해서 진행. Method: (Teacher's representation)Target 와 (Student representation)source 의 Mutual information $I(T;S)$ 를 최대화하도록 critic 목적함수를 $\arg\max$ 하는 f_s 와 \max h 를 찾게끔 함. 이렇게 함으로서 MI bound 를 Maximize 하는 representation 을 학습함. 이렇게 Contrastive 학습이 된 임베딩을 student 에게 제공.

류정현: Don't Stop Pretraining: Adapt Language Models to Domains and Tasks

- 목적: 언어모델에서의 domain, task 의 adaptation 을 고찰 Adaptively 하게 domain 에서 pretrain 을 진행, Task-Adaptive pretrain 을 진행, Task-adaptive pretraining 된 학습데이터 augmenting. Method: RoBERTa 를 기준모델로 두어 실험 진행, 4 domain, 8 tasks 로 실험 셋팅. 실험: DAPT 가 RoBERTa 보다 모든 면에서 더 좋은 성능을 보임. 도메인에서 adaptive 하게 pretrain 의 중요성을 환기시킴. TAPT 는 DAPT 보다 적은 데이터셋 사용하여 연산이 적은데 비슷한 성능을 보임. Human curated-TAPT 는 사람이 진행, Automated Data selection for TAPT 는 task relative adaptation 과 관련된 것임.

박재현, image Deraining : Task of removing rain from single image

- Unlabeled real-world 데이터를 training 과정에 활용함으로써 semi-supervised learning 수행하고, Gaussian process 기반 pseudo labeling 도 활용하여, synthetic conditional generation 문제에서, generalization 성능 끌어 올리는 연구 확인할 수 있었음.

박진우, Barlow Twins: Self-Supervised Learning via Redundancy Reduction (Facebook AI research, ICML 2021)

- Self-supervised learning 상황에서, model 이 trivial solution(constant representation)으로 collapsed 되는 문제를 막기 위해, invariance term 과 redundancy reduction term 을 통해 문제를 해결하는 과정에 대하여 소개함.

박희강, GraphCodeBERT: Pre-Training Code Representations with Data Flow

- Source code 에 대한 language model 을 만드는 code representation 분야에서, CodeBERT 와 data flow 를 통해 'semantic-level structure 에 대한 정보'까지도 학습할 수 있는 GraphCodeBERT 제시

서승현, Human Pose Regression with Residual Log-likelihood Estimation

- Human pose estimation 을 heat map-based 대신 최초로 'regression-based'로 수행하는 연구에 대해 소개 (Method 적 특이사항으로는, Normalizing Flow 와 MLE 방식, reparametrization 활용한 점이 있음).

송우혁, SimPLE: Similar Pseudo Label Exploitation for Semi-Supervised Classification

- semi supervised classification 을 이용한 모델로, 레이블이 없는 데이터를 이용하여 계산한 pair loss 를 main contribution 을 가지는 SimPLE 을 제안했습니다.

신주현, FlexMatch: Boosting Semi-Supervised Learning with Curriculum Pseudo Labeling

- Unlabeled data 를 dynamically leveraging 함으로써, semi-supervised learning 에서 학습에 활용될 수 있는, Curriculum Pseudo Labeling (CPL)를 제시하고, 그에 대한 강점 보여주었음 (수업에서 배웠던 entropy minimization 를 통해 pseudo labeling 수행, consistency regularization term 을 loss 에 활용).

심수연, GraphSMOTE: Imbalanced Node Classification on Graphs with Graph Neural Networks

- SMOTE 는 imbalanced data learning task 로 유명한 모델로, 이를 graph 도메인에 적용한 모델로, generator 를 이용하여 node, edge 를 생성하는 모델을 제안했습니다.

안진형, TENT: Fully Test-Time Adaptation by Entropy Minimization

- TENT 는 학습이 완료된 모델을 test 데이터에 generalize 시킴으로서 성능을 높여주는 adaptation framework 를 말하고, 이때 entropy minimization 을 이용하여 availability 와 efficiency, accuracy 를 높일 수 있다고 합니다.

양낙영, Towards Understanding and Mitigating Social Biases in Language Models

- Language model 에 대한 social bias 에 대한 측정하는 방법(다른 sentiment 모델 및 KL divergence 활용)과, 그 bias 를 mitigating 하는 방법(Bias-sensitive words 를 dynamical 하게 찾아서 학습시에 고려하는 AutoRegressive INLP)에 대한 연구 소개

양은석, Testing DNN-based Autonomous Driving Systems under Critical Environmental conditions

- 위험한 환경에서 MUNIT(Image-to-image translation 모델)과 search based 방법을 결합하여 DNN 기반 자율주행시스템을 평가하는 framework 인 TACTIC 제안한 논문 소개.(GANs 와 optimization 이 수업과 관련).

유혜원, Energy-Based Learning for Scene Graph Generation

- 보통 scene graph generation 은 객체 검출 모델, 객체간 관계를 찾는 모델 두가지로 구성되는데, 많이 사용되는 cross entropy 가 object 와 relation 을 independent 하게 취급하는 성질이 있어서, 소개된 논문에서는, **energy-based modelling** 을 통해 극복하려하였음(수업에서 배운 볼츠만 분포 활용)..

윤호영, Fast and Memory-Efficient Tucker decomposition for Answering Diverse Time Range Queries

- 시간관계를 가지는 tensor 에 대해 분석을 수행할 수 있는 method 로, 시계열 tensor 데이터를 분해시, 유저가 원하는 시간 간격만 decomposition 수행이 가능한 **Zoom-Tucker** 소개하고, 얼마나 효과적인지, 얼마나 효율적인지에 대해서 보여준 논문 소개.

이근택, Self-Supervised Image Prior Learning with GMM from a Single Noisy Image

- Noisy image 로부터 image prior 를 학습함으로써 clean image 없이도 denoising 수행할 수 있는 **GMM 기반 self-supervised denoising 모델**인 self-supervised EM-GMM 의 학습 방식(eigenvector 의 sparsity 에 대한 constraint 를 부여 하고, noisy, clean pair images 의 관계에 대한 constraints 부여한 loss function 으로 학습) 소개하였음.

이동혁, Robust Long-Tailed Learning Under Label Noise

- Long-tailed(class-imbalanced) & noisy data setting 에 대해서도, embedding space 에서 class 의 prototype 및 GMM 을 통해 **pseudo-labeling** 을 한 후 (long-tail 문제를 잘 해결한다고 알려진 nearest class mean 활용), 이를 통해 기존 모델대비 더욱 잘 작동할 수 있는 **RoLT** 라는 방법 소개하고, 실험 결과를 통해 우수성 확인시켜줌.

이상준, What Should Not Be Contrastive In Contrastive Learning

- 최근 self-supervised setting 에서 많이 사용되는 contrastive learning 의 기존 문제점(특정 데이터에 적용되는 augmentation 에 따라 학습 성능 차이가 많이 다름, 유익하지 않을 수도 있는 데이터 셋팅이 있음)에 대해, 하나의 embedding space 에서 contrastive learning 하는 것 보다는, **sub-embedding space 에서 contrastive learning** 을 수행하는 **LooC(Leave-One-Out Contrastive learning)**을 통해 좋은 성능을 얻는 모델을 학습한 논문 소개.

이정섭, Graph Autoencoder for graph compression and representation learning

- Original graph 의 information 손실 없이 잘 압축하여 representation 을 학습할 수 있는지에 대한 **GNN 도메인에서의 auto-encoder 기반 모델** 학습 방법인 Multi-kernel Inductive Attention Graph Autoencoder(MIAGAE, first kernel 과 final feature 를 aggregate 함)에 대한 논문 소개 (특이사항 : decoder 에서 노드와 엣지를 개별적으로 compression 하는 대신에, node similarity 와 graph 전체의 topology 를 활용).

이태경, Deep Burst Super-Resolution

- 저화질의 영상, 이미지를 고화질로 바꾸어주는 super-resolution 문제 상황에 대해, multi-frame super-resolution 기법을 활용한 **Burst super resolution** (alignment module, warp, attention-based fusion 활용한 auto-encoder 구조) 제시 (학습 시에, raw image 얻고, translation, rotation 을 통해 burst shot 을 construct 한 후, 학습에 활용).

이휘준, Momentum Contrast for Unsupervised Visual Representation Learning

- Self-supervised learning 의 일종인 contrastive learning(contrastive loss 는 비슷한 pair 는 가깝게, 비슷하지 않은 것은 멀도록 representation 학습)을 vision task 에 적용하되, 기존 contrastive loss mechanism 의 end-to-end 방식과 memory bank 방식의 문제점(negative sample 의 encoder 의 inconsistency, updated negative sample 이 encoder 에 반영되지 않는다는 점)을 **dynamic dictionary 기반으로 극복한 방법인 MoCo** 에 대한 소개.

전지형, Decision Transformer: Reinforcement Learning via Sequence Modeling

- 주변환경과 상호작용하여 나온 결과가 주어졌을 때, 보상(Return)을 maximize 하는 parametrized policy 를 찾는 강화학습(RL) 문제에 대해, sequence 기반 model 이 기존 traditional RL 알고리즘대비 유리하다는 점에 대해 언급하고, RL 에서 trajectory representation 을 정의한 뒤, **transformer 를 활용하여 autoregressive 한 trajectory 모델을 학습하는 방법** 제시한 논문 소개.

정민준, Decision Transformer: Reinforcement Learning via Sequence Modeling,

- 기존 language 모델들은 visual 과 word 토큰을 align 시키기 위한 방법들을 제시하였는데, 이들이 주로 image 의 region feature 에 의존하고 있었다는 특징이 있고, 이로부터 야기되는 한계점 (vision-language interaction 학습 어려움, object detector 학습 cost, image-text dataset 이 noisy)을 극복하기 위한 **ALign BEfore Fuse(ALBEF)** 제시하였고, 여기서는 noisy 문제도 해결하여 improve 하기 위해 momentum distillation(MoD)도 사용하였음.(Downstream task 성능도 좋은 것 확인 가능).

정윤기, Swin Transformer: Hierarchical Vision Transformer using Shifted Windows

- NLP 에서 사용되던 vision transformer 가 computer vision task 에서 널리 사용되던 CNN 기반 구조를 뛰어 넘는 성능을 보일 수 있는 가능성을 제시한 ViT (패치단위로 나눈 이미지에 class token 을 더하고, self-attention 활용)의 문제점을 극복하기 위해서, KD 를 활용한 Data efficient image Transformer(DeiT)라는 방법 제시되었는데, DeiT 에도 구조적 문제가 있었고, 이를 해결하기 위해 **shifted windows 를 활용한 hierarchical vision transformer** 인 Swin Transformer 를 소개함.

채윤기, Diff-TTS: A Denoising Diffusion Model for Text-to-Speech

- Non-autoregressive TTS 에 **denoising diffusion probabilistic model(DDPM)을 적용한** Diff-TTS 의 학습 방법 소개하였고, 이 모델의 강점 (Parameter 의 수가 Tacotron2 나 Glow-TTS 의 반 밖에 되지 않는다는 점, Accelerated sampling 을 적용하여 inference speed 를 높였다는 점, Temperature term 을 이용하여 pitch variability 를 조정할 수 있다는 점) 들을 강조함.

한동훈, Learning Transferable Visual Models From Natural Language Supervision

- Contrastive pre-training, label text 로 부터 얻은 dataset classifier, zero-shot prediction 을 통해, 기존 standard image model 은 feature extractor 와 label classifier 가 joint 하게 학습 되었던 반면, CLIP 은 image encoder 와 text encoder 를 joint 하게 학습하여 correct pair 에 대한 예측 수행하며, 이를 통해 **NLP 에서 transferable visual model 을 얻을 수 있음을 보임** (CLIP 을 통한 multi-modal embedding 학습).

현수민, A Practical Bayesian Framework for Structural Model Updating and Prediction_review

- Structural model update 에서, Bayesian regularization 기반 prediction 을 수행할 수 있는 practical 한 statistical framework 에 대해 소개하고, 더 나아가 measured modal data 를 활용한 calibrated structural FE model 를 통해 natural frequency 와 mode shapes 에 대한 posterior predictive distribution 을 얻을 수 있음을 보이고, calibrated FE model 를 통해 noise 에 대한 robustness 를 획득할 수 있음을 보임.

홍상민, Learning Continuous Image Representation with Local Implicit Image Function

- Implicit neural representation 을 super resolution 등 2-dimensional image 에서 적용한 논문 소개하고, 이 implicit function 이 continuous 한 real-world 문제와 밀접하다는 것과 great scalability 등을 통해 이 논문의 장점을 강조함(implicit image neural representation 에 대해 알 수 있었음).

권정현, *Normalizing Kalman Filters for Multivariate Time Series Analysis*

- non-Gaussian multivariate time series data 에 대해, Kalman-like recursive filtering 을 과정을 가지는 generative model 과 normalizing flow 를 통해, **Normalizing Kalman Filters** 기반 방법 제시하고, 이 모델을 활용한 application(forecasting 과 missing value) 실험 결과를 통해 훌륭한 성능을 보임을 공유함.

권준형, Learning Continuous Image Representation with Local Implicit Image Function

- Super-resolution task 에서, 기존 DNN 기반 모델의 문제를 해결하기 위해, 3-dimensional data 의 reconstruction 에서 주로 사용되던 **implicit neural image representation** (input 으로 이미지 좌표값, output 으로 RGB 값을 출력하도록 학습하면, network 가 이미지를 잘 표현할 수 있음)을 활용하여 image 를 continuous domain 에서 정의된 함수로 모델링해서, **restore** 및 **generation** 이 가능하다는 것을 보임.

기호건, Goal-conditioned Reinforcement Learning with Imagined Subgoals

- Goal-conditioned RL 문제에 대해, **subgoal-reaching policy** 를 **guide**(KL-divergence constraints 부여를 통한 guide)로 활용하여 **goal-reaching policy** 를 Actor-critic 방법 기반으로 학습하는 방법 제시하고, Ant navigation, Vision-based robotic manipulation 실험 수행한 논문 소개.

김근우, Deep repulsive clustering of ordered data based on order-identity decomposition

- DNN 기반 supervised learning 을 통한 age estimation 에서 발생하는 문제(인종, 성별 별로 노화속도가 다름)를 해결하기 위한 방법 (data clustering(k-means clustering 에 repulsive term 추가) 및 cluster 안에서 비교를 통한 age 추정)인 **deep repulsive clustering (DRC)**를 제시한 논문 소개.

김도형, Catch & Carry: Reusable Neural Controllers for Vision-Guided Whole-Body Tasks

- Motion capture data 에 대해, **inverse dynamics**(data 를 individually tracking 하는 RL agent 학습)와 supervised learning 을 통해 low-level controller(Auto-encoder 구조 활용)를 학습한 뒤, egocentric vision 과 model-free RL(Maximum a posteriori policy optimization 기반 학습)을 통해 high-level controller 를 학습하는 방식을 취하는 방법 소개.

황예린, CHoRaL: Collecting Humor Reaction Labels from Millions of Social Media Users

- 다양한 문화권에 맞게 Humor detection 을 수행하기 위해 Facebook Post 를 사용하여 여기에 달린 humor score 를 활용하여 humor label 을 생성하고 humor score, non-humor score 를 어떻게 설정하는지, 그리고 해당 CHoRaL 방식을 사용하여 다양한 문서에 적용할 수 있고, humor 뿐만 아니라 분노나 슬픔에 관한 것도 용이하게 분류하는 등 여러 분야로 확장가능성이 있음을 알게 되었습니다.

김동욱, LiDAR Place recognition OverlapNet

- SLAM(Simultaneous Localization and Mapping)의 의미 및 용도에 대해 알았고, 시간이 지날수록 drift error accumulation가 발생하므로 에러를 보정하기 위해 loop closure가 필요한 점과, 현재 자신의 위치(Localization)와 주변 사물의 위치(Mapping)를 동시에 파악하고 매 시간마다 연속된 frame을 이어붙이는 registration을 적용하며, 오버랩의 정도를 정의하여 training함으로써 해당 에러를 정확히 보정할 수 있음을 알게 되었습니다.

김민우, Sketch Your Own GAN

- 다른 자세를 취하는 네트워크 수많은 학습 이미지를 바탕으로 네트워크 파라미터를 학습, 사용자의 일부 스케치를 바탕으로 네트워크 파라미터를 변경함으로써 달성 가능하다는 사실을 알았고, 네트워크 아키텍처를 Latent vector, Photo, Sketch, Discriminator 순으로 구성하는 방식에 대해, image space regularization 접목할 수 있는 점, Latent space interpolation 및 image editing 이 제한된 네트워크만으로 가능하고 latent vector 의 분포를 잘 이해할 수 있다는 것을 알게 되었습니다.

김범준, Online Knowledge Distillation via Collaborative Learning

- Distillation 은 복잡해지는 teacher 네트워크를 그렇지 않은 student network 로 학습시키는 방법이라는 사실, DML, ONE, KDCL 각각의 특징들을 확인하였고, 동시에 사이즈가 다른 다양한 네트워크들 간에 학습이 가능하고 심지어 작은 사이즈의 네트워크가 큰 사이즈의 네트워크 학습에 기여할 수 있다는 점, 좋은 성능을 가져오기 위해 logit 의 설정이 중요하다는 점을 알게 되었습니다.

김보은, Uncertainty-aware Score Distribution Learning for Action Quality Assessment

- AQA 는 비디오로부터 특정한 action 이 얼마나 잘 수행되었는지를 자동으로 평가하는 task 라는 점, 평가 점수를 직접 regression 하는 것보다, 평가 점수의 uncertainty 를 고려하여 score distribution learning 을 통해 추정하겠다는 아이디어에 대해 알게 되었고, 하나의 점수가 아니라 다양한 점수들로부터 distribution 을 구하는 것을 통해 여러 심판들이 매기는 점수들에 대한 내재적인 uncertainty 를 학습함으로써 기존 scheme 보다 성능이 향상될 수 있음을 알게 되었습니다.

김세민, Conditional Variational Autoencoder with Adversarial Learning for End-to-End Text-to-Speech

- VIT 는 ene-to-end 로 합성이 가능하면서 naturalness 가 높다는 점, 아직 딥러닝 과목을 수강하지는 않았지만 Generative model 로서 GAN, VAE, Flow-based generative model 3 개가 있고, GAN 은 generator 와 discriminator 두 가지를 사용하여 기존의 dataset 을 구분하지 못하는 샘플을 generate 한다는 것, VAE 는 학습된 Z 를 활용하여 generate 가 가능해지고, flow-based 의 경우 단순한 분포에서 샘플링한 데이터를 앞선 함수의 역함수에 통과시켜 data 를 generate 하는 방식이라는 점, 그리고 VIT 는 이상 세가지 방식을 복합적으로 사용하여 구성한다는 점을 알게 되었습니다.

김시환, Normalization Matters in Weakly Supervised Object Localization

- CAM 즉 Class Activation Map 은, Classification 을 CNN 방식으로 end-to-end 로 수행할 경우 convolution layer 에서 산출되는 feature map 을 1 차원 배열로 압축하고 최종 결과 산출직전의 학습된 weight 값을 각 채널별 feature map 에 곱하여 weighted sum average 를 구하며, 최종 class activation map 을 구할 수 있다는 사실을 알게 되었고, CAM 기반의 WSOL method 관련하여 CAM, HaS, ADL, CutMix 등을 살펴보고, 이것을 normalize 해주는 함수 H 를 사용하여 최종 score map 을 구할 수 있다는 사실과, 이러한 H 의 종류에 따라 정성적 및 정량적 normalize 값이 다를 수 있게 한다는 점을 알게 되었습니다.

김영국, Efficient and Scalable Bayesian Neural Nets with Rank-1 Factors

- Bayesian Neural Network 에 대해 수업시간에 배웠던 내용들을 다시 확인해 볼 수 있었고, Batch ensemble 의 경우 여러 개의 모델들을 조합하여 좀 더 나은 성능을 얻도록 하는데 모델들을 더하면서 파라미터도 같이 증가한다는 점과, batch ensemble 의 경우 파라미터를 모두 공유한다는 점, 추가적인 파라미터 벡터들만 구하고, 새로운 weight 를 만든다는 점을 알게 되었고, 모델이 증가함으로써 weight parameter 가 많이 증가하지 않게 되는 장점 및, BNN 에 batch ensemble 을 어떻게 적용할 것인가에 대해 알 수 있었고, BNN 의 loss 함수를 약간 변형하여 적용하면 성능개선을 이룰 수 있다는 가능성에 대해 알게 되었습니다.

김준엽, Bias-Variance Reduced Local SGD for Less Heterogeneous Federated Learning

- 빅데이터 분야에서 고차원, non-convex 성질의 경향이 큰데, 이런 데이터에 더욱 적합한 SGD 를 찾기 위해 보다 완화된 조건에서 로컬 알고리즘을 찾아야 하고, 이를 위한 Bias-Variance Reduced Local SGD 를 위해 localization, Bias Reduction, Stochastization, Variance Reduction 등을 적용할 수 있으며 이를 통해 특수한 상황에서 새로운 방식을 적용해 볼 수 있다는 점을 알았고, 빅데이터의 딥러닝 작업시 최적화에서 SGD 대신 Bias-Variance Reduced Local SGD 를 사용할 경우 complexity 를 낮추고 속도와 정확도를 높일 수 있음을 알았습니다.

김지환, Learning Disentangled representation via product manifold projection

- Feature 를 latent space 에서 다루는 연구에 대해 확인하였고, latent space 에서 데이터를 어떻게 효율적으로 분리할 수 있을지, 그리고 product metric space 에 대해 알아보았고, 기하학적 관점에서 disentanglement 를 해석하면 현재 사용되는 방법론을 일반화할 수 있다는 사실을 알았으며 Sample distribution 의 관점에서, 기존 신경망 모델에 weakly-supervised recipe 를 쉽게 적용할 수 있을 것이라는 가능성을 알았습니다.

김진우, k-Nearest Neighbors by Means of Sequence to Sequence Deep Neural Networks and Memory Networks

- 수업시간에 배운 KNN 의 연장선상에서, Sequence to Sequence DNN 과 Memory Network 와의 연계가 어떻게 되는지 살펴볼 수 있었는데, 먼저 sequence to sequence 모델이 무엇인지, 그리고 그 sequence to sequence model 에 KNN model 이 올라갈 수 있다는 내용과, V2LS 및 V2VS 는 각각 주어진 feature 로부터 nearest neighbor 들의 label 이나 feature vector 를 예측하고, V2VSLS 는 최종 모델로서 이 두가지 예측을 모두 수행한다는 사실, 그리고 각각의 구성 및 특징을 알 수 있었습니다.

김태형, Empirical Study of the Benefits of Overparameterization in Learning Latent Variable Models

- 필요보다 더 많은 모델 파라미터를 가지는 Overparameterization 의 예로서 샘플로부터 역으로 모델을 학습할 때 학습되는 모델은 ground truth model 과 latent variable 수를 비교할 때 더 많은 latent variable 를 갖는 경우가 바로 overparameterized model 이라는 사실, Noisy OR Network 는 1 개의 binary latent variable 과 1 개의 binary observed variable layer 라는 점, ground truth 모델에서 인공적으로 생성된 데이터 샘플을 갖고 모델링을 할 때, 모델의 latent variable 숫자를 늘려가며 recovered 되는 (즉 ground truth model 에서의 latent variable 과 같은 파라미터를 갖는 것이 학습된 모델에 존재하는 것을 의미) ground truth latent variables 수를 측정하여 Larger model 일수록 recovered ground truth latent variable 이 증가하게 됨을 알 수 있었습니다.

김태훈, Multisource Domain Factorization Network for Cross-domain fault diagnosis of Rotating Machinery : An Unsupervised Multisource Domain Adaptation Method

- source domain 데이터와 label 이 없는 산업 데이터인 target domain 데이터에 대한 대응으로 domain adaptation 을 적용하며 복잡한 multi source domain adaptation 에 대한 연구가 어떤 식으로 될 수 있는지에 대한 방향성, 그리고 그 과정에서 shared space 가 각 domain 들로부터의 noise 에 대해 취약한 점을 극복하기 위해 private feature encoder (domain 과 관련)와 shared feature encoder(domain 과 관련 없음)를 factorization 을 적용해 따로 학습하도록 만든 모델, 그리고 특정 domain 에 속하는 sample 이 target domain 과 큰 관련성이 없는 경우 발생하는 negative transfer 의 영향을 극복하기 위해 IET loss 가 사용될 수 있다는 점을 알게 되었습니다.

김현기, Prototypical Contrastive Learning (PCL) of Unsupervised (self-supervised) representation

- unsupervised representation learning 은 이미지 데이터 입력이 들어올 때 이것의 의미적 정보를 가장 잘 표현할 수 있는 representation vector 를 구하는 것이고 여기에 Prototype 이라는 latent variable 을 도입하여 MLE 를 EM 알고리즘으로 훈련하는 방식인데, 이 프로토타입은 의미적으로 유사한 이미지들이 하나의 클러스터를 이루고 있을 때 이를 대표할 수 있는 embedding representation 을 의미한다는 사실을 알게 되었습니다.

남승우, Image Restoration for Under-Display Camera

- under display 카메라의 단점인 패널을 통과해서 촬영하여 영상 품질이 저하되는 단점을 극복하고 원상 이미지로 복구하기 위해 제안한 UNET structure 에 대해 알게 되었는데, 구체적으로는 Diffraction-limited point spread function 의 의미, image formation 및 restoration 의 과정, 그리고 제안된 UNET structure 는 encoder 를 2 개로 나뉘어 residual detail 과 content encoding 을 분리하고 다시 새로운 레이어에 집어넣는 구조에 대해 알게 되었습니다.

문성준, Representation Learning via Invariant Causal Mechanisms

- 인과관계를 이용해 모델을 설명하는 causality 에 대해 알게 되었고, causal mechanism 은 원인과 결과의 관계를 살펴보는 메커니즘이라는 사실, do 연산을 하여 원인 X 를 바꾸어가며 Y 를 관찰하는 것이 intervention 이며, causality 와 representation learning 이 두 개념을 결합하여 causality 관점에서 contrastive learning 을 설명해 볼 수 있는 가능성에 대해 알게 되었습니다.

민대원, DEAR: Deep Reinforcement Learning for Online Advertising Impression in Recommender Systems

- 강화학습의 실제 사례를 접하면서 state, action, reward 를 잘 정의하는 것이 매우 중요하다는 사실을 알았고, 강화학습을 사용하면 본 연구의 내용과 같이 지속적으로 advertising strategy 를 업데이트하고 장기적으로 reward 를 극대화 시킬 수 있다는 점, 아울러 DQN 의 구성시 어떠한 방식으로 내부 task 를 정의할지에 대한 원리를 알 수 있었습니다.

민철희, The continuous Bernoulli: fixing a pervasive error in variational autoencoders

- MNIST 데이터셋은 1~255 의 픽셀값을 0~1 의 float 값으로 변환하여 0~1 사이의 continuous 한 값으로 주로 활용되고 있다는 사실, 기존 Bernoulli 분포가 갖고 있는 normalization 문제를 해결하기 위해 0 에서 1 사이의 support 를 가진 continuous Bernoulli 분포를 통해 기존 Bernoulli 분포의 normalization 문제를 해결할 수 있으며, 이 모델로 VAE 를 형성할 경우 기존의 Bernoulli, gaussian 으로 모델링한 VAE 보다 훨씬 sharp 한 이미지를 생성할 수 있어 0 과 1 사이의 데이터를 더 잘 모델링 할 수 있음을 알게 되었습니다.

박기범, Fast Bayesian Uncertainty Estimation and Reduction of Batch Normalized Single Image Super-Resolution Network

- 일반적인 딥러닝에서 실패리스크가 클 수 있는 부분에 대해 수업시간에 배운 Bayesian 을 통해 해결을 모색하는 과정을 보았고, 구체적으로 Super Resolution Deep Learning 모델은 output 의 신뢰성을 결정하기 위해 더욱 높은 수준의 투명성을 필요로 한다는 점, training 단계에서 평균과 표준편차를 미리 구하고 input 을 넣을 때 같은 이미지들을 concatenation 해 놓고 patch 를 한꺼번에 계산하고 앞에서 구한 평균과 표준편차를 재활용 할 수 있다는 사실, 그리고 Monte Carlo Batch normalization 이 무엇인지, 여기에 Bayesian Network layer 를 적용하여 uncertainty 를 예측할 수 있고, 이렇게 uncertainty 를 감소시켜 super resolution deep learning 에서의 성능을 향상시킬 수 있음을 알게 되었습니다.

박상하, BAYES-TREX: a Bayesian Sampling Approach to Model Transparency by Example

- 딥러닝 적용을 위해 여러 XAI 방법론이 제안되어 왔는데 주로 test set 에 대한 모델의 decision 을 설명하는데 중점을 두었지만, 그 test set 자체가 가지는 대표성에 대한 연구도 필요하다는 사실, 그리고 Bayes-Trex 라는 방법론으로 특정 confidence 값을 가지는 input 을 sampling 하여 해당 모델에 대한 전반적인 이해를 도울 수 있다는 사실, Posterior 에서 샘플링 하는 대표적인 MCMC 방식의 두 가지 단점, 이를 극복하기 위해 level set 을 넓히고 generative model 을 사용한다는 점, 그리고 Bayes-Trex 의 장점은 training distribution 과 상관없이 다양한 distribution 에서도 classifier 에 대한 평가를 할 수 있고, 모델이 어떤 sample 에 대해 사용자가 정의하는 confidence level 을 충족하는지 알려주어서 모델의 정확성을 올릴 수 있는 가능성을 알게 되었습니다, .

박선재, Disentangling Label Distribution for Long-tailed Visual Recognition

- 실제 데이터들은 동일한 분포를 갖지 않고, tail class 는 수량이 부족한 데이터의 경우에 해당하며, long tail dataset 으로 얽힘 현상이라는 문제가 발생하는 점, 이런 경우 label loss 를 사용하여 학습하면 성능을 향상시킬 수 있음을 알았고, 먼저 Post-Compensated Softmax 를 이용하여 fully connected layer 의 결과 값에서, 학습 데이터에서 해당 클래스의 비율에 로그를 취한 값을 빼주고, 시험 데이터에서 해당 클래스의 비율에 로그를 취한 값을 더해주는 간단한 방법으로 쉽게 추론 성능을 올릴 수 있고, 그 다음 loss 함수로 regularizer 항을 설계하는데 Long tail 데이터 셋의 항을 제거하고 학습 데이터 셋의 분포를 사전 균일 분포로 대체한 후 함수를 Donsker varadhan representation 을 이용하여 최적의 형태로 변형하여 구하고, cross-entropy loss 함수에 더하여 구한 최종 loss 함수를 모델의 학습에 이용할 수 있음을 알게 되었습니다.

박선지, Explaining in Style: Training a GAN to explain a classifier in StyleSpace

- CNN Classification 에 영향을 미치는 이미지 attribute 를 찾고 이를 사람이 이해할 수 있는 방식으로 설명하는 것과 반례를 들어 판단할 확률을 높이는 방식을 알게 되었고(counterfactual explanation), Style space 는 특정 임베딩 공간을 통해 입력데이터가 들어가서 Detail 한 feature 를 알 수 있도록 한다는 사실 (마치 강아지 사진이 고양이 사진으로 분류될 확률을 구함, 상위 3 개의 특성을 찾음), 그리고 StyEx 의 구조는 classifier 를 추가하여 classifier based training 을 실시하여 이미지의 classification 에 영향을 미치는 특성을 찾고 이를 설명하기 위해 반례를 드는 개념을 도입한다는 사실 (KL divergence 를 취하여 최대한 같도록 함) 생성된 이미지에서 Classifier 출력과 원 입력 이미지 간의 KL-divergence 를 통해 양자를 같은 class 로 분류, 그리고 StyleGAN2 에 대한 Classifier 기반의 학습을 통해 주어진 이미지의 classification 을 설명할 수 있음을 알게 되었습니다.

박정호, CoPhy: Counterfactual Learning of Physical Dynamics

- 인공지능이 만약의 경우에 대해 상상 및 추론을 하는 것은 어렵는데 환경의 물리적인 변수를 이해하고 미래를 예측하는 것을 목표로 하며 이러한 요소들을 confounder 라고 하고 이에 대한 이해가 중요하다는 점, CoPhy Dataset 의 의미 및 이 dataset 을 사용하여 실제로 task 를 풀 수 있는 Cophynet 의 구조 및 전체 flow, 마지막 timestep 때 나오는 hidden state 가 바로 환경의 특성을 담은 confounder 역할을 한다는 점, CoPhynet 을 통해 t-step 동안의 미래를 모두 예측할 수 있다는 점, CoPhyNet 을 통해 추출되는 confounder 값들이 실제 환경에 대한 정보를 많이 담을 수 있다는 사실, 그리고 환경속에 잠재된 confounder 를 잘 학습하여 변수들의 인과관계를 이해하고 이를 바탕으로 물리적 시스템의 미래를 잘 예측할 수 있음을 알게 되었습니다.

박준수, End-to-End Semi-Supervised Object Detection with Soft Teacher

- semi-supervised learning 은 labeled data 와 unlabeled 된 나머지 데이터 모두를 학습하는 것이고 크게 consistency based method 와 pseudo-labeling based method 로 나뉘는데 후자는 라벨링된 이미지로 teacher 모델을 미리 학습시키고 라벨이 없는 이미지에 teacher 모델을 도입해서 pseudo label 을 생성하고 이를 기반으로 모델을 학습시키는 multi-stage 형태로 구성되어 있다는 사실을 알게 되었고, 여기에 학습 구조를 변경하여 pseudo labeling 을 하는 teacher 모델 또한 학습 과정 중에 성능이 향상되도록 하고, Student 모델은 라벨링 된 데이터와 pseudo-labeling 된 데이터 모두를 이용하여 학습하고, 이를 EMA(지수이동평균) 하여 teacher 모델을 업데이트하는 점, Multi-state 인 기존 프레임워크를 end-to-end 구조로 변경하고 새로운 구조를 형성하면서 soft teacher 와 box jittering 기법을 도입하여 soft teacher 는 background box 에 대해, box jittering 은 box regression 에 대해 reliability 를 평가할 수 있는 scheme 을 알게 되었습니다.

박준영, End-to-End Semi-Supervised Object Detection with Soft Teacher

- model scaling 은 기존의 baseline network 를 사용하여 model 의 크기를 키우는 것이며 depth, width, resolution 등을 조정하여 accuracy 향상을 얻을 수 있다는 점, 관련 기존연구로 AutoML, EfficientNET, RegNet 이 무엇인지, EfficientNET, RegNet 을 기반으로 보다 간단하고 빠른 model scaling 방법이 어떻게 될 수 있는지, run time 은 activation 과 큰 상관관계가 있다는 점, width 중점으로 depth, resolution 도 같이 증가시키는 fast scaling 을 통해 accuracy 및 runtime 모두 향상시킬 수 있는데 가중치 alpha 값의 조정을 통해 width 만 scaling 할지 아니면 depth 와 resolution 을 scaling 할지 정의하며 좋은 성능을 보이는 값을 확인하였고, runtime 에 가장 큰 영향을 미치는 요인은 activations 의 수이고 이것을 최소화하며 좋은 성능을 만족하기 위해 width scaling 을 중점으로 하는 fast scaling 이 효과적임을 알게 되었습니다.

박찬웅, GS-WGAN: A Gradient-Sanitized Approach for Learning Differentially Private Generators

- Differential Privacy (DP)는 프라이버시 수준을 수치화하는 방법이고, 이것을 만족하는 generator 를 만들 경우 노이즈가 들어가서 training 이 어려워지는 문제를 해결하기 위해 GAN 에서 generator 의 gradient 에만 노이즈를 인가하여 안정적으로 training 될 수 있게 하고, generator 에 용이한 training 이 가능한 점, 그리고 가우시안 메커니즘은 어떤 함수에 가우시안 노이즈를 더하는 것인데 DP 가 되며 DP 메커니즘에 어떤 함수를 취하든 여전히 DP 가 된다는 사실을 알게 되었습니다.

박충현, Simple and Principled Uncertainty Estimation with Deterministic Deep Learning via Distance Awareness

- OOD 는 특정 분야의 데이터에 대해 학습시킨 모델에 새로운 데이터가 들어올 때 그 데이터가 기존에 학습한 영역에 속하는 것인지 아니면 다른 영역에 있는 것인지 구별하는 분야이고, 학습단계에 개입해서 입력에 대한 uncertainty 를 적절히 계산하여 OOD 여부를 판단하며, OOD data 를 uniform distribution output 으로 만들고 이러한 학습결과를 실행하기 위해 SNGP 를 적용하여 모델의 representation 과 전체 학습데이터 간의 거리로 측정할 수 있다는 사실, Gaussian process 를 활용하면 Input distance awareness 를 확보할 수 있다는 점, SNGP 는 residual-based DNN 에 대한 Distance awareness 를 부여할 수 있는 방식이며, 이를 만족하기 위해 distance aware output layer 와 distance preserving hidden layers 를 만족해야 한다는 점, 학습시 각 step 마다 weight matrix 에 spectral normalization 을 적용하여 spectral normalization 을 하는 방안을 알게 되었습니다.

박태언, Differentiable Convex Optimization Layers

- 복잡한 Convex optimization 문제도 자동으로 솔루션을 구할 수 있는 differentiable neural network layer 를 설계하여 최종적으로 연구자들의 convex optimization layer 에 대한 진입장벽이 낮아질 수 있다는 점, CVXPY 가 무엇인지, 사용자가 Slover 를 사용할 수 있는 형태로 구현하여 최종적으로 솔루션을 CVXPY 로 구하는 방법, CVXPY 가 어떻게 CVXPY LAYER 로 embedding 되어 convex optimization function 에 대한 미분을 수행하는지 파악하였고, 더 나아가 어떻게 NN 으로 convex optimization function 을 푸는 과정이 embedding 될 수 있는지, 구체적으로 모든 convex 프로그램은 cone program 으로 canonicalized (규범화) 될 수 있고, conic solver 의 derivative 를 사용하고 canonicalization 을 위한 chain-rule 을 사용한다는 사실, batched inputs 및 parallelization 으로 효율적으로 forward, backward 가능하다는 점을 알게 되었습니다.

서지민, Your Classifier is Secretly an Energy Based Model and You Should Treat it Like One

- 기존에 많이 이용되던 Classifier 모델을 새롭게 해석하여 Label 과 Data 를 Joint 하게 모델링하는 구조에 대해 알 수 있었고, 기존의 Generative 및 Discriminative 모델의 Hybrid 모델들에 비해 성능이 좋도록 두 종류의 모델 장점들을 취할 수 있다는 사실, 특징적으로 EBM 을 접목한 부분을 확인하였고, 구체적으로 Classifier 의 출력으로 나오는 Logit 을 기존에는 Softmax 연산을 이용해서 x 를 알 때 각 Class 의 확률을 계산할 때 이용했는데, 이 Logit 을 Energy 로 해석하자라는 아이디어를 알 수 있었고 이를 통해 구성된 확률 모델을 이용하면 기존의 Discriminative 모델을 Generative 모델과 같이 확장할 수 있다는 사실, 그리고 고려해야 할 점으로 Likelihood 를 정확하게 계산하기 위해서는 Normalizing factor 를 알아야 할 필요도 있지만 계산할 수 없는 점이 문제가 된다는 것과, 하이퍼파라미터를 제대로 fine-tune 하지 못한다면 Gradient estimator 가 불안정하여 결과가 diverge 할 수 있다는 사실을 알게 되었습니다.

손이인, A Simple Framework for Contrastive Learning of Visual Representation

- 이미지 데이터의 정답 label 이 없는 상황에서 효과적으로 visual representation 을 추출하는 SimCLR 이라는 unsupervised learning 을 알게 되었는데, 이것은 data augmentation 을 통해 얻은 positive/negative sample 에 대해 contrastive learning 을 이용하여 unsupervised learning 을 하는 것이고, contrastive learning 을 이용하여 Visual Representation 을 효과적으로 학습할 수 있으며 VAE 처럼 이미지를 나타내는 feature 를 잘 뽑아낼 수 있고, unsupervised learning 의 성능을 높일 수 있다는 사실을 알게 되었습니다.

오설아, NORMFORMER: IMPROVED TRANSFORMER PRETRAINING WITH EXTRA NORMALIZATION

- 기존 트랜스포머 구조는 각 sublayer 의 residual connection 이후에 variance 를 줄이기 위해 layer normalization 을 적용하는 Post-LN 를 수행하는데 이 경우 앞단의 layer 에 비해 뒷단의 layer 들의 gradient 의 magnitude 가 상대적으로 큰 경향이 있는 점을 개선하기 위해 layer norm operation 을 각 sublayer 의 시작 단계로 옮기는 Pre-LN, 그리고 이것은 Post-LN 에 비해 학습 시 안정성을 향상시키지만 오히려 앞단 layer 의 gradient 가 뒷단 layer 에 비해 커지는 단점이 있는데 이것은 각 layer 에 3 가지 normalization 을 추가한 NormFormer 방식을 적용하여 gradient magnitude mismatch 문제를 완화할 수 있다는 점, 구체적으로 attention head 를 scaling 하고, attention 모듈 이후에 layernorm 을 추가한 후, 첫 번째 fully connected layer 이후에 layernorm 을 추가하는 방식에 대해 알게 되었습니다.

오영진, Deblurring using Analysis-Synthesis Networks Pair

- 분석하는 네트워크와 합성하는 네트워크 페어를 사용하여 deblurring 을 어떻게 수행할지, 그리고 deblurring 의 개념, non-blind deblurring 의 케이스가 blind deblurring 보다 용이하고 효율적으로 deblurring 을 할 수 있다는 점, analysis network 는 blur kernel 알아내서 synthesis network 로 보내주어 문제를 non-blind 하게 만든다는 점, 이 과정에서 cross correlation layer 를 새로 만들어 analysis network 를 만든다는 점, blur 에 대한 정보를 담고 있는 correlation 을 네트워크에 layer 로 구현하여 효과적으로 kernel 을 구할 수 있다는 사실을 알게 되었습니다.

오우석, Decision Transformer: Reinforcement Learning via Sequence Modeling

- 강화학습 문제를 sequential modeling 문제로 바꿔 해결하는 방법에 대해 알아보았는데 구체적으로 Sequential modeling 을 할 때 NLP 에서 좋은 성능을 보이는 Transformer 구조를 사용할 수 있다는 점, Bootstrapping 과 credit assignment 로 인한 문제를 어떻게 해결할지에 대해 알았고, offline RL 는 기존의 RL 에서 interaction 의 비용이 매우 크기 때문에 데이터셋을 미리 모아 놓고 추가 interaction 없이 학습하는 방법이라는 사실, 하지만 학습하는 policy 와 데이터를 수집한 policy 가 달라지기 때문에 distribution shift 으로 인해 학습이 힘들어진다는 점, sequence modeling 을 할 때 사용하는 Transformer 구조 중 GPT 라는 자연어 처리 분야에서 개발된 auto regressive 한 모델을 사용하고, 이것은 예측된 action 들 중에서 가장 최근의 action 을 사용하고, 이 action 을 사용해서 얻은 state 와 reward 및 action 을 사용해서 다시 모델에 넣어주는 과정을 반복하면 된다는 사실을 알게 되었습니다.

온광현, Learning Continuous Image Representation with Local Implicit Image Function

- super resolution 은 low resolution image 를 받아서 high resolution 으로 바꾸어 주는 image processing method 이고, coordinated based method 를 취하여 super resolution 을 할 경우 pixel 과 pixel 간 예측을 하고, Coordinate pixel value, 즉 integer value 가 아닌 Float value 에 coordinate 를 넣어주면 픽셀값을 알고 있기 때문에 임의의 scale factor 로 super resolution 이 가능하다는 사실, LIIF 는 representation 을 잘 학습하고 그 representation 을 통해 복원해낸 high resolution 역시 high fidelity 를 보여준다는 사실을 알게 되었습니다.

유승찬, EfficientNetV2: Smaller Models and Faster Training

- network 의 depth, width, resolution 을 coefficient 파이에 맞춰서 uniformly scaling 하는 compound scaling, 그리고 해당 파이는 CPU 나 메모리 등 리소스에 맞춰서 바뀔 수 있는 coefficient 이고, image net 에서 accuracy 를 향상시킬 수 있으며, learning 에서 크기에 상관 없이 모든 input 에 같은 정규화를 적용하는 것은 성능이 저하되고 작은 사이즈의 input 에는 작은 augmentation, 큰 사이즈 input 에는 강한 augmentation 이 좋은 accuracy 를 보이는 점을 알게 되었고, 그래서 초기에는 작게, 뒤로 갈수록 큰 크기의 input 에 강한 정규화를 적용하는 adaptive regularization 방식, 이상 각기 다른 아키텍처에 progressive learning 을 도입하면 파라미터와 training efficiency 에서 도움을 얻을 수 있다는 사실을 알게 되었습니다.

유진오, Gradient - starvation : A Learning Proclivity in Neural Networks

- Gradient Starvation 은 Cross entropy 를 사용하여 분류문제를 다룰 경우 신경망의 일부 dominant 한 feature 들만 추출하고 다른 feature 들은 찾지 못하는 문제가 있고 그 이유는 신경망이 경사하강법을 통해 학습할 때 피상적이고 강한 feature 들에만 집중하기 때문이라는 사실, 이를 극복하기 위해 spectral decoupling 이라는 간단한 regularizer 를 이용한 cross-entropy 방식을 적용하여 Gradient Starvation 를 해결하고, 그 사례로 MNIST dataset 을 다룬 내용까지 알게 되었습니다.

윤영석, Training Independent Subnetworks for Robust Prediction

- ensemble learning 은 Bayesian 와 함께 Uncertainty estimation 및 out-of-distribution robustness 를 위한 방식중 하나로 쓰이며 여러 개의 모델을 학습시키고 각각의 결과를 조합하여 좀 더 robust 한 결과를 얻는 방식이라는 사실, 여러 개의 모델을 학습시켜야 해서 computational cost 가 커지는 경향이 있다는 점, neural network 의 sparsity 로부터 MIMO 에 대한 아이디어를 얻게 되었다는 점, MIMO 에서는 neural network architecture 에서 input layer 와 output layer 를 대체하는데 training 시에는 input x_1, x_2, x_3 가 서로에게 간섭 없이 y_1, y_2, y_3 에게만 영향을 주어야 하며 하나의 body 안에 여러 개의 subnet 에 대한 학습을 진행하고, test 시에는 하나에 여러 개의 그림을 겹쳐서 같은 그림을 forward 하여 평균을 내거나 voting 하는 방식을 사용한다는 것을 알게 되었습니다.

윤정환, Re-distributed Biased Pseudo Labels for Semi-supervised Semantic Segmentation

- Semi-supervised learning 의 Pseudo label 방식은 먼저 labeled data 로 모델을 학습시키고, 여기에 unlabeled data 를 적용하여 pseudo label 을 부여한 후, 이것과 labeled data 를 혼합하여 model 을 다시 학습시키는 방식이라는 사실, 구체적으로는 먼저 labeled data 로 student model 을 훈련시키고 해당 student model 을 teacher 로 삼아 저자들이 제안한 DARS 를 적용하여 unlabeled data 에 unbiased pseudo label 을 생성하고 기존의 training data 와 혼합하여 다시 모델을 재학습시키는 과정을 iterative 하게 반복하여 모델을 향상시킬 수 있다는 사실과, unbalanced dataset 에는 bias 가 있기 때문에 tail class 에 대해 학습에 많은 나쁜 영향을 미치게 되는데 DARS 를 사용하여 pseudo labeled distribution 을 가급적 true distribution 과 맞추어 주는 방향으로 학습시킬 수 있는 가능성에 대해 알 수 있었습니다.

이동렬, SituatedQA

- 사용자가 같은 질문을 다시 하더라도 주어진 시간, 공간과 같은 상황에 따라 다르게 답을 해야하는 경우, 즉 context aware 해야 하는 상황에서 적절한 답을 하는 문제인 temporal / geographical context problem 을 어떻게 해결할지 살펴보고 이를 위해 하나의 question 을 context type, context value, answer 로 구성하고, data collection 은 identification, (context/answer) collection, validation 의 세 단계로 나누고, context 변화에 따른 QA 시스템을 training 하고 평가할 수 있는 dataset 이 어떻게 만들어지는지 알게 되었습니다.

이동준, On the Out-of-distribution Generalization of Probabilistic Image Modeling

- OOD detection 은 데이터중에서 in distribution 과 out of distribution 둘 중에서 OOD 를 잘 구분해 낼 수 있는 task 이며 distance-based method 와 density-based method 가 있는데 후자는 generative model 을 활용하여 직접 확률 분포를 modeling 하고 likelihood 를 얻어서 OOD 여부를 판별하는 방법이라는 사실, in distribution 보다 OOD 가 오히려 더 높은 likelihood 를 부여 받는 경우가 생길 수 있는데 그 이유 중 generative model 로 확률모델을 만들 때 image dataset 간에 서로 공유되는 local feature 들이 존재하고 likelihood score 에 큰 가중치를 부여해서 OOD 임에도 불구하고 높은 score 를 갖게 될 수 있다는 사실, 이를 극복하기 위해 Product of Experts 를 도입하여 model 을 local expert 와 non-local expert 로 나누고 이렇게 나누어진 모델을 이용하여 새로운 likelihood score 를 제안한 점, product of expert 는 확률분포는 여러 개의 분포의 곱으로 나타낼 수 있고 다시 renormalization 해주면 된다는 점, OOD 에서는 일반적으로 AUROC 라는 score measure 를 사용한다는 사실을 알게 되었습니다.

이동진: Co-Mixup

- data augmentation 은 입력 데이터에 모델의 generalization 성능을 높이는 기법으로, mixed up 은 두 이미지를 합쳐 하나의 이미지를 만드는 것이며, 어떤 비율로 어떻게 섞을지에만 초점을 두어 각 이미지의 정보가 제대로 보존이 안되는 단점이 있는데, batch 단위의 mixed-up framework 를 통해 각 이미지의 중요한 영역이 잘 보존될 수 있고 다양한 조합의 augmented image 생성이 가능하다는 사실, Saliency 라는 말은 뉴럴 네트워크가 객체를 인식할 때 중요하다고 여기는 관심 영역이며 이를 수치화 할 때 Background 보다는 class 의 label 과 직접적인 관련이 있는 부분이 더 높게 나온다는 점, Background 보다는 class 의 label 과 직접적인 관련이 있는 부분, patch 별로 어느 인풋이 얼마만큼 반영될지 비율을 결정하는 multinomial distribution 가 있고 ' 개의 (x) matrix 구하는 것이 목표임, 여기서 z 가 충족해야 할 부분은 saliency 를 최대화 하고 augmented image 의 smoothness 이라는 점, ' 개의 동일한 augmented image 가 생성되는 문제는 objective function 에 compatibility term 을 추가하여 패널티를 주어 해결한다는 점, 그리고 과하게 penalization 되는 것을 방지하기 위한 lower bound 로 구성한다는 점, 다른 augmentation 은 pairwise 인데, Co-Mixup 은 batch 단위로 mix 하기 때문에 성능이 좋다는 의문에 대해서도 비교를 통해 다른 augmentation 은 pairwise 인데, Co-Mixup 은 batch 단위로 mix 하기 때문에 성능이 좋다는 사실을 알게 되었습니다.

이동휴: Efficient and Scalable Bayesian Neural Nets with Rank 1 Factors

- 수업시간에 배웠던 VI를 다시 한 번 복습할 수 있었고, batch ensemble의 목적 및 개념, 즉 Ensemble 네트워크는 수가 증가함에 따라 비용이 선형적으로 증가하는 경향이 있으며, 이를 해결하기 위해 batch ensemble을 제안하였고 ensemble 모델이 공유하는 가중치 W 에 각 모델의 rank1 행렬을 elementwise 하게 곱하여 가중치 행렬을 계산한다는 점, 이를 통해 모델 학습시 가중치의 행렬이 아닌 두 개의 vector를 업데이트하여 학습비용을 상당히 줄일 수 있다는 점, 특히 가중치 분포에 대한 학습비용을 줄이기 위해, 두 가지 컨셉을 결합한 개념을 제시한 점, BNN 가중치 분포를 rank1 parameterization 한다는 내용, 가중치간의 사후분포 correlation을 제거하는 효과를 주어 학습의 성능을 향상시키는 부분을 알게 되었습니다.

이성주: Free Lunch For Few-Shot Learning: Distribution Calibration

- No Free Lunch Theorem는 모든 상황에 대해 만능인 best classifier model은 존재하지 않는다는 의미이고, 과거 Learning 연구들은 주로 Stronger model을 만드는 것을 목표로 했는데, 데이터 개수가 증가하면 Ground Truth distribution에 가깝게 근사할 수 있기 때문이라는 사실, 반면 데이터 개수가 적은 경우 (Few-shot) 모델이 학습 샘플에만 overfitting 되는 경향이 있다는 점, 이렇게 biased distribution을 estimation하면 모델의 일반화 성능에 큰 타격을 주게 된다는 점, 만약 biased된 distribution을 GT(Ground Truth) distribution과 비슷하게 Calibration할 수 있다면 좋은 성능을 보일 것이라는 점을 알게 되었습니다.

이성호: Efficiently Modeling Long Sequences with Structured State Space

- Hippo는 High-Order Polynomial Projection Operators의 약자로, HIPPO의 목적은 Computational efficiency가 개선되고, Stable한 gradient flow가 가능하다는 점, Timescale이 robust하고 Approximation error에 bound가 보장된다는 점, linear state space model은 훈련할 때 recurrent view 혹은 convolutional view로 볼 수 있는데 computationally efficient하지는 않다는 점, 그리고 Deep State Space Model 및 S4 Network의 효율성을 확인하였고 학습 및 추론은 recurrent 및 convolutional representation 모두에서 수행하는 방식에 대해 알게 되었습니다.

이승윤: Relevance-CAM: Your Model Already Knows Where to Look

- conventional CAM 의 한계를 어떻게 극복하고 성능을 향상시킬지에 대해 알아보았는데, 먼저 CAM 에서 모델 구조의 한계를 겪고 마지막 레이어에서만 잘 동작하며 gradient 사용시 gradient 가 with noise 처럼 변동이 심하게 되는 문제가 있는데, LRP 와 CAM 을 결합한 relevance CAM 을 고안하여 같은 수준의 레이어에서도 특정 class 의 정보를 잘 추출할 수 있다는 사실, CAM 은 모델의 입력이 어느 부분을 보고 class 에 대해 판단했는지 확인하는 기법이라는 점, LRP 는 layer 단위로 역전파를 통해 상관성을 보는 기법이라는 점, 역전파 수행시 class 판단에 영향을 주는 상관성을 node 마다 구하는 것, 이 둘을 합쳐서 기존 CAM 방식에서 weight 파라미터로 상관성을 사용하여 국한되지 않고 gradient 변동문제도 해결할 수 있으며, layer 가 깊어지면서 채널이 많아지고 점차 목표한 인식을 제대로 할 수 있는 것을 알게 되었습니다.

이아영: COVID_AL: The diagnosis of COVID-19 with deep active learning

- data annotation 의 시간과 비용이 많이 드는 문제를 해결하기 위해 weakly-supervised deep active learning framework 을 적용한 COVID-Active learning 이 무엇인지 확인하였으며, Active Learning 은 학습모델이 data instance 에 대한 label 을 query 하여 원하는 출력으로 data point 에 레이블을 지정할 수 있는 머신러닝의 한 종류이며, model 이 label 할 query 를 생성하는 Membership Query Synthesis, 모델이 label 이 필요한지 여부를 판단하는 Stream-based Selective Sampling, 그리고 큰 데이터 pool 이 존재하여 중요한 data point 를 선택하는 Pool-based Sampling 이상 세 가지로 구성된다는 점, active learning 에서는 model 이 중요한 data 를 찾을 수 있도록 어떠한 query 전략을 선택하는지가 중요하다는 점을 알았고, 패턴인식 과목에서 배운 내용이 바이오 공학에 접목되는 실사례로 COVID-AL framework 의 구성을 흥미롭게 보았고, COVID 진단을 위한 네트워크는 Loss prediction network 와 classification network 로 구성되는 부분, 다른 AI 애플리케이션들에 대한 labeling cost 를 줄이는 방법은 weak label 로 딥러닝 모델을 훈련시키고, 불필요한 labeling 노력을 줄이기 위해 가장 정보가 많이 담긴 data 샘플을 찾아내는 것이며, 이러한 data sample 이 유의미한지 판단하기 위해 input data 의 Uncertainty 와 Diversity 를 사용한다는 사실을 알게 되었습니다.

이원도: Re-envisioning the comparison between Neural Collaborative Filtering and Matrix Factorization

- 최근에 기존의 선형적 추천시스템 모델이 딥러닝 모델보다 예측성능이 뛰어나다는 주장이 나오며 새로운 각도에서 머신러닝을 바라보는데, 구체적으로 선행연구의 실험을 재현해 선형적 추천시스템 모델(i.e. Matrix Factorization)과 딥러닝 모델(i.e. Neural CF)의 성능을 재차 비교하고, '예측성능' 지표 외에 '추천다양성' 지표에서 선형적 추천시스템 모델과 딥러닝 모델을 비교함으로써 선형적 추천시스템 모델이 딥러닝 모델에 비해 예측성능이 뛰어날 수 있음을 알게되었고 (하지만 추천다양성의 지표에서는 선형적 추천시스템 모델보다 딥러닝 모델의 성능이 우수하다는 사실도 같이 알게 됨), 기존 선형적 모델이 딥러닝 모델보다 뛰어난 이유로는 딥러닝 모델이 간단한 내적 함수를 학습하기 위해 많은 파라미터와 학습 데이터를 필요로 하기에 예측하고자 하는 함수가 내적으로 효과적으로 모델링된다면, 딥러닝 모델은 예측에 어려움을 느낄 수 있다는 점, 그리고 비정형 데이터가 아닌 정형적 수치데이터를 최적화에 있어 딥러닝 모델보다 선형모형이 효과적일 수 있다는 사실을 알게 되었습니다.

이종민: Neural Reprojection Error: Merging Feature Learning and Camera Pose Estimation, CVPR, 2021

- visual localization 의 과정, R 과 T 의 reflection error 로 표현된다는 점, matching distribution, 즉 주어진 descriptor 을 통해 matching 이 이루어 질 때 3D 포인트를 2D 어느 픽셀에 대응시킬 것인지, 그리고 두 가지 distribution 이 동일해야 하며 object function 에도 minimum 이 존재한다는 사실, object function 에 Gaussina Kernel 을 적용하여 전체적인 object function 을 smooth 하게 해주고 t 의 근사를 사용하는 방법으로 소개된 Gaussian-homotopy-like method 를 알게 되었습니다.

이주원: Bayesian Triplet Loss: Uncertainty Quantification in Image Retrieval,

- image retrieval 은 어떤 이미지를 넣었을 때 그와 비슷한 이미지를 찾아주는 function 이라는 사실, 공분산이 커짐에 따라 거리제곱의 평균이 커져서 uncertainty 도 덩달아 커진다는 사실을 기반으로 새로운 image embedding 방식을 고안한 점, image 가 stochastic feature 로 보고 prior, likelihood, posterior 를 정의한다는 점, 정규분포를 따르는 '타우'의 expectation 등의 파라미터를 구하면서 likelihood 를 구하는 과정, embedding 에 L2 normalization 을 쓰는 이유 등을 알게 되었습니다..

이주현: KiloNeRF,

- NeRF 는 Neural radiance field 의 약자로, 광선에 기반하여 3 차원 오브젝트를 학습하는 딥러닝 베이스의 렌더링 기법이라는 점과 어떤 3 차원 오브젝트를 랜덤한 뷰에서 찍은 이미지를 가지고 딥러닝 네트워크를 학습시켜, 훈련셋에 없던 카메라 위치에서의 이미지를 합성하게 하는 뉴럴 렌더링 기법이라는 사실, 한 레이어 위의 각 샘플들에 대한 컬러와 밀도 값을 뽑아내는 네트워크라는 점, 이것을 사용하면 훈련되지 않은 뷰에 대한 이미지를 렌더링할 수 있다는 장점이 있지만, 한 이미지에 대해 모든 픽셀을 광선 기반으로 추측해야 하므로 속도가 느려진다는 점, 그래서 속도 향상을 위해 기존의 NeRF 에서 네트워크의 크기만 줄이는 방식을 적용하면 비주얼 퀄리티가 감소되기 때문에, KiloNeRF 에서는 오브젝트를 구획별로 나누고, 각 구획에 NeRF 보다 더 작은 네트워크를 각각 할당하여 속도 증진을 이룰 수 있다는 점, 오브젝트 구획화에 이은 KiloNeRF 의 훈련 전략은 distillation 인데 KiloNeRF 의 파라미터 값을 초기화 하고 이후 이미지의 픽셀값에 대해 photometric loss 를 적용함으로써 파인튜닝을 할 수 있다는 점을 알게 되었습니다.

이진규: Fast Projection Onto Convex Smooth Constraints

- constraint 가 존재하는 최적화 문제를 푸는 다양한 방법중 SGD 가 있고 대표적인 SGD 로 Newton Projection algorithm 이 있다는 점, 이러한 projection operation 은 quadratic optimization 을 풀게 만들기 때문에 constraint 를 approximation 하여 optimization problem 을 tractable 하게 만드는 방법, 이것은 dual function 을 approximated gradient 를 활용하는 것이며 original objective function 의 lagrangian 을 먼저 구하고, (입실론 그라디언트 값을 구할 때 approximation 을 구하는 것이 핵심), dual function 의 solution 을 구하고, 비교하여 original optimization solution 을 구하는 것을 알게 되었습니다.

이태경: AdaSGN: Adapting Joint Number and Model Size for Efficient Skeleton-Based Action Recognition

- SGN 을 이용하여 동작인식에 필요한 각 skeleton 에 대한 최적의 joint 개수와 크기를 예측하는 방안 및 효율적으로 feature 를 추출하고 action 을 예측하는 것을 목표로 하여 각 skeleton 에 대해 optimal 한 joint number 및 optimal network size 를 선택하도록 하는 점, 기존대비 SGN 차이는 action recognition 을 할 때 최적의 joint 개수와 network size 를 사용하여 정확성과 효율성 사이의 균형을 제어할 수 있는 점, classification 단계에서는 각 skeleton 에 대한 액션을 모델 액션과 조인트 넘버 액션으로 분리하고 classification feature 를 계산하면 된다는 점, loss function 에서 accuracy loss 는 모델의 분류 정확도에만 연관되고 efficiency loss 는 computational cost 에만 연관된다는 사실 등을 알게 되었습니다.

장주영: Emerging Properties in Self-Supervised Vision Transformers

- DINO 는 간단한 Self-supervised method 로서 momentum encoder 및 Multi-crop training 이 해당되며 ViT 는 Vision Transformer 로서 convnets 과 경쟁관계라는 사실과 Computationally 하게는 더욱 demanding 하며 더 많은 training data 를 필요로 한다는 점을 알게 되었고, DINO 를 기존 transformer 계열에 적용하여 (DINO 와 ViT 의 결합) 기존에 ViT 나 convnets 에서 못 보았던 scene layout 이나 object boundary 를 확인할 수 있고, label 이 없는 self-supervised 로 적용할 수 있다는 점, KNN 에서 높은 accuracy 를 달성할 수 있다는 점, 그리고 collapse 를 피하기 위해 centering 이나 sharpening 의 방법이 있는데 centering 은 batch 에 대해 덜 의존적이고, uniform distribution 을 encourage 한다는 사실과, Sharpening 은 teacher softmax normalization 를 위해 temperature 에 대해 낮은 값을 사용한다는 사실을 알게 되었습니다.

전창범: Amortized Conditional Normalized Maximum Likelihood -Reliable Out of Distribution Uncertainty estimation

- 머신러닝 모델에 distribution shifted test data 입력이 발생할 경우 정확성에 높은 수준의 overconfident mistake 가 발생할 수 있다는 점, out-of-distribution 입력 발생시 calibration 에 관한 기존의 다른 uncertainty estimation 방법들과 비교하여 개선시킬 수 있는 포인트로 Distribution shift 로 인해 accuracy 및 uncertainty estimation 이 열화 되고 이로 인해 높은 수준의 overconfident mistake 가 발생하게 되는 점, 에러가 더욱 발생할 만한 out-of-distribution test input 들에 대해 uncertainty estimation 을 좀 더 보수적으로 하겠다는 아이디어, CNML(Conditional Normalized Maximum Likelihood)에서 목표로 하는 것은 Regret 이 최소화되는 distribution p 를 찾는 것인데, 동일한 query point 에 대해 각기 다른 label 을 할당하게 되어 이 두 가지 모델이 모두 training data 와 일치하게 되는 경우가 발생하는 경우가 있으며, 이를 위해 CNML 을 generalization 하여 regularized estimator 를 사용하고, 각 label 에 대한 MLE 대신 normalized maximum a posteriori (NMAP)를 취하는 방법, 그리고 training loss 를 simple approximate Bayesian posterior Density 로 대체하고, Training 중에 approximate posterior density 를 학습할 수 있어 test time 에 좀 더 간단하게 문제 해결이 가능하게 된다는 점 등을 알게 되었습니다.

정구엽: Characterising the Role of Pre-Processing Parameters in Audio-based Embedded Machine Learning

- embedded ML 의 제한된 조건에서, 신뢰할 만한 성능을 보장하기 위해 model-centric 의 접근법이 아닌 embedded ML 의 data-centric view 에 대해 생각해 볼 수 있었으며 pre-processing 파라미터가 embedded ML 시스템에서의 다양한 performance metrics 의 균형을 맞추는데 역할을 할 수 있다는 사실을 알게 되었고, Embedded ML 에서 포함하는 reliable performance 의 측면은 Accuracy, Inference latency, Energy consumption, Model fairness 가 있다는 것을 알 수 있었고, Pre-processing parameter 들이 매우 중요한데 model-centric 이외에도 data-centric 접근법이 embedded ML 에 효과가 있을 것이라는 점, embedded ML 의 제한 조건에 따라 절대적으로 우위에 있는 파라미터들도 존재한다는 점을 알게 되었고, Audio KW 에 대한 내용 위주이지만 data centric 접근법을 살펴보면 다른 분야에도 적용이 가능할 것 같다고 생각하게 되었습니다.

정만수 : SeFa - Closed-Form Factorization of Latent Semantics in GANs,

- GAN 에서 문제점은 다른 generative 모델중 Variational autoencoder 가 의미를 갖는 경우가 있는데, Random input 은 아무런 의미가 없는 latent space 가 아니냐는 의문점이 있다는 사실, Latent space 가 있을 때 영상의 객체인 고양이들의 전체적인 포즈는 동일하면서 종이 달라지는 것을 볼 수 있었고, PCA 에 축이 여러 개가 있는데 다른 축을 더해보면 고양이의 종은 유지하며 포즈, 앵글이 바뀌는 현상이 나온다는 것은 latent space 에도 어떤 의미 있는 방향이 있다는 사실을 추정해 볼 수 있었으며, 기존에는 empirically sampling 된 데이터를 가지고 의미 있는 direction 을 찾을 수 있었는데, 이번에는 Generator 자체만 갖고 의미 있는 방향을 찾아보는 시도가 수업시간에 배웠던 PCA 와 연계되어 흥미롭게 볼 수 있었으며, 결국 첫번째 레이어로부터 가장 많이 바뀌는 direction 을 구해보는 과정과, 핵심 포인트로 Lagrange multiplier 를 사용하여 풀어보면 N_i 는 $A(\text{transpose})A$ 의 eigenvector 라는 사실, (즉 필요로 하는 latent space 에서 가장 meaningful 한 방향은 간단하게 weight 의 $A(\text{transpose})A$ 의 eigen vector 이라는 사실), contents 는 유지하면서 특정 feature 만 바꿀 수 있다는 사실 등을 알게 되었습니다.

정명훈: Posterior Network: Uncertainty Estimation without OOD Samples via Density-Based Pseudo-Counts

- classification task 에서 잘 학습된 NN(Neural Network)모델이 있다고 가정할 때 제대로 답변할 수 없는 input 이 있을 때에도 NN 은 임의의 답변을 내놓게 된다는 단점, 그리하여 NN 가 답변에 대해 얼마나 uncertainty 를 modeling 할 수 있는냐는 포인트, uncertainty 의 두 종류중 Aleatoric Uncertainty (즉 data uncertainty)는 class 가 overlap 되거나 label 에 noise 가 끼어 있는 경우에 해당하며 데이터로부터 줄일 수 있고, Epistemic uncertainty (즉 model uncertainty)는 unseen 데이터에 대해 모델이 학습한 적이 없기 때문에 생긴다는 점, uncertainty estimation 은 reliable 한 시스템을 구축하는데 중요하고, Posterior Network 는 normalizing flow 를 사용하여 어떤 입력 샘플이 들어오더라도 Prediction 된 확률에 대한 posterior 분포를 잘 학습하여 uncertainty 를 잘 반영할 수 있다는 점을 알게 되었습니다.

정옥철: CReST: A Class-Rebalancing Self-Training Framework for Imbalanced Semi-Supervised Learning

- 핵심 아이디어로 recall 과 precision 의 차이점 및 둘 간의 trade-off 가 존재한다는 사실, self-training 의 방식, minority class 의 precision 이 높고 training set 에 포함시키기에는 덜 위험하다고 판단한다는 점, 그리고 minority class 로부터 sample 을 추가함으로써 training set 이 전체적으로 balance 를 갖추게 되어 classifier 가 덜 bias 된다는 내용, Progressive distribution alignment 는 class balancing 에서 조정을 취하는 것인데 회차를 거듭할수록 class rebalancing 의 영향력을 증대시키는 것이라는 사실 (generation 에 대해 t 를 줄이면서 class-rebalancing 의 강도를 progressive 하게 증가시키는 것임), model 의 bias 를 활용하여 minority class 에 대한 성능 저하를 감소시킬 수 있다는 점, CReST(Class rebalancing self-training framework)는 labeled sector 에 매우 높은 quality 의 pseudo label 을 제공함으로써 minority class 가 더 자주 그리고 공격적으로 업데이트 된다는 점, distribution alignment 기법을 사용하여 class rebalancing strength 를 더 강화시키고 이를 통해 만든 CReST+라는 개념에 대해 알게 되었습니다.

정재호: Energy-based Out-of-Distribution Detection,

- Energy-based model 은 input space 의 각 point x 를 energy 라 불리는 non-probabilistic scalar 에 매핑시키는 점, 신경망을 학습하는 방법은 에너지를 최소화하여 확률값을 크게 하는 것과 같다는 사실, 그리하여 logit 값을 에너지로 정의하고 해당 에너지를 score 로 사용하여 OOD (out of distribution detection)를 수행하는 것이 핵심이라는 점, ID 샘플에 대해서는 확률값이 높게 나오고 OOD 샘플에 대해서는 확률값이 낮게 나온다는 사실에 착안하여 OOD 샘플을 detection 한다는 점 Softmax 대비 에너지 스코어로 비교할 때 ID 와 OOD 의 차이가 확연함을 파악할 수 있다는 점. Threshold 기준으로 크면 in-distribution, 낮으면 out-of-distribution 로 취급하는 점을 알 수 있었고, OOD input 을 detection 하기 위해 에너지 score 를 활용한 통합된 framework 가 어떤지, logit 을 energy 로 정의하고 에너지 스코어를 사용하여 in distribution 은 에너지를 낮게, out of distribution 은 에너지를 높게 하여 OOD 를 수행한다는 점, (즉 in-distribution 과 out of distribution 간의 에너지 차이가 효과적인 differentiation 을 가능케 할 수 있다는 점), training 중에는 적은 에너지 값을 in distribution 데이터에 인가하고 더 높은 에너지 값을 OOD training 데이터에 인가하는 점, inference time 에는 에너지는 어떠한 pre-trained NN 을 위한 softmax confidence 를 대체할 수 있다는 점, shifted logit space 대신에 원래 logit space 에서의 작업은 각 샘플에 대한 더욱 유용한 정보를 제공한다는 점 등을 알게 되었습니다.

정지은: Sharpness-Aware Minimization for Efficiently Improving Generalization

- generalization 을 효율적으로 향상시키기 위한 Sharpness-Aware Minimization 을 어떻게 하는지 알게 되었는데 먼저 현재 사용되는 모델에 파라미터의 수가 매우 많고 복잡하기에 train loss 로만 훈련하는 방법은 local minima 에 빠지기 쉬워 generalization 에 대한 보장을 하지 못하는 경향이 있다는 점, 이를 해결하기 위해 flat minima 를 통해 generalization 성능을 이끌어내는 방법, flat minima 및 sharp minima 의 개념, loss 가 적으면서 flat 한 점에서의 minima 를 찾는 것을 목표로 하는 이유, 그리고 flat minima 를 찾기 위해 objective function 에 sharpness 와 관련된 항을 추가하여 loss 를 minimize 하는 값을 찾게되면 기존의 loss 가 작으면서 sharpness 또한 작은 optimal 지점을 찾을 수 있다는 점, SAM 알고리즘의 원리, 이를 통해 Sharpness 를 반영하면서 원래 loss 또한 minimize 시킬 수 있다는 점 등을 알게 되었습니다.

조성민: Deconfounded Recommendation for Alleviating Bias Amplification

- 핵심사항으로 Bayesian Network 에 do-operator 를 추가할 경우 Causal Reasoning 에 도움이 될 수 있다는 점을 알게 되었고, 구체적으로 먼저 Causal Reasoning 의 3 가지 Action 중 Observation 은 확률변수의 값을 관측하고 이를 토대로 나머지 변수들의 확률을 추정하는 것이고, Intervention 은 확률변수의 값들이 특정한 값이 되도록 외부적 요인을 더하는 것이며, Counterfactual 은 어떠한 상황을 먼저 observation 하고 그 상황에서 어떤 변수들이 다른 값을 가질 경우 결과가 어떻게 바뀔지 추론해 보는 것이라는 사실, observation 과 intervention 의 다른 점, intervention 에 해당하는 do operator 는 특정 확률변수의 값을 고정시키고, 단순한 observation 과 달리 causal graph 를 변형시킨다는 점, 즉 값을 외부적으로 고정시켰기 때문에 해당 변수의 원인에 해당하는 노드와 링크를 모두 없앤다는 사실, confounding variable 의 의미, 해당 추천시스템의 기본적인 formulation, bias amplification 문제의 의미, backdoor adjustment 기법 도입, predictor 가 갖추어 졌을 때 bias 를 무조건 제거하는 것이 아니라 user 의 상황에 맞게 제거하는 것도 방법이 될 수 있음을 알게 되었습니다.

주재석: Do 2D GANs Know 3D Shape? Unsupervised 3D Shape Reconstruction from 2D Image GANs

- 핵심사항으로 2D image GAN 으로부터 만들어진 image 는 물체의 3D shape 를 recover 하기에 충분한 정보를 이미 포함하고 있다는 가정에서 연구를 출발하였는데, 이러한 과정을 보며 연구 초기에 통찰력 있는 가정을 세우는 것이 중요하다는 사실을 다시금 깨달았고, 구체적으로 먼저 GAN 의 핵심요소는 generator 와 discriminator 이며 generator 는 random 하게 만들어진 vector 로부터 input image 를 만들어내며 discriminator 가 input image 를 data set 에 있는 실제 이미지와 비교하여 진위여부를 구별하여 더 이상 구별되지 않을 때까지 training 을 계속하면 결국 실제 이미지와 매우 유사한 이미지를 만들어 낸다는 사실, GAN2Shape model 은 photo-geometric Autoencoding Design 을 사용하는데 렌더링 과정을 거쳐 이미지를 나타내는지, 그리고 어떻게 2D image 로부터 물체의 3D shape 로 recover 하는 과정은 첫째로 input image 의 물체가 대부분 convex 형태라는 점을 이용하여 초기 depth 를 타원체로 설정하는 등 weak shape prior 를 사용하는 단계이고, 둘째는 렌더링 과정을 거쳐 다양한 시점과 빛의 방향으로 pseudo sample 을 만들어내고 encoder 와 GAN 을 통해 projected sample 로 reconstruct 한다는 점, 이를 통해 비지도적인 방법으로도 물체의 3D shape 를 recover 할 수 있다는 점, view/light/depth/albedo 모두를 reconstruction loss 를 최소화 하는 방향으로 학습을 시키면 view/light/depth/albedo 가 모두 물체의 3D shape 를 올바르게 inference 를 하는 방향으로 네트워크 training 이 가능하고 반복하면 좋은 성능을 낼 수 있다는 점, Ellipsoid Prior 및 GAN 을 이용하여 생성한 Pseudo Projected Sample 들을 활용하여 물체의 3D Shape

Reconstruction 을 비지도학습의 방법으로도 잘 수행할 수 있다는 점을 알게 되었습니다.

채민석: Learning to Balance: Bayesian Meta-Learning for Imbalanced and Out-of-distribution Tasks

- 핵심사항으로 imbalance 한 문제를 Meta-learning Bayesian 으로 어떻게 푸는지에 대한 과정을 알 수 있었는데, 구체적으로 먼저 Meta-learning 은 학습하는 방법을 배우는 스마트 러닝이며, 파라미터와 데이터셋이 두 가지라는 점과 이전 task 의 데이터나 경험이 주어진 가운데 새로운 task 를 더욱 빠르고 효율적으로 학습하기 위한 방법이라는 점, 메타러닝의 imbalance 문제로 task imbalance 는 support set 의 각 class 별 task 의 수가 다른 문제, class imbalance 는 추출하는 모집단의 class 분포가 다르다는 점, out of distribution task 는 평가를 할 때 query set 에 처음보는 class 가 들어오는 상황이며, 이러한 문제의 해결을 위해 Loss function 에 적절한 weight 를 주거나, task 간 imbalance 는 learning rate 에 weight 를 주어 task 간 학습 정도를 적절하게 고르게 하며, out of distribution 에는 처음보는 클래스에 대해서는 입실론 포인트가 다르기 때문에 맞추어 준다는 사실, 가정과 추정방법을 통해 p 의 분포를 구할 때 어떻게 하면 q 의 분포를 좀 더 좋은 분포로 가정하고 풀 것인지에 대해 statics pooling 을 통해 중복도를 고려하여 분포를 만들고, mean 과 Var 를 구하고 각각의 분포를 추출하여 메타러닝 프로세스 진행을 한다는 사실을 알게 되었습니다.

최석훈: Adversarial Self-Supervised Learning for Semi-Supervised 3D Action Recognition

- 핵심사항으로 3D action 인식을 위해 adversarial self-supervised learning 이라는 'self-supervised learning 과 supervised learning 을 결합한' semi-supervised learning 및 이를 통한 action recognition 을 알게 되었고, 구체적으로 adversarial regularization 및 neighborhood consistency 를 고려한 loss 를 활용하여 효율적 semi-supervised learning 방식을 어떻게 구성하는지에 대해 알게 되었으며, Action recognition 은 video 으로부터 어떤 행동을 하는지 인식하는 것으로 Skeleton based 로 접근하는데 joint 부분을 찾고 해당 부분들의 연결움직임으로부터 어떤 동작을 하는지 classification 하는 방법이라는 사실, Semi-supervised learning 은 대부분 unlabeled 된 데이터로부터 learning 을 진행하는 것인데, VAE/GAN 과 같은 generative 모델이 대표적이라는 사실, 그리고 Self-supervised learning 은 Network 스스로 'pre-text'에서 label 을 만들어서 이후 task 에 적용하므로 데이터 레이블링이 필요 없기에 시간 및 비용을 줄이고 사용자에 의한 편향성을 줄일 수 있다는 사실, 알고리즘에서 labeled data 는 supervised learning, unlabeled data 에 대해서는 self-supervised learning 을 하여 전체적으로 semi-supervised learning 을 진행하는 구조, Neighborhood consistency 의 개념 등을 알게 되었습니다.

최웅철 : MD-kNN: An instance-based Approach for Multi-Dimensional Classification

- 핵심사항으로 Instance based learning 을 통해 다차원 분류를 어떻게 할 것인지, 가장 좋은 classifier 를 어떻게 찾을 것인지에 대해 알게 되었으며, 구체적으로 먼저 Instance based learning 은 새로운 데이터가 들어오는 경우, 학습된 데이터와 새로운 데이터의 similarity 를 측정하여 새로운 데이터를 가장 유사도가 높은 기존 학습 데이터의 class 로 분류하는 것이고 대표적인 예로 categorization 가 있다는 점, MD-kNN 의 개요, Multi-dimensional classification 여러 feature 를 동시에 보유한 자료들을 잘 분류하는 것이 목적이고 여기서 새로운 변수인 를 잘 분류해 주는 classifier 를 선택하는 과정, 그리고 instance-based approach 를 이용하여 multi-dimensional classification 을 시도하는 과정, MD-kNN 은 KNN counting statistics 에 기반한 MAP 을 활용하여 unseen instance 의 class vector 를 예측하는 것이 목표라는 점 등을 알게 되었습니다.

최진혁: LiTANIM2

- 핵심사항으로 LiTAMiN 의 속도와 accuracy 의 trade-off 경향을 극복하고 computational cost 도 줄일 수 있는 방안을 알게 되었는데, 구체적으로 먼저 LiTAMIN SLAM 이 무엇인지, SLAM 은 Lidar 지도를 작성하는 동시에 차량의 위치를 작성된 지도 안에서 인식하는 기법으로 미지의 환경에 대한 지도를 작성할 수 있어서 자율주행을 가능하게 할 수 있다는 사실, 이를 위해 Slam 을 사용하는데 high computational cost 와 accuracy 가 trade off 로 작용하는 분야라는 점, Lidar 에서 개선해야 할 중요 이슈인 두 포인트 클라우드를 registration 할 때 둘 간의 거리를 최소화하는 기법인 ICP 를 이용하여 computational cost 를 줄이면서 accuracy 도 SOTA 급으로 유지할 수 있다는 사실, 포인트 클라우드에서 edge point / planner point 등을 뽑아 feature 매칭을 기반으로 registration 해주는 feature based method 도 있다는 점, 기존 LiTAMIN 과 새로 제안하는 LiTAMIN2 의 차이점은 mean 뿐만 아니라 covariance 를 이용하고 Point to normal distribution mapping 에서 normal distribution to normal distribution mapping 이 되는 과정 및 Voxel size 를 늘리고 포인트 수를 줄여서 computational cost 를 줄이고, Distribution shape 을 고려하여 accuracy 를 높인다는 점, 그리하여 LiTAMiN 의 속도와 accuracy 의 trade-off 경향을 극복하고 computational cost 도 줄일 수 있음을 알게 되었습니다.

편준호: Contrastive Learning for Compact Single Image Dehazing

- 핵심사항으로 autoencoder 를 사용한 image dehazing 에 관해, dehazing 의 의미(이미지 촬영시 희뿌옇게 보이는 현상) 및 Deep Generative Model 의 일종인 autoencoder 에 제안된 contrastive regularization 을 추가한 방식으로 dehazing 을 어떻게 해결하는지 알 수 있었고, 구체적으로 먼저 해당 구조는 dehazing network 와 contrastive regularization 의 두 부분으로 구성되어 있으며 deformable convolution 은 기본 convolution 연산에 offset 을 추가하여 이미지의 부분부분 중요도에 따라 변화된 커널을 적용하고, 커널의 크기를 좌우하는 offset 값은 train 시 자동으로 학습된다는 점, encoding 단계에서 얻어진 feature 를 매칭되는 decoding layer 와 mix 하여 다음 decoding layer 로 전달하는데 이 mixed up 관계는 convex combination 이며 user 가 그 비율을 조절할 수 있다는 점, 이를 반복하여 low layer 의 정보를 최종 단계까지 전달할 수 있다는 점, 이렇게 이미지를 복원하고 이어서 loss 를 줄이도록 하여 clean image 와 복원 image 간의 차이를 줄이고 L1 loss 를 사용하고 복원 이미지를 anchor 로 사용하고 이를 기준으로 clean image 와는 당기고 hazing image 는 미는 과정 및 contrastive regularization 을 통해 보다 자연스러운 복원 이미지를 얻는 점, mix-up module 을 적용하여 low level feature 의 정보를 최종 레이어까지 전달할 수 있다는 점, encoding 으로 얻어진 latent space 에서 다양한 이미지 enhancement 모듈을 적용시켜 decoding 시 성능 향상에 기여할 수 있다는 점을 알았습니다.

한영웅: Extracting Training Data from Large Language Models

- 핵심사항으로 training data extraction attack 이슈의 개선을 위해 decaying temperature 로 sampling 하여 다양한 문자열을 생성하거나 conditioning on internet text 로 직접 인터넷에서 적당한 문자열을 찾아서 사용하는 것을 고려할 수 있다는 점, 구체적으로 먼저 훈련된 language 모델로부터 training data 를 역으로 추출하는 방법, 그리고 실제로 overfitting 이 일어나지 않았음에도 불구하고 language model 로부터 training data 가 추출될 수 있다는 점, 딥러닝 기반의 Language model 은 그동안 입력 시퀀스 전체에서 정보를 추출하는 방향으로 발전했다는 점, likelihood 가 비정상적으로 큰 것을 추려내는 법, 몇 개의 문서에서만 등장했음에도 개인정보가 추출되는 이슈 해결을 위한 몇 가지 방법, 즉 Training Set 에서 개인적인 정보를 제거하거나 Training Set 자체를 잘 선정하거나, 아니면 Differential Privacy 를 기반으로 모델을 훈련하거나 Fine-tuning 을 통해 training data 를 잊게 만드는 방법이 있음을 알게 되었습니다.

한형록: Reducing Information Bottleneck for Weakly Supervised Semantic Segmentation

- 핵심사항으로 DNN 에서 후반부에 있는 sigmoid 나 softmax 와 같은 bidirectional 한 activation function 이 DNN 의 information bottleneck 을 일으키는데 단순히 activation map 을 없애고 training 을 진행하면 information bottleneck 을 없앨 수 있다는 점, 구체적으로 확인해보면 전체 픽셀 대비 gradient 가 0.3 이상인 픽셀의 비율을 HGR 로 놓고 확인을 해보면 마지막 layer 에서 information bottleneck 이 급격히 일어나고 localization map 이 target object 에 국소적으로 생겨서 모델의 성능을 저하시킬 수 있다는 점, final layer 에 activation function 을 없애고 fine tuning 을 하면 마지막 layer 에서 gradient 값이 fine tuning 을 더 진행 할수록 커지며 실제 영상에서 사물의 segmentation 이 좋아지고 information loss 를 줄일 수 있음을 알게 되었습니다.

황지훈: Memory Enhanced Global-Local Aggregation for Video Object Detection

- 핵심사항으로 local 과 global 정보를 동시에 사용하는 모델을 base model 이라 지칭했고 여기에 Long Range Memory 를 추가한 방식을 MEGA 라고 제안하여 이전 frame 에서 연산된 feature 들을 memory 에 저장하고 다시 사용하는 방식으로 연산량을 최대한 늘리지 않으면서 Connection 자체를 늘려 정보부족 문제를 해결하는 방안, 구체적으로 video object detection 은 single image 만 받아서 Detection 을 수행하는 것이 아니고 연속된 frame 으로부터 얻는 정보들을 이용해 Detection 을 수행하려는 task 라는 점, Single frame 만을 이용했을 때 발생하는 문제점으로 Motion blur, occlusion, out of focus 가 있고, 사람이 모호한 물체 발견 시 이전 기억속에서 high semantic similarity 를 가지는 object 와 연결하고 Object 의 존재 유무를 확신 못하는 경우 연속된 motion 정보로 추정하는 원리에 기반한 점, 마지막에는 최종 feature 를 faster RCNN 같은 object detection 모델을 사용하여 classification 과 regression 을 수행하는 구조에 대해 알게 되었습니다.