

# KAKAO

Vo1.03

# AI

2017.05

AI CODE

```
vector<pairif_t> c_dmf_topk(...)
{
    Map<const VectorXf> vv(v, M), AA(A, N);
    Map<const MatrixXf> XX(X, N, M);
    vector<pair<int, float>> result;
    RowVectorXf dot = hidden_layers[0].transpose() * vv;
    for(int j=1; j < (int)hidden_layers.size(); ++j)
        dot = dot * hidden_layers[j];
    dot = dot.array().max(0.0f);
    VectorXf ex = fast_softmax(XX, dot, AA);
    get_fast_knn(ex, k, result);
    return result;
}

vector<pairsf_t> prediction_by_vector(const float* v, int k)
{
    auto ret = pool_.enqueue([](const float* _v,
                                vector<FactorType>& hidden_layers,
                                const float* sfixb, const float* sfix,
                                int N, int M, int _k){
        return c_dmf_topk(_v, hidden_layers, sfixb, sfix, N, M, _k);
    }, v, hidden_layers_, (const float*)softmax.data(), k);
    auto val = move(ret.get());
    vector<pairsf_t> result = as_result(val);
    return result;
}
```

kakao



카카오에서 매일 발행하는 리포트입니다.

# KAKAO AI REPORT

Vo1.03

**발행일** | 2017년 5월 19일

**발행처** | (주)카카오

**발행인** | 카카오 정책지원팀

**총괄** | 정혜승

**편집** | 김대원, 김명수, 정수현

**디자인** | 김성현

**메일** | kakaoireport@kakaocorp.com

## COVER

카카오 AI 리포트의 표지에선 AI와 관련된 의미 있는 코드들을 매월 소개할 예정입니다.

Vol.03 코드 | 김광섭 [lucaskim@kakaocorp.com](mailto:lucaskim@kakaocorp.com)

표지 코드는 다음뉴스의 '나를 위한 추천'

서비스를 구성하는 토로스 추천시스템의

알고리즘 일부를 발췌한 것입니다. '나를 위한

추천'은 딥러닝 기술을 이용하여 뉴스 이용자의

소비 패턴과 소비한 기사의 내용을 학습하고,

개인 별로 맞춤형 기사를 제공하는 서비스입니다.

## contents

### preface

카카오 AI 리포트 3호를 내며 02

### research

#### AI 연구 지금까지의 발자취

정수현, 김대원 | ICML과 NIPS 발표논문 6,163건 메타 분석 06

김진화 | AI연구자들이 열광하는 ICML, NIPS 24

엄태웅 | 딥러닝 연구의 현재와 미래 part 1.(1/2) 28

### industry

#### AI, 혁명의 주역

노명철, 이주영 | 사진에서 사람을 읽다 AI in Kakao 36

최예림 | 산업 현장 속으로 들어간 AI 42

### information

AI 유튜브 강의 모음 46

### closing

마치며 48

# 카카오 AI 리포트 3호를 내며

저희는 왜 ‘카카오 AI 리포트’를 내고 있을까요. 이번 호는 이 같은 질문에 대해 또 다른 답을 찾아가는 과정 같습니다. 카카오는 인공지능(AI) 기술을 활용해 새로운 혁신에 도전하고자 하는 기업이지만, 저희 힘 만으로는 부족합니다. 저희는 세상을 바꾸는 AI 기술 동향과 흐름에 대해 국내에서도 더 많은 관심과 논의가 필요하다고 생각합니다. 훌륭한 인재들이 AI 연구에 더 적극적으로 뛰어들어 서로 지식을 교환하면서 새로운 가능성을 탐색할 수 있지 않을까요.

AI 연구가 활발한 해외에서는 대체 어느 방향으로 학술적 탐험이 진행되고 있을까요. AI 연구의 트렌드를 세밀하게 살펴보는 것도 의미가 있다고 판단했습니다. 지난 30여 년 척박한 연구 환경에서도 꾸준히 AI 연구의 중심이 되어 온 두 학회에서 2005~2016년 발표된 6,163건의 논문에 대해 메타 분석을 시도했습니다. 대체 어떤 키워드로 연구가 발전해왔는지, 연구자들이 가장 많이 인용하는 논문은 어떤 것인지, AI 사대천왕을 잇는 핵심인물은 누구인지 구경하시죠. 이런 작업의 의미는 실제 학회에 논문을 발표하신 서울대 박사과정 김진화 님을 통해 살펴봤습니다. 워낙 발전 속도가 빠른 딥러닝 분야에서는 1년 된 논문이 ‘고전’이 되고, 알고리즘이 2개월 만에 구식이 되어버린다는데, 캐나다 워털루공대의 엄태웅 님이 연구 동향을 조목조목 분석했습니다. 사실 이 글들은 인문학적, 사회학적 인간, 기술 비전문가들에게는 난이도가 매우 높습니다. 그럼에도, AI를 본격 연구하고자 하는 분들에게는 참고가 될 것이라 기대해봅니다. AI가 제조업 현장에 미치는 영향은 경기대 최예림 교수의 기고에서 확인할 수 있습니다. ‘4차 산업혁명’이라 부를게 아니라 ‘AI 혁명’이라 부르자는 이유가 뭔지 살펴보시죠.

카카오의 각 서비스에도 AI 기술이 다양하게 활용되고 있는데요. 이번에는 카카오맵의 ‘로드뷰’에서 사용하는 기술을 자세하게 소개했습니다. 카카오에서 조슈아, 마이클이라는 영어 이름으로 불리는 노명철, 이주영 님이 정리해줬습니다. 이렇게 구체적인 기술 문서를 저희도 오랜만에 읽어봤는데요, 그저 존경할 따름입니다. AI 연구자들에게는 분명 도움이 되리라 믿어봅니다.

대중에게 공개된 AI 동영상 강의도 직접 다 들어보고 골라봤습니다. 쉽지 않지만, 저희도 계속 도전중입니다. 함께 관심 가져보시죠.

2017년 5월 19일  
카카오 정책지원팀 드림

# AI 연구

## 지금까지의

## 발자취

research

정수현, 김대원 | ICML과 NIPS 발표논문 6,163건 메타 분석

06

김진화 | AI연구자들이 열광하는 ICML, NIPS

24

엄태웅 | 딥러닝 연구의 현재와 미래 part 1. (1/2)

28

자율주행차가 도로를 주행하고, AI가 환자를 진단하고, 바둑과 텍사스 홀덤 포커 게임까지 하는 등 AI는 지금 다양한 분야에서 눈부신 성과를 보여주고 있습니다. 이러한 결과들은 수많은 연구자들이 오랜 기간 각고의 노력으로 이루어낸 결실들입니다. 이번 글에서는 AI 연구가 현재 수준으로 발전하기까지 가장 중요한 역할을 해온 두 학회인 NIPS와 ICML을 통해 12년간 발표된 논문들을 메타 분석한 결과를 살펴봅니다. 논문 참여 저자들 간의 관계, 연도별 주요 논문들의 피인용횟수(citations), 논문 제목에 등장한 핵심어(keyword)들을 중심으로 분석했습니다. 두 학회가 AI 연구에 차지하는 의미와 메타 분석의 결과에 대해선 지난해 12월 NIPS를 통해 논문을 발표한 바 있는 서울대 인지과학 박사과정 김진화 님이 문답 형식으로 친절하게 설명해주셨습니다. 최근 5년간 양과 질적으로 크게 성장한 딥러닝에 대해서는 캐나다 워털루 공대 박사과정 재학 중인 엄태웅 님의 글을 통해 딥러닝 연구의 현재와 미래를 확인하실 수 있습니다.

# ICML과 NIPS 발표논문 6,163건 메타 분석

2005년~2016년

12년간 발표 논문 분석

최근 인공지능(AI) 연구는 다양한 분야에서 눈부신 성과들을 보여주고 있다. 이는 오랜 기간 이 분야를 연구해온 AI 연구자들의 노력의 결과다.

AI 연구 분야에서 가장 대표적인 학회는 인용 횟수, 참가자 수, 발표

논문 수, 역사 등을 종합할 때 ICML(International Conference on Machine Learning, 머신러닝 국제컨퍼런스)과 NIPS(Neural Information Processing Systems, 신경정보처리시스템 학회)

두 곳이 꼽힌다. ICML이 머신러닝 분야에 집중해왔다면 NIPS는

인지과학과 머신러닝 응용분야 등 좀 더 넓은 분야를 다루고 있다.

ICML은 37년 전인 1980년 첫 학술대회를 개최했고, 올해 8월 호주

시드니에서 제34차 학술대회를 개최한다.

1987년 출범한 NIPS는 2017년 12월 캘리포니아에서 제30차'

학술대회를 개최할 예정이다. 본 연구팀은 최근 AI 연구 동향을

살펴보기 위해 2005년부터 2016년 사이 두 학회를 통해 발표된 논문

6,163건에 대한 메타 분석을 시도했다. 과거 12년간 가장 많이 인용된

논문들을 최다 누적 피인용 횟수, 저자 간 공동작업 관계망, 논문 키워드

변화 등 다각도로 분석해 AI 연구의 최근 흐름과 변화상을 살펴볼 수

있었다.

글 | 정수현 noah.jung@kakaocorp.com

숫자의 숨은의미? 머니볼의 힘? 숫자들에서 남들이 찾지 못한 숨겨진 의미를 찾아내는 걸 좋아한다. 어릴적 모바일 게임을 개발하던 코딩 실력으로 딥러닝을 쫓아가려고 허덕거리며 달리고 있다. 하지만, 오래 달리는건 자신있다. 마라톤을 달리며 나와의 싸움을 이기는데 익숙해 있기 때문이다. 비록 늦게 출발했지만 언젠가는 트랙에서 함께 뛰고 있기를 기대하며, 지금도 비록 느리지만 달리고 있다.

글 | 김대원 ive.kim@kakaocorp.com

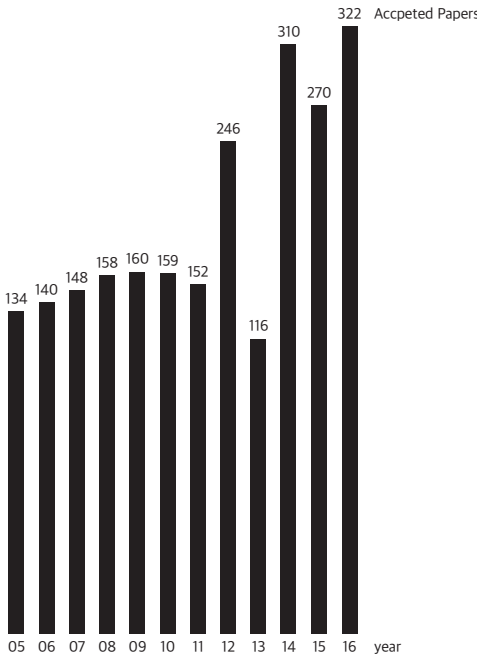
로봇저널리즘을 시발로, 로봇 그리고 인공지능이 사회에 어떻게 확산될 지, 그리고 앞선 새로운 기술이 확산되기 위한 조건이 무엇인지를 살펴보고 있다. 로봇 그리고 인공지능의 건전한 발전을 위해서는, 사회과학도 공학에 버금가는 날개를 확보해야 한다고 생각한다. 두 날개로 날아야 새는 멀리 갈 수 있기에.

2005년부터 2016년까지 ICML 학회를 통해 발표된 논문은

총 2,315건이다. 지난 2016년 발표 논문(accepted paper)은

322편으로, 11년 전인 2005년 134편 대비 2배 이상 증가했다.

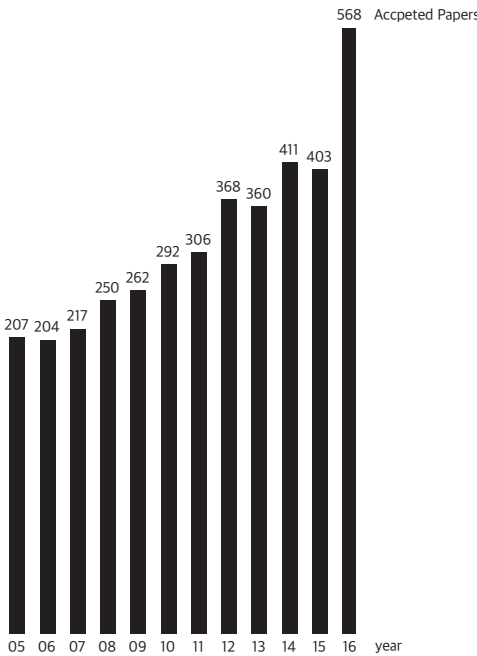
[ ICML 학회 12년간 발표논문 수 변화 추이 ]



NIPS의 경우, 2005년 207편에 머물던 발표 논문의 수가 2016년

568편으로 역시 2배 이상 증가했다.

[ NIPS 학회 12년간 발표논문 수 변화 추이 ]



[ 분석방법론 ]

1. 자료 수집

ICML Conferences <http://www.machinelearning.org/icml.html>

NIPS Proceeding <https://papers.nips.cc/>

2005년 부터 2016년까지 12년간 발표 논문(accepted papers) 자료를 활용하였다.

2. 표본 데이터

ICML에서 2005년 부터 2016년까지12년간 발표된 총 2,315개의 논문, 같은 기간

NIPS에서 발표된 총 3,848개의 논문 데이터를 활용하였다.

3. 논문 인용횟수 조회

<https://scholar.google.co.kr/> 사이트를 통해서 각 논문 제목별 피인용횟수(citations)

를 조회했다. 시간이 지남에 따라 새로운 논문들이 발표되면서 기존 논문들의 인용회 수가 조금씩 증가하고 있다. 분석을 위해 특정 시점인 4월 21일 조회한 피인용횟수를 기준으로 하였음을 알려 둔다.

4. 자료 수집

[https://github.com/giallo41/Data\\_Science/tree/master/Conf](https://github.com/giallo41/Data_Science/tree/master/Conf)

수집한 데이터 파일과 분석을 위해 사용한 파이썬(python) 소스를 올려 두고 있다.

- ICML 과 NIPS의 12년간 논문제목, 저자, 논문별 피인용횟수 파일을 엑셀에 저장하고 이를 파이썬의 판다스 데이터 프레임(pandas dataframe)을 통해 분석 작업을 진행 하였다.

- 각 학회의 연도별 가장 높은 피인용횟수를 기준으로 dataframe.sort() 함수를 사용하여 가장 높은 순위의 10개 논문을 선택하였다.

- 저자 이름별로 12년간 누적피인용횟수를 모두 합산하여 그중 가장 높은 횟수를 기록 한 저자들을 역시 dataframe.sort() 함수를 사용하여 선택하였다.

- 논문제목 분석을 위해 제목을 단어별로 분리하고 모두 소문자로 변환하였다. 이들 중 (',', '?', 'for', 'a', 'an', 'in', 'of', 'with', 'and', 'the', 'to', 'on', 'from', 'by', 'using', 'very', 'via', 'it', 'that', 'as', ',', 'which', '-', 'through', 'without', 'while', 'is', 'than', 'where', 'much', 'many', 'or', 'so') 단어들은 분석에서 제외하였다.

- 연도별 논문 제목에 가장 많이 등장하는 핵심어(keyword) 비율의 상대적 크기를 이 미지로 표현한 워드클라우드(wordcloud) 분석은 파이썬에서 제공하는 워드클라우드 패키지를 사용하였다.

5. 네트워크 분석 방법의 개념과 방법

- 개념 : 연구는 단독으로 이뤄지기도 하고, 여럿이 협업 형태로 진행되기도 한다. 연구자 A 씨가 있다고 가정해보자. A 씨는 (가)라는 연구에도 참여하지만, (나) 연구에 동참할 수도 있다. 이렇게 되면, A 씨는 학계에서 두 연구 모두를 잇는 역할을 할 수도 있다. 이렇게, 하나의 네트워크 안에서 매개의 위치는 서로 다른 정보 혹은 지식을 연결하는 주체로서의 잠재력을 가질 수 있다. 조직 사회학 등을 비롯하여 네트워크 구조에 관심을 갖는 학문에서는 '매개 중심성 (betweenness centrality)'이란 개념을 통해, 구조 내에서 주요 연결자의 역할을 하는(할 수 있는) 개체를 탐색해왔다. 이 방식을 이번 분석을 위해 차용하였다. 이 밖에, 각 개체에 연결된 정도를 확인할 수 있는 연결정도 중심성(degree centrality)도 분석 지표로서 활용됐다. 연결 중심성은 연결망 내에서 해당 주체가 연결 객체(들)에 대해 갖는 영향력의 크기를 의미한다.

- 방법 : 2005년부터 2016년까지 NIPS와 ICML에 발표된 논문의 저자를 한 데 정리했다. 연구자의 수는 ICML이 3,949명, NIPS가 5,878명이었다. 이렇게 만들어진 저자 목록을 행과 열로 세워, 정사각형 행렬을 만들었다. ICML은 3,949 × 3,949 행렬, NIPS는 5,878 × 5,878 행렬이 된 것이다. 연구자 별로 다른 연구자와 논문을 함께 쓴 횟수를 계산했다. 만약, A와 B가 4회에 걸쳐 논문을 함께 썼다면, 행이 A이고 열이 B인 칼럼의 값은 4가 되는 것이다. 이렇게 만들어진 행렬을 네트워크 분석 수단인 UCINET 6.0을 통해 분석했다. 분석 결과로 도출된 연결 중심성과 매개 정도 중심성에서 각각 상위 20위까지를 발췌했다.

### 연도 별 발표 논문의 피인용 횟수

#### ICML

##### [ 2005년 ]

순위	논문	인용	저자들
1	Learning to Rank using Gradient Descent	1,507	Greg Hullender, Chris Burges 외 5인
2	Fast Maximum Margin Matrix Factorization for Collaborative Prediction	707	Nati Srebro, Jason D. M. Rennie
3	A Support Vector Method for Multivariate Performance Measures	706	Thorsten Joachims
4	Comparing Clusterings - An Axiomatic View	522	Marina Meila
5	Semi-supervised Graph Clustering A Kernel Approach	442	Raymond J. Mooney, Brian Kulis 외 2인
6	Learning Structured Prediction Models A Large Margin Approach	441	Carlos Guestrin, Ben Taskar 외 1인
7	Beyond the Point Cloud from Transductive to Semi-supervised Learning	411	Mikhail Belkin, Vikas Sindhwani 외 1인
8	Learning from Labeled and Unlabeled Data on a Directed Graph	346	Bernhard Schalkopf, Dengyong Zhou 외 1인
9	Near-Optimal Sensor Placements in Gaussian Processes	341	Ajit Pauk Singh, Carlos Guestrin 외 1인
10	Non-Negative Tensor Factorization with Applications to Statistics and Computer Vision	337	Tamir Hazan, Amnon Shashua

##### [ 2006년 ]

순위	논문	인용	저자들
1	The Relationship Between Precision-Recall and ROC Curves	1,807	Mark Goadrich, Jesse Davis
2	Dynamic Topic Models	1,440	John D. Lafferty, David M. Blei
3	An Empirical Comparison of Supervised Learning Algorithms	1,132	Alexandru Niculescu-Mizil, Rich Caruana
4	Topic Modeling Beyond Bag-of-Words	623	Hanna M. Wallach
5	Label Propagation through Linear Neighborhoods	569	Changshui Zhang, Fei Wang
6	Cover Trees for Nearest Neighbor	548	John Langford, Alina Beygelzimer 외 1인
7	Pachinko Allocation DAG-Structured Mixture Models of Topic Correlations	512	Andrew McCallum, Wei Li
8	Connectionist Temporal Classification Labelling Unsegmented Sequence Data with Recurrent Neural Networks	426	Juergen Schmidhuber, Alex Graves 외 2인
9	Fast Time Series Classification using Numerosity Reduction	383	Chotirat Ann R, Xiaopeng Xi 외 2인
10	Accelerated Training of Conditional Random Fields with Stochastic Gradient Methods	344	Kevin P. Murphy, S. Vishwanathand 외 2인

##### [ 2007년 ]

순위	논문	인용	저자들
1	Pegasos Primal Estimated sub-Gradient Solver for SVM	1,523	Shai Shalev-Shwartz, Yoram Singer 외 1인
2	Information-Theoretic Metric Learning	1,271	Jason V. Davis, Brian Kulis 외 3인
3	Self-taught Learning Transfer Learning from Unlabeled Data	1,003	Andrew Y. Ng, Honglak Lee 외 3인
4	Learning to Rank From Pairwise Approach to Listwise Approach	993	Zhe Cao, Tao Qin 외 3인
5	Restricted Boltzmann Machines for Collaborative Filtering	818	Geoffrey E. Hinton, R. Salakhutdinov 외 1인
6	Boosting for Transfer Learning	781	Wenyuan Dai, Qiang Yang 외 2인
7	Combining Online and Offline Knowledge in UCT	507	Sylvain Gelly, David Silver
8	An Empirical Evaluation of Deep Architectures on Problems with Many Factors of Variation	452	Yoshua Bengio, Hugo Larochelle 외 3인
9	Trust Region Newton Methods for Large-Scale Logistic Regression	426	Chih-Jen Lin, Ruby Chiu-Hsing Weng 외 1인
10	Scalable Training of L1-regularized Log-linear Models	413	Galen Andrew, Jianfeng Gao

##### [ 2008년 ]

순위	논문	인용	저자들
1	A Unified Architecture for Natural Language Processing Deep Neural Networks with Multitask Learning	1,519	Jason Weston, Ronan Collobert
2	Extracting and Composing Robust Features with Denoising Autoencoders	1,466	Yoshua Bengio, Hugo Larochelle외 2인
3	Bayesian Probabilistic Matrix Factorization using Markov Chain Monte Carlo	836	Andriy Mnih, Ruslan Salakhutdinov
4	Efficient Projections onto the L1-Ball for Learning in High Dimensions	687	Tushar Chandra, John Duchi 외 2인
5	A Dual Coordinate Descent Method for Large-scale Linear SVM	598	S. Sundararajan, Cho-Jui Hsieh 외 3인
6	Training Restricted Boltzmann Machines using Approximations to the Likelihood Gradient	480	Tijmen Tieleman
7	Classification using Discriminative Restricted Boltzmann Machines	437	Yoshua Bengio, Hugo Larochelle
8	Fast Support Vector Machine Training and Classification on Graphics Processors	368	Kurt Keutzer, Bryan Catanzaro 외 1인
9	Deep Learning via Semi-Supervised Embedding	360	Ronan Collobert, Jason Weston외 1인
10	Listwise Approach to Learning to Rank - Theory and Algorithm	345	Hang Li, Fen Xia외 3인

##### [ 2009년 ]

순위	논문	인용	저자들
1	Convolutional Deep Belief Networks for Scalable Unsupervised Learning of Hierarchical Representations	1,379	Andrew Y. Ng, Honglak Lee 외 2인
2	Online Dictionary Learning for Sparse Coding	1,185	Guillermo Sapiro, Julien Mairal 외 2인
3	Group Lasso with Overlaps and Graph Lasso	558	Jean-Philippe Vert, Laurent Jacob 외 1인
4	Learning Structural SVMs with Latent Variables	547	Thorsten Joachims , Chun-Nam Yu
5	Curriculum Learning	453	Jason Weston, Yoshua Bengio 외 2인
6	Evaluation Methods for Topic Models	427	David Mimno, Hanna Wallach 외 2인
7	Feature Hashing for Large Scale Multitask Learning	386	Josh Attenberg, Kilian Weinberger 외 3인
8	Learning with Structured Sparsity	378	Dimitris Metaxas, Junzhou Huang 외 1인
9	Identifying Suspicious URLs An Application of Large-Scale Online Learning	340	Geoffrey M. Voelker, J. Ma 외 2인
10	More Generality in Efficient Multiple Kernel Learning	339	Rakesh Babu, Manik Varma

##### [ 2010년 ]

순위	논문	인용	저자들
1	Rectified Linear Units Improve Restricted Boltzmann Machines	1,484	Geoffrey E. Hinton, Vinod Nair
2	3D Convolutional Neural Networks for Human Action Recognition	747	Kai Yu, Shuiwang Ji 외 2인
3	Robust Subspace Segmentation by Low-Rank Representation	686	Yong Yu, Guangan Liu 외 1인
4	A Theoretical Analysis of Feature Pooling in Visual Recognition	455	Yann LeCun, Y-Lan Boureau 외 1인
5	Gaussian Process Optimization in the Bandit Setting378 No Regret and Experimental Design	378	Matthias Seeger, Niranjan Srinivas 외 2인
6	Detecting Large-Scale System Problems by Mining Console Logs	376	Michael I. Jordan, Wei Xu 외 3인
7	Deep learning via Hessian-free optimization	349	James Martens
8	Proximal Methods for Sparse Hierarchical Dictionary Learning	301	Francis Bach, Rodolphe Jenatton 외 2인
9	Tree-Guided Group Lasso for Multi-Task Regression with Structured Sparsity	282	Eric P. Xing, Seyoung Kim
10	Large Graph Construction for Scalable Semi-Supervised Learning	269	Shih-Fu Chang, Wei Liu 외 1인

##### [ 2011년 ]

순위	논문	인용	저자들
1	Multimodal Deep Learning	665	Andrew Y. Ng, Honglak Lee 외 4인
2	Parsing Natural Scenes and Natural Language with Recursive Neural Networks	554	Andrew Y. Ng, Chris Manning 외 2인
3	Domain Adaptation for Large-Scale Sentiment Classification A Deep Learning Approach	532	Yoshua Bengio, Xavier Glorot 외 1인
4	The Importance of Encoding Versus Training with Sparse Coding and Vector Quantization	466	Andrew Y. Ng, Adam Coates
5	Hashing with Graphs	451	Shih-Fu Chang, Wei Liu 외 2인
6	Contractive Auto-Encoders Explicit Invariance During Feature Extraction	436	Yoshua Bengio, Salah RIFAI 외 3인
7	Manifold Identification of Dual Averaging Methods for Regularized Stochastic Online Learning	415	Stephen Wright, Sangkyun Lee
8	Generating Text with Recurrent Neural Networks	376	Geoffrey E. Hinton, Ilya Sutskever 외 1인
9	On optimization methods for deep learning	356	Andrew Y. Ng, Quoc Le 외 4인
10	Minimal Loss Hashing for Compact Binary Codes	345	David Fleet, Mohammad Norouzi

##### [ 2012년 ]

순위	논문	인용	저자들
1	Building high-level features using large scale unsupervised learning	1,147	Andrew Y. Ng, Marc Aurelio Ranzato 외 5인
2	Conversational Speech Transcription Using Context-Dependent Deep Neural Networks	563	Gang Li, Dong Yu 외 1인
3	Conditional Likelihood Maximization A Unifying Framework for Information Theoretic Feature Selection	408	Mikel Lujan, Gavin Brown 외 2인
4	A fast and simple algorithm for training neural probabilistic language models	198	Yee Whye Teh, Andriy Mnih
5	Modeling Temporal Dependencies in High-Dimensional Sequences Application to Polyphonic Music Generation and Transcription	197	Yoshua Bengio, Pascal Vincent 외 1인
6	Marginalized Denoising Autoencoders for Domain Adaptation	195	Fei Sha, Minmin Chen 외 2인
7	Revisiting k-means New Algorithms via Bayesian Nonparametrics	191	Michael Jordan, Brian Kulis
8	High Dimensional Semiparametric Gaussian Copula Graphical Models	190	Larry Wasserman, Han Liu 외 3인
9	Learning Task Grouping and Overlap in Multi-task Learning	161	Hal Daume III, Abhishek Kumar
10	Making Gradient Descent Optimal for Strongly Convex Stochastic Optimization	148	Karthik Sridharan, Alexander Rakhlin 외 1인

##### [ 2013년 ]

순위	논문	인용	저자들
1	Revisiting Frank-Wolfe Projection-Free Sparse Convex Optimization	322	Martin Jaggi
2	Block-Coordinate Frank-Wolfe Optimization for Structural SVMs	169	Patrick Pletscher , S. Lacoste-Julien 외 2인
3	A Practical Algorithm for Topic Modeling with Provable Guarantees	142	Michael Zhu , Sanjeev Arora 외 6인
4	Connecting the Dots with Landmarks Discriminatively Learning Domain-Invariant Features for Unsupervised Domain Adaptation	135	Fei Sha , Boqing Gong 외 1인
5	Adaptive Task Assignment for Crowdsourced Classification	126	Jennifer W. Vaughan, Chien-Ju Ho 외 1인
6	Stochastic Gradient Descent for Non-smooth Optimization Convergence Results and Optimal Averaging Schemes	122	Tong Zhang , Ohad Shamir
7	Fast dropout training	109	Christopher Manning , Sida Wang
8	Making a Science of Model Search Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures	108	David Cox , James Bergstra 외 1인
9	Stochastic Alternating Direction Method of Multipliers	101	Alexander Gray , Hua Ouyang 외 2인
10	Combinatorial Multi-Armed Bandit General Framework and Applications	98	Yang Yuan , Wei Chen 외 1인

##### [ 2014년 ]

순위	논문	인용	저자들
1	DeCAF A Deep Convolutional Activation Feature for Generic Visual Recognition	1,257	Trevor Darrell, Jeff Donahue 외 5인
2	Distributed Representations of Sentences and Documents	1,005	Tomas Mikolov, Quoc Le
3	Stochastic Backpropagation and Approximate Inference in Deep Generative Models	354	Daan Wierstra, Danilo J. Rezende 외 1인
4	Towards End-To-End Speech Recognition with Recurrent Neural Networks	285	Navdeep Jaitly, Alex Graves
5	Recurrent Convolutional Neural Networks for Scene Labeling	243	Ronan Collobert, Pedro Pinheiro
6	Multimodal Neural Language Models	165	Rich Zemel, Ryan Kiros 외 1인
7	Neural Variational Inference and Learning in Belief Networks	159	Karol Gregor, Andriy Mnih
8	An Asynchronous Parallel Stochastic Coordinate Descent Algorithm	146	Srikrishna Sridhar, Ji Liu 외 3인
9	Deep Generative Stochastic Networks Trainable by Backprop	139	Jason Yosinski, Yoshua Bengio 외 2인
10	Austerity in MCMC Land Cutting the Metropolis-Hastings Budget	114	Max Welling, Anoop Korattikara 외 1인

##### [ 2015년 ]

순위	논문	인용	저자들
1	Batch Normalization Accelerating Deep Network Training by Reducing Internal Covariate Shift	1,356	Christian Szegedy, Sergey Ioffe
2	Show Attend and Tell Neural Image Caption Generation with Visual Attention	663	Yoshua Bengio, Kelvin Xu 외 6인
3	DRAW A Recurrent Neural Network For Image Generation	339	Daan Wierstra, Karol Gregor 외 3인
4	Unsupervised Learning of Video Representations using LSTMs	224	Ruslan Salakhutdinov, N. Srivastava 외 1인
5	An Empirical Exploration of Recurrent Network Architectures	186	Ilya Sutskever, Rafal Jozefowicz 외 1인
6	Deep Learning with Limited Numerical Precision	124	Pritish Narayanan, Suyog Gupta 외 2인
7	Trust Region Policy Optimization	119	Philipp Moritz, John Schulman 외 3인
8	Learning Transferable Features with Deep Adaptation Networks	115	Michael Jordan, Mingsheng Long 외 2인
9	Unsupervised Domain Adaptation by Backpropagation	111	Victor Lempitsky, Yaroslav Ganin
10	Compressing Neural Networks with the Hashing Trick	92	Yixin Chen, Wenlin Chen 외 3인

##### [ 2016년 ]

순위	논문	인용	저자들
1	Ask Me Anything Dynamic Memory Networks for Natural Language Processing	142	Richard Socher, Ankit Kumar 외 7인
2	Asynchronous Methods for Deep Reinforcement Learning	142	Koray Kavukcuoglu, Volodymyr Mnih 외 6인
3	Deep Speech 2 End-to-End Speech Recognition in English and Mandarin	138	Zhenyao Zhu, Dario Amodei 외 10인
4	Pixel Recurrent Neural Networks	105	Koray Kavukcuoglu, Aaron V. Oord 외 1인
5	Dropout as a Bayesian Approximation Representing Model Uncertainty in Deep Learning	96	Zoubin Ghahramani, Yarin Gal
6	Dynamic Memory Networks for Visual and Textual Question Answering	85	Richard Socher, Caiming Xiong 외 1인
7	Train faster generalize better Stability of stochastic gradient descent	62	Yoram Singer, Moritz Hardt 외 1인
8	Low-rank Solutions of Linear Matrix Equations via Procrustes Flow	61	Ben Recht, Stephen Tu 외 3인
9	Autoencoding beyond pixels using a learned similarity metric	60	Hugo Larochelle, Ole Winther 외 2인
10	Dueling Network Architectures for Deep Reinforcement Learning	58	Nando de Freitas, Ziyu Wang 외 4인



## NIPS

## [ 2005년 ]

순위	논문	인용	저자들
1	Distance Metric Learning for Large Margin Nearest Neighbor Classification	1,419	Kilian Q. Weinberger, John Blitzer 외 1인
2	Saliency Based on Information Maximization	942	Neil Bruce, John Tsotsos
3	Laplacian Score for Feature Selection	811	Xiaofei He, Deng Cai 외 1인
4	Correlated Topic Models	725	John D. Lafferty, David M. Blei
5	Sparse Gaussian Processes using Pseudo-inputs	663	Edward Snelson, Zoubin Ghahramani
6	Infinite latent feature models and the Indian buffet process	628	Z. Ghahramani, Thomas L. Griffiths
7	Metric Learning by Collapsing Classes	590	Amir Globerson, Sam T. Roweis
8	Multiple Instance Boosting for Object Detection	578	Cha Zhang, John C. Platt 외 1인
9	Location-based activity recognition	518	Lin Liao, Dieter Fox 외 1인
10	Bayesian Surprise Attracts Human Attention	508	Laurent Itti, Pierre F. Baldi

## [ 2006년 ]

순위	논문	인용	저자들
1	Graph-Based Visual Saliency	2,261	Jonathan Harel, Christof Koch 외 1인
2	Greedy Layer-Wise Training of Deep Networks	2,076	Yoshua Bengio, Hugo Larochelle
3	Efficient sparse coding algorithms	2,030	Andrew Y. Ng, Honglak Lee 외 2인
4	Map-Reduce for Machine Learning on Multicore	1,216	Andrew Y. Ng, Sang K. Kim 외 5인
5	Multi-Task Feature Learning	797	Andreas Argyriou, T. Evgeniou 외 1인
6	Efficient Learning of Sparse Representations with an Energy-Based Model	689	Yann LeCun, Marc'Aurelio Ranzato 외 2인
7	Correcting Sample Selection Bias by Unlabeled Data	632	Jiayuan Huang, Arthur Gretton 외 3인
8	Sparse Representation for Signal Classification	531	Ke Huang, Selin Aiyente
9	A Kernel Method for the Two-Sample-Problem	508	Arthur Gretton, K. M. Borgwardt 외 3인
10	Learning to parse images of articulated bodies	489	Deva Ramanan

## [ 2007년 ]

순위	논문	인용	저자들
1	Supervised Topic Models	1,724	Jon D. McAuliffe, David M. Blei
2	Probabilistic Matrix Factorization	1,573	Andriy Mnih, R. R. Salakhutdinov
3	Random Features for Large-Scale Kernel Machines	844	Ali Rahimi , Benjamin Recht
4	The Tradeoffs of Large Scale Learning	791	Olivier Bousquet, Leon Bottou
5	Sparse deep belief net model for visual area V2	688	Andrew Y. Ng, Honglak Lee 외 1인
6	Direct Importance Estimation with Model Selection and Its Application to Covariate Shift Adaptation	470	Masashi Sugiyama, S. Nakajima 외 3인
7	Sparse Feature Learning for Deep Belief Networks	459	Yann LeCun, Marc'Aurelio Ranzato 외 1인
8	Multi-task Gaussian Process Prediction	402	Edwin V. Bonilla, Kian M. Chai 외 1인
9	Predicting human gaze using low-level saliency combined with face detection	359	Moran Cerf, Jonathan Harel 외 2인
10	SpAM Sparse Additive Models	341	Han Liu, Larry Wasserman 외 2인

## [ 2008년 ]

순위	논문	인용	저자들
1	Spectral Hashing	1,365	Yair Weiss, Antonio Torralba 외 1인
2	Mixed Membership Stochastic Blockmodels	1,147	Edo M. Airoldi, David M. Blei 외 2인
3	Supervised Dictionary Learning	827	Julien Mairal, Jean Ponce 외 3인
4	A Scalable Hierarchical Distributed Language Model	471	Andriy Mnih, Geoffrey E. Hinton
5	Offline Handwriting Recognition with Multidimensional Recurrent Neural Networks	377	Alex Graves, Juergen Schmidhuber
6	Dynamic visual attention searching for coding length increments	370	Xiaodi Hou, Liqing Zhang
7	Policy Search for Motor Primitives in Robotics	360	Jens Kober, Jan R. Peters
8	DiscLDA Discriminative Learning for Dimensionality Reduction and Classification	329	Simon Lacoste-Julien, Fei Sha 외 1인
9	Sparse Online Learning via Truncated Gradient	315	John Langford, Lihong Li 외 1인
10	Near-optimal Regret Bounds for Reinforcement Learning	249	Peter Auer, Thomas Jaksch 외 1인

## [ 2009년 ]

순위	논문	인용	저자들
1	Robust Principal Component Analysis Exact Recovery of Corrupted Low-Rank Matrices via Convex Optimization	828	John Wright, Arvind Ganesh 외 3인
2	Reading Tea Leaves How Humans Interpret Topic Models	780	Jonathan Chang, Sean Gerrish 외 3인
3	Nonlinear Learning using Local Coordinate Coding	613	Kai Yu, Tong Zhang 외 1인
4	Fast Image Deconvolution using Hyper-Laplacian Priors	607	Dilip Krishnan, Rob Fergus
5	A unified framework for high-dimensional analysis of M-estimators with decomposable regularizers	570	P. K. Ravikumar, S. Negahban 외 2인
6	Whose Vote Should Count More Optimal Integration of Labels from Labelers of Unknown Expertise	552	Jacob Whitehill, Ting-fan Wu 외 3인
7	Unsupervised feature learning for audio classification using convolutional deep belief networks	539	Andrew Y. Ng, Honglak Lee 외 2인
8	Efficient Learning using Forward-Backward Splitting	447	Yoram Singer, John C. Duchi
9	Learning to Hash with Binary Reconstructive Embeddings	425	Brian Kulis, Trevor Darrell
10	Locality-sensitive binary codes from shift-invariant kernels	417	Maxim Raginsky, Svetlana Lazebnik

## [ 2010년 ]

순위	논문	인용	저자들
1	Online Learning for Latent Dirichlet Allocation	773	Matthew Hoffman, Francis R. Bach 외 1인
2	Object Bank A High-Level Image Representation for Scene Classification Semantic Feature Sparsification	733	Li-jia Li, Hao Su 외 2인
3	Efficient and Robust Feature Selection via Joint 21-Norms Minimization	579	Feiping Nie, Heng Huang 외 2인
4	The Multidimensional Wisdom of Crowds	452	Peter Welinder, Steve Branson 외 2인
5	Parallelized Stochastic Gradient Descent	448	Martin Zinkevich, Markus Weimer 외 2인
6	Monte-Carlo Planning in Large POMDPs	375	David Silver, Joel Veness
7	Learning Convolutional Feature Hierarchies for Visual Recognition	348	Yann LeCun, Koray Kavukcuoglu 외 4인
8	Robust PCA via Outlier Pursuit	334	Huan Xu, Constantine Caramanis 외 1인
9	Kernel Descriptors for Visual Recognition	272	Liefeng Bo, Xiaofeng Re 외 1인
10	Label Embedding Trees for Large Multi-Class Tasks	235	Samy Bengio, Jason Weston 외 1인

## [ 2011년 ]

순위	논문	인용	저자들
1	HogWild A Lock-Free Approach to Parallelizing Stochastic Gradient Descent	632	Benjamin Recht, Christopher Re 외 2인
2	Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials	562	Philipp Krahenbuhl, Vladlen Koltun
3	Dynamic Pooling and Unfolding Recursive Autoencoders for Paraphrase Detection	401	Andrew Y. Ng, C. D. Manning 외 3인
4	Linearized Alternating Direction Method with Adaptive Penalty for Low-Rank Representation	388	Zhouchen Lin, Risheng Liu 외 1인
5	Algorithms for Hyper-Parameter Optimization	320	James S. Bergstra, Remi Bardenet 외 2인
6	An Empirical Evaluation of Thompson Sampling	293	Olivier Chapelle, Lihong Li
7	Im2Text Describing Images Using 1 Million Captioned Photographs	279	Vicente Ordonez, Girish Kulkarni 외 1인
8	Co-regularized Multi-view Spectral Clustering	266	Abhishek Kumar, Piyush Rai 외 1인
9	Iterative Learning for Reliable Crowdsourcing Systems	254	David R. Karger, Sewoong Oh 외 1인
10	Semantic Labeling of 3D Point Clouds for Indoor Scenes	238	Hema S. Koppula, Abhishek Anand 외 2인

## [ 2012년 ]

순위	논문	인용	저자들
1	ImageNet Classification with Deep Convolutional Neural Networks	11,057	Ilya Sutskever, Geoffrey E. Hinton 외 1인
2	Large Scale Distributed Deep Networks	845	Andrew Y. Ng, Marc'Aurelio Ranzato 외 10인
3	Practical Bayesian Optimization of Machine Learning Algorithms	748	Jasper Snoek, Hugo Larochelle 외 1인
4	Learning to Discover Social Circles in Ego Networks	577	Jure Leskovec, Julian J. McAuley
5	Multimodal Learning with Deep Boltzmann Machines	564	Nitish Srivastava, R. R. Salakhutdinov
6	Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images	340	Dan Ciresan, Alessandro Giusti 외 2인
7	A Stochastic Gradient Method with an Exponential Convergence Rate for Finite Training Sets	262	Nicolas L. Roux, Mark Schmidt 외 1인
8	Convolutional-Recursive Deep Learning for 3D Object Classification	223	C. D. Manning, Andrew Y. Ng 외 3인
9	Image Denoising and inpainting with Deep Neural Networks	181	Junyuan Xie, Linli Xu 외 1인
10	A Simple and Practical Algorithm for Differentially Private Data Release	155	Moritz Hardt, Katrina Ligett 외 1인

## [ 2013년 ]

순위	논문	인용	저자들
1	Distributed Representations of Words and Phrases and their Compositionality	3,600	Jeff Dean, Ilya Sutskever 외 3인
2	DeViSE A Deep Visual-Semantic Embedding Model	381	Jeff Dean, Marc'Aurelio Ranzato 외 5인
3	Accelerating Stochastic Gradient Descent using Predictive Variance Reduction	378	Rie Johnson, Tong Zhang
4	Reasoning With Neural Tensor Networks for Knowledge Base Completion	327	C.D. Manning, Andrew Y. Ng 외 2인
5	Deep Neural Networks for Object Detection	287	Christian Szegedy, Alexander Toshev 외 1인
6	Translating Embeddings for Modeling Multi-relational Data	272	Antoine Bordes, Nicolas Usunier 외 3인
7	Zero-Shot Learning Through Cross-Modal Transfer	225	C.D. Manning, Andrew Y. Ng 외 2인
8	Learning a Deep Compact Image Representation for Visual Tracking	223	Naiyan Wang, Dit-Yan Yeung
9	Phase Retrieval using Alternating Minimization	182	Naiyan Wang, Dit-Yan Yeung 외 1인
10	Deep content-based music recommendation	174	Aaron v. Oord, Sander Dieleman 외 1인

## [ 2014년 ]

순위	논문	인용	저자들
1	Sequence to Sequence Learning with Neural Networks	1,473	Ilya Sutskever, Oriol Vinyals 외 1인
2	Learning Deep Features for Scene Recognition using Places Database	697	Bolei Zhou, Agata Lapedriza 외 3인
3	Two-Stream Convolutional Networks for Action Recognition in Videos	636	Karen Simonyan, Andrew Zisserman
4	Generative Adversarial Nets	574	Yoshua Bengio, Ian Goodfellow 외 6인
5	How transferable are features in deep neural networks?	515	Jeff Clune, Yoshua Bengio 외 2인
6	Deep Learning Face Representation by Joint Identification-Verification	337	Yi Sun, Yuheng Chen
7	Neural Word Embedding as Implicit Matrix Factorization	300	Omer Levy, Yoav Goldberg
8	Recurrent Models of Visual Attention	288	Volodymyr Mnih, Nicolas Heess
9	Joint Training of a Convolutional Network and a Graphical Model for Human Pose Estimation	288	Yann LeCun, Christoph Bregler 외 2인
10	Depth Map Prediction from a Single Image using a Multi-Scale Deep Network	241	David Eigen, Christian Puhrsch 외 1인

## [ 2015년 ]

순위	논문	인용	저자들
1	Faster R-CNN Towards Real-Time Object Detection with Region Proposal Networks	964	Shaoqing Ren, Kaiming He 외 2인
2	Spatial Transformer Networks	263	Max Jaderberg, Karen Simonyan 외 2인
3	End-To-End Memory Networks	240	Sainbayar Sukhbaatar, Arthur szlam 외 2인
4	Skip-Thought Vectors	233	Ryan Kiros, Yukun Zhu 외 5인
5	Teaching Machines to Read and Comprehend	214	Karl Moritz Hermann, Tomas Kocisky 외 5인
6	Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks	214	Emily L. Denton, Soumith Chintala 외 2인
7	Grammar as a Foreign Language	196	Ilya Sutskever, Geoffrey E. Hinton 외 4인
8	Learning both Weights and Connections for Efficient Neural Network	168	Song Han , Jeff Pool 외 2인
9	Character-level Convolutional Networks for Text Classification	159	Junbo Zhao, Yann LeCun 외 1인
10	Training Very Deep Networks	154	Rupesh K. Srivastava, Klaus Greff 외 1인

## [ 2016년 ]

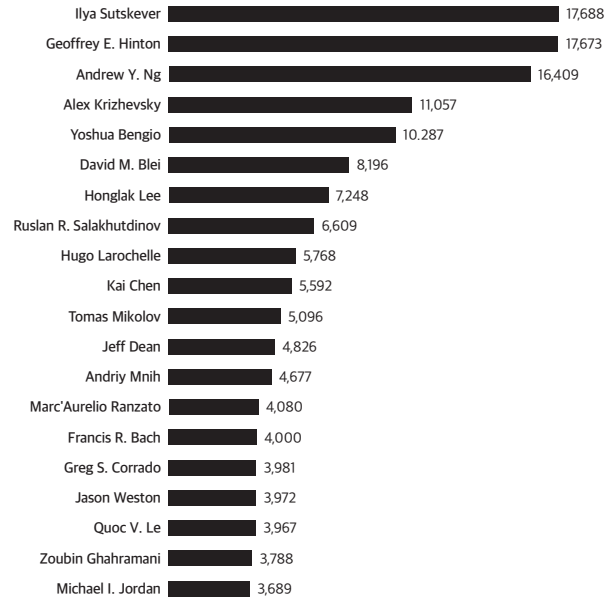
순위	논문	인용	저자들
1	Improved Techniques for Training GANs	123	Tim Salimans, Ian Goodfellow 외 4인
2	InfoGAN Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets	59	Xi Chen, Ilya Sutskever 외 5인
3	Conditional Image Generation with PixelCNN Decoders	46	Aaron Oord, Nal Kalchbrenner 외 4인
4	A Theoretically Grounded Application of Dropout in Recurrent Neural Networks	43	Yarin Gal, Zoubin Ghahramani
5	Matrix Completion has No Spurious Local Minimum	42	Rong Ge, Jason D. Lee 외 1인
6	Generating Images with Perceptual Similarity Metrics based on Deep Networks	42	Alexey Dosovitskiy, Thomas Brox
7	R-FCN Object Detection via Region-based Fully Convolutional Networks	42	jifeng dai, Yi Li 외 2인
8	Deep Exploration via Bootstrapped DQN	40	Ian Osband, Charles Blundell 외 2인
9	Residual Networks Behave Like Ensembles of Relatively Shallow Networks	37	Andreas Veit, Michael J. Wilber 외 1인
10	Hierarchical Question-Image Co-Attention for Visual Question Answering	37	Jiasen Liu, Jianwei Yang 외 2인

피인용(citation)

피인용횟수 기준으로 상위 20위의 저자

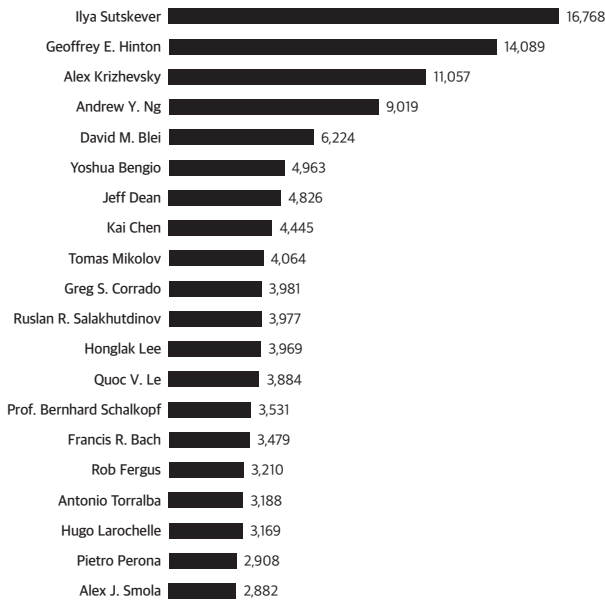
1. NIPS와 ICML 합계

[ NIPS 와 ICML의 2005년 부터 2016년까지 누적 인용횟수가 가장 높은 저자 Top 20 ]



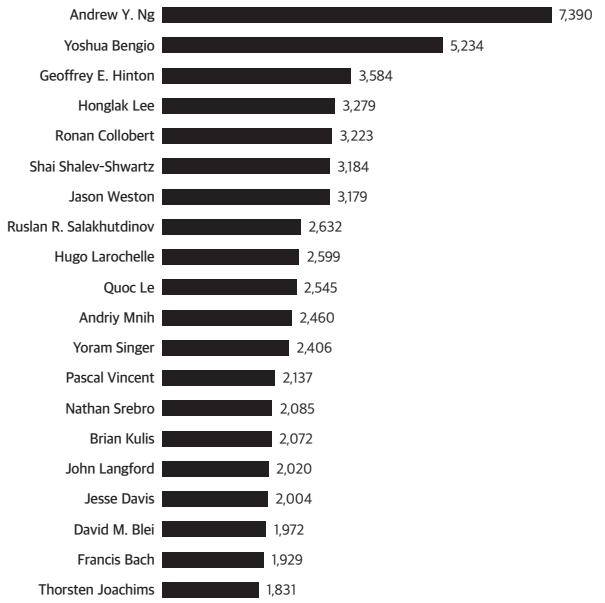
2. NIPS

[ NIPS의 2005년 부터 2016년까지 누적 인용횟수가 가장 높은 저자 Top 20 ]



3. ICML

[ ICML의 2005년 부터 2016년까지 누적 인용횟수가 가장 높은 저자 Top 20 ]



누적 피인용 횟수 상위 30위의 저자들<sup>2</sup>

[ Total ]

순위	저자	인용횟수	논문수
1	Ilya Sutskever	17,688	17
2	Geoffrey E. Hinton	17,673	30
3	Andrew Y. Ng	16,409	46
4	Alex Krizhevsky	11,057	1
5	Yoshua Bengio	10,287	50
6	David M. Blei	8,196	23
7	Honglak Lee	7,248	24
8	Ruslan R. Salakhutdinov	6,609	41
9	Hugo Larochelle	5,768	17
10	Kai Chen	5,592	3
11	Tomas Mikolov	5,096	5
12	Jeff Dean	4,826	3
13	Andriy Mnih	4,677	11
14	Marc'Aurelio Ranzato	4,080	10
15	Francis R. Bach	4,000	24
16	Greg S. Corrado	3,981	2
17	Jason Weston	3,972	17
18	Quoc V. Le	3,967	14
19	Zoubin Ghahramani	3,788	58
20	Michael I. Jordan	3,689	57
21	Tong Zhang	3,617	34
22	Rajat Raina	3,534	4
23	Prof. Bernhard Schalkopf	3,531	29
24	Yoram Singer	3,464	14
25	Eric P. Xing	3,378	34
26	Ronan Collobert	3,366	11
27	Inderjit S. Dhillon	3,346	35
28	Shai Shalev-Shwartz	3,343	25
29	John Langford	3,299	26
30	Rob Fergus	3,272	14

[ ICML ]

순위	저자	인용횟수	논문수
1	Andrew Y. Ng	7,390	19
2	Yoshua Bengio	5,324	21
3	Geoffrey E. Hinton	3,584	9
4	Honglak Lee	3,279	10
5	Ronan Collobert	3,223	9
6	Shai Shalev-Shwartz	3,184	20
7	Jason Weston	3,179	13
8	Ruslan R. Salakhutdinov	2,632	9
9	Hugo Larochelle	2,599	10
10	Quoc Le	2,545	4
11	Andriy Mnih	2,460	7
12	Yoram Singer	2,406	5
13	Pascal Vincent	2,137	4
14	Nathan Srebro	2,085	8
15	Brian Kulis	2,072	6
16	John Langford	2,020	12
17	Jesse Davis	2,004	6
18	David M. Blei	1,972	6
19	Francis Bach	1,929	9
20	Thorsten Joachims	1,831	11
21	Mark Goadrich	1,807	1
22	Inderjit S. Dhillon	1,713	2
23	Trevor Darrell	1,703	5
24	Rich Caruana	1,700	5
25	Jean Ponce	1,640	2
26	Julien Mairal	1,630	5
27	Yong Yu	1,610	4
28	Kai Yu	1,596	9
29	Chris Burges	1,507	1
30	Nicole Hamilton	1,507	1

[ NIPS ]

순위	저자	인용횟수	논문수
1	Ilya Sutskever	16,768	8
2	Geoffrey E. Hinton	14,089	8
3	Alex Krizhevsky	11,057	8
4	Andrew Y. Ng	9,019	8
5	David M. Blei	6,224	8
6	Yoshua Bengio	4,963	8
7	Jeff Dean	4,826	8
8	Kai Chen	4,445	8
9	Tomas Mikolov	4,064	8
10	Greg S. Corrado	3,981	8
11	Ruslan R. Salakhutdinov	3,977	8
12	Honglak Lee	3,969	8
13	Quoc V. Le	3,884	8
14	Prof. Bernhard Schalkopf	3,531	8
15	Francis R. Bach	3,479	8
16	Rob Fergus	3,210	8
17	Antonio Torralba	3,188	8
18	Hugo Larochelle	3,169	8
19	Pietro Perona	2,908	8
20	Alex J. Smola	2,882	8
21	Eric P. Xing	2,844	8
22	Marc'Aurelio Ranzato	2,770	8
23	Andrew Zisserman	2,741	8
24	Tong Zhang	2,720	8
25	Christof Koch	2,626	8
26	Jonathan Harel	2,620	8
27	Zoubin Ghahramani	2,550	8
28	Michael I. Jordan	2,533	8
29	Pradeep K. Ravikumar	2,469	8
30	Yann LeCun	2,344	8

ICML과 NIPS 두 학회에 12년간 누적 피인용 횟수가 가장 높은 저자

15명의 12년간 추이를 나타냈다. 15명을 선정한 이유는 데이터

시각화(data visualization)를 통해 가장 효과적으로 보여줄 수 있는

인원수이기 때문이다. 누적피인용회수 전체 리스트는 GitHub를

통해 확인할 수 있다.

15인 저자들의 피인용횟수가 가장 많은 논문들<sup>3</sup>

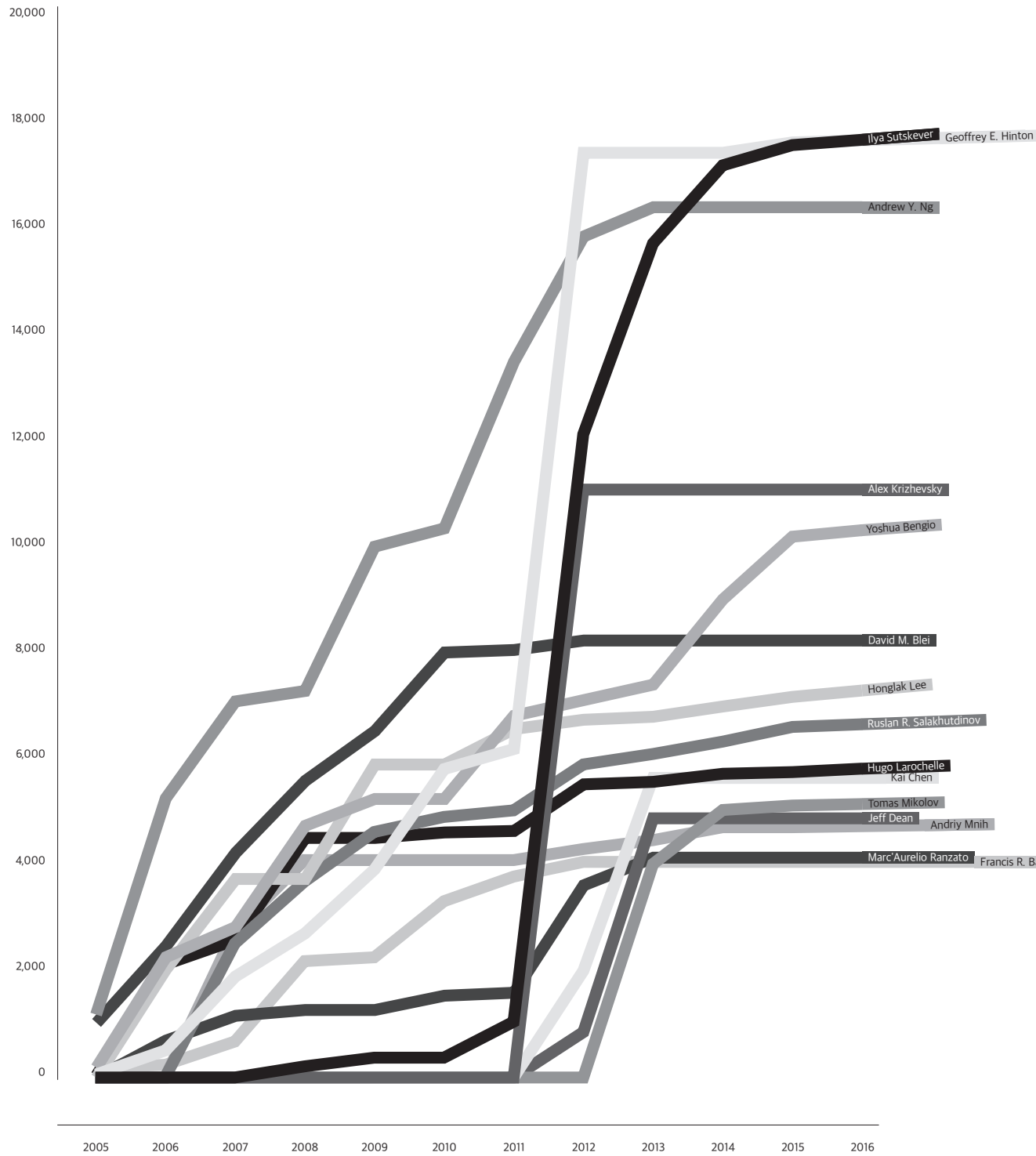
[ ICML & NIPS Cumulative Citations ]

저자	학회	연도	피인용	공동저자
Ilya Sutskever	NIPS	2012	11,057	Geoffrey E. Hinton, Alex Krizhevsky ImageNet Classification with Deep Convolutional Neural Networks
				2013
Geoffrey E. Hinton	ICML	2010	1,484	Vinod Nair Rectified Linear Units Improve Restricted Boltzmann Machines
Andrew Y. Ng	NIPS	2006	2,030	Honglak Lee, Alexis Battle, Rajat Raina Efficient sparse coding algorithms
				2009
Yoshua Bengio	NIPS	2006	2,076	Pascal Lamblin, Dan Popovici, Hugo Larochelle Greedy Layer-Wise Training of Deep Networks
				2008
David M. Blei	NIPS	2007	1,724	Jon D. McAuliffe Supervised Topic Models
				2006
Honglak Lee	NIPS	2006	2,030	Alexis Battle, Rajat Raina, Andrew Y. Ng Efficient sparse coding algorithms
				2009
Ruslan R. Salakhutdinov	NIPS	2007	1,573	Andriy Mnih Probabilistic Matrix Factorization
				2008
Hugo Larochelle	NIPS	2006	2,076	Yoshua Bengio, Pascal Lamblin, Dan Popovici Greedy Layer-Wise Training of Deep Networks
				2008
Kai Chen	ICML	2012	1,147	Andrew Y. Ng, Quoc Le, Marc'Aurelio Ranzato, Rajat Monga 외 2인 Building high-level features using large scale unsupervised learning
Tomas Mikolov	ICML	2014	1,005	Quoc Le Distributed Representations of Sentences and Documents
Jeff Dean	NIPS	2012	845	Greg Corrado, Rajat Monga, Kai Chen, Matthieu Devin, Mark Mao 외 6인 Large Scale Distributed Deep Networks
				2008
Francis R. Bach	NIPS	2010	773	David M. Blei, Matthew Hoffman Online Learning for Latent Dirichlet Allocation



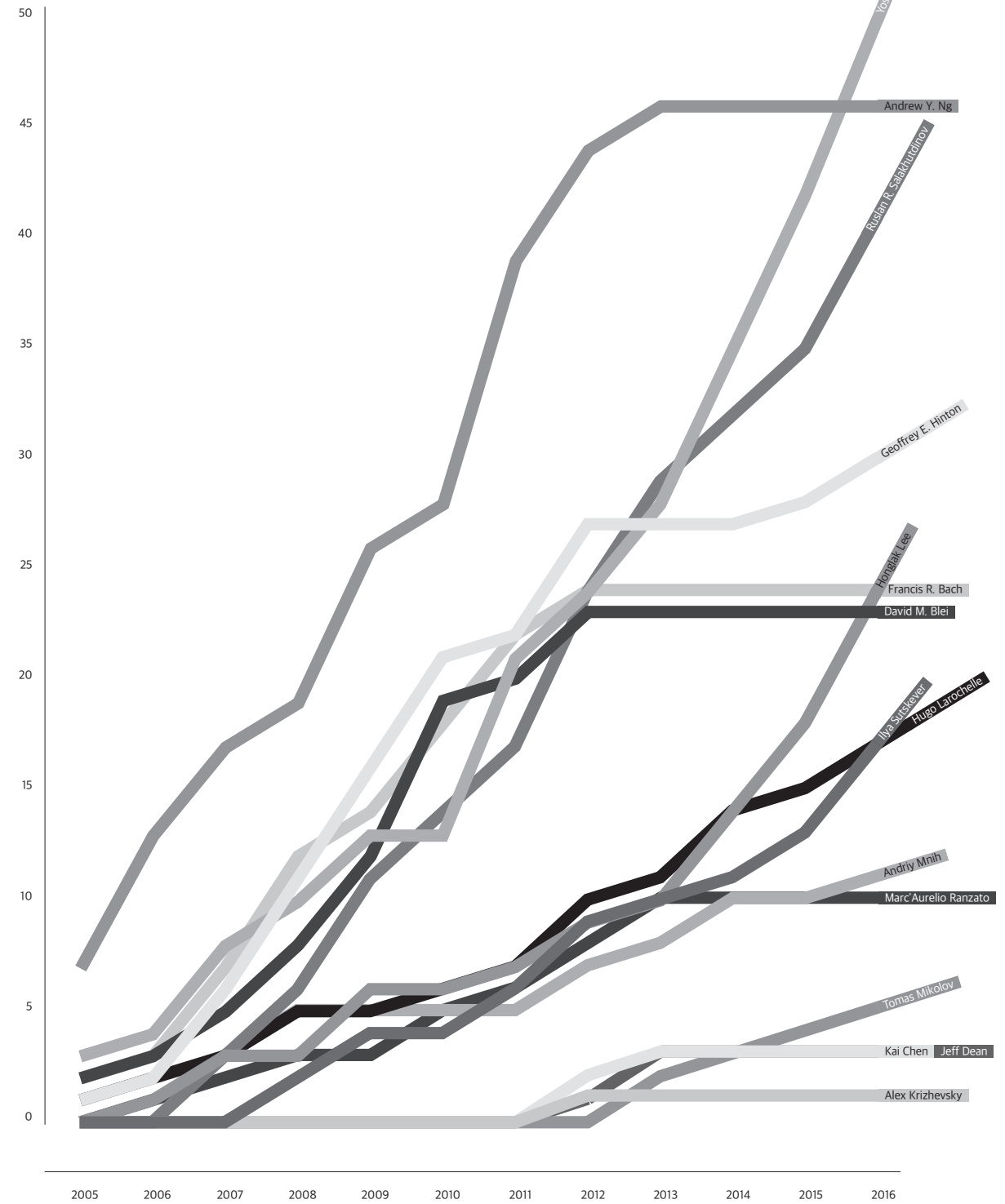
ICML 과 NIPS 두 학회 피인용 논문 횟수가 많은 상위 15명 저자의  
연도별 피인용 횟수 추이

[ ICML & NIPS Cumulative Citations ]



피인용 논문 횟수가 많은 상위 15명 저자의  
연도별 논문 발표 현황<sup>4</sup>

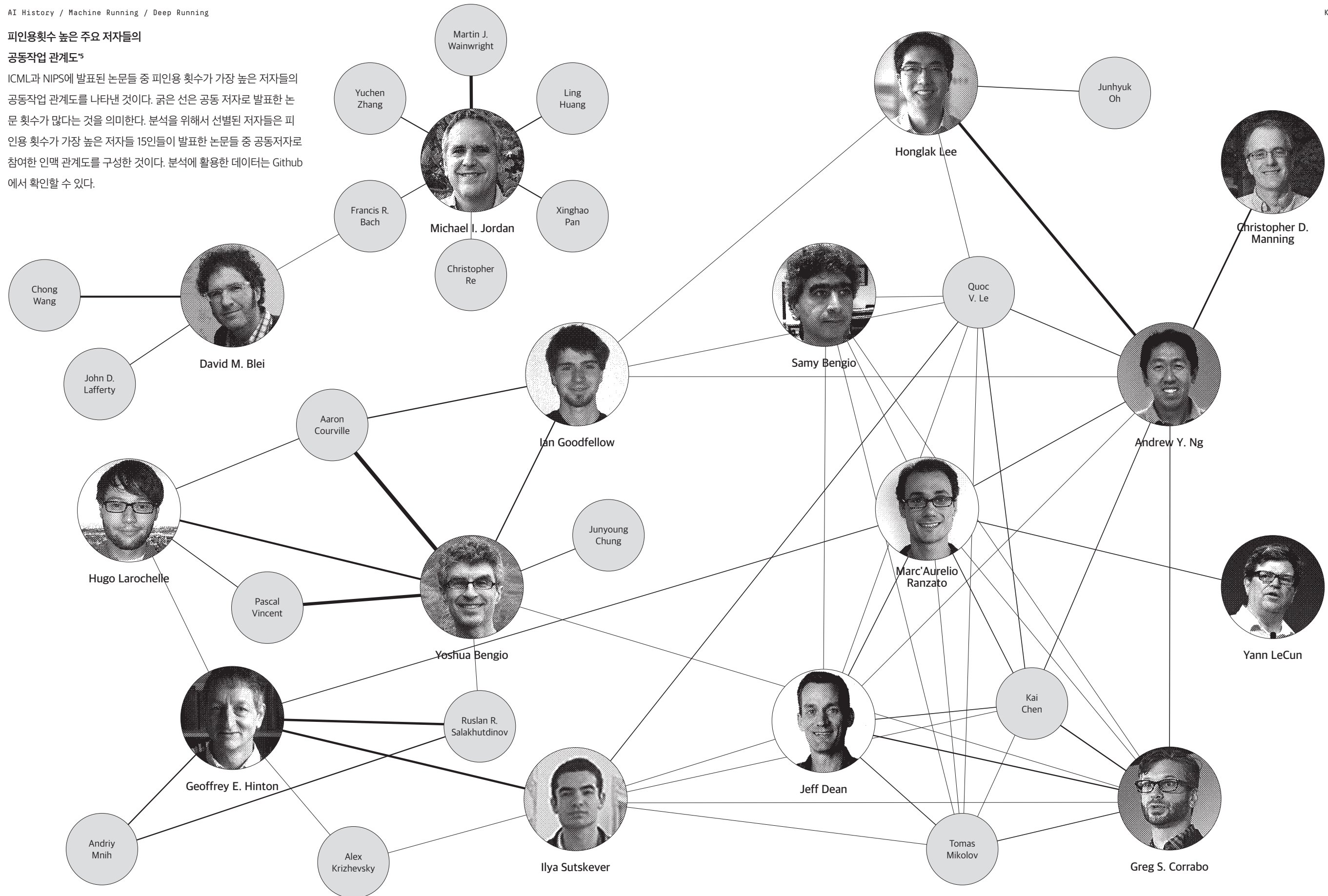
[ ICML & NIPS Cumulative Accepted Papers ]



피인용횟수 높은 주요 저자들의

공동작업 관계도<sup>5</sup>

ICML과 NIPS에 발표된 논문들 중 피인용 횟수가 가장 높은 저자들의 공동작업 관계도를 나타낸 것이다. 굵은 선은 공동 저자로 발표한 논문 횟수가 많다는 것을 의미한다. 분석을 위해서 선별된 저자들은 피인용 횟수가 가장 높은 저자들 15인들이 발표한 논문들 중 공동저자로 참여한 인맥 관계도를 구성한 것이다. 분석에 활용한 데이터는 Github에서 확인할 수 있다.



논문 제목에 가장 많이 등장한 핵심어

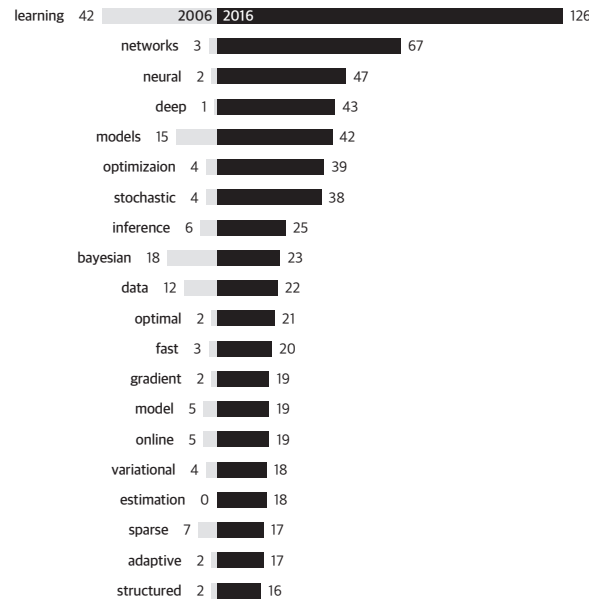
지난 12년간 AI 연구 주제 변화를 간접적으로 알아보는 방법으로 논문 제목에 등장한 핵심어(keyword)의 변화를 분석하였다. 12년간 변화 양상을 개괄적으로 보기 위해 우선 워드 클라우드(word cloud)를 활용했다. 두 학회에서 12년간 평균적으로 가장 많이 등장한 논문 제목의 핵심어는 'Learning'과 'Model'이다. 이 두 단어가 모든 기간에 걸쳐서 제목에 가장 많이 등장한 핵심어였기에, 이 두 단어를 제외한 나머지 단어가 연구의 변화 양상을 나타냈다고 판단했다. 이에 해당 연도별 변화를 살펴보기 위해서, 워드 클라우드를 만드는 과정에서 'Learning'과 'Model' 두 단어를 배제했다. 기존 연도로는 2006년, 2011년, 2016년을 택했다. 5년이라는 시간 간격을 둔 것은 변화를 보다 분명하게 확인하기 위한 조치였다.

2006년과 2016년 논문 제목에 가장 많이 나타난

20개의 핵심어

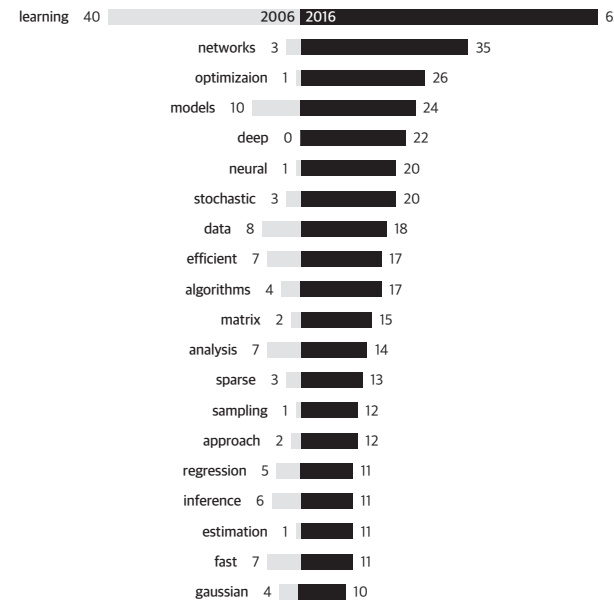
NIPS

[ 20 Most appeared Title Keywords ]



ICML

[ 20 Most appeared Title Keywords ]



ICML

[ 2006년 ]



[ 2011년 ]



[ 2016년 ]



NIPS

[ 2006년 ]



[ 2011년 ]



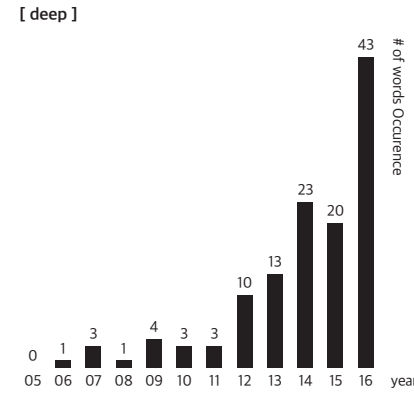
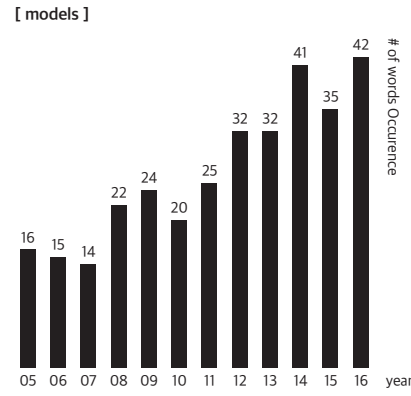
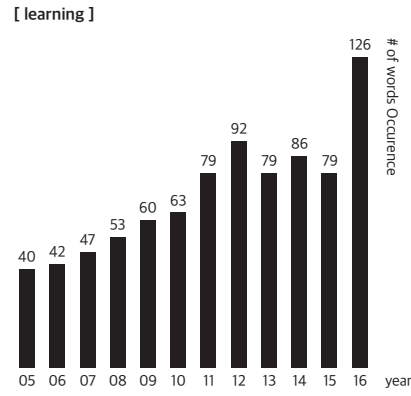
[ 2016년 ]



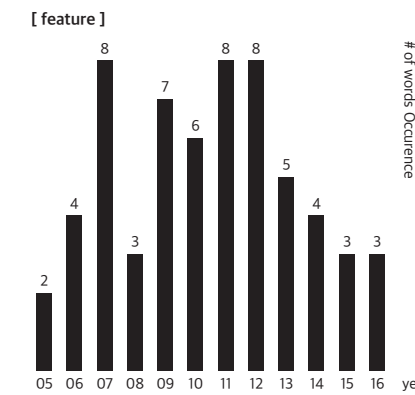
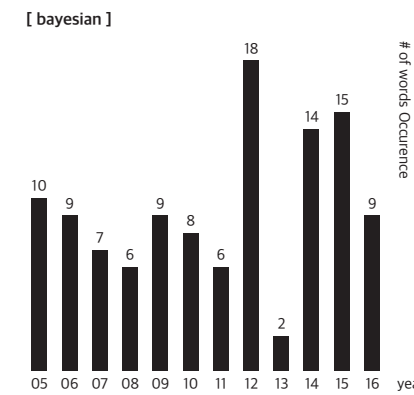
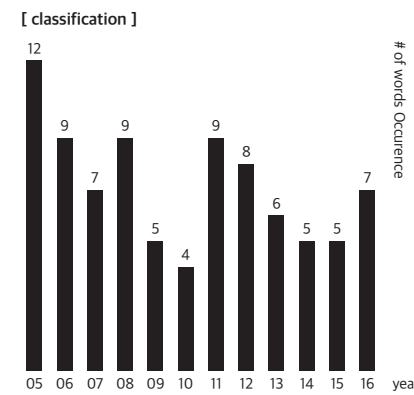
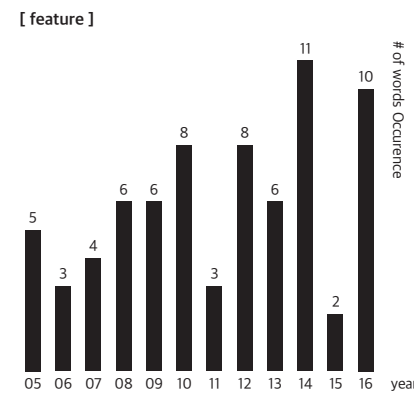
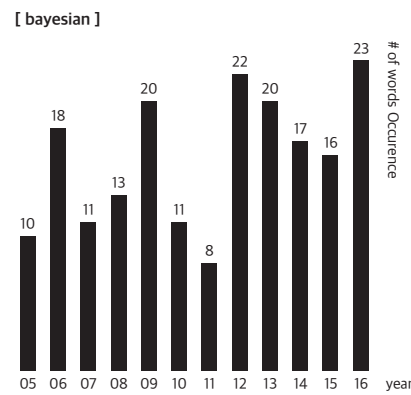
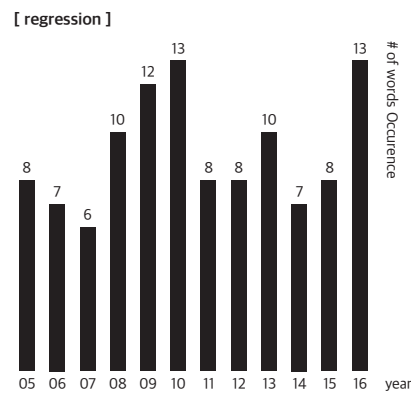
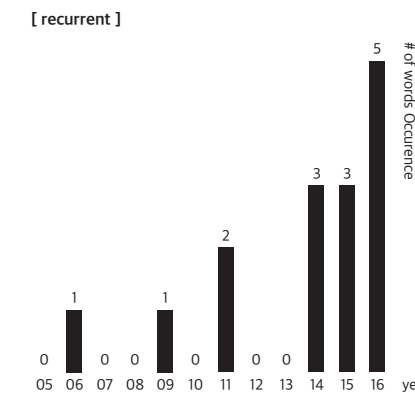
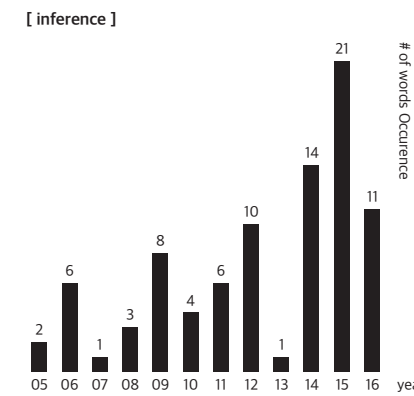
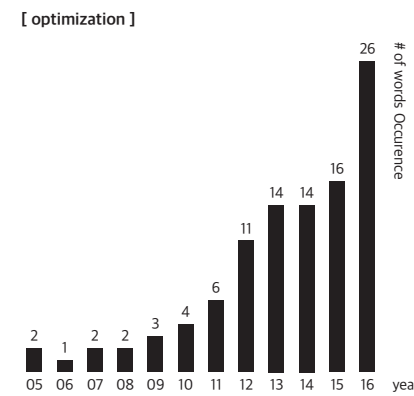
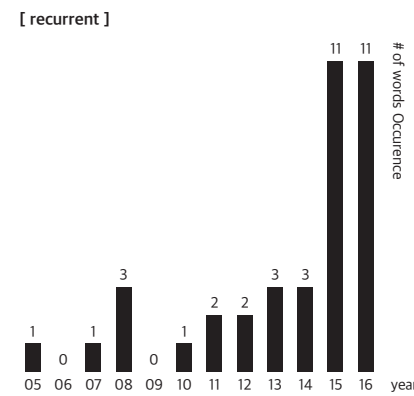
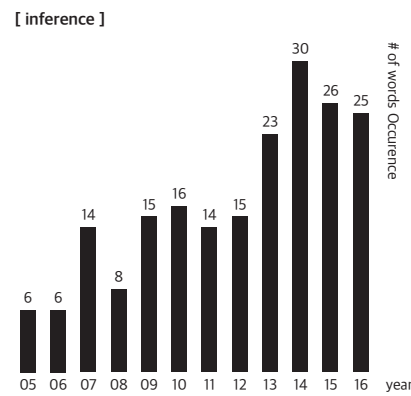
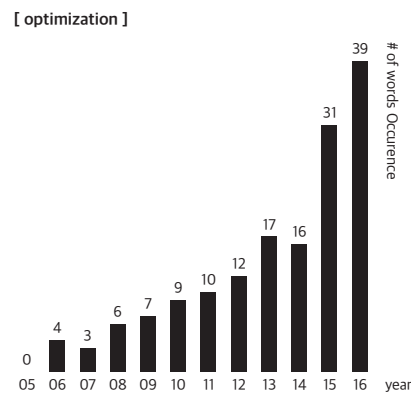
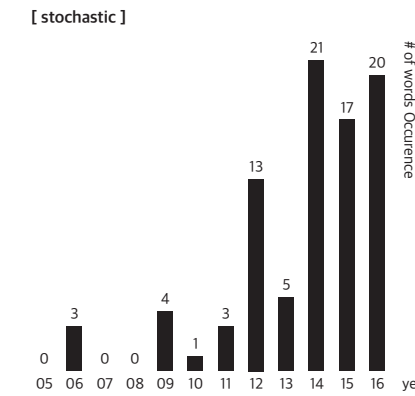
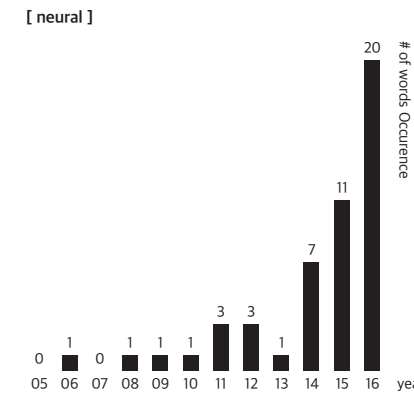
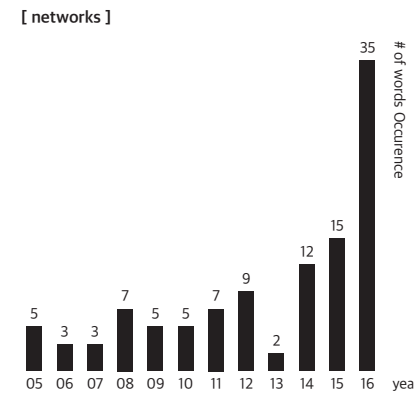
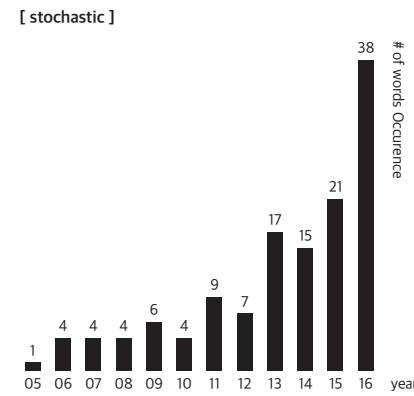
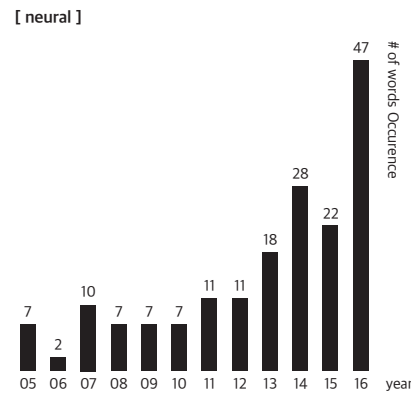
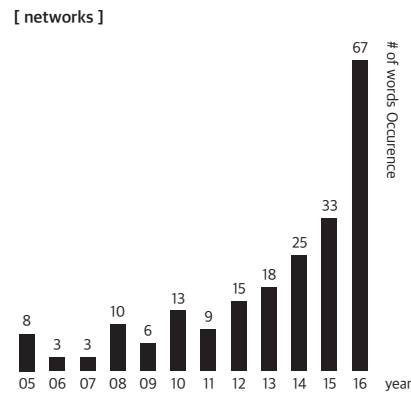
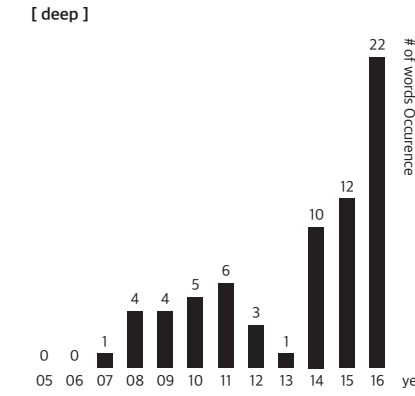
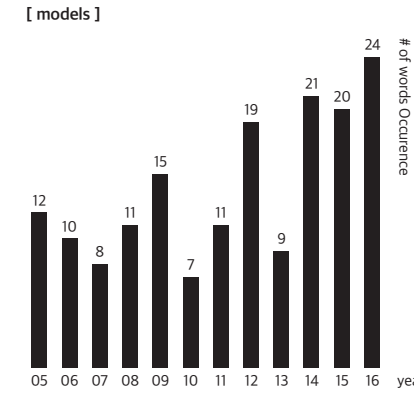
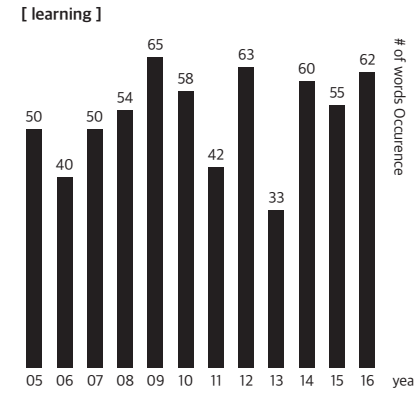


연도 별 주요 핵심어의 논문 등장 횟수

NIPS



ICML



AI 학계의 연구자 네트워크

AI 연구 네트워크의 핵심 중개자는 마이클 조던 교수

NIPS와 ICML의 논문을 기준으로 연구자 네트워크를 탐색해 봤다. 특히, 네트워크 내에서 연구자의 위상이 어느 정도인지 살펴보는데 중점을 뒀다. 기준 지표로는 네트워크 내에서 개체 간 직접적 연결 정도를 의미하는 연결 중심성과 네트워크 내에서 다른 주체들을 연결해 주는 매개자, 달리 말하면 브로커(broker)로서의 역할 능력을 뜻하는 매개 중심성을 택했다. 두 개념을 학계 네트워크에 대입하여 설명하자면, 연결 중심성은 학자들과의 직접적인 연결의 정도를 의미하며, 매개 중심성은 학자들을 연결하는, 학문의 교환자 혹은 중개자로서의 능력을 의미한다. 네트워크 구조를 구성한 연구자 수의 경우, NIPS는 5,878명이었으며, ICML은 3,949명이었다. 네트워크 분석 결과를 기준으로, 연결정도 중심성과 매개 중심성 상위 20위를 아래와 같이 정리했다. 표안의 지표는 1위의 절댓값을 분모로 하여 표준화한 수치다. 이는 상대적 비교를 용이하게 하기 위한 조치다.

NIPS의 경우, 연결 중심성과 매개 중심성 모두에서 마이클 조던(Michael I. Jordan)\*6 UC버클리대 교수가 1위였다. 조던 교수는 현재 AI 연구의 4대 천황으로 불리는 앤드류 응(Andrew Y. Ng)의 지도 교수다. 조던 교수는 ICML에서도 매개 중심성에서 1위를 기록했다. 이 결과는 마이클 조던 교수가 AI 연구 네트워크의 허브(hub) 역할을 하고 있음을 나타낸다. 두 학회의 연구 네트워크 내에서 조던 교수가 연구자들을 연결하는 핵심 주체인 셈이다. 일반적으로 중개자는 네트워크 내의 의사소통을 제어할 수 있는 통제력을 확보하고 있으며, 네트워크 내 다른 구성원들은 이 중재자에 의존한다. 이번 네트워크 분석의 결과는 최근 10년간 AI 연구의 중심에는 마이클 조던 교수가 있었음을 시사한다.

조던 교수 외에 조던 교수의 제자인 앤드류 응, 제프리 힌튼(Geoffrey E. Hinton), 그리고 요슈아 벤지오(Yoshua Bengio) 등의 AI 그루로 불리는 이들도 네트워크 내 상위 위상을 점유하고 있는 것으로 확인됐다. 한국인 중에서는 미시간대 컴퓨터공학과와 이홍락 교수\*7가 ICML의 연결중심성에서 19위에 올랐다. 이 교수는 앤드류 응 교수로부터 박사 논문 지도를 받았다.

[ NIPS의 연구자 네트워크에서 연결중심성과 매개중심성 상위 20위 ]

순위	이름	연결 중심성
1	Michael I. Jordan	1.00
2	Yoshua Bengio	0.75
2	Andrew Y. Ng	0.75
4	Zoubin Ghahramani	0.73
4	Berhard Schalkopf	0.73
6	Lawrence Carin	0.69
7	Ruslan R. Salakhutdinov	0.68
8	Eric P. Xing	0.62
8	Pradeep K. Ravikumar	0.62
10	Klaus-Robert Muller	0.59
11	Csaba Szepesvari	0.58
12	Le Song	0.56
13	Remi Munos	0.53
14	Ryan P. Adams	0.47
15	Tong Zhang	0.46
15	Quoc V. Le	0.46
17	Inderjit S. Dhillon	0.44
17	Geoffrey E. Hinton	0.44
19	Pieter Abbeel	0.43
19	Dale Schuurmans	0.43

순위	이름	매개 중심성
1	Michael I. Jordan	1.00
2	Ruslan R. Salakhutdinov	0.73
3	Zoubin Ghahramani	0.66
4	Csaba Szepesvari	0.47
5	Berhard Schalkopf	0.46
6	Le Song	0.43
7	Ryan P. Adams	0.37
8	Xi Chen	0.36
8	Eric P. Xing	0.36
10	Francis R. Bach	0.35
11	Stefanie Jegelka	0.32
12	Yoshua Bengio	0.30
13	Daniel J. Hsu	0.28
14	Tong Zhang	0.27
15	Pradeep K. Ravikumar	0.26
16	Andrew Y. Ng	0.26
17	Lawrence Carin	0.24
18	Jakob H. Macke	0.23
19	Andreas Krause	0.23
20	Klaus-Robert Muller	0.21

[ ICML의 연구자 네트워크에서 연결중심성과 매개중심성 상위 20위 ]

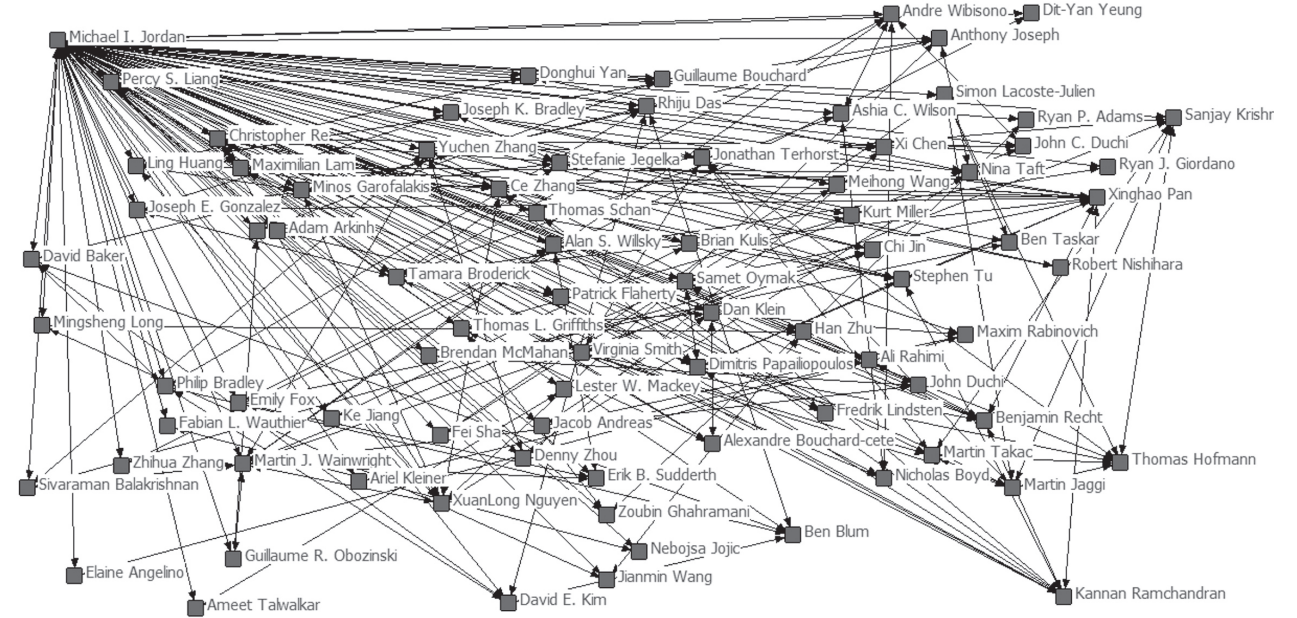
순위	이름	연결 중심성
1	Lawrence Carin	1.00
2	Yoshua Bengio	0.87
3	Andrew Y. Ng	0.72
4	Ryan Adams	0.70
5	Michael Jordan	0.68
6	Csaba Szepesvari	0.68
7	Zoubin Ghahramani	0.66
8	Andreas Krause	0.64
9	Shie Mannor	0.64
10	Berhard Schalkopf	0.64
11	Alex Smola	0.62
12	Rong Jin	0.60
13	Inderjit Dhillon	0.58
14	Shai Shalev-Shwartz	0.57
15	Le Song	0.57
16	David Silver	0.53
17	John Langford	0.51
18	Masashi Sugiyama	0.49
19	Honglak Lee	0.49
20	Michael I. Jordan	0.49

순위	이름	매개 중심성
1	Michael Jordan	1.00
2	Alex Smola	0.77
3	Zoubin Ghahramani	0.72
4	Ryan Adams	0.64
5	John Langford	0.61
6	Ohad Shamir	0.60
7	Michael I. Jordan	0.59
8	Kilian Weinberger	0.59
9	Le Song	0.58
10	Yoshua Bengio	0.57
11	Suvrit Sra	0.55
12	Berhard Schalkopf	0.54
13	Sham Kakade	0.51
14	Alekh Agarwal	0.49
15	Rong Jin	0.48
16	Hugo Larochelle	0.47
17	Shie Mannor	0.45
18	Shai Shalev-Shwartz	0.45
19	Andreas Krause	0.44
20	Ben Taskar	0.44

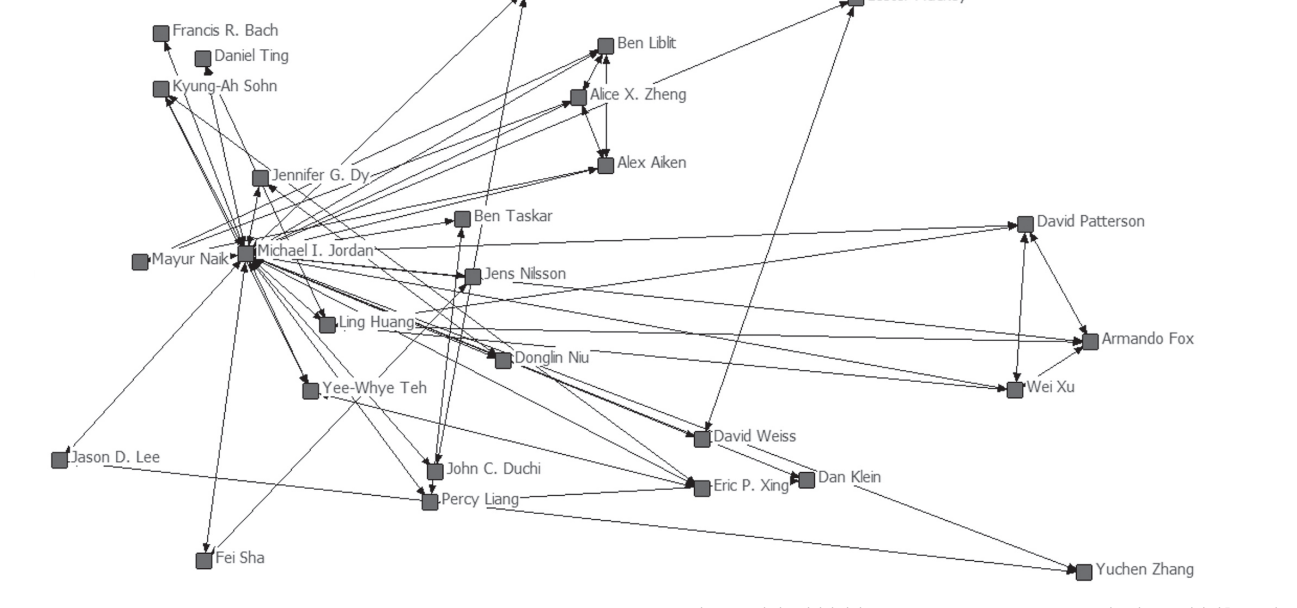
AI 연구자 간 네트워크의 중심에 있는 마이클 조던 교수를 중심으로

학자 간 관계도\*8를 학회 별로 그려봤다.

NIPS



ICML



\*1 참고 | 1987년 처음 개최된 학회는 Neural Information Processing Systems라는 이름으로 개최 이후 1988년부터 학회 명칭이 Advances in Neural Information Processing Systems로 변경되어 1회가 개최 되었다. 하지만, 학회 약칭은 여전히 NIPS( Neural Information Processing Systems)를 사용하고 있다. \*2 설명 | NIPS와 ICML 발표된 논문에 한정된 분석 \*3 설명 | Alex Krizhevsky은 논문을 1개만 발표 하였음. Andriy Mnih는 Ruslan R. Salakhutdinov과 공저로 발표한 2개의 논문이 상위 피인용횟수를 기록함. Marc'Aurelio Ranzato 공저로 포함된 2개의 논문이 모두 피인용횟수 상위를 기록하고 있음. \*4 설명 | NIPS와 ICML 발표된 논문에 한정된 분석 \*5 설명 | 본 구조도는 피인용 횟수가 높은 논문의 저자 간 관계도를 보여주기 위해 인위적으로 수립된 것이다. 네트워크 내의 각 저자 간의 거리는 임의적으로 조정된 것이며, 해당 거리가 개별적으로 의미하는 바는 없다. \*6 참고 | 조던 교수의 홈페이지 URL: <https://people.eecs.berkeley.edu/~jordan/> \*7 참고 | 이홍락 교수의 홈페이지 URL: <http://web.eecs.umich.edu/~honglak/> \*8 설명 | 마이클 조던 교수를 중심으로 한 자아 네트워크(ego network)를 그린 것이다. 자아 네트워크는 네트워크 내에서 특정 주제와 친밀한 관계를 가진 주체를 확인하는데 효과적인 수단이다. 두 개의 네트워크 그림의 원본인 UCINET 결과값에는 일부 주체가 겹쳐져 있었다. 필자는 개별 주체의 위치 및 존재를 명확하게 하기 위해, 일부 주체의 네트워크 위치를 미세조정했다. 이 조치는 그림의 정보성을 높이기 위함이었으며, 그 외 결과값은 최대한 보존하여 최초 결과를 그대로 전달하려 노력했다.



# AI 연구자들이 열광하는 ICML, NIPS

학회 논문 메타 분석에서 나타나는 연구 흐름

AI 연구자들이 집중하는 ICML(International Conference on Machine Learning)과 NIPS(Neural Information Processing Systems)에 대한 발표 논문 메타 분석을 통해 우리는 어떤 시사점을 구할 수 있을까. 실제 ICML과 NIPS에 논문을 발표한 경험이 있는 김진화 님에게 이 같은 연구 과정과 맥락에 대해 묻고 답했습니다.

글 | 김진화 jnhwkim@snu.ac.kr

서울대학교에서 인지과학전공 박사수료를 하였다. 바이오지능 연구실의 장병탁 교수님의 지도를 받으며 인공지능과 기계학습을 연구하고 있다. 구체적으로는 멀티모달과 주의 모델에 대한 관심을 바탕으로 시각적 질의 응답과 대화 모델을 연구하고 있다. 현재는 Facebook AI Research(Menlo Park, CA)에서 연구 인턴으로 근무하고 있으며 오는 5월 말 귀국 예정이다. 가끔씩 Torch7.rn contributor로 활동한다.

**Q. 안녕하세요. 일단 소개부터 부탁드립니다.**

**A.** 안녕하세요? 서울대학교 바이오지능 연구실 박사과정의 김진화입니다. 현재는 박사 수료 후 캘리포니아에 있는 페이스북 AI Research에서 Yuandong Tian의 지도로 4개월 간 연구 인턴을 수행하고 있습니다. 이제 연구 인턴 과정을 마무리하는 단계에 있으며, 박사 졸업 준비를 위해 5월 말 귀국할 예정입니다.

사람의 인지 과정과 인공지능에 대한 관심에서 대학원 과정을 시작했습니다. 시각 정보와 언어 정보를 종합적으로 판단하고 추론하는 멀티모달 학습과, 시각 주의를 연구 주제로 공부했습니다. 자연스럽게 시각 질의 응답 모델을 연구했습니다. 작년 CVPR 2016 학회에서 열린 Visual Question Answering Challenge에서 4위를 차지했고, 해당 연구는 NIPS 2016에서 발표했습니다. 후속 연구로 최고 점수를 갱신하여 새로운 논문으로 올해 ICLR(International Conference on Learning Representations) 2017에서 발표했습니다.

최근에는 시각 질의 응답에서 목표 중심 시각 대화로 연구를 확장하여 새로운 데이터 수집과 함께 다음 논문을 준비하고 있습니다.

**Q. NIPS 혹은 ICML에서 발표하신 연구는 어떤 내용인가요.**

**A.** 2016년도 NIPS에서 1저자로 발표한 논문의 제목은 "Multimodal Residual Networks for Visual QA" 입니다. 서울대 장병탁 교수의 바이오지능 연구실에서 함께 연구한 이상우, 광동현, 허민오 박사과정과 당시 네이버 랩스의 김정희, 하정우 님께서 공저자로 참여했습니다.

Kaiming He 등(2015)의 연구인 'Deep Residual Networks'에 영감을 받아 단일 이미지 입력으로 학습하는 unimodal 학습에서 이미지와 자연어를 동시에 고려하여 학습하는 multimodal 학습으로 확장하여 시각 질의 응답 모델을 제안했습니다. 당시 시각 질의 응답 모델로 다양한 비전 처리 기술과 자연어 처리 기술이 소개됐는데, 이 연구는 잘 알려진 전처리 기술을 이용하여 이미지와 자연어에 대한 표상을 구한 후 ResNet(He et

al., 2015)과 같이 지름 연결선들을 이용하여 상대적으로 깊은 층의 단순한 모델 구조를 탐색했다는 의의를 갖습니다. 정성 평가를 위해 새로운 시각화 기법을 소개하여 각 층에서 자연어로 주어진 질의문에 의해 시각 주의가 어떻게 변하는지 관찰하기도 하였습니다.

**Q. NIPS와 ICML라는 학회에 대한 개괄적인 소개 부탁드립니다.**

**A.** 두 학회 모두 30년 정도 진행된 성숙기에 진입한 학회입니다. 이 전에는 AAAI와 IJCAI라는 인공지능 학회에서 주목할 만한 연구들이 많이 나왔었는데, 기계학습이라는 세부 분야가 해당 산업의 성장과 더불어 커지면서 NIPS와 ICML로 많은 연구자들과 대중의 관심이 옮겨간 듯합니다.

특히 NIPS는 그 학회 이름에서 추측할 수 있듯이 계산신경과학 등과 같은 분야도 포함하는 인간 지능과 인공지능에 대한 포괄적 관심을 가진 연구자들이 많았습니다. 단지, 최근에는 기계학습에 대한 연구의 비율이 높아져 ICML과 분위기가 많이 비슷해진 것 같지만 일정한 비율로 다양한 연구들이 발표되고 있습니다.

올해 5회를 맞는 ICLR은 기계학습 중에서 특히 모델을 이용하여 표상 학습을 강조하는 deep learning에 초점을 맞춘 Top-tier 학회로 급성장하고 있습니다. 구글, 페이스북 등의 후원을 바탕으로 올해 학회 규모가 두 배 이상 커졌으며 등록자가 천 명을 넘어섰습니다.

**Q. 카카오는 2005년부터 2016년 사이에 NIPS와 ICML에서 발표된 논문을 메타 분석했습니다. 논문 제목을 중심으로 연구 트렌드와 주요 저자에 대한 분석을 진행했습니다.**

**A.** 양적 분석을 바탕으로 핵심 키워드를 살펴보고 중요 논문들을 목록화했습니다. 눈에 띄는 것은, 피인용횟수가 가장 많은 논문들 중에서 deep neural network를 학습하는 방법과, 이러한 네트워크를 잘 학습할 수 있게 하는 보완 기술이나 분석 논문들을 관찰할 수 있었습니다.

하지만 연구의 범위와 주제가 다양하고, 연구자들의 네트워크 또한 회사 조직처럼 강한 연결이 아니므로 무겁게 다룰 내용은 아닌 것 같습니다.

**Q. NIPS와 ICML에서 학회마다 발표되는 논문의 수가 가파르게 증가하는 경향을 보이고 있습니다. 이 추세 배경은 무엇인가요? NIPS와 ICML 모두 2012년이 논문 수가 증가하는 변곡점으로 나타납니다.**

**A.** 여러 가지 복합적인 요인에 의해서가 아닐까 합니다. 미국 대학들에서 컴퓨터 관련 전공 학생 수를 늘리고 대학원 과정의 규모를 키워왔습니다. 학회에서 발표되는 논문 상당수는 연구 조직을 갖춘 글로벌 기업들에서도 비롯됩니다. 연구 조직의 공격적인 채용과 재원 투자가 그 배경이 됩니다.

이러한 현상에는 설명 가능한 많은 요인이 있겠지만 개인적으로 주목하는 것은 GPU를 이용한 기술의 발전이 아닐까 합니다. GPU 컴퓨팅의 선두 주자인 NVIDIA의 주가 폭등을 살펴보면 그 맥락을 같이 하며 지난 2016년 NIPS가 열린 12월에 주가가 급성장한 것은 우연의 일치만은 아닐 것 같습니다.

**Q. NIPS 논문의 제목 단위로 분석하면, 2005년부터 2011년까지 learning 이 가장 높은 빈도를 보였습니다. 그러나, 2012년 deep, networks, neural, classification, imagenet, convolutional 이 learning 보다 높은 빈도를 기록했습니다. 논문 제목에만 한정된 분석이지만, “deep”의 빈출 정도는 2012년 이전과 이후가 확연한 차이를 보입니다. 이 결과가 의미하는 바는 무엇인가요?**

**A.** 기존의 인공 신경망과 달리 더 많은 층을 쌓아서 학습할 수 있는 다양한 방법이 제안되고 있습니다. 기본적인 단위 기술들의 반복적인 적용으로 성능을 비약적으로 향상할 수 있다는 생각은 GPU computing 기술의 발달과 더불어 정부와 기업들에게 많은 연구 투자를 받게 된 계기가 되었습니다. 또한 기본적인 단위 기술들이 중요해지면서

개발 방법론을 응용한 deep learning framework들의 등장도 촉진하게 된 배경이 됐습니다.

**Q. 전반적으로 NIPS 연구의 중심은 learning에 있습니다. AI 연구에서 learning에 대한 주목도가 높은 이유는 무엇인가요?**

**A.** 전통적인 인공지능은 전문 지식을 가진 전문가 그룹이 규칙들을 개발하여 자동 계산하는 방식을 선호했습니다. 하지만 인공지능이 풀고자 하는 문제가 점점 복잡해지면서 대용량의 데이터를 수집하여 자동적으로 규칙을 학습하게 하는 learning에 주목하게 된 것 같습니다.

**Q. NIPS와 ICML에서 발표된 논문 중 인용 상위 논문들 중심으로, AI 연구의 흐름을 살펴본다면 어떨까요.**

**A.** Krizhevsky et al. (NIPS 2012)의 연구는 이미지 분류 문제를 풀기 위해 전통적인 전문 지식을 활용한 특징 추출 방법을 떠나 deep convolutional neural networks를 제안하여 자동적으로 그러한 특징을 추출할 수 있는 방법을 제안했습니다. Mikolov et al. (NIPS 2013)의 연구는 자연어 언어 처리에서 단어를 나타내는 특징 벡터를 자동으로 추출할 수 있는 방법을 neural networks의 틀 안에서 소개하였습니다. 두 연구 모두 learning에 대한 주목도를 보여줍니다. Nair & Hinton (ICML 2010)에서 소개된 ReLU는 일반적인 deep neural networks에서 gradient diminishing 문제를 일부 해소하는 단순한 방법으로 현재까지 많이 이용되고 있는 활성화 함수(activation function)입니다.

**Q. 특정 논문 그리고 특정 단어 중심의 연구가 이뤄지는 경향이 확인됩니다. AI 연구 역시도 일종의 유행이 있는 것으로 판단됩니다. 학계에서의 이러한 유행은 어떤 결과를 가져올까요.**

**A.** 학계에서 유행이라는 표현은 조심스럽게 다뤄야 할 필요가 있지만 일반적인 관점에서 보자면, 해당 분야에 대한 활발한 논의와 연구를 토대로 빠르게 탐색할 수 있다는 장점이 있습니다. 그러나 자칫 속도 경쟁에 치우치게 되면 incremental 연구 결과들이 나오고 혁신성이 떨어질 수 있어 주의해야 합니다. 공개 코드는 이러한 현상을 가속화할 수 있습니다. 공개 코드는 관심 있는 연구자들이 빠르게 SOTA(state-of-the-art, 현재 최고 수준) 결과를 낼 수 있도록 하여 해당 연구를 활성화시킨다는 장점이 있지만, 상대적으로 모델의 일부나 학습 변수를 수정하여도 손쉽게 성능을 올릴 수 있는 가능성이 있습니다. 이러한 유행은 정부나 기업의 연구 지원 규모를 늘리는 계기가 되기도 하지만, 연구 주제에 편향을 가져와 연구의 다양성을 해칠 수 있다는 단점 또한 있으므로 주의할 필요가 있습니다.

**Q. NIPS와 ICML에서 발표되는 연구 논문이 다소 차이는 보이는 까닭은 무엇이며, 두 학회에서 발표된 연구 결과는 각각 어떤 분야 그리고 어떤 산업에 중요한 영향력을 행사하며, 그 양상은 어떠한 지 구체적으로 설명 부탁드립니다.**

**A.** NIPS는 인간 지능과 인공 지능에 대한 포괄적 주제를 다루는 학회였지만, 최근에는 기계학습에 대한 비중이 높아져 ICML과 비슷한 연구들이 발표되고 있습니다. 적어도 산업에서는 NIPS 학회에 대한 관심이 더 큰 것 같이 느껴집니다. 아마도 인간 지능과 인공 지능에 대한 포괄적 탐색이 대중이나 산업에 더 선호도가 있던 것이 아니었나 생각합니다.

**Q. NIPS와 ICML의 지난 10년간 나온 연구 논문 제목의 주요 단어를 종합하면, AI 연구를 “신경망 모델을 통해 학습을 최적화하기 위한 체계를 찾는 연구”로 정리될 수 있을 것 같은데, 어떻게 보시는지요.**

**A.** 정리하신 자료에서 살펴보면 신경망 모델이 아닌 기계학습 방법론도 관찰됩니다. NIPS 2016 제출 논문의 연구 주제를 분석해보면 약 25%가 deep learning 또는 신경망 모델에 대한 연구임을 알 수 있습니다. 따라서 언론이나 대중매체를 통해 접하는 기계학습과 실제 연구자들의 생각은 조금 차이가 있을 수 있다는 생각이 듭니다.<sup>1)</sup>

**Q. 앤드류 응, 제프리 힌튼, 안 르쿤, 요슈아 벤지오, 데미스 하시비스와 같은 대중에게도 널리 알려진 연구자 외에 최근 AI학계에서 주목받는 학자는 누구이며, 어떤 연구로 인해 주목을 받고 있는지 알려주세요.**

**A.** 자료에서 언급된 이홍락 교수님을 제외하면 뉴욕대의 조경현 교수님이 계시지 않을까요? 개인적으로는 순환신경망의 일종인 GRU의 고안자로, 그리고 자연어 처리, 기계 번역 연구의 선구자로 알려진 분입니다.

\*1 참고 | <http://www.tml.cs.uni-tuebingen.de/team/luxburg/misc/nips2016/index.php>



# 딥러닝 연구의 현재와 미래

## part 1. (1/2)

가장 많이 인용된 딥러닝 논문 100개로 살펴본다.

바야흐로 딥러닝의 시대, 그리고 인공지능(AI)의 시대이다. 구글, 페이스북, 마이크로소프트, 바이두와 같은 세계 최고의 IT 기업들이 인공지능 기술을 “새 시대의 전기”<sup>1</sup>에 비유하며 핵심기술 확보에 위해 총력을 기울이고 있으며, 이를 위해 각 기업들은 인재 영입 전쟁과 공격적인 R&D 투자를 마다하지 않고 있다. 세계에서 이름을 날리던 머신러닝 대가들은 대부분 이들 기업에 영입된 지 오래이고, 떠오르는 샛별들 역시 마치 ‘FC바르셀로나가 어린 메시를 다루듯’ 기업에 의해 키워지고, 또 영입되고 있으니 말이다.



글 | 엄태웅 terry.tum@gmail.com

서울대 기계항공공학부에서 로봇의 모션플래닝을 전공한 뒤 회사에선 영동하게 엑소스켈레톤을 개발하다가 서른이 되어 돌연 머신러닝계로 이적한 연구계의 박쥐, 낭인, 저니맨, 능력자 친구의 친구. 현재는 캐나다 워털루공대에서 딥러닝을 이용해 모션분석 연구(a.k.a. 틈새시장 공략)를 하고 있으며, 라임이 분유값을 벌 수 있던 피임에 솔깃해 카카오프트에 글을 쓰게 되었다고 한다. 폐북에 서식하며 관심 먹고 살아가는 미물이지만, 막상 컨택하면 비싼척 하는 미청년, 아니 미야재.

딥러닝 관련 논문의 수 역시 폭발적으로 증가하고 있다. 딥러닝계의 라이징 스타 중 한 명인 OpenAI의 Andrej Karpathy의 간단한 조사<sup>2</sup>에 따르면 공개 논문 저장소 arXiv<sup>3</sup>를 통해 매월 공개되는 머신러닝 논문의 수가 5년 사이 100배 이상 늘었다고 한다. 다시 말하면 10년 동안 나올 머신러닝 논문들이 지금은 한 달 안에 쏟아지고 있는 셈이다. 일례로 Wassertein GAN<sup>4</sup>이란 페이스북 시리서치의 논문은 올해 1월 26일에 arXiv에 공개되었는데, 곧 3월 31일 공개된 Improved Wassertein GAN<sup>5</sup>이란 알고리즘으로 두 달 만에 구식 알고리즘이 되었다. 이는 논문이 저널에 실리는 데만 몇 달이 걸리던 기존의 프로세스를 생각한다면 양과 속도 면에서 모두 놀라운 발전이라 할 수 있을 것 같다.

발전 속도가 빠른 딥러닝 학계에선 이미 공개된 지 1년이 지나면 오래된 논문에 속하고, 2-3년이 지나면 “고전”이라 불리고 있긴 하지만, 지난 5년간의 연구 추세를 살펴보는 일은 딥러닝 연구에 대한 복습의 차원에서고, 그리고 현재를 진단하고 미래를 가능해보는 차원에서도 의미 있는 일이 될 것이다. 필자가 깃허브(Github)를 통해 공개한 “가장 많이 인용된 딥러닝 논문 리스트 Top 100”<sup>6</sup>를 기초로 하여 과거 5년 간(2012-2016)의 딥러닝 연구를 돌아보고, 이와 함께 딥러닝 연구가 앞으로 나아가갈 방향에 대해 조심스럽게 예측해보도록 하겠다.

[ Contents ]<sup>7</sup>

**Part 1.**

1. Convolutional Network Models
2. Image Segmentation / Object Detection
3. Image / Video / Etc
4. Natural Language Processing / RNNs
5. Speech / Other Domains

**Part 2.**

6. Understanding / Generalization / Transfer
7. Unsupervised / Generative Models
8. Optimization / Training Techniques
9. Reinforcement Learning / Robotics
10. More Papers from 2016, 2017

위의 카테고리 분류를 통해서도 알 수 있듯이, 딥러닝 알고리즘은 이미지는 물론, 자연어 처리, 음성인식, 로봇 등 매우 다양한 분야에서 활발히 연구 중이다. 딥러닝 연구의 폭발적인 성장에는 이러한 ‘연구분야의 통합’도 큰 기여를 하고 있는데, 예전에는 각각의 도메인에서 따로 연구를 했을 학자들이 ‘딥러닝’이라는 한 가지 주제에 대해 파고들다 보니, 늘어난 연구자의 수와 다양성만큼, 연구의 양과 속도 역시 유례없는 발전을 이루고 있다. 그것의 산업적 효용성은 AI분야 연구의 파이를 키우는 촉매제가 되고 있다.

비록 이 글에서 그 많은 양의 연구를 모두 따라잡을 수는 없겠지만, 필자의 논문 리스트 100개 중 최대한 많은 논문들의 기본 아이디어를 커버하며 딥러닝 알고리즘의 발전과 응용을 한눈에 바라보고자 한다. 따라서 이 정리는 필연적으로 깊이 있는 토론(in-

depth study)을 배제하고 있으며, 독자께서 관심이 있는 주제가 있다면 그 논문으로 독자를 이끌 수 있도록 하는 가이드가 되었으면 한다. 리포트는 이번 달과 다음 달, 2부로 나누어 발간하며, 위의 콘텐츠의 순으로 논문들을 검토 하도록 하겠다.

### Convolutional Neural Networks Models

딥러닝을 이끄는 양대 알고리즘이라고 한다면 이미지 인식에 주로 사용되는 CNN(convolutional neural networks)과 자연어 처리, 음성인식 등에 주로 사용되는 RNN(recurrent neural networks)을 들 수 있을 것이다. 그중 CNN은 데이터로부터 자동으로 피쳐(feature)를 학습하는 대표적 알고리즘이라고 할 수 있다. 머신러닝을 통해 데이터를 학습하기 위해선 먼저 날 것의 데이터(e.g. 픽셀 단위의 데이터)를 조금 더 추상적 레벨이 높은 피쳐(e.g. 선, 면, 모서리)로 가공하는 과정이 필요한데, 딥러닝, 특히 CNN은 이러한 피쳐를 데이터로부터 매우 효율적으로 학습한다.<sup>8</sup> 현재 사용되고 있는 CNN은 기본적으로 르쿤이 1989년에 개발한 구조<sup>9</sup>를 토대로 하고 있는데, 2012년 ILSVRC이미지인식 대회<sup>10</sup>에서 힌튼 교수 팀의 AlexNet[1-1]이 놀라운 성능 개량을 보임으로써 현재까지 CNN의 폭발적인 연구 성장이 이어져 왔다. 깊은 학습 구조(deep learning)가 복잡한 문제를 다루는 해결의 열쇠라는 점이 밝혀지면서 이후 다양한 형태의 깊은 학습구조에 대해 연구가 이루어져 왔는데, VGGNet[1-2], GoogLeNet[1-3], ResNet[1-4] 등이 2011년 26% 수준의 인식오차율을 3.6%까지 낮춘 개량된 CNN의 주인공들이었다.

[ 소개된 주요 논문들 ]

[1-1] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[1-2] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[1-3] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1-9).

[1-4] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 770-778).

### Image: Segmentation / Object Detection

3.6%라는 인식오차율(error rate)은 데이터 레이블 자체의 결함이나 사람의 인식오차율을 고려해 보았을 때 더 이상의 개선이 무의미할 정도의 높은 성능이라 할 수 있다. 따라서 “잘

정의된”(구분하고자 하는 물체에 주목할 수 있게 여백 없이 잘 잘라져 있고, 배경의 방해가 별로 없는) 이미지들에 대한 단순 분류는 이미 딥러닝에 의해 정복되었다고 할 수 있으므로, 연구자들은 나아가 더 어려운 문제, 예를 들어 다양한 배경이 있는 이미지 안의 사물인식(object recognition)이나 픽셀 단위 이미지 영역구분(image segmentation)에 도전을 하고 있다.

사물 인식이나 영역 구분은 이미지 안의 물체 분류뿐만 아니라 그것의 위치까지도 특정해야 한다는 도전적 과제가 있다. 사물 인식은 물체의 위치를 나타내는 바운딩 박스, 영역 구분은 각 픽셀 별 분류를 목적으로 한다. 분류하고자 하는 물체 이외 다른 배경 이미지들의 존재는 타겟 작업을 더욱 어렵게 하며, 때론 더욱 많은 데이터, 더욱 고도화된 CNN 구조를 요구하기도 한다. 따라서 기본적인 CNN을 변형한 다양한 뉴럴넷(neural networks) 구조들이 제안됐는데, 각각의 대표적인 방법론으로는 R-CNN[2-1]과 FCN(fully convolutional networks)[2-4]을 들 수 있다.

R-CNN은 이미지 내에서 물체가 있을 법한 영역 후보들을 먼저 제안하고 이들의 스코어를 매겨 물체를 인식하는 방법인데, Fast R-CNN[2-2], Faster R-CNN[2-3]의 후속 연구는 계산 성능과 정확도를 획기적으로 개선했다. 최근에는 페이스북의 Mask R-CNN[2-4]이 사물인식과 영역구분을 동시에 하는 놀라운 성능을 보여주었다. 영역 구분은 이미지를 CNN 레이어들을 통해 작은 사이즈로 추상화 한 뒤 그들을 다시 풀어헤치며(unroll) 픽셀에 대한 단순화 추론을 가능케하는 FCN[2-5]을 통해 이뤄지는데, 이후 FCN은 영역 구분 뿐만 아니라 CNN의 완전연결층(fully-connected layer)를 생략하게 하는 주요한 CNN 구조의 발전으로 남아있다. 구분(segmentation)은 자율주행차량의 주행 상황 인식에도 매우 중요한 역할을 하는데, 이러한 역할로는 단순하고 빠른 것으로 알려진 YOLO[2-6]가 기본 알고리즘으로 많이 쓰이고 있다.



**Image / Video / Etc.**

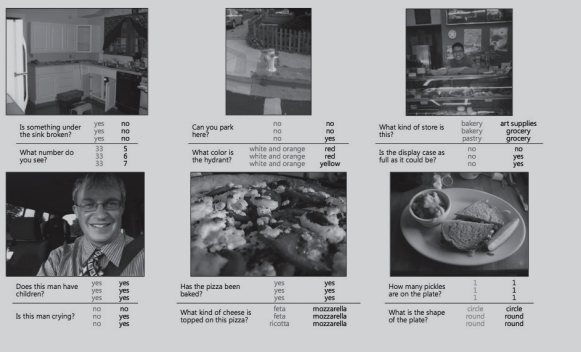
이 외에도 CNN의 응용분야는 너무나 많다. 저해상도의 이미지를 고해상도의 이미지로 복원하는 super resolution 문제에서는 CNN을 통해 더 실제와 가깝게 복원하기 좋은 피쳐를 자동으로 학습함으로써 그 성능을 크게 개선했다[3-1]. 이미 일반인들도 모바일 앱을 통해 많이 알고 있는 “화가풍으로 이미지를 바꿔주는 알고리즘”[3-2]은 CNN을 통해 학습한 피쳐들을 블렌딩 함으로써 새로운 형태의 이미지로 바꿀 수 있다는 딥러닝의 예술적 쓰임 가능성을 보여주었으며, 이제 CNN은 정적인 이미지를 넘어 비디오 속에서 내용 분류를 하고[3-3], 사람의 움직임을 인식하는 것[3-4]에도 도전하고 있는 중이다. 비록 비디오는 이미지와 달리 특정 동작들의 시작과 끝을 정확히 구분하기 어렵고, 이들을 수작업으로 레이블링(labeling)하는 작업마저 쉽지 않아 아직 뚜렷한 승자가 나타나지 않은 상태이지만, 최근 RNN과 같은 시계열(time-series) 기반 알고리즘과의 결합이 시도되고 있는 만큼 머지않아 이미지 못지않은 성능을 볼 수 있지 않을까 기대해본다.

특히 자연어 처리에 많이 쓰이는 RNN과의 결합은 이미지 자동 자막 생성[3-5]이나 사진 속 내용에 대한 문답[3-6]과 같은 재미있는 응용 가능성을 보여주는데, 기존의 이미지 분류(classification) 문제가 사진 속 사물에 대해 단순한 “단어”를 뱉어내는 수준이었다면, 이들 문제는 나아가 “문장”을 생성해 냄으로써 인식(perception)과 이해(understanding)에 대한 인공지능의 수준을 한 단계 높이고 있다. 비록 아직까지는 인간수준에는 한참 미치지 못하지만, 비디오 인식 분야에서 CNN이 많은 발전을 이루고, 문장 생성과 관련하여 RNN이 큰 진보를 거듭해 이들 성과물이 완결성 있게 결합할 수 있다면, 지금과는 차원이 다른 “인식”과 “이해”의 수준을 갖춘 인공지능의 출현도 미래에 기대해볼 수 있을 것이다.

**[ 소개된 주요 논문들 ]**

[3-1] Dong, C., Loy, C. C., He, K., & Tang, X. (2016). Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence, 38(2), 295-307.  
 [3-2] Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576.  
 [3-3] Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (pp. 1725-1732).  
 [3-4] Toshev, A., & Szegedy, C. (2014). Deeppose: Human pose estimation via deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1653-1660).  
 [3-5] Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2015). Show and tell: A neural image caption generator. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3156-3164).  
 [3-6] Antol, S., Agrawal, A., Lu, J., Mitchell, M., Batra, D., Lawrence Zitnick, C., & Parikh, D. (2015). Vqa: Visual question answering. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2425-2433).

**[ 그림 2 ] 주어진 질문에 대해 주어진 그림을 볼 때(그레이)와 보지 않을 때(블랙) 인공지능의 대답들[3-6]**



**Natural Language Processing / RNNs**

CNN과 더불어 딥러닝의 또 다른 핵심 축을 이루는 알고리즘을 꼽으라면 RNN을 꼽을 수 있다. 기본적인 뉴럴넷을 각 시간 순으로 연결하여 매우 깊은 구조를 만든 RNN은 보통 LSTM(long short-term memory)과 같은 게이트 유닛(gate unit)을 임베딩하여 학습한다. 이 게이트 유닛은 마치 “과거의 일을 기억하는 메모리”와 같이 취급할 수 있다. 이러한 기억 기능을 통해 ‘순차적으로 입력된 단어의 기억’인 문장을 이해하여 이에 대해 답변을 할 수도 있고[4-1], 입력된 문장을 다른 언어로 번역할 수도 있는 능력을 보여주는데[4-2], 특히 번역은 뉴럴넷 기반 구글 번역기[4-3] 등에서 그 압도적인 성능을 선보인 바 있다.

자연어를 이와 같이 뉴럴넷을 통해 처리할 수 있게 된 배경에는 단어를 벡터화하여 수학적 공간에 매핑시킨 Word2Vec[4-4], GloVe[4-5]와 같은 워드 임베딩(word embedding) 기술의 역할이 컸다. 특히 이러한 워드 임베딩은 ‘왕-남자=왕비’와 같은 관계를 만들어줄 정도로 단어 간 거리와 실제

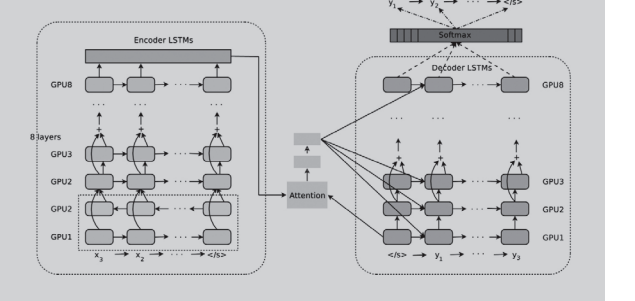
의미적 차이의 관계를 매우 실제와 가깝게 모델링했는데, 이러한 훌륭한 메트릭(metric, 거리를 나타내는 방법)의 제공은 자연어를 숫자를 이용해 다룰 수 있게 함으로써 그 발전을 더욱 가속화시키는 촉진제가 됐다.

보통 RNN은 시계열 또는 순서(sequence)가 있는 데이터에, CNN은 정적인(static) 데이터에 많이 사용되는데, 서로 반대의 영역에 RNN과 CNN을 적용해보려는 시도도 이뤄지고 있다. 그 대표적인 예로는 픽셀을 순서 데이터로 취급하여 이미지를 생성하는 딥마인드의 Pixel RNN[4-6]과 문장의 분류를 CNN을 이용해서 하는 뉴욕대의 연구[4-7]를 들 수 있다. 과거 다른 영역이라 여겨졌던 CNN 중심의 이미지 연구영역과 RNN 중심의 자연어 연구영역이 점차 융합되면서 이들의 장점을 두루 합친 새로운 형태의 뉴럴넷 탄생도 기대해 볼 수 있을 것 같다.

**[ 소개된 주요 논문들 ]**

[4-1] Weston, J., Chopra, S., & Bordes, A. (2014). Memory networks. arXiv preprint arXiv:1410.3916.  
 [4-2] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.  
 [4-3] Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Klingner, J. (2016). Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. arXiv preprint arXiv:1609.08144.  
 [4-4] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.  
 [4-5] Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global Vectors for Word Representation. In EMNLP (Vol. 14, pp. 1532-1543).  
 [4-6] Oord, A. V. D., Kalchbrenner, N., & Kavukcuoglu, K. (2016). Pixel recurrent neural networks. arXiv preprint arXiv:1601.06759.  
 [4-7] Kim, Y. (2014). Convolutional neural networks for sentence classification. arXiv preprint arXiv:1408.5882.

**[ 그림 3 ] 현재까지 최고의 성능을 보이고 있는 구글의 번역기 구조. 8층으로 쌓인 LSTM 레이어들을 Attention모델과 결합하여 좋은 성능을 보여줬는데, 이 과정을 8개의 GPU로 분산처리함으로써 대량 요청의 동시 처리도 가능하도록 했다.[4-3]**



**Speech**

음성인식 역시 딥러닝을 통해 크게 발전된 분야라고 할 수 있다. 기존의 고전적 음성인식은 GMM(gaussian mixture model)을 이용해 각각의 음소(음성 상의 최소 단위)를 모델링하고, 이들의



연속적 다이내믹스를 HMM(hidden markov model)으로 포착하는 형태가 기본이었는데, 모델의 표현력(express power)에 있어 한계가 드러나 사람에 따라 변화무쌍한 인간의 음성을 이해하기에는 부족하다는 의견들이 많았다.

하지만 딥러닝은 거대한 모델과 많은 양의 데이터를 통해 이러한 표현력 부족 문제를 해결했다. 기존 GMM을 deep belief network와 같은 비지도학습(unsupervised) 모델로 대체해 성능을 개선시키기도 했고[5-1], HMM을 이용해 연속적 음성의 변화를 모델링하던 것을 표현력이 더욱 풍부한 RNN으로 대체[5-2]시킴으로써 End-to-End 학습[5-3]을 달성하기도 했다. 이렇게 뉴럴넷 기반의 음성인식이 가능해진 배경에는 2006년 고안된 CTC(connectionist temporal classification)<sup>12</sup> 방법이 자연어처리에 있어서의 “워드 임베딩”과 같이 기반 기술의 역할을 해준 기여가 컸다.

최근에는 음성 인식을 넘어 음성 합성에도 깊이 있는 연구가 이루어지고 있다. 구글 답마인드는 Pixel RNN과 유사한 방식을 이용해 이미지 생성이 아닌 음성 합성을 구현했는데, WaveNet[5-4]이라 불리는 이 알고리즘은 딥러닝을 기반으로 매우 고품질의 음성을 생성해낸다. (비록 샘플을 순차적으로 생성해내는 한계 때문에 1초의 음성 생성에 몇 분의 시간이 소요되긴 하지만 말이다.) 최근 구글은 TTS(text-to-speech) 분야에서도 TACOTRON이란 알고리즘[5-5]을 통해 놀라운 결과를 보여주었는데, 심포의 위치에 따라 문장을 읽는 높낮이와 속도가 달라지고, 같은 철자라도 문맥에 따라 발음을 달리하는 등 (e.g. read의 현재와 과거형 발음 구분) 기존 컴퓨터의 어색한 TTS와는 전혀 다른, 사람과 같이 자연스러운 TTS를 구현해 놀라움을 선사한 바 있다. WaveNet과 Tacotron의 음성 생성 결과는 각각 구글 답마인드 홈페이지(<https://deepmind.com/blog/wavenet-generative-model-raw-audio/>)과 구글의 깃허브 페이지(<https://google.github.io/tacotron/>)에서 확인할 수 있다.

**[ 소개된 주요 논문들 ]**

[5-1] Mohamed, A. R., Dahl, G. E., & Hinton, G. (2012). Acoustic modeling using deep belief networks. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(1), 14-22.

[5-2] Graves, A., Mohamed, A. R., & Hinton, G. (2013, May). Speech recognition with deep recurrent neural networks. In *Acoustics, speech and signal processing (icassp)*, 2013 *IEEE international conference on* (pp. 6645-6649). *IEEE*.

[5-3] Bahdanau, D., Chorowski, J., Serdyuk, D., Brakel, P., & Bengio, Y. (2016, March). End-to-end attention-based large vocabulary speech recognition. In *Acoustics, Speech and Signal Processing (ICASSP)*, 2016 *IEEE International Conference on* (pp. 4945-4949). *IEEE*.

[5-4] van den Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., ... & Kavukcuoglu, K. (2016). Wavenet: A generative model for raw audio. *CoRR abs/1609.03499*.

[5-5] Wang, Y., Skerry-Ryan, R. J., Stanton, D., Wu, Y., Weiss, R. J., Jaitly, N., ... & Le, Q. (2017). Tacotron: A Fully End-to-End Text-To-Speech Synthesis Model. *arXiv preprint arXiv:1703.10135*.

**[ 그림 4 ] Tacotron은 사람처럼 문장을 이해한 후 읽는 듯한 모습을 보여준다.<sup>13</sup>**

**Tacotron works well on out-of-domain and complex words.**  
 "Generative adversarial network or variational auto-encoder."  
 00:00 ■■■■■ 00:05 [M] ■■■■■  
 "Basilar membrane and otolaryngology are not auto-correlations."  
 00:00 ■■■■■ 00:05 [M] ■■■■■

**Tacotron learns pronunciations based on phrase semantics.**  
 (Note how Tacotron pronounces "read" in the two phrases.)  
 "He has read the whole thing."  
 00:00 ■■■■■ 00:05 [M] ■■■■■  
 "He reads books."  
 00:00 ■■■■■ 00:05 [M] ■■■■■

**Tacotron is somewhat robust to spelling errors.**  
 "Thiss isrealy awhsome."  
 00:05 ■■■■■ 00:10 [M] ■■■■■

**Tacotron is sensitive to punctuation.**  
 (Note how the comma in the first phrase changes prosody.)  
 "This is your personal assistant, Google Home."  
 00:00 ■■■■■ 00:05 [M] ■■■■■  
 "This is your personal assistant Google Home."  
 00:00 ■■■■■ 00:05 [M] ■■■■■

**Other Domains**

사실 필자가 제작한 딥러닝 논문 리스트[자료 기]는 논문 인용수에 기반해 작성한 자료이기 때문에 연구 커뮤니티가 비교적 큰 비전, 자연어, 음성의 연구 분야에 비하여 다른 분야의 어플리케이션 논문들이 많이 포함되지 못한 측면이 있다. 하지만 딥러닝 알고리즘 자체에 대한 개발 못지않게, 개발된 딥러닝을 실제 생활에 적용할 수 있는 다양한 어플리케이션의 개발 역시 중요한 만큼, 인용수와 관계없이 이들 연구들도 함께 소개해보고자 한다.

먼저 이미지 인식이 직접적으로 쉽게 적용될 수 있는, 하지만 파급력이 매우 큰 분야로는 의료영상 분석[6-1]을 꼽을 수 있다. 의료영상은 대부분 통제 가능한 환경에서 촬영되기 때문에 비교적 이미지 속 타겟의 형태가 일정하고 다른 물체들(e.g. 배경이미지)의 방해가 적다는 점에서 딥러닝 기반의 이미지 인식이 잘 적용될 수 있는 분야다. 의료영상 분석의 목적은 전체 이미지보다는 주로 특정 장기(organ)나 부분의 이상(lesion)에 주목하기 때문에, CT 영상에서의 영역 구분을 통한 장기/해부 구조의 이해와[6-2] 환부의 정확한 위치 파악을 위한 랜드마크 검출[6-3] 등이 활발히 연구되고 있다. 최근에는 스탠포드 대학 연구진이 약 13만 장의 피부암 사진을 CNN을 통해 학습하여 피부과 의사들의 진단에 버금가는 수준의 결과를 얻었다는 사실이 네이처를 통해 발표된 바 있다[6-4].

딥러닝을 통한 인간 모션의 이해는 인공지능의 상황 인식에 있어 매우 중요하게 다루어지는 분야이다. 예를 들어 비디오 분석 연구에선 단지 ‘사람이 있다’가 아니라 ‘사람이 무얼 하고 있다’라는 이해가 매우 중요하게 다루어지는데 (e.g. 감시카메라 속 행위

분석, 유튜브 영상 주제 분석), 이는 보행자의 행동을 이해하며 안전 운전하는 자율주행차량 개발의 기본 조건이기도 하다. HAR(human activity recognition)이라 불리는 이 문제는 영상에서 직접 HAR 분석을 수행하기도 하고[6-5], 인간의 골격 포즈를 영상으로부터 먼저 예측한 후[6-6] 골격 데이터를 기반으로 분석하기도 하는데, 모션 캡처 데이터를 이용해 골격 데이터 분석에만 집중하는 HAR 연구[6-7]도 활발히 진행 중이다. 한편, 헬스케어 분야에서는 프라이버시나 간편함을 이유로 영상이나 모션 캡처보다 웨어러블 센서를 통한 모션 분석을 종종 선호하는데, 필자는 이와 관련하여 50가지 트레이닝 동작 구분[6-8]과 파킨슨병 증상 감지 알고리즘을 상용 웨어러블 디바이스를 이용해 개발한 바 있다.

**[ 소개된 주요 논문들 ]**

[6-1] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *arXiv preprint arXiv:1702.05747*.

[6-2] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 234-241). Springer International Publishing.

[6-3] Zheng, Y., Liu, D., Georgescu, B., Nguyen, H., & Comaniciu, D. (2015, October). 3D deep learning for efficient and robust landmark detection in volumetric data. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 565-572). Springer International Publishing.

[6-4] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115-118.

[6-5] Simonyan, K., & Zisserman, A. (2014). Two-stream convolutional networks for action recognition in videos. In *Advances in neural information processing systems* (pp. 568-576).

[6-6] Toshev, A., & Szegedy, C. (2014). Deeppose: Human pose estimation via deep neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1653-1660).

[6-7] Du, Y., Wang, W., & Wang, L. (2015). Hierarchical recurrent neural network for skeleton based action recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1110-1118).

[6-8] Um, T. T., Babakeshizadeh, V., & Kulic, D. (2016). Exercise Motion Classification from Large-Scale Wearable Sensor Data Using Convolutional Neural Networks. *arXiv preprint arXiv:1610.07031*.

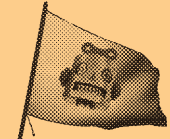
**[ 그림 5 ] DeepPose를 이용한 이미지 속 인물의 pose estimation 결과 [6-6]**

**[ 그림 6 ] DeepPose를 이용한 이미지 속 인물의 pose estimation 결과 [6-6]**

<sup>11</sup> 참고 | Andrew Ng: Why AI is the new electricity, <http://news.stanford.edu/thedish/2017/03/14/andrew-ng-why-ai-is-the-new-electricity/>, March 2017. <sup>12</sup> 참고 | Andrej Karpathy, A Peek at Trends in Machine Learning, <https://medium.com/@karpathy/a-peek-at-trends-in-machine-learning-ab8a1085a106>, April 2017. <sup>13</sup> 참고 | <https://arxiv.org/>, Categories: cs.AI,cs.LG,cs.CV,cs.CL,cs.NE,stat.ML. <sup>14</sup> 논문 | Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein gan. *arXiv preprint arXiv:1701.07875*. <sup>15</sup> 논문 | Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., & Courville, A. (2017). Improved Training of Wasserstein GANs. *arXiv preprint arXiv:1704.00028*. <sup>16</sup> 참고 | Terry. T. Um, "Most-cited Deep Learning Papers", <https://github.com/terryum/awesome-deep-learning-papers> <sup>17</sup> 설명 | 이 분류는 필자에 의한 임의적 분류라는 것을 밝혀 둔다. 이번 1부에서는 앞의 5가지 분야를, 2부에서는 뒤의 5가지 분야와 미래 연구에 대해 다룰 예정인데, 논문의 full-list를 위해선 [자료 기]를 참고하시기 바란다. <sup>18</sup> 참고 | Um, T., "Convolutional Neural Networks", <http://t-robotics.blogspot.ca/2016/05/convolutional-neural-netw> <sup>19</sup> 논문 | LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551. <sup>20</sup> 참고 | Deng, J. et al, ILSVRC-2012, 2012, <http://www.image-net.org/challenges/LSVRC/2012/> <sup>21</sup> 논문 | Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780. <sup>22</sup> 논문 | Graves, A., Fernández, S., Gomez, F., & Schmidhuber, J. (2006, June). Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks. In *Proceedings of the 23rd international conference on Machine learning* (pp. 369-376). *ACM*. <sup>23</sup> 참고 | <https://google.github.io/tacotron/>



# AI, 혁명의 주역



industry

노명철, 이주영 | 사진에서 사람을 읽다

36

최예림 | 산업 현장 속으로 들어간 AI

42

AI에 대한 '기대'와 '걱정'이 공존하고 있습니다. 'AI가 과연 우리 삶에 변화를 줄까?' 하는 불안한 의심과 '앞으로의 삶이 어떻게 변할까?' 하는 기대가 동시에 존재합니다. 앞서 발간한 '카카오 AI 리포트' 1,2호는 변화의 주역에 초점을 맞춰 AI를 소개하였습니다. 이번 호 'Industry' 코너에 실린 두 건의 글은 기업들이 실제 서비스와 제조업 현장에서 어떻게 AI를 활용하고 있는지를 다뤘습니다. 첫 번째 글에서는 카카오 맵의 '로드뷰'에서 활용되고 있는 AI 기술을 확인할 수 있고, 두 번째 글에서는 AI가 전통 제조업에 가져올 새로운 혁신의 가능성을 엿볼 수 있습니다.

# 사진에서 사람을 읽다

카카오맵 로드뷰(Roadview)의  
AI 객체 검출 기술

카카오맵 로드뷰는 전국 각지의 실제 거리 모습을 확인할 수 있도록 해준다. 로드뷰로 제공되는 화면은 DSLR 카메라를 활용하여 고해상도 파노라마 방식으로 거리를 촬영한 것이다.<sup>1)</sup> 촬영은 특정 장소를 360도를 찍는 방식으로 진행된다. 차가 다닐 수 있는 도로에서는 차량 위에 고정된 특수 촬영장비로, 차가 다닐 수 없는 좁은 도로나 공원, 아파트 단지 등은 특수 제작된 세그웨이(segway)<sup>2)</sup>나 파노캠을 사용하여 촬영이 이뤄진다.



[그림 1] 서울, 로드뷰 예시

이렇게 촬영된 화면을 그대로 서비스에 활용할 수는 없다. 촬영된 화면 안의 사람 얼굴이나 자동차 번호판을 공개하는 것은 사생활 침해가 될 수 있기 때문이다. 카카오는 프라이버시를 보호하기 위해 사람의 얼굴과 차량의 번호판 식별을 어렵게 하기 위해 흐릿하게 처리(blurring)한다. [그림 1]에도 사람의 얼굴과 차량의 번호판이 흐릿하게 처리된 모습을 확인할 수 있다. 문제는 여기에서 발생한다. 전국 곳곳에서 찍은 수많은 양의 사진을 사람이 일일이 확인하고, 흐리게 처리하는 것은 엄청난 비용과 시간이 소요되는 일이다.

이를 카카오에서는 객체 검출(object detection) 기술로 해결하고 있다. 이 기술은 영상 내에서 얼굴과 차량 번호판을 인공지능(AI)이 스스로 찾아낸 뒤, 일반인이 해당 개체(object, 사물)를 식별이 불가능하게 흐릿하게 만드는 것이다. 객체 검출 기술의 도입으로 기존의 수동 방법 대비 90% 이상의 작업량 감소 효과가 발생했다. 부연하면, 한 장의 영상에 100개의 얼굴 또는 번호판이 있다면, 이전에는 100개의 위치를 일일이 표시하고 확인해야 했다. 객체 검출 방법 도입 후에는 90여 개 이상이 자동으로 처리되고, 10개 이하의 객체만 사람이 수동으로 확인하면 되어 그만큼 작업량과 시간을 절약할 수 있게 됐다.<sup>3)</sup>

## 객체검출의 개념과 특징

객체 검출은 사람, 차량, 건물 등과 같이 특정한 클래스(class)<sup>4)</sup>에 속하는 객체를 영상에서 자동으로 찾는 컴퓨터 비전(computer vision)과 영상처리(image processing)의 컴퓨터 기술이다.

스마트폰, 로봇, 출입통제 시스템 등에서 널리 사용되고 있는 얼굴 검출, 자동 차량 주행에 사용되는 보행자 검출의 목적으로 객체 검출 기반의 기술이 최근 널리 활용되고 있다.

객체 검출의 수준을 대폭 개선시킨 것은 딥러닝(deep learning)<sup>5)</sup>이었다. 딥러닝은 객체 검출 방식의 패러다임을 바꾸었다. 딥러닝 이전의 객체 검출 방법에서는 여러 클래스(multi-class)의 객체를 검출하고자 할 때, 각 클래스에 사용되는 특징(feature)을 다르게 사용해야 좋은 결과를 얻을 수 있기 때문에 각 클래스 별로 검출기를 설계하고 학습해야 했다. 즉, N개 클래스의 객체를 검출하기 위해서는 N개의 검출기가 필요로 했고 그와 비례하는 비용과 수고가 필요했다. 딥러닝은 이러한 수고를 덜어줬다. 딥러닝의 도입으로 영상 내에서 찾고자 하는 특징을 자동적으로 학습할 수 있게 되었고, 하나의 검출기 모델로 여러 클래스를 구분하고 검출할 수 있게 됐다.

딥러닝의 단점도 존재한다. 딥러닝은 연산 시간의 측면에서 오랜 시간 동안 잘 다듬어져 왔고 상대적으로 훨씬 가벼운 기존의 객체 검출 방법에 비해 많은 연산 양을 필요로 한다. 이로 인해, 피라미드 표현(pyramid representation)<sup>6)</sup>이나 슬라이딩 윈도우즈 기법(sliding windows)<sup>7)</sup>과 같은 객체 검출용으로 사용됐던 전통적인 기법들은 딥러닝을 활용한 객체 검출 기법에는 적용하기 어렵다. 이러한 한계를 극복하기 위해, 효율성 높은 기존의 방법을 차용하여 객체가 있을만한 후보 영역을 먼저 찾아내고, 그 영역에 대해 딥러닝 기법을 적용하는 방법들이 제안되었다. 대표적인 예가 R-CNN(region with convolutional neural network)<sup>8)</sup>이다. 이는

글 | 노명철 joshua.roh@kakaocorp.com

오랫동안 관련 분야에 몸 담아 온 사람으로, 최근 기술의 엄청난 급변을 보는 것이 흥미롭기도 하고 두렵기도 하네요. 이런 변화 사이에서 (치킨집 창업으로 떠밀리지 않기 위해) 열심히 연구하고 있습니다. 좋은 환경에서 훌륭한 동료들과 연구 할 수 있는 것에 감사하고 있고, 아이와 아내에게 자랑스러운 사람이 되고자 합니다.

글 | 이주영 michael.lee@kakaocorp.com

놀이 좋아하고 궁금한거 많은 아주 평범한 이웃집 개발자 같은 개발 + 연구하는 사람입니다. 능력 좋고 성격 좋은 분들과 오랜 시간을 함께 하며 많이 배우고 또 앞으로 배울 것 같습니다. deitel & deitel's 책 저자들 처럼 아이가 커서 성인이 되었을 때에 개발 + 연구하는 사람이 되는데 목표입니다.

영상에 선택적 탐색(selective search)<sup>9</sup>이라는 방법을 적용하여, 객체가 있을 법한 후보 영역(region proposal, RP)을 찾고, 각 RP에 컨볼루션 신경망(convolutional neural network, CNN)을 적용하여 객체를 분류한 뒤, 객체의 위치를 보정하는 방법이다. R-CNN은 많은 연산 시간의 소모가 필요한 CNN이 각 RP마다 적용되어야 하기에, 전체적으로 검출을 위해 소요되는 시간이 길다는 단점이 있다. 반복되고 병렬 연산이 가능한 CNN 연산 속도는 그래픽스 처리장치(graphics processor unit, GPU) 덕에 개선됐지만, RP를 계산하는 부분은 여전히 계산 속도가 느려서 병목(bottleneck) 현상이 발생한다.

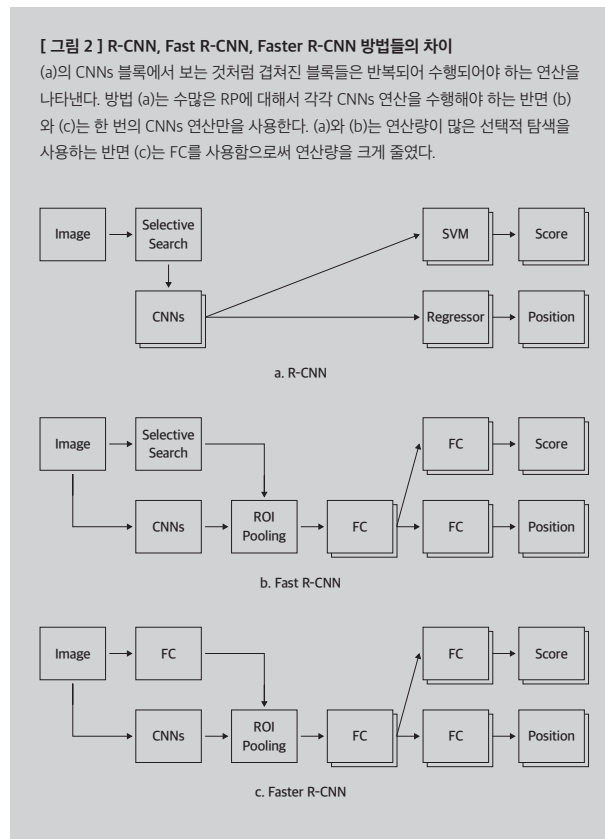
### Fast R-CNN & Faster R-CNN

R-CNN의 시간문제를 개선한 방법이 Fast R-CNN<sup>10</sup>이다. 이 방법은 ROI Pooling (region of interest pooling)이라는 계층(layer)을 도입하여 CNN에서 얻어진 특징 지도(feature map)의 일부 영역으로부터 정규화된 특징을 추출한다. Fast R-CNN은 각 RP에 CNN을 반복적으로 적용하는 대신, 입력 영상에 CNN을 한 번만 적용하고 ROI Pooling으로 객체를 판별하기 위한 특징을 추출한다. 이 방법으로 Fast R-CNN은 CNN에 비해 특징을 추출하기 위한 시간을 대폭적으로 줄였다.

더 나아가, 선택적 탐색 등을 이용하여 RP를 추출하였던 기존 방법을 딥러닝 방법으로 대체하는 방안을 제안하여 속도를 획기적으로 개선한 Faster R-CNN<sup>11</sup>이 있다. 현재 카카오맵 로드뷰에서 사용되고 있는 딥러닝 기법이 Faster R-CNN이다. 이 방법은 RPN(regional proposal network)을 제안하고 CNN 이후 단계에 연결하여 이전의 방법들에서 사용했던 선택적 탐색을 대신하도록 하였다. RPN은 CNN에서 얻은 특징 지도로부터 객체가 있을 것으로 보이는 영역을 제안하고 해당 영역은 Faster R-CNN과 동일하게 ROI Pooling을 통해서 정규화된 특징을 추출하고 그 특징으로 객체를 판별하고 위치와 크기를 보정한다.

[그림 2]는 R-CNN, Fast R-CNN, Faster R-CNN의 구조 간 차이를 보여준다. R-CNN(그림 2.a)은 선택적 탐색을 통해 얻어진 각 영역에 대해서 CNN 연산을 하고 CNN에서 얻어진 특징을 이용하여 SVM(support vector machine)<sup>12</sup>과 regressor<sup>13</sup>로 객체 여부를 판단하고 위치를 보정한다. 연산량이 많은 CNN 연산이 반복되는 단점이 있어서 이를 극복하기 위해서 Fast R-CNN(그림 2.b)은 영상 전체에서 CNN 연산을 한번 하고, 선택적 탐색으로 얻어진 영역에 대해서 ROI pooling 기법을 이용하여 특징을 추출한다. 추출된 특징은 FC(fully connected layer)를 거쳐서 객체 여부와 위치를 얻는다. 그러나 선택적 탐색 방법 자체가 느린 단점이 있어서 Faster

R-CNN(그림 2.c)과 CNN 으로부터 얻은 특징 값을 이용하여 RP를 얻을 수 있는 RPN를 설계함으로써 속도를 크게 높였다. Faster R-CNN의 또 하나의 장점은 End-to-End<sup>14</sup>학습이 가능하다는 것이다. Faster R-CNN이 일반적으로 우수한 성능을 보이지만, 큰 객체와 작은 객체를 같이 학습할 경우 작은 객체의 검출 성능이 상대적으로 떨어지는데, 이는 Faster R-CNN의 단점이다.

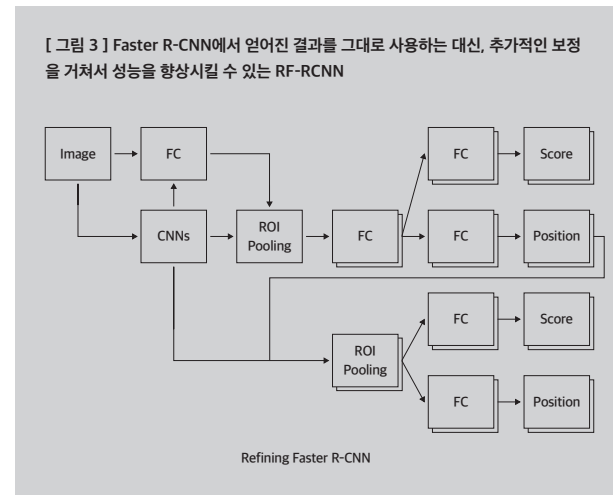


### 로드뷰로의 Faster R-CNN 적용

크기가 크고 특징이 많은 자동차, 사람 등의 객체가 번호판과 같이 작은 객체에 비하여 '객체 여부(objectness)'가 더 뚜렷하므로 일반적으로 큰 객체에 대해서 RPN이 더 잘 동작한다. 그러므로 번호판의 위치에 RP가 잘 주어질 가능성보다는 자동차 전체로 주어질 가능성이 더 크다. 자동차 전체 영상에서 번호판을 판단하는 것은 잘 추정된 얼굴이나 자동차의 RP에서 얼굴과 자동차를 판단하는 것보다 어려운 문제이므로 번호판의 성능이 상대적으로 떨어지게 된다. 이러한 단점을 개선하기 위해서 카카오는 정제 네트워크(refining network)<sup>15</sup>를 사용하였다.

정제 네트워크는 Faster R-CNN의 검출 결과로 나오는 객체 위치정보를 이용하여 다시 한번 검증하고 위치를 보정하는 역할을 한다. Faster R-CNN 부분으로부터 얻은 객체의 영역

정보에 해당하는 정규화된 특징을 특징 지도로부터 추출하고 그 특징으로 객체를 판단하고 위치를 재보정하여 성능을 높인다. 그림 3은 제안한 RF-RCNN(refining faster RCNN)의 네트워크 구조를 보여준다. Faster R-CNN (그림 2.c)와 비교하면, Faster R-CNN의 CNN 이후의 네트워크와 비슷한 구조가 다시 한번 반복되는 구조를 보여준다. Faster R-CNN부분에서 얻어진 위치로부터 다시 한번 ROI pooling과 FC 연산으로 이루어진다. 이때 Faster R-CNN에서 얻은 위치 정보는 한번 보정된 값이므로 보다 정확한 위치에 대해서 객체 여부를 판단할 수 있다.



### 로드뷰 객체검출기의 성능

로드뷰 객체 검출기의 성능을 평가하기 위해서 실제로 카카오맵 로드뷰에 사용되는 900장의 초고해상도 영상을 사용했다. [표 1]은 Faster R-CNN과 제안한 RF-RCNN의 결과를 보여준다. 수치는 EER(equal error rate)로, 작을수록 오류가 작음을 나타낸다. 표에서 보는 바와 같이 번호판 검출의 오류가 크게 줄어들었음을 알 수 있다. 사람 검출 오류가 1%가량 증가하였으나, 전체적인 오류 감소에 비해 미미하다고 할 수 있다.

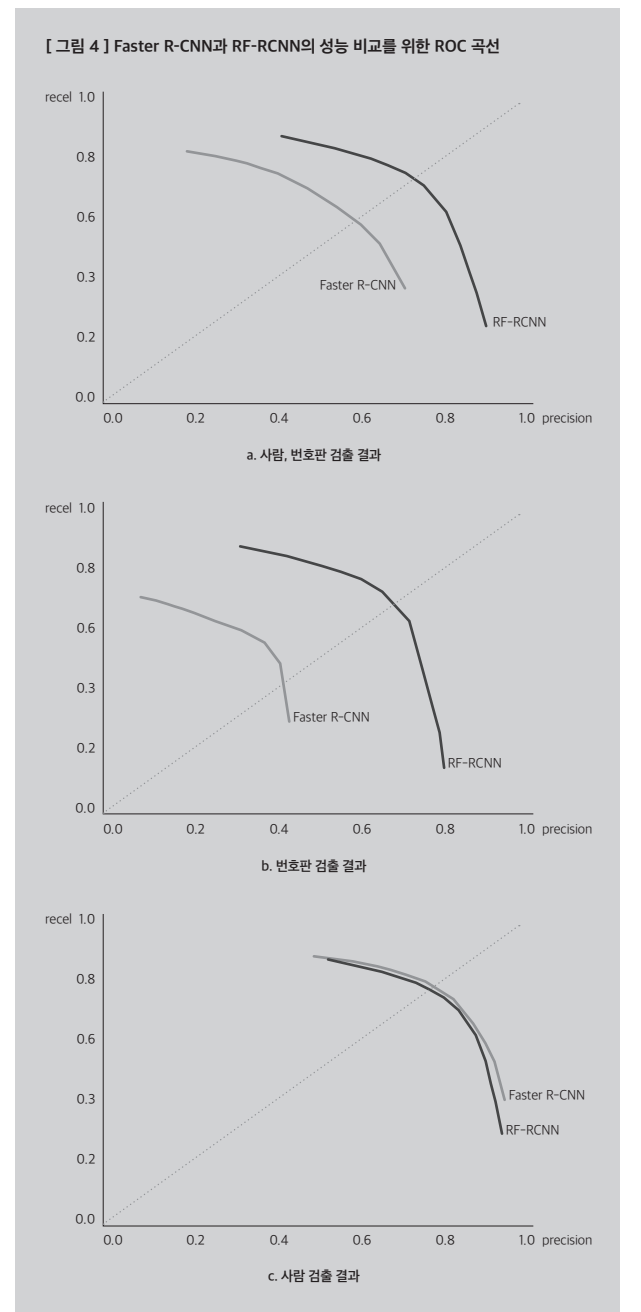
**[ 표 1 ] Faster R-CNN과 RF-RCNN의 검출 오류 비교 실험.**

구분	전체	번호판	사람
Faster R-CNN	0.39	0.57	0.21
RF-RCNN	0.25	0.30	0.22

주: 숫자는 검출 오류를 나타내는 것으로 0-1사이의 값을 가진다

[그림 4]는 Faster R-CNN과 RF-RCNN의 성능을 ROC 곡선(receiver operating characteristic curve)으로 나타낸 것이다. 곡선이 우측 상단으로 붙을수록 좋은 성능을 나타낸다. [그림 4]에서 x-축은 precision<sup>16</sup>을, y-축은 recall<sup>17</sup>을 나타낸다.

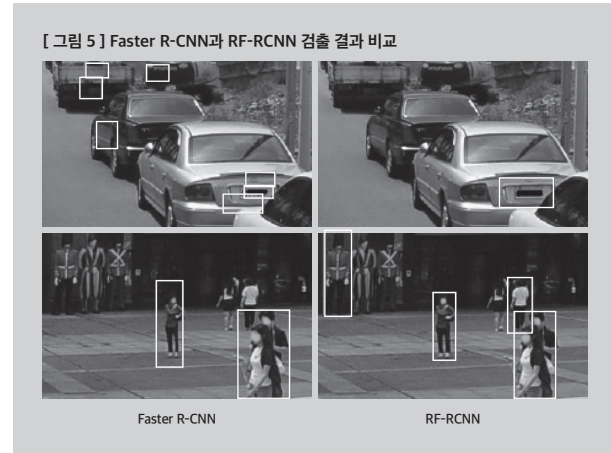
회색 곡선은 Faster R-CNN을, 검정색 곡선은 RF-RCNN을 각각 나타낸다. 그림에서 보듯이, RF-RCNN이 사람 검출(그림 4.c)에서는 비슷하거나 미세하게 떨어지지만 번호판 검출(그림 4.b)에 있어서는 매우 우수한 성능을 보여주고 있다. 결과적으로 전체적인 성능 즉, 사람과 번호판 검출의 성능은 (그림 4.a)와 같이 상승하는 것을 볼 수 있다.



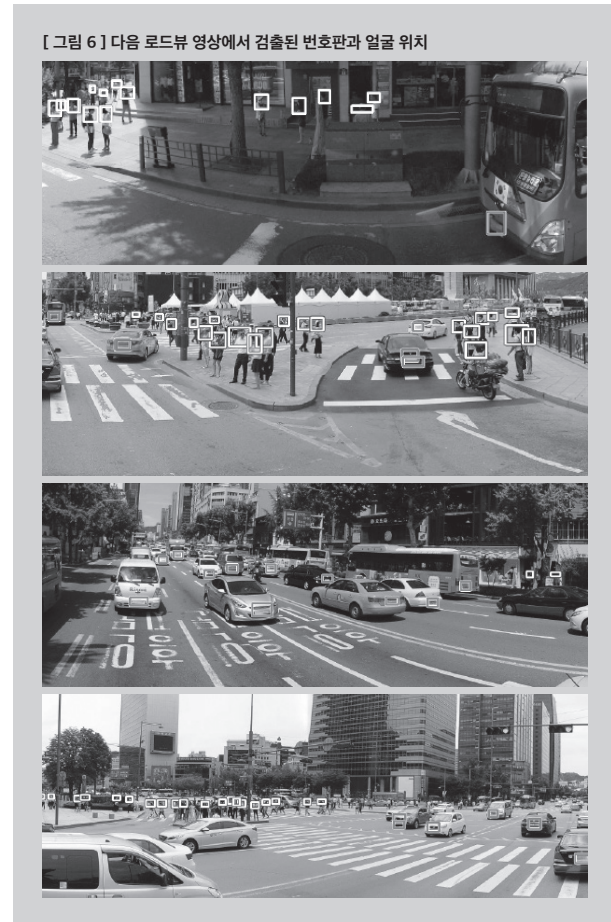
[그림 5]는 Faster R-CNN과 RF-RCNN 검출 결과를 비교하여 보여준다. 차량이 있는 영상 (위쪽)에서, Faster R-CNN의 번호판 검출 결과(왼쪽)에서는 오 검출(false positive)이 여러 개 보이는 반면 RF-RCNN(오른쪽)에서는 그런 오류가 나타나지 않음을



보여준다. 사람이 있는 영상(아래쪽)에서, Faster R-CNN(왼쪽)은 미검출(false negative)이 두 명 있는 반면 RF-RCNN(오른쪽)에서는 미검출이 한 명 있는 결과를 보여준다. RF-RCNN(오른쪽)왼쪽 상단에 있는 마네킹은 오 검출의 사례이다.



[그림 6]은 결과 영상의 예를 보여준다. 하얀색 사각형은 검출된 사람의 윗부분 (얼굴 영역에 해당하는 부분)을 나타내고 회색은 검출된 자동차 번호판을 나타낸다. 사람들과 자동차 수가 상당히 많은 영상에서도 높은 수준의 검출 결과를 보여준다.



검출 성능 평가는 사전에 수동으로 만들어 놓은 정답의 위치와 검출 결과를 비교하여 자동평가를 수행한다. 이때, 해상도가 낮거나 작아서 식별이 불가능한 번호판과 얼굴은 정답에 포함시키지 않았다. 그러나, 검출기가 검출한 결과에는 이렇게 제외한 객체도 검출되는 경우가 종종 발생하고, 이는 자동평가 시에 오 검출로 평가된다. 실환경에서는 이렇게 검출된 객체도 틀렸다고 할 수는 없기 때문에, 오 검출의 개수가 줄어들어서 체감 성능은 더 좋다고 할 수 있다. 실제로, 수동으로 검증하여 평가하였을 경우 96%가량의 검출률을 보여주는 것으로 집계됐다.

**결론**

카카오맵 로드뷰 서비스에서 사생활 침해 방지를 위하여 사람 얼굴과 차량 번호판을 찾아서 흐릿하게 하는 작업이 필요한데, 이를 위하여 최근 AI에서 각광을 받고 있는 딥러닝을 이용하여 검출기를 개발한 과정을 공개했다. 최근에 가장 좋은 성능을 보이고 있는 Faster R-CNN 기반으로 하고, 작은 객체의 검출 성능을 향상할 수 있는 refining Faster R-CNN을 제안했다. 카카오맵 로드뷰의 흐림 처리 서비스에 적용하여 기존 대비 96% 이상의 효율 개선 효과를 볼 수 있었다. AI가 실제 서비스로 구현될 때, 어떤 역할을 할 수 있는지 참고가 되기 바란다.

\*1 설명 | 로드뷰는 국내 포털 지도 서비스에서는 최초로 시도된 서비스이다. \*2 설명 | 전기를 동력으로 사용하는 1인용 소형 이륜차. \*3 설명 | 로드뷰로 인한 사생활 침해가 발생할 가능성을 차단하고자, 영상은 자동 처리 된 이후에도 사람에 의해 최종적으로 검수된 이후에야 서비스 된다. \*4 설명 | 객체를 지칭하는 추상적인 의미의 부류, 여기서의 검출하고자 하는 객체의 종류를 의미한다. \*5 설명 | AI에 딥러닝 적용이 본격적으로 시작된 시점은 2012년 ILSVRC(ImageNet Large Scale Visual Recognition Challenge) 대회 이후이다. 딥러닝은 기존의 학습 방법보다 2배 이상의 성능을 보여줬다. \*6 설명 | 하나의 영상으로부터 여러 단계의 크기로 늘리거나 줄인 영상들을 생성하는 영상처리 기법으로, 객체의 크기 변화에 강한 검출을 위해서 흔히 사용되는 방법. \*7 설명 | 시각 형 영역(윈도우)을 영상 전체로 옮겨가면서, 찾고자하는 객체가 그 안에 존재하는지를 확인하여 검출하는 기법. \*8 논문 | Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587). \*9 논문 | Uijlings, J. R., Van De Sande, K. E., Gevers, T., & Smeulders, A. W. (2013). Selective search for object recognition. International journal of computer vision, 104(2), 154-171. \*10 논문 | Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1440-1448). \*11 논문 | Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems (pp. 91-99). \*12 설명 | 기계학습의 분류(classification)문제에 많이 사용되는 알고리즘. \*13 설명 | regression이라는 기법으로 박스의 위치와 크기를 예측하는 방법. \*14 설명 | 여러 다른 기능들(RP, 분류기, regressor)을 각각 학습하는 대신, 처음부터 끝까지 한꺼번에 학습시키는 기법. \*15 논문 | M.-C. Roh, J.-Y. Lee, (2017). "Refining Faster-RCNN for Accurate Object Detection." In Proceedings of 15th IAPR International Conference on Machine Vision Applications. \*16 설명 | 검출한 객체 중에서 제대로 검출된 객체의 비율 \*17 설명 | 검출해야할 객체 중에서 검출 성공한 객체의 비율

# 산업 현장 속으로 들어간 AI

## 4차 산업혁명과 AI가 가져올 변화

2016년, 과학기술분야의 가장 핫한 키워드가 알파고(AlphaGo)의 충격으로 인한 인공지능(artificial intelligence, AI)이었다면, 2017년에는 4차 산업혁명(The Fourth Industrial Revolution)이 주요 화두로 떠오르고 있다. 4차 산업혁명이라는 단어를 처음으로 사용한 세계경제포럼(World Economic Forum)의 창립자이자 회장인, 클라우스 슈밥(Klaus Schwab) 교수는 그의 저서 ‘The Fourth Industrial Revolution’에서 4차 산업혁명을 인간이라는 존재의 의미까지도 다시 생각해볼게 할 변화라고 서술하며 중요성을 강조했다.<sup>1)</sup> 본 글에서는 4차 산업혁명에서 인공지능의 역할을 살펴보고, 대표적인 사례로 스마트제조에 대해 소개한다. 그리고 인공지능이 이끄는 4차 산업혁명 대한 우려와 기대를 이야기하고자 한다.

글 | 최예림 yrchoi@kgu.ac.kr

서울대학교 산업공학과에서 학사, 박사를 받았다. 선견지명이 뛰어나신 지도교수님 덕분에 SVM만 돌려도 박사받던 시절부터 머신러닝과 인공지능을 공부했다. 네이버랩스에서 최초의 Data Scientist로 근무하였고, 현재는 경기대학교 산업경영공학과에서 학생들을 가르치며 Information Management 연구실을 운영하고 있다.

## 4차 산업혁명과 인공지능

2016년 1월, 스위스 다보스(Davos)에서 열린 세계경제포럼에서 클라우스 슈밥 교수가 처음으로 4차 산업혁명을 언급했다. 4차 산업혁명은 3차 산업혁명을 기반으로 디지털과 바이오산업, 물리학 등을 융합하는 차세대 기술 혁명을 의미하며, 대표적인 구동요인(driven factor)으로 유비쿼터스(a.k.a. 연결성)와 AI를 꼽는다. 약 1만 년 전 인류는 농업혁명을 통해 수렵·채집 활동을 마감하고 정착, 농경생활로 진입했다. 18세기 중반 증기기관의 발명으로 인한 1차 산업혁명과 19세기 말 전기의 발명으로 인한 2차 산업혁명은 기계를 이용한 생산, 대량 생산을 가능하게 했다. 20세기 중반 컴퓨터와 인터넷의 발명으로 디지털 혁명으로 불리는 3차 산업혁명이 일어나 자동화를 통한 생산성 향상에 기틀을 마련했으며, 21세기 초 연결성, AI와 머신러닝(machine learning)으로 대표되는 4차 산업혁명이 시작되고 있다.

3차 산업혁명과 4차 산업혁명의 핵심은 자동화·연결성으로 둘의 구분이 모호해 4차 산업혁명은 3차 산업혁명의 연장선이며 버즈워드(buzzword)에 지나지 않는다는 의견도 존재한다. 4차 산업혁명을 3차 산업혁명으로부터 구분 짓는 특징은 오랜 기간산업을 구성하던 기존의 기준과 법칙들을 뒤집는 파괴적(disruptive) 변화라는 점이다. 저장·운송·복제에 비용이 없는 정보 재화(information goods)가 생산품에서 차지하는 비중이 높아지면서 기존 산업에서 당연하게 받아들여진 수확 체감의 법칙이 무너졌고, 한계비용이 사라져 과거의 생산성은 산업을 평가하는 지표로 더 이상 유효하지 않게 되었다. 예를 들어, 카카오택시와 같은 서비스는 지불 의사 금액과 실제 지불 금액의 차이를 의미하는 소비자 잉여를 발생시켜 소비자가 편리함을 느끼지만 직접적인 수익을 창출하지는 않아 생산성의 관점에서 평가될 수 없다. 또한, 단순 생산을 위한 노동의 가치가 하락하고 있다. IBM의 CEO 지니 로메티(Ginni Rometty)는 2016년 11월 미국의 대통령 당선자 도널드 트럼프(Donald Trump)에게 쓴 편지에서 블루칼라나 화이트칼라가 아닌 새로운 직업의 출현을 이야기했다. 앞으로의 노동자 계급은 대학과 같은 장기 직업교육이 아닌 단기 교육으로 업무역량을 획득하여 빠르게 변화하는 흐름을 따라갈 수 있어야 한다.



4차 산업혁명을 가능하게 하는 요인으로 언급되는 AI는 기계를 인간과 유사하거나 더 뛰어난 지능을 갖게 하기 위한 기술로 다양한 방식으로 구현되어 왔다. 최근에는 머신러닝, 특히 딥러닝(deep learning)의 발전과 함께 데이터 분석 방식에 기반한 AI가 가능성을 인정받고 있다. 예를 들어, 자율주행 또는 무인이동체 분야에서는 수집된 센서 데이터를 분석하여 상황을 판단하거나 행위를 예측하기 위한 AI 기법이 도입되고 있으며, 로봇공학 분야에서는 이미지 기반의 사람/사물 인식, 음성 인식을 위한 기법이 사용되고 있다. 특히, 오랜 기간 정체되어 있던 제조업에 AI 기술이 융합된 스마트제조(smart manufacturing) 분야가 주목받고 있다.

## 스마트제조 분야

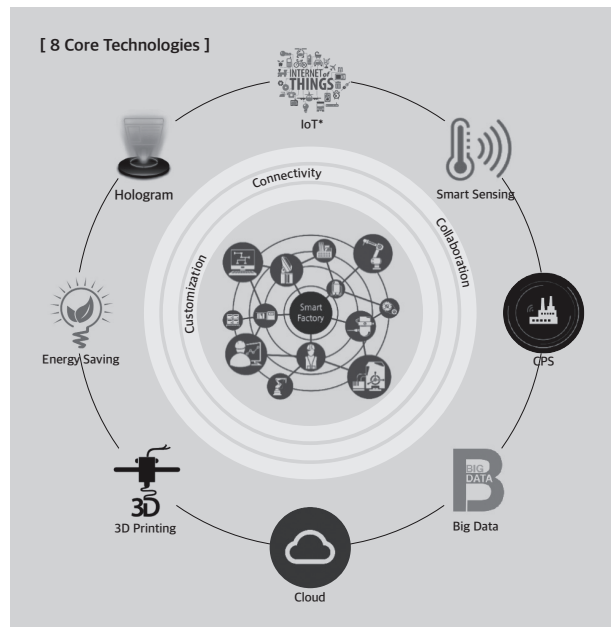
수업시간에 학생들이 진행하고 있는 과제를 하나 소개할까 한다. 고객의 주문 취소로 인한 위험을 최소화하기 위해 MTO(make to order, 미리 생산하여 재고를 저장하지 않고 주문이 들어오면 생산하는 방식) 생산 방식을 취하고 있는 작은 규모의 공장이 있다. 그런데 매달 월말에 주문이 몰려 lead time(상품의 생산 시작부터 완성까지 걸리는 시간)이 지연되는 문제가 발생하고 있다. 학생들은 이 문제를 해결하고자 다양한 전통적인 방식을 떠올리며 고심하고 있다. 만약 AI의 발달로 월말 주문량을 미리 정확하게 예측할 수 있다면 MTO 방식 사용 시 초과되는 주문량에 대해 MTS(make to stock, 미리 생산하여 재고를 저장해 두는 방식) 방식을 도입하여 잉여시간(idle time)에 미리 생산함으로써 문제가 쉽게 해소될 수 있다. 이와 같은 제조업에서 발생하는 생산성을 저하시키는 다수의 문제는 고도화된 AI 기술로 손쉽게 해결할 수 있다.

4차 산업혁명과 종종 혼용되는 단어 중에 ‘인더스트리 4.0’이 있다. 2011년 독일 하노버박람회(Hannover Fair)에서 처음 등장한 용어로 4차 산업혁명이 다양한 산업을 아우르는 개념이라면 인더스트리 4.0은 제조업에 국한되어 스마트제조 of the 활성화를 목적으로 한다. 국내에서는 2020년까지 스마트공장 1만 개 구축을 목표로 하는 제조혁신 3.0이 진행되고 있으며, 해외에서는 독일 외에도 미국의 첨단제조시스템, 일본의 제조업 중흥 프로그램, 중국제조 2025 등이 진행되고 있다.

스마트제조는 밸류체인(value chain) 상의 모든 구성요소(원자재, 제품, 기계, 사람 등)들이 정보기술을



통해 실시간으로 연결되고 데이터 기반의 시스템 제어를 이용하여 고객 맞춤형 제품 및 서비스를 창출하는 플랫폼으로 스마트제조에 기반한 지능형 공장을 스마트 공장이라고 부른다. 고객 맞춤성(customization), 연결성(connectivity), 협업성(collaboration)을 의미하는 3C로 스마트제조 속성을 나타낼 수 있다.<sup>\*2</sup> 이상적인 스마트 공장은 로봇에 의한 자동화와 데이터 기반 AI에 의한 지능화를 지향한다. 스마트제조를 가능하게 하는 8대 기술에는 빅데이터, CPS(cyber-physical system), 사물인터넷(internet of things), 스마트센싱, 클라우드, 3D 프린팅, 에너지절약, 홀로그램이 있다. 이 중 CPS는 가상의 공간에서 수행되는 작업들이 결국은 실제 물리적 공장 내에서 수행되도록 하는 개념을 의미하고, 빅데이터는 수집된 데이터로부터 유의미한 지식을 추출하는 기법으로 향후에 널리 적용될 수 있는 가능성이 있다.<sup>\*3</sup>



스마트제조 분야에서 가장 눈에 띄는 국내 스타트업은 딥러닝 기반의 스마트 공장 솔루션 업체인 수아랩(SUALAB)이다.<sup>\*4</sup> 실제 제조업 현장에 딥러닝 기술을 적용하여 머신비전(machine vision) 기반의 표면 불량 검출 기능을 높은 정확도로 제공한다. 딥러닝 기술을 스마트제조 분야에 상용화한 점을 인정받아 세계 최대 자동화 컨퍼런스 Automate 2017의 Innovators Award에서 대상을 수상했다. 또한, 신성이앤지는 생산라인에서 기계들이 자동으로 제품을 만들어 포장·적재까지 하는 수준의 스마트 공장으로 변화 후 생산량이 2배로 증가했다. 대기업들도 스마트공장 솔루션 사업에 뛰어 들고 있다. SK는 산업용 빅데이터 분석 플랫폼을 개발 중에 있으며, 현대위아는 장비원격모니터링 및 진단시스템을 개발했고, 삼성SDS와 LG CNS는 각각 스마트공장 솔루션을 개발 중이다.



**우려점들**

AI, 특히 딥러닝이 이끄는 4차 산업혁명에 대해 방법론적인 측면과 사회 규범적 측면에서 일부 우려들이 존재하고 있다.

딥러닝이 각광받기 이전 ‘shallow learning’ 방법론을 이용하는 연구를 수행하던 시절에는 주위 사람들의 머신러닝이 실제 세계에 적용 가능하다고 생각하나는 물음에(우선 머신러닝이 러닝머신과는 다르다고 설명해야 했다.) 확신을 갖고 불가능하다고 이야기하곤 했었다. 그 당시 실험을 통해 얻은 정확도는 실제 적용이 불가능한 수준이었고, 답이 있는 데이터를 수집하는 것은 매우 어려웠다. 사용하던 방법론은 데이터의 세밀한 특징까지 잡아내지 못하고 대략적인 특성만 반영했다. 이후 딥러닝이 빅데이터의 출현, 학습 기법의 개선, 컴퓨팅 파워의 증진으로 ANN(artificial neural network)이 갖고 있던 기존의 한계를 풀어나가며 다양한 문제들에 대해 획기적인 성능 향상을 이뤄내고, 관련 분야 연구자 모두가 딥러닝으로 수렴하는 분위기를 지켜보면서 머신러닝이 세상을 바꿀 수 있겠다는 생각을 갖게 됐다. 이처럼 강력한 딥러닝을 기반으로 한 인공지능을 통해 4차 산업혁명이 일어나고 있다.

하지만 딥러닝 또한 머신러닝이 갖는 약점을 그대로 갖고 있다. 머신러닝을 간단하게 설명하자면, 데이터의 관계를 설명할 수 있는 수학적 모델을 미리 구성해 두고 이 모델의 파라미터를 데이터로부터 추정(학습)하는 방식을 의미한다. 따라서 데이터의 양과 질이 머신러닝 방법의 성능을 좌우하고, 답이 있는 데이터를 이용하여 학습하는 것(지도 학습, supervised learning)이 답이 없는 데이터를 이용하여 학습하는 것(비지도 학습, unsupervised learning) 보다 좋은 성능을 제공한다.

지도 학습은 사용을 위해 답이 있는 데이터를 구해야 한다는 치명적인 한계를 갖고 있다. 예를 들어, 스마트 공장에서 머신비전 기술을 도입하여 생산품 표면의 오류를 높은 정확도로 검출하기 위해서는 오류의 여부와 위치가 정확하게 명시된 생산품 표면 이미지 데이터가 (그것도 매우 많이) 있어야 한다. 특히, 양이 많을수록 정확도가 향상될 뿐만 아니라, 복잡한 구조의

딥러닝 모델을 사용하기 위해서는 더욱 많은 데이터가 필요하다. 장기간 지속적인 수집을 통해 데이터를 확보한다고 해도 인간의 실수(human error)로 인해 데이터의 편향(bias)이 발생하면 그 데이터를 이용해 학습한 모델의 성능에도 문제가 발생한다. 따라서 머신러닝과 딥러닝 분야의 권위자 중 한 명인 앤드류 응(Andrew Ng) 교수는 이제는 비지도 학습 방식을 발전시켜나가야 한다고 주장한다. 실제로 강화 학습(reinforcement learning), one-shot learning, GAN(generative adversarial network) 등 새로운 학습 방식을 이용하는 기법들이 속속 출현하고 있다.

또한, 데이터 기반 인공지능 방법론의 성능의 불안정성, 낮은 설명력 때문에 연구나 위험이 비교적 적은 예제 문제가 아닌 실제 산업에의 적용 가능성은 떨어진다는 의견이 많다. 그만큼 실제 문제를 해결하기 위해 인공지능 기법을 도입하는 것은 까다롭다. 알파고 역시도 단순히 딥러닝 모델만을 사용한 것이 아니라, 기본적으로는 시뮬레이션 기반의 검색 기법인 몬테카를로 트리 탐색(monte carlo tree search)<sup>\*5</sup> 알고리즘을 사용했다. 앞서 언급한 수아랩에서도 제조 이미지 처리 전문가들의 지식을 불량검사 모델에 반영하여 실제 적용 가능 한 수준의 성능을 확보했다. 아직 모두가 상상하는 완전한 지능형 자동화까지는 지속적인 연구와 투자가 요구된다.

4차 산업혁명의 대표적인 키워드는 지능형 자동화이다. 즉, AI는 데이터 기반의 제어를 통해 고객-생산 간의 관계를 밀접하게 연결하고 다품종 소량 생산을 가능하게 하여 제조에서 인간의 개입을 감소시킨다. 예를 들어, 제조업에서 생산량 예측이나 오류 검토 등 기존에는 인간의 개입이 반드시 필요했던 작업에서 인간의 참여가 불필요하게 된다. 따라서 고용이 줄어들고 직업이 사라지게 된다.

자연스럽게 자동화로 인한 AI의 인력 대체를 걱정하는 목소리가 크다. 실제로 옥스퍼드대 마틴스쿨(Oxford Martin School)에서 진행한 ‘컴퓨터화에 민감한 직업 조사 연구’에서 10년에서 20년 사이에 2010년 미국에 존재했던 직업의 47%가 자동화될 것이라고 예측했다.<sup>\*6</sup> 반복적인 일들은 이미 더 이상 인간이 할 필요가 없어졌고 고급 숙련도를 요구하는 일들조차도 AI가 대체하게 될 것이다. 우리나라에서도 콜센터 및 은행의 창구는 시스템으로 대체되고 있으며, IBM이 개발한 AI 왓슨은 변호사나 의사의 업무도 대체하기 시작했다. 또한, 플랫폼 효과를 우려하는 목소리도 있다. 플랫폼 효과란 시장을 지배하는 강력한 소수의 플랫폼으로의 집중 현상을 의미한다. 예를 들어, 구글은 웹에서의 검색 플랫폼, 이메일 플랫폼, 스마트기기 OS의 플랫폼으로써 막대한 양의 데이터를 수집하고 있다. 결국 이 데이터는 더 나은 성능의 AI 개발 원동력이 되어 다른 소규모 기업과의 격차를 점차 늘린다.

앞서 언급한 우려들에 대해서는 긍정과 부정 의견이 모두 존재한다. 스스로를 실용적 긍정론자라고 이야기하는 슈왈 교수는 결과적으로는 소비자가 받는 혜택이 제일 크다고 주장한다. 하지만 자본이 노동자를 대체하여 막대한 양의 부를 축적함에 따라 노동자와 자본가 사이의 격차가 점점 커질 수밖에 없고 이를 보정할 수 있는 사회적 제도가 필요하다. 빌 게이츠(Bill Gates)는 로봇을 소유한 사람에게 세금을 부과하는 방식의 로봇세를 제안했고, 유럽연합에서는 이 세금을 기본소득의 재원으로 활용해야 한다는 주장도 나왔다. 이처럼 가치와 힘이 소수에게 집중되는 것을 막기 위해 제도를 정비하고 데이터의 개방성과 다양한 기회를 보장해야 한다.

**그리고 앞으로의 기대**

클라우드 슈왈 교수는 4차 산업혁명이 이전의 산업혁명들과 다른 양상으로 사회를 탈바꿈시키고 있다고 설명한다. 같은 맥락에서 AI가 인류에 미치는 영향은 본질적으로 인간의 존재 가치에 대한 진보이며, 우리가 알고 있는 ‘산업혁명’ 수준을 넘어선다고 생각한다. 따라서 4차 산업혁명보다는 ‘AI 혁명’이라고 부르는 것이 적절해 보인다.

초기 인류의 목적은 모든 동물들이 그러하듯 생명유지에 있었다. 도구의 발명과 진보를 통해 인류는 다른 동물들과는 다른 한 차원 위의 존재가 될 수 있었다. 이후 지금까지의 인류의 목표는 노동을 통한 생산에 있어 왔다. 학습과 노동을 통해 사회를 위한 가치를 창출해야만 생존할 수 있었으며, 학생들도 이 목표를 달성하기 위한 교육을 받고 있다. 하지만 AI의 발달과 함께 앞으로의 인류는 노동·학습·생산이라는 개념에서 탈피하게 될 것이며, 지금보다 한 차원 더 높은 차원의 가치를 추구하게 될 것이다.

물론, 우리가 살고 있는 시대는 그 변화의 가장 초기로 직접적인 변화를 느끼게 되기까지는 긴 세월이 필요할 것이다. 하지만 이 엄청난 변화와 진보의 과정에서 인류가 길을 잃지 않기 위해서는 다가오는 변화를 받아들이고 이해하여 정책·교육·복지 측면에서 함께 준비하는 과정이 필요하다.

<sup>\*1</sup> 참고 | Schwab, K. (2017). The Fourth Industrial Revolution. Crown Publishing Group. <sup>\*2</sup> 참고 | 신동민, 정봉주, 조현보. (2017). 스마트제조(제4차 산업혁명의 예술). 이프레스. <sup>\*3</sup> 참고 | 정재윤, 장태우, 최예림, 이정철. (2017). 스마트공장을 고려한 제조혁신방법론(KPS) 발전방안. 한국전자거래학회 춘계학술대회. <sup>\*4</sup> 참고 | http://www.sualab.com/ <sup>\*5</sup> 참고 | AlphaGo의 알고리즘과 모델, http://sanghyukchun.github.io/97/ <sup>\*6</sup> 참고 | Frey, C., & Osborne, M. (2017) The future of employment: How susceptible are jobs to computerisation?, Technological Forecasting and Social Change, 114, 254-280.

# AI 유튜브강의 소개

## 이론에서부터 활용까지 다양한 AI 유튜브 강의 소개

‘카카오 AI 리포트’ 2호에서는 AI를 공부하고자 하는 사람들을 위한 초급에서 고급까지의 온라인 수업과정들을 제공했다.

([http://1boon.kakao.com/kakao-it/ai\\_online\\_lecture](http://1boon.kakao.com/kakao-it/ai_online_lecture)) 이번에는 AI와 관련된 이론뿐 아니라 활용까지 다양한 분야를 접할 수 있는 유튜브 강의를 소개하고자 한다. AI 연구의 사대천왕이 말하는 AI와 딥러닝, 인공지능경망에 대한 내용들과 연구자들의 연구 분야들, 최신 논문이 다루는 주제들에 대해 수준 높은 강의들을 접할 수 있다.

## Beginner

### How we teach computers to understand pictures

어린이가 사진을 볼 때, '고양이', '책', '의자'와 같이 단순한 것을 식별할 수 있다. 시발점으로 컴퓨터도 시각 정보를 구분할 수 있게 되었다. 어떻게 AI가 이미지를 구분하게 되는지에 대해 쉬운 설명을 통해 이해할 수 있다. Caltech 연구자로서 컴퓨터 비전(computer vision) 발전에 기여한 노력들을 전해 들을 수 있다. 한글 자막이 제공된다.

**강연자** | Fei Fei Li (Professor at Stanford University)  
**분량** | 18분  
**수준** | Beginner  
**주소** | <https://www.youtube.com/watch?v=40riCqvRoMs>

## Intermediate

### A Path to AI

2017년 1월에 개최된 아실로마 컨퍼런스(Asilomar conference)에서 Yann LeCun 교수가 강연한 내용이다. AI가 현재까지 오게 된 계기와 유익하고 완벽한 AI (beneficial AI, complete AI)가 되기 위한 가능성에 대한 내용을 다루고 있다. 기술적인 내용보다는 개념적인 내용이 많아서 비교적 쉽게 들을 수 있다.

**강연자** | Yann LeCun (Professor at New York University)  
**분량** | 22분  
**수준** | Intermediate  
**주소** | <https://www.youtube.com/watch?v=bub58oYJTm0>

### Foundations and Challenges of Deep Learning

딥러닝의 학습방법인 비지도학습, 지도학습 등이 실제 문제 해결을 위해서 어떻게 활용되고 있는지 강의하고 있다. 또한, 학습과정 중에 만나는 문제들을 어떻게 헤쳐나갈 수 있는지에 대해서 오랜 연구 경험을 바탕으로 설명하고 있다. 머신러닝 학습의 기본 개념을 이해하고 있으면 좀더 쉽게 들을 수 있는 강의이다.

**강연자** | Yoshua Bengio ( Head of the Montreal Institute for Learning Algorithms )  
**분량** | 72분  
**수준** | Intermediate  
**주소** | [https://www.youtube.com/watch?v=11rsu\\_WwZTc](https://www.youtube.com/watch?v=11rsu_WwZTc)

## Advanced

### Deep Reinforcement Learning

강화학습(reinforcement learning)의 개념을 차분한 설명으로 학습할 수 있다. 강화학습이 기계학습의 다른 문제들을 해결하는데 어떻게 활용될 수 있는지, 딥강화학습(deep reinforcement learning)을 활용한 장점과 단점을 이해할 수 있다.

**강연자** | John Schulman ( Research Scientist at OpenAI )  
**분량** | 87분  
**수준** | Advanced  
**주소** | <https://www.youtube.com/watch?v=PtAlh9KSnjo>

### Introduction to Generative Adversarial Networks

2016년 겨울 개최된 NIPS 학회의 tutorial로 진행된 강의 내용이다. 최근 많이 연구되고 있는 GANs(generative adversarial networks)에 대한 소개와 개념을 이해하고, 간단한 사례들을 다루고 있다. GANs을 통해 어떤 문제들을 해결할 수 있는지에 대해서 알아볼 수 있다.

**강연자** | Ian Goodfellow ( Research Scientist at Google Brain )  
**분량** | 31분  
**수준** | Advanced  
**주소** | <https://www.youtube.com/watch?v=RvgYvHyT15E>

### Feedforward neural network

딥러닝의 기본이 되는 인공신경망(artificial neural networks)에 대한 강의. 강의자료를 PDF로 제공하고 있으며, 뉴런(neuron)의 개념과 신경망을 통한 신호의 전달 개념 등을 설명하고 있다. DNN(deep neural network) 구조에 대한 개념 설명과 수식을 통해 구현할 수 있는 방법을 배울 수 있다.

**강연자** | Hugo Larochelle ( Research Scientist at Google Brain )  
**분량** | 58분  
**수준** | Advanced  
**주소** | <https://www.youtube.com/watch?v=SGZ6BttHMPw>

## Advanced

### Neural Networks for Machine Learning

제프 힌튼 교수가 Coursera를 통해서 제공하는 강의이지만, 유튜브로 회원가입 없이 시청할 수 있다. 강의는 총 60개로 인공지능경망의 개념, 구현 방법, 응용사례 등을 배울 수 있다. 2013년 강의 자료로 빠르게 변하는 AI연구 분야에서 조금 시간이 지난 내용일 수 있지만, 기본기를 탄탄하게 다질 수 있는 좋은 내용들로 구성되어 있다.

**강연자** | Geoffrey Hinton ( Professor at Toronto University )  
**분량** | 610분  
**수준** | Advanced  
**주소** | [https://www.youtube.com/watch?v=\\_Cl6U\\_O6OY](https://www.youtube.com/watch?v=_Cl6U_O6OY)

### Deep Learning and Neural Net short course

현재는 존스홉킨스대학 연구자로 있는 Kevin Duh 일본인 학자가 NAIST(Nara Institute of Science and Technology)교수로 재직 중일 때 강의한 내용이다. 총 4개의 동영상으로 구성되어있고, 유창한 영어로 딥러닝의 구조와 DBN(deep belief networks)에 대해서 설명하고 있다. AI 연구에 중요하게 기여한 논문들의 내용을 소개하며 이해를 돕고 있다.

**강연자** | Kevin Duh ( Research Scientist at Johns Hopkins University )  
**분량** | 360분  
**수준** | Advanced  
**주소** | <https://www.youtube.com/watch?v=DLItuVVKJOW>

### Deep Learning for Self-Driving Cars

딥러닝을 활용한 자율주행차(self-driving car)를 만드는 방법을 알 수 있는 강의들이다. 6개로 구성되어 있고, MIT에서 진행된 수업으로 딥러닝의 기본개념, 자율주행차에 응용하는 방법들에 대해서 깊이 있게 학습할 수 있다.

**강연자** | Lex Fridman ( Postdoctoral Associate at the MIT AgeLab. )  
**분량** | 370분  
**수준** | Advanced  
**주소** | <https://www.youtube.com/watch?v=1L0TKZQcUtA&t=5s>

# 마치며

桑田碧海(상전벽해) 세상이 하루아침에 바뀌지 않듯, 지금의 AI 기술의 발전은 수많은 연구자들의 오랜 노력의 결실입니다. 이번 리포트를 통해 AI 발전의 과거와 현재 돌아보고 미래를 엿볼 수 있었기를 기대합니다. 카카오 AI 리포트는 벌써 세 번째 발간하고 있습니다. 앞으로도 좋은 내용을 제공하기 위해 많은 분들의 참여와 관심을 기다리고 있습니다.

