cvpaper.challenge

ー CVの動向とこれからの問題を作るために ー

片岡 裕雄, *Ph.D.*

産業技術総合研究所

知能システム研究部門 コンピュータビジョン研究グループ (CVRG)

http://www.hirokatsukataoka.net/

cvpaper.challenge

CV分野の論文サーベイチャレンジ

- 産総研，東京電機大，筑波大，東京大，慶應大による約30名
- 論文を「読む」から「書く」まで
  - 読む：多読・体系化・共有
  - 書く：洗練された問題設定を論文化





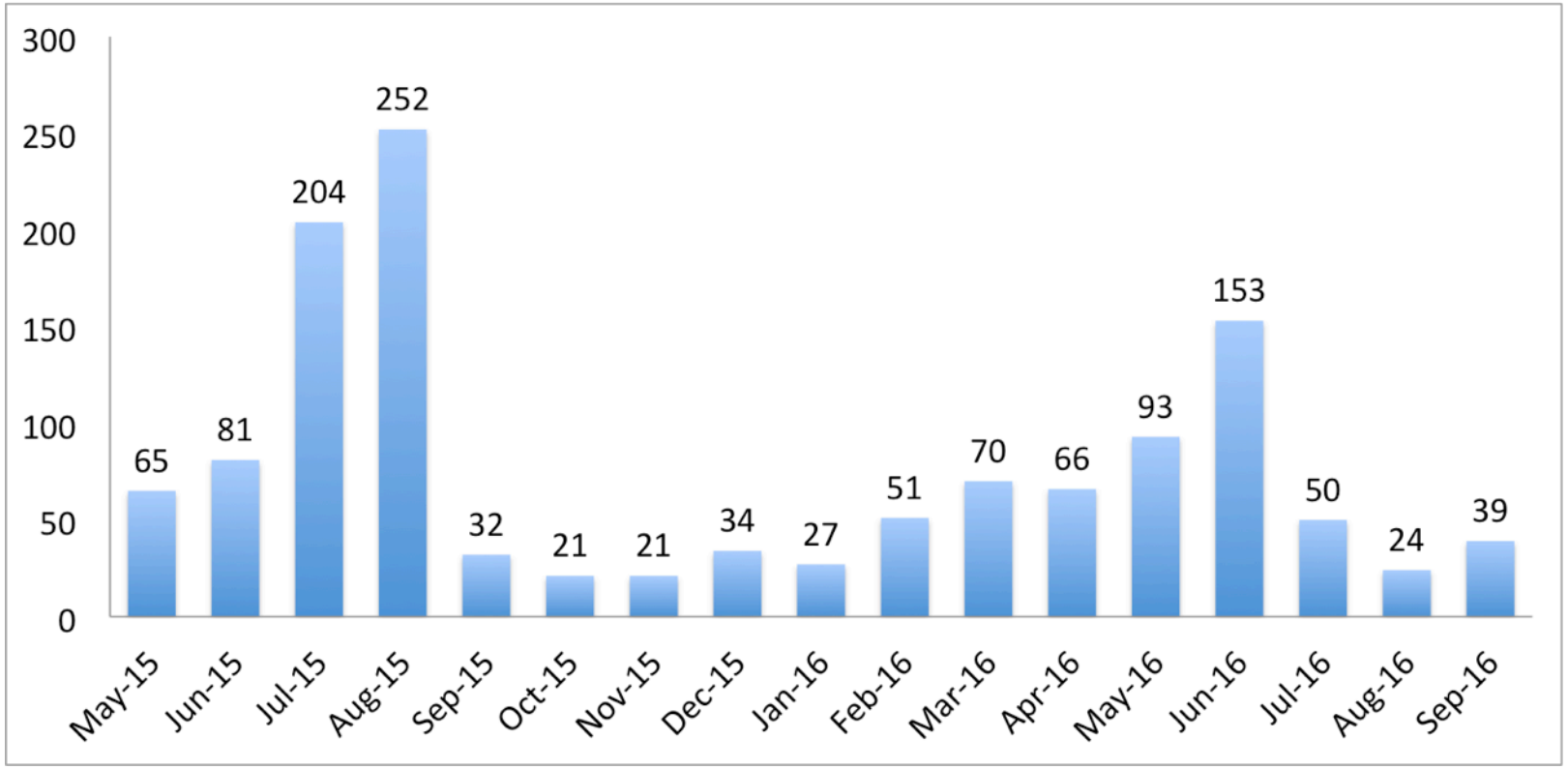
※ Twitter, SlideShareもご覧ください
Twitter: @CVpaperChalleng
SlideShare: @cvpaperchallenge

1,283

論文を「読む」

現在(2016年9月26日)までに読んだ本数：1,283

- 2015年8月迄：CVPR2015の全602本論文を読破
- ~現在：研究のトレンドを把握すべく幅広く調査
- 全てのまとめスライドを共有中



なぜ、サーベイするのか？

世界水準の研究のため

トレンドの把握

- 知識がないと既存研究の劣化版を作りかねない
- トレンドを知らないと(天才でない限り)最先端の研究を生み出すことは難しい

自身の研究の立ち位置を確認

- 何が違う？なぜやる？どこが良いのか？という哲学

究極的には次のトレンドを作るため

- 分野の方向性を自ら定める
- より良く，正しい方向へ導く

なぜ，そこまでサーベイするのか？

日本人研究者はサーベイが足りない？

- よく言われるが，そうは思わない（読んでる人は読んでいる）
- しかし海外の（大学院生含めた）研究者は年間数百本読んでる

この合同プロジェクトは好例になれるか？

- 日本のプレゼンスを少しでも高める
- 一足飛びに追いつくには思考が飛んだことをやらないといけない
- 論文を膨大に読んで動向を網羅的に把握する

関連の取り組み

http://hci.tokyo/seminar/chi2016/

CHI勉強会2016

The ACM CHI Conference on Human Factors in Computing Systemsの論文（Proceedingsに入っているPapers&Notes）、全545本を6時間で一気に読破します！

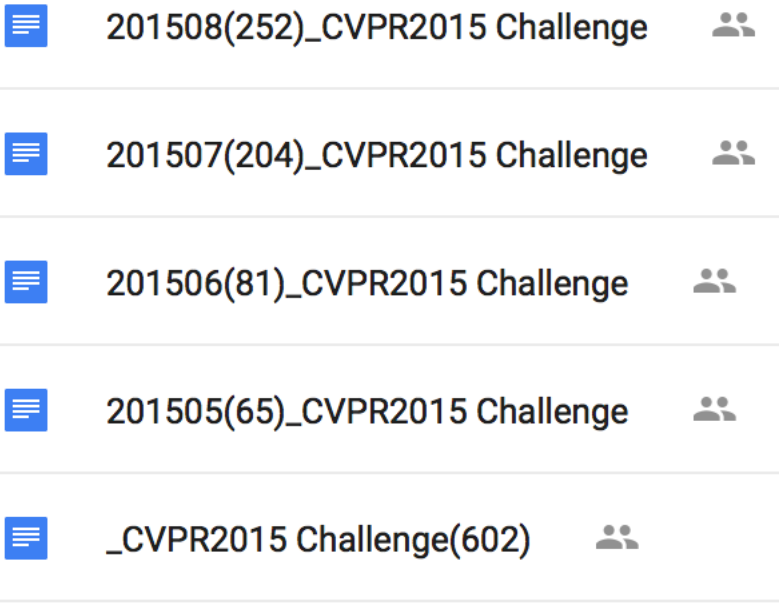
- CHIはユーザインタフェース界のトップ会議
- 2006年から11年間続く取り組み
- 日本のHCI分野の好調を支えている

cvpaper.challengeはCV分野をさらに盛り上げるための起爆剤になれるか？

どうやって読んでいるのか？

クラウドにて資料を共有し，高効率化

- Google Driveをメンバーで同期
- 読んだ/読んでない 論文を可視化
- 成果・進捗を月別に管理





	Erhan, Vincent Vanhoucke, Andrew Rabinovich		Anton van den Hengel, Chris Russell, Anthony Dick, John Bastian, Daniel Pooley, Lachlan Fleming, Lourdes Agapito
1505	**Understanding Image Representations by Measuring Their Equivariance and Equivalence** Karel Lenc, Andrea Vedaldi	1505	**SUN RGB-D: A RGB-D Scene Understanding Benchmark Suite**
			Shuran Song, Samuel P. Lichtenberg, Jianxiong Xiao
1505	**Deep Neural Networks Are Easily Fooled: High Confidence Predictions for Unrecognizable Images** Anh Nguyen, Jason Yosinski, Jeff Clune	1505	**Small-Variance Nonparametric Clustering on the Hypersphere** Julian Straub, Trevor Campbell, Jonathan P. How, John W. Fisher III

Monday June 8, 10:10am-12:30pm

	Poster Session
1505	**Going Deeper With Convolutions** Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich
1506	**Propagated Image Filtering** Jen-Hao Rick Chang, Yu-Chiang Frank Wang
1506	**Web Scale Photo Hash Clustering on A Single Machine** Yunchao Gong, Marcin Pawlowski, Fei Yang, Louis Brandy, Lubomir Bourdev, Rob Fergus
1506	**Expanding Object Detector's Horizon: Incremental Learning Framework for Object Detection in Videos** Alina Kuznetsova, Sung Ju Hwang, Bodo Rosenhahn, Leonid Sigal
1506	**Supervised Discrete Hashing** Fumin Shen, Chunhua Shen, Wei Liu, Heng Tao Shen
1505	**What do 15,000 Object Categories Tell Us About Classifying and Localizing Actions?** Mihir Jain, Jan C. van Gemert, Cees G. M. Snoek
1508	**Landmarks-Based Kernelized Subspace Alignment for Unsupervised Domain Adaptation** Rahaf Aljundi, Rémi Emonet, Damien Muselet, Marc Sebban

_CVPR Challenge (602)

どうやってまとめている？

テンプレートに従って１論文１ページ

- 素早く研究の肝をつかむ
- 【概要】【新規性・差分】【手法】【結果】【Links】

【1】 

Jianxiong Xiao, James Hays, Krista A. Ehinger, Aude Oliva, Antonio Torralba, “SUN Database: Large-scale Scene Recognition from Abbey to Zoo”, in CVPR2010.

Keywords: Dataset, Scene Categorization, Benchmark, Recognition

概要

コンピュータビジョンにおいてシーン認識のデータベースである Scene UNderstanding (SUN) databaseを提案．シーン認識の裾野を広げた．

データセットの概要

シーン認識に関する397クラス，130,519枚の画像が含まれる．
画像例は次ページ．比較した特徴量は，HOG, denseSIFT, self-similarity (ssim), LBP, GIST, textonなど．

新規性・差分

それまでの物体認識のデータセットでは数百クラスの識別クラスが用意されていたが，シーン認識では15種類程度しか含まれていなかった．SUN databaseでは，それまでのデータセットをさらに拡大させ，397クラスのシーンを含む，大規模なデータセットである．

結果

次ページの図の通り．全ての特徴量を統合するのが最も精度が高いことが判明した(38.0%)．次いでHOG2x2 (27.2%), geometry texton hist (23.5%), ssim (22.5%), dense SIFT (21.5%)であった．

Links

論文ページ：
http://cs.brown.edu/~hays/papers/sun.pdf

プロジェクトページ：
http://vision.princeton.edu/projects/2010/SUN/

HOG https://hal.archives-ouvertes.fr/inria-00548512/document
GIST http://cvcl.mit.edu/scene_understanding.html
SSIM
http://www.researchgate.net/profile/Eli_Shechtman/publication/221362526_Matching_Local_Self-Similarities_across_Images_and_Videos/links/02e7e520897af25746000000.pdf
DenseSIFT http://www.vision.caltech.edu/Image_Datasets/Caltech101/cvpr06b_lana.pdf
LBP http://www.outex.oulu.fi/publications/pami_02_opm.pdf
Sparse SIFT http://www.robots.ox.ac.uk/~vgg/publications/papers/sivic04b.pdf
Texton http://www.ics.uci.edu/~fowlkes/papers/mftm-iccv01.pdf

まとめ資料はどのように使えば良い？

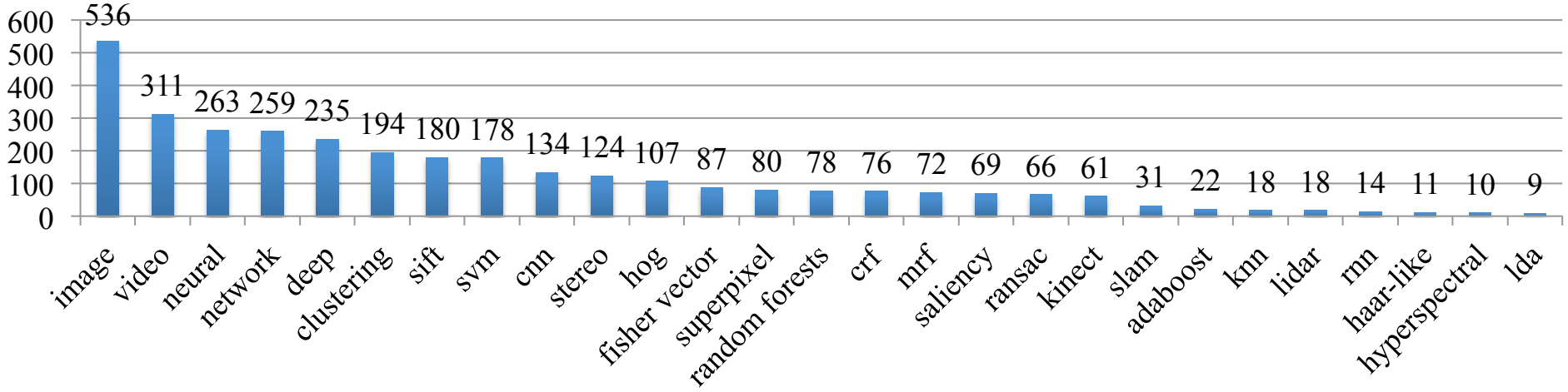
基本的には自由

- とにかく読む
- 気になった論文があれば原本を読んでみる
- ブレインストーミングに使う
- 考案したアイディアへの関連付けに使う
- 引用論文を探す
- など

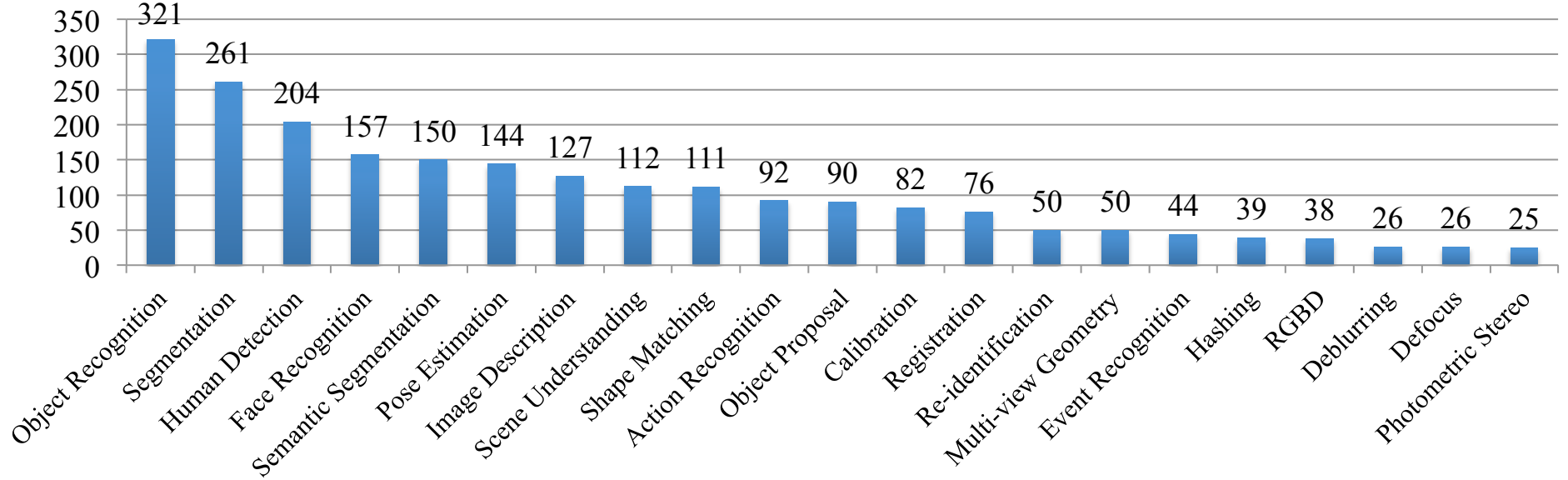
近年の画像認識分野の動向

CVPR2015の統計

ワードでヒットした論文数

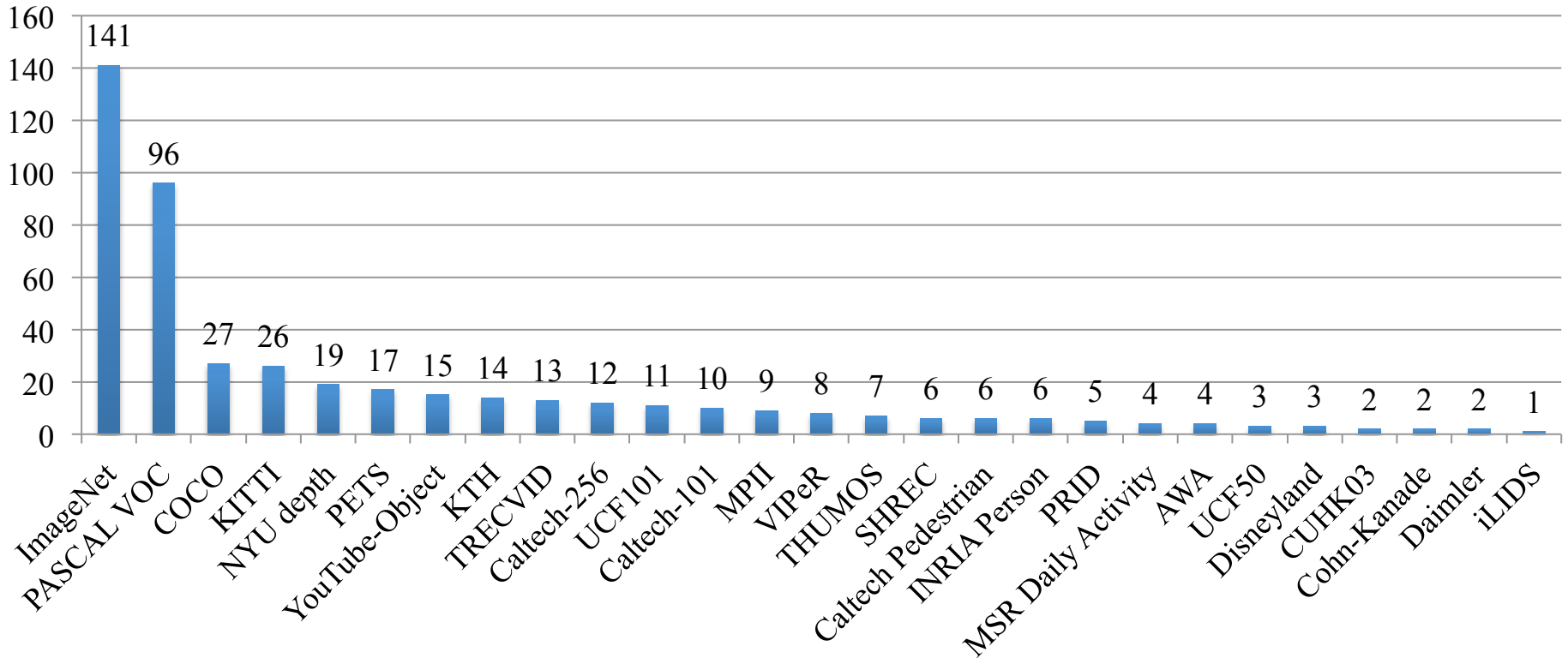


研究部門でヒットした論文数



CVPR2015の統計

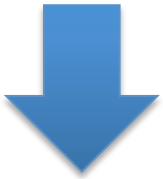
データセットでヒットした論文数



CVPR2015 最大の波

Convolutional Neural Network (CNN)の隆盛！

- 約250件 (全体の約40%)が深層学習の研究に関連



- 精度を出すための知識が整った
  - 勾配を保存しつつ，より深い構造に
  - 学習のためのデータ数を膨大に
- ライブラリ/Big Data/計算リソースが整った
  - Caffe/Torch/Pylearn2/Theano/Chainer/TensorFlow
  - ImageNet, Microsoft COCO, …

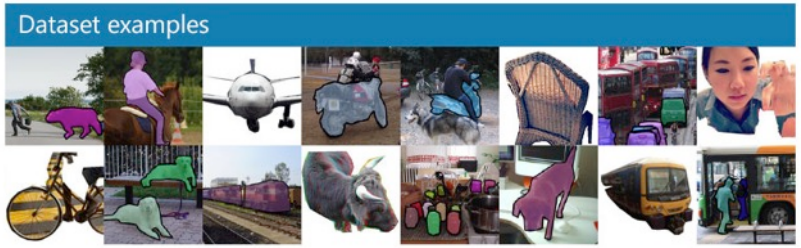
ImageNet and around

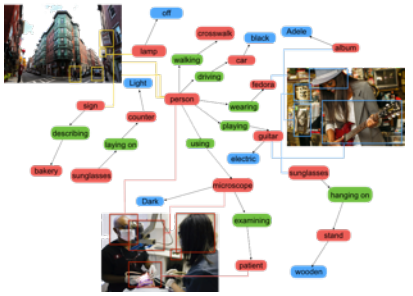
Caltech101/256, ImageNet, MSCOCO, Visual Genome





IMAGENET

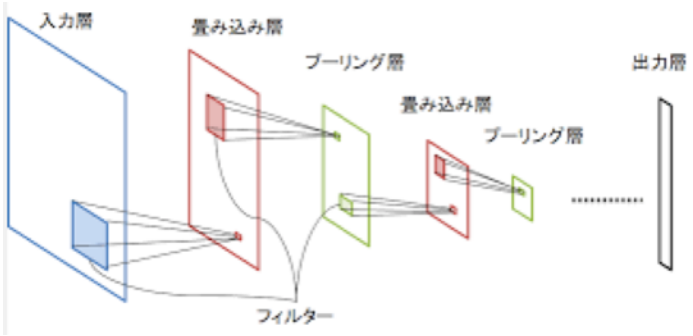


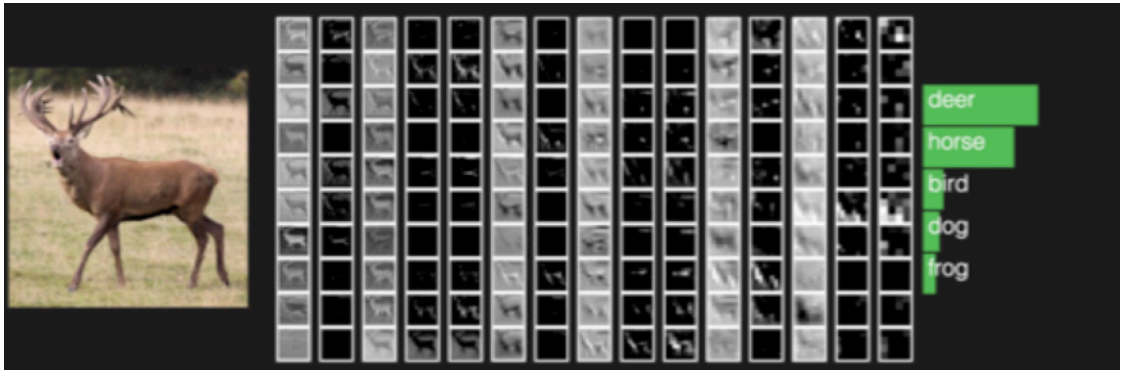


深層学習

深層畳み込みニューラルネット (CNN)

- 畳み込みと呼ばれるフィルタリングにより特徴抽出
- プーリングにより位置ずれの影響を最小限に抑える
- 畳み込み・プーリングの繰り返しにより画像から概念へ
- 全結合層により特徴を高次化



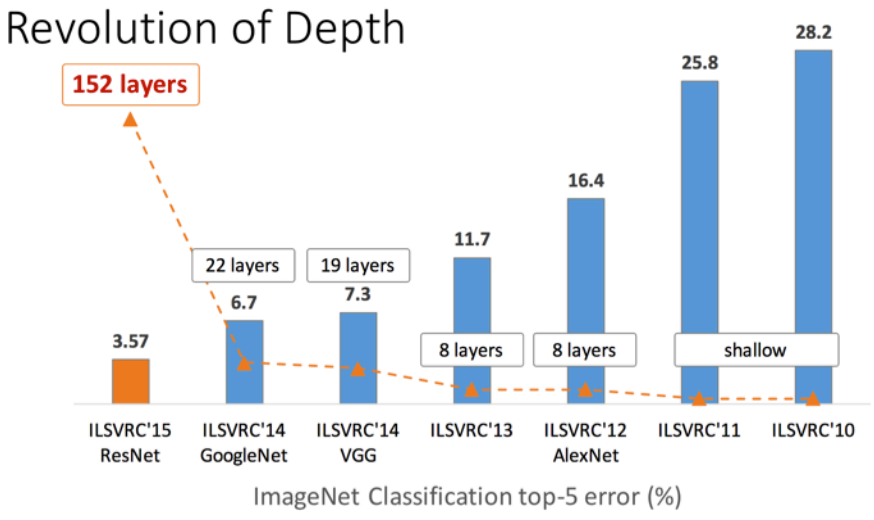


https://www.youtube.com/watch?v=zj_JlVqWK1M

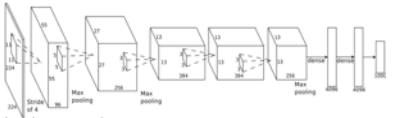
深層学習の衝撃

画像認識の精度が格段に向上

- 2012年のILSVRCにてDeep Learningが大勝利
- 2016年現在1,000クラス分類にて3.08%のエラー率
- 画像認識のみならず，物体検出/ セグメンテーション/ 時系列解析/ 画像生成/ 可視化/ 3次元/ カメラ幾何などに応用



Kaiming He, Xiangyu Zhang, Shaoqing Ren, & Jian Sun. "Deep Residual Learning for Image Recognition". CVPR 2016.



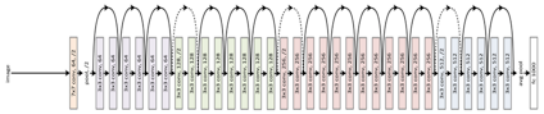
AlexNet [Krizhevsky+, ILSVRC2012]
ILSVRC2012 winner，DLの火付け役



VGGNet [Simonyan+, ILSVRC2014]
16/19層ネット，deeperモデルの知識



GoogLeNet [Szegedy+, ILSVRC2014/CVPR2015]
ILSVRC2014 winner，22層モデル

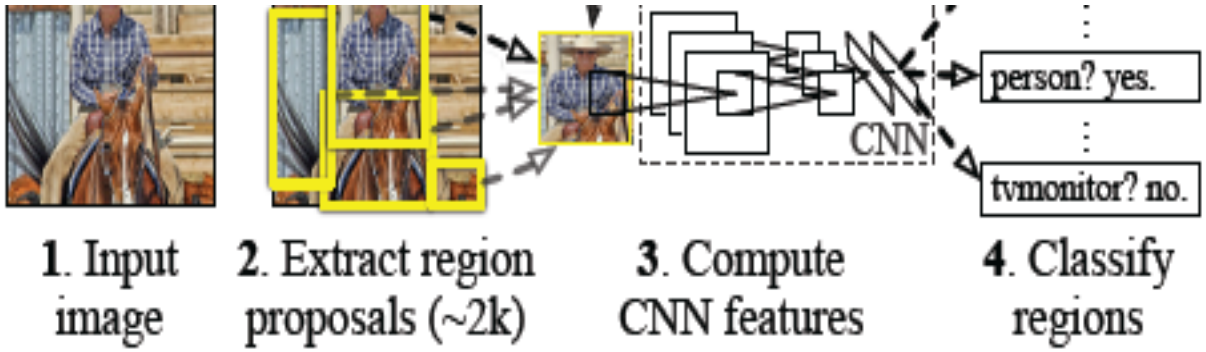


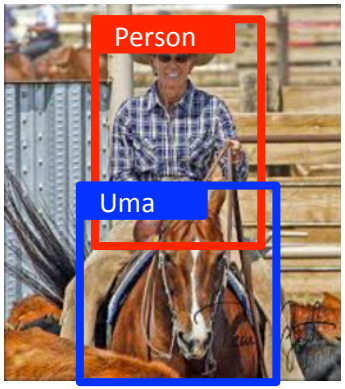
ResNet [He+, ILSVRC2015/CVPR2016]
ILSVRC2015 winner，152層！(実験では10^3+層も)

物体検出の仕組みと高速化

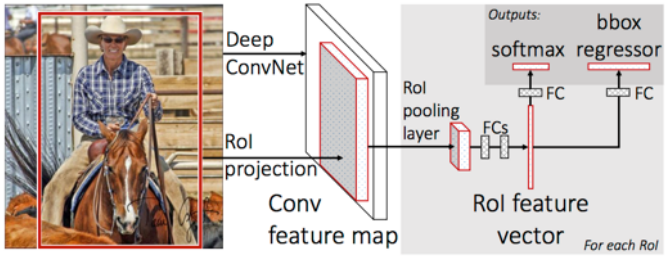
Region with CNN (R-CNN)が物体検出の先駆け

- 物体候補領域を任意の手法で抽出，各領域にてCNNの識別
- Fast R-CNNでは畳み込みマップを共有して識別
- Faster R-CNNでは候補領域抽出自体もCNN内にて実行

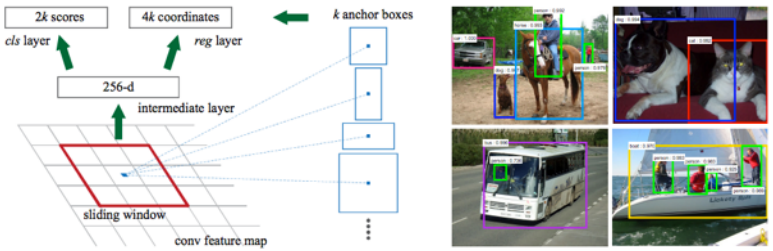




R-CNN [Girshick+, CVPR14] 0.021fps



Fast R-CNN [Girshick, ICCV15] 3.1fps

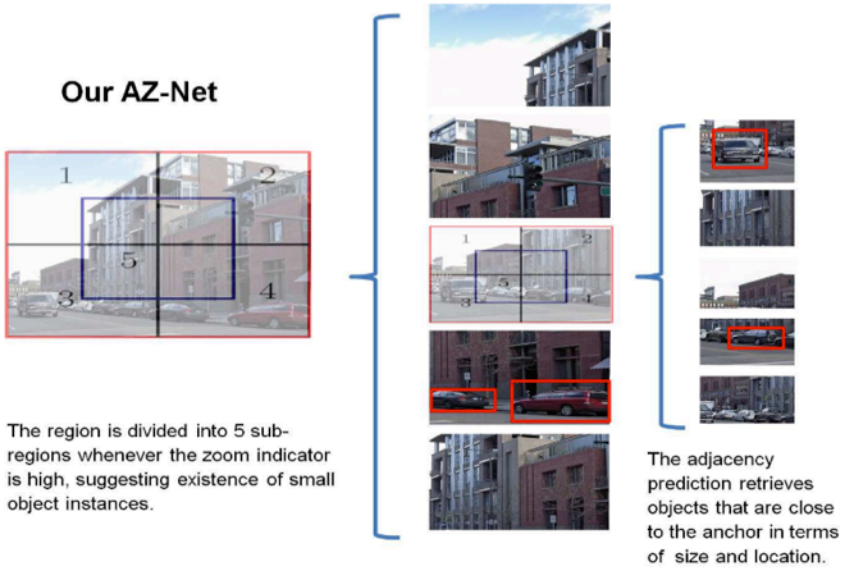


Faster R-CNN [Ren+, NIPS15] 5.0fps
候補領域をニューラルネット内で実行

物体検出のリアルタイム化

候補領域抽出の簡略化による高速化

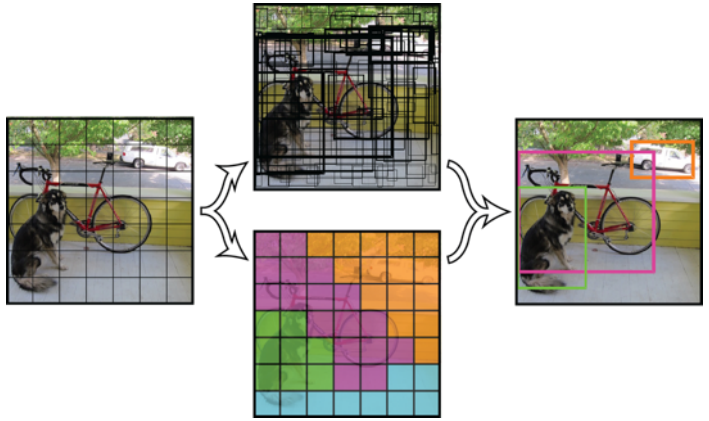
- AZ-Net (~30fps): 領域選択と拡大，評価により高速化
- YOLO (30 ~ 150fps): ブロック分割による評価



AZ-Net [Lu+, CVPR16] 30fps



https://www.youtube.com/watch?v=Gl1rxvEvgEs



YOLO [Redmon+, CVPR16] 30~150fps

驚くべきことに・・・

高速化と同時に高精度化も実現！

- @PASCAL VOC
- 動画は運転シーン



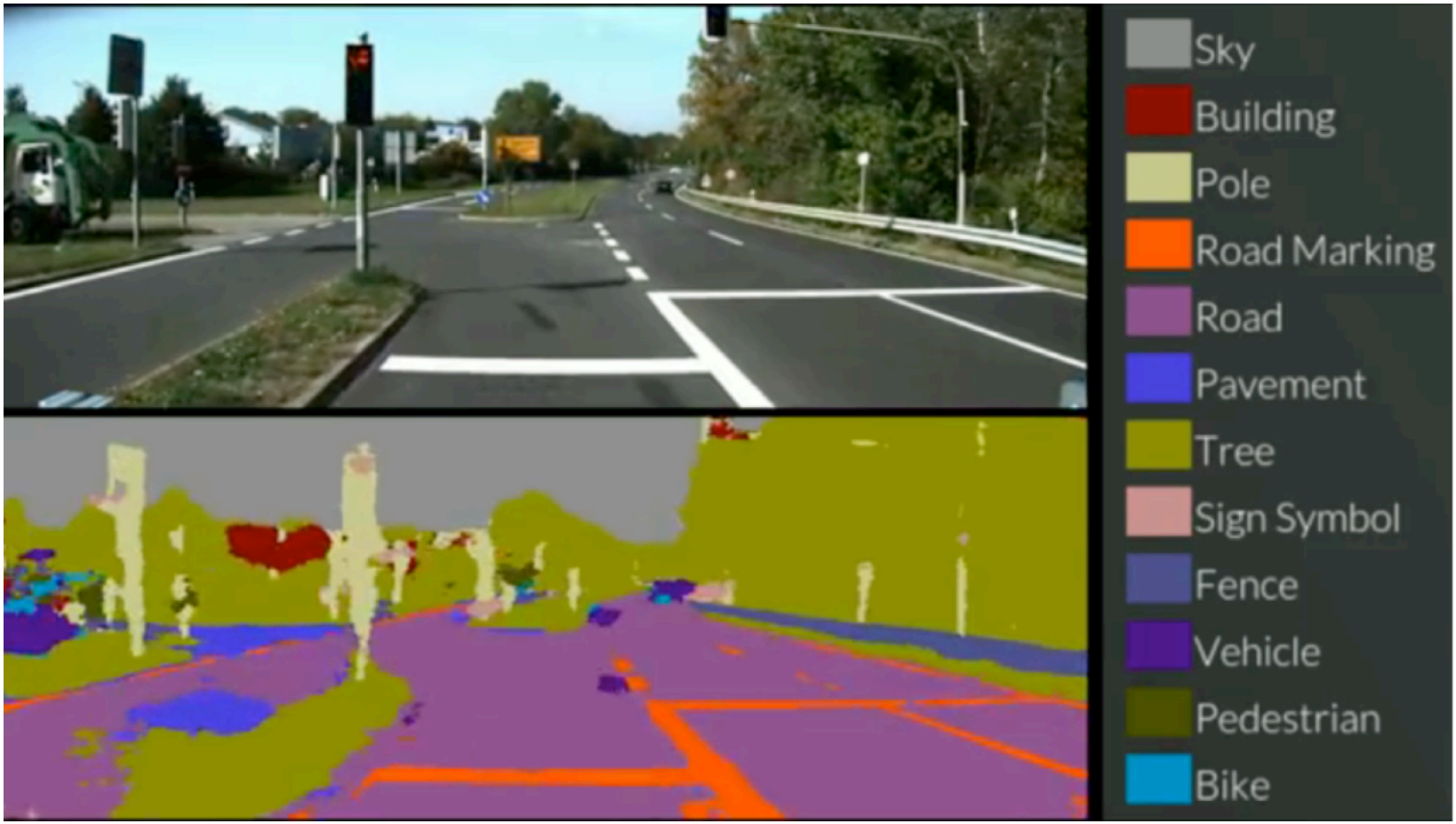
https://www.youtube.com/watch?v=WZmSMkK9VuA

	2007	2010	2012
R-CNN [CVPR14] 0.021 fps	58.8	53.7	53.3
R-CNN+ [PAMI15]	66.0	62.9	62.4
Fast R-CNN [ICCV15] 3.1 fps	70.0	68.8	68.4
Faster R-CNN (VGGNet) [NIPS15] 5.0 fps	73.2	-	70.4
Faster R-CNN (Resnet) [CVPR16] 5.0 fps?	**85.6**	-	**83.8**
YOLO (Model Fusion) [CVPR16] 30 ~ 150 fps	75.0	-	70.7

セマンティックセグメンテーション

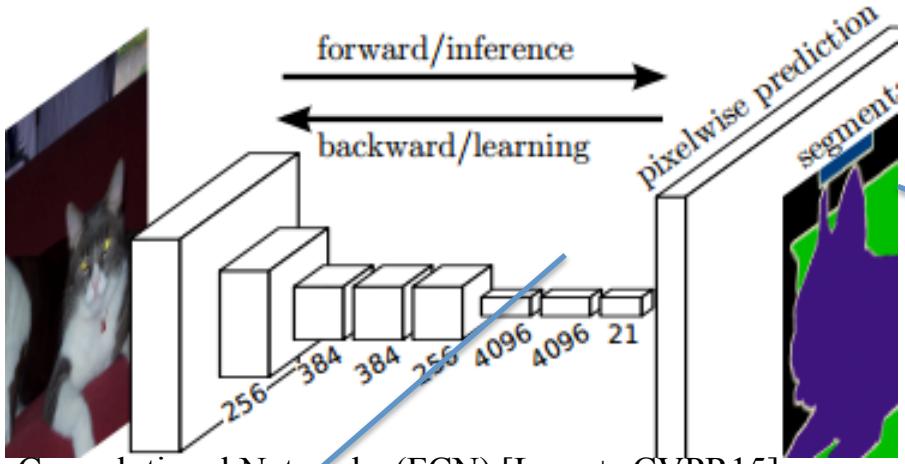
クラス分類と領域セグメンテーションを同時実行

– 近年最も急激に進んだ技術である



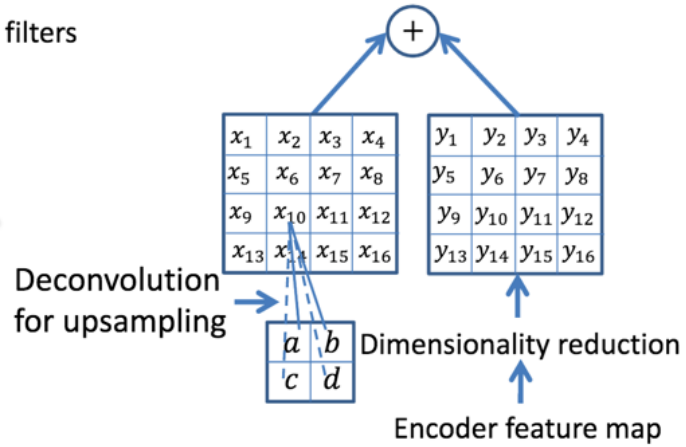
セマンティックセグメンテーションの仕組み

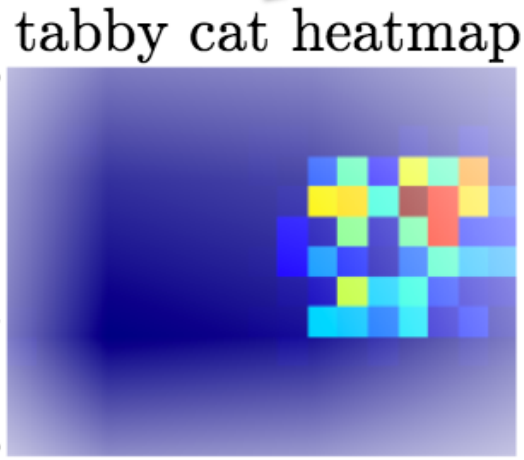
概念レベルの特徴からセマンティック化，画像位置に投影



Fully Convolutional Networks (FCN) [Long+, CVPR15]

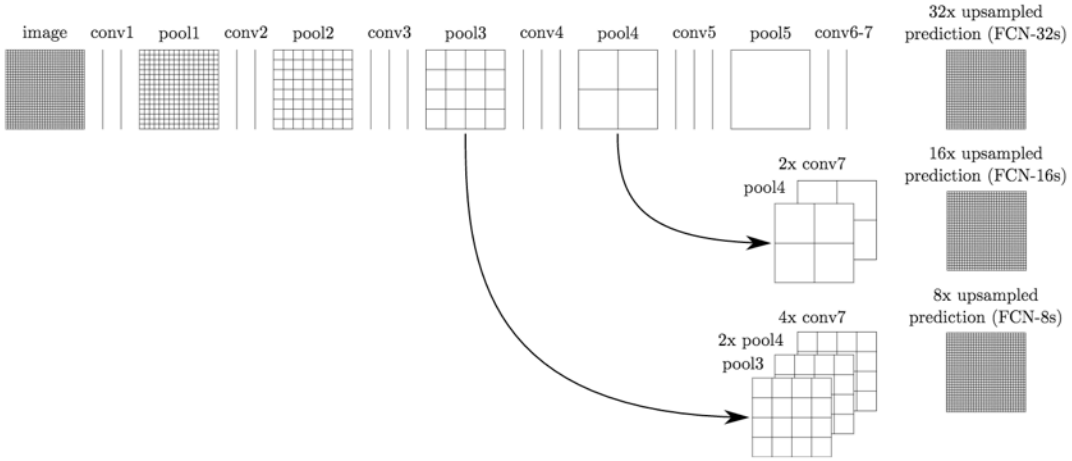
画像空間へ投影するフィルタ





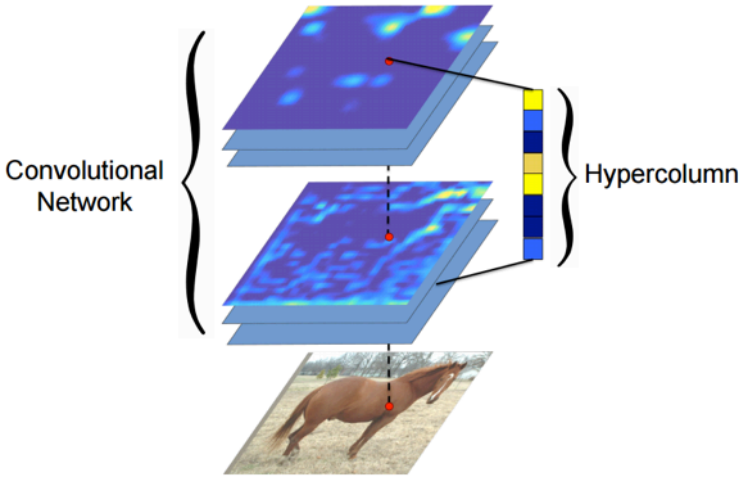
物体の概念レベルの特徴マップ

階層の異なる特徴マップを参照

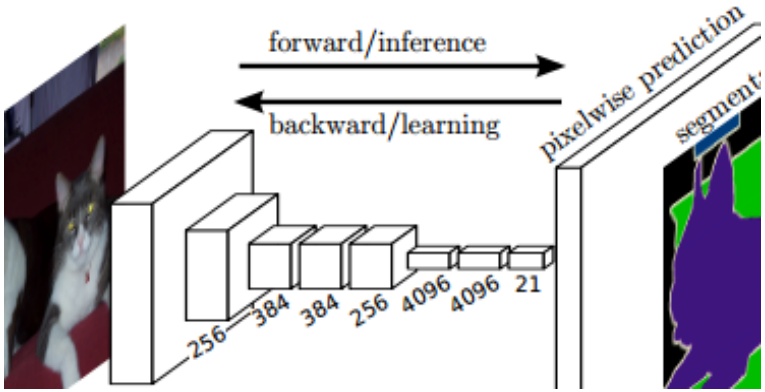


最近の有力な手法

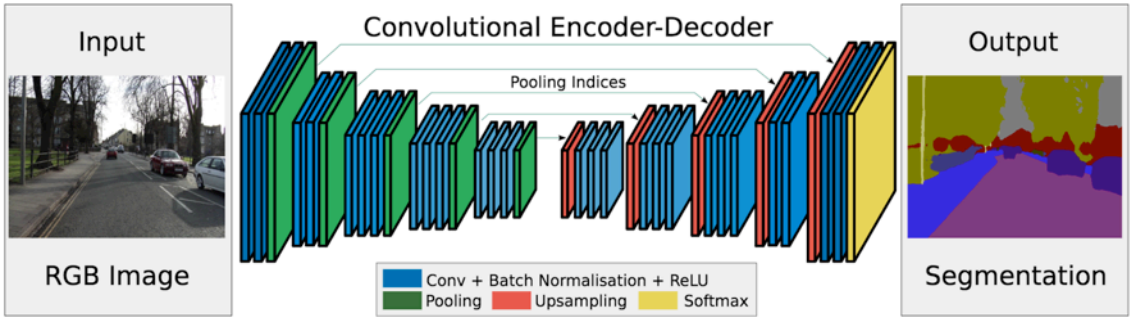
Hypercolumns, FCN, SegNet

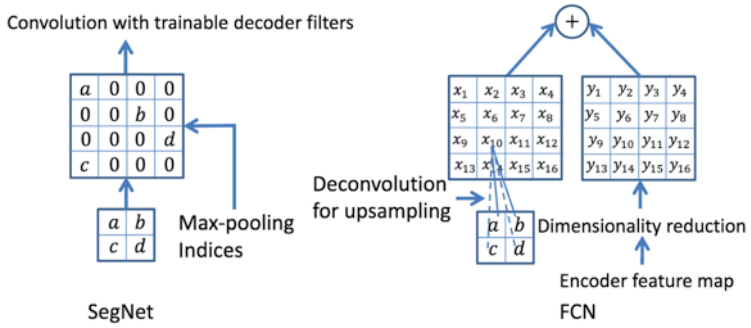


Hypercolumns [Hariharan+, CVPR15]
異なる階層の畳み込みマップを結合



Fully Convolutional Networks (FCN)
[Long+, CVPR15]
畳み込みマップからのセグメント結果復元



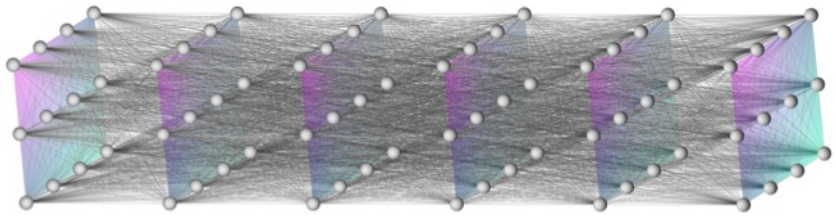


SegNet [Kendall+, arXiv15]
畳み込みと同じ階層分の復元層

セマンティックセグメンテーション

自動運転の実用レベル？





3D Dense CRF [Kundu+, CVPR16]
空間的・時間的に密な結合による領域評価

時系列解析

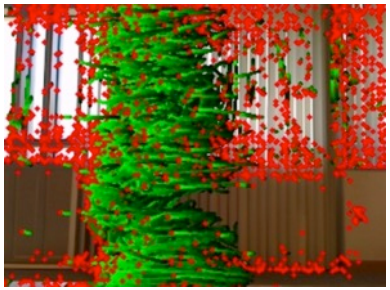
メディア解析のための時系列特徴抽出

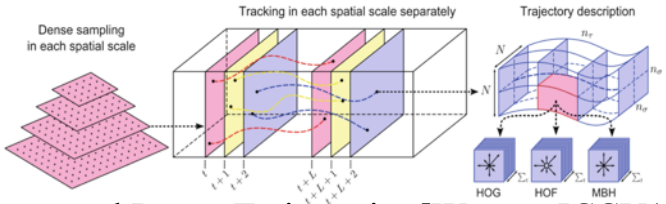
- YouTubeなど動画共有サイトの動画像を解析
- 1動画に対して1ラベルを付与する問題設定が主流



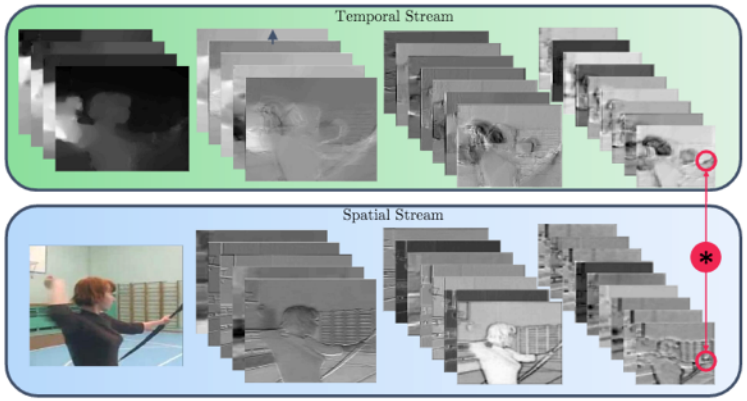
時系列解析

ハンドクラフト特徴 + 深層学習による表現

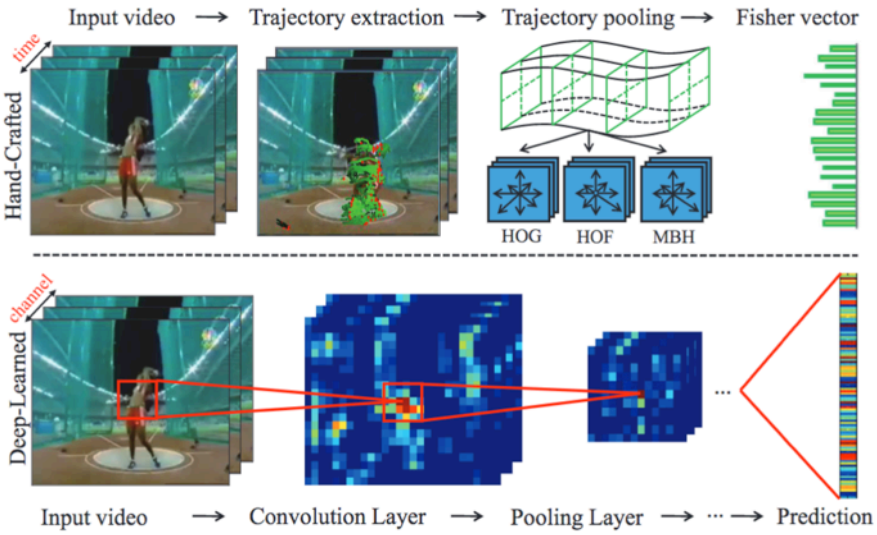




Improved Dense Trajectories [Wang+, ICCV13]
密な動線抽出とベクトル量子化による記述



Two-Stream CNN [Simonyan+, NIPS14]
RGBとFlow画像を入力とした動画像表現



Trajectory-pooled Deep-convolutional Descriptors (TDD)
[Wang+, CVPR15]
IDT + Two-Stream CNN
動線位置を用いてTwo-Streamの特徴マップ参照

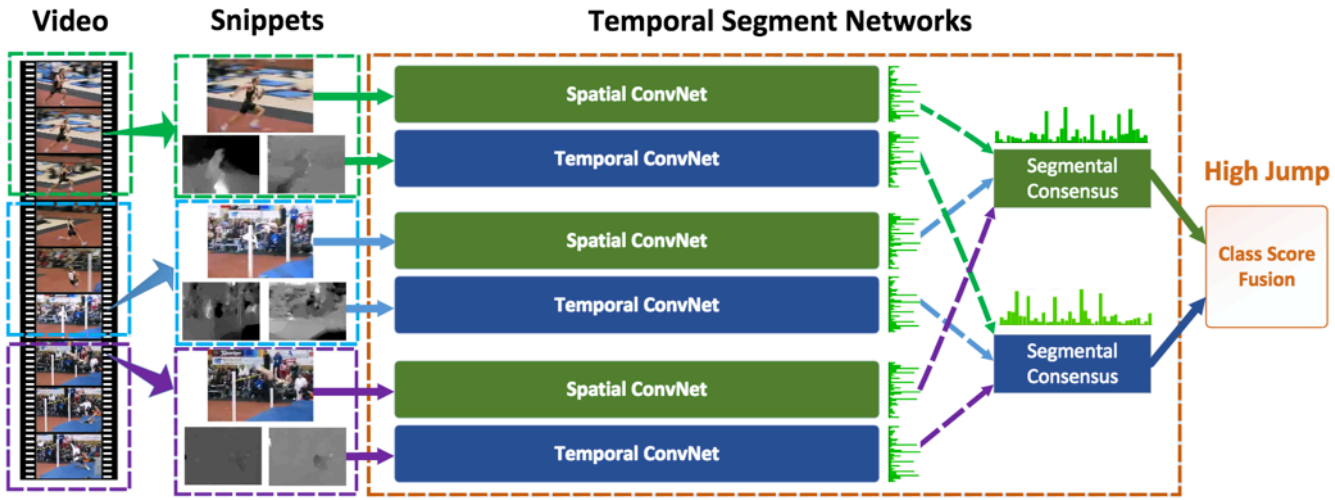
ActivityNet Challenge

ACTIVITYNET
Large Scale Activity Recognition Challenge

200	100	1.54	648
CLASSES	UNTRIMMED VIDEOS PER CLASS	ACTIVITY INSTANCES PER VIDEO	VIDEO HOURS

動画認識の世界選手権

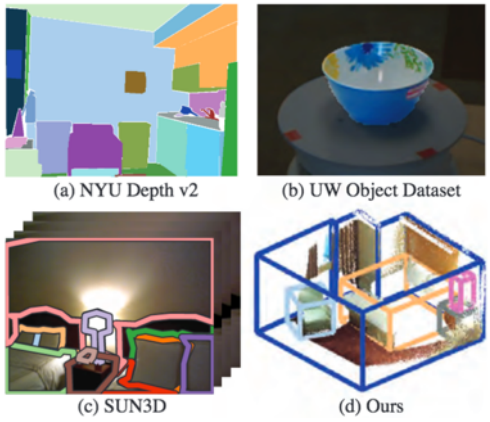
- 200クラス分類問題にてTop-1が88%, Top-3が96%
- Temporal Segmental Networks [Wang+, ECCV2016]により特徴抽出

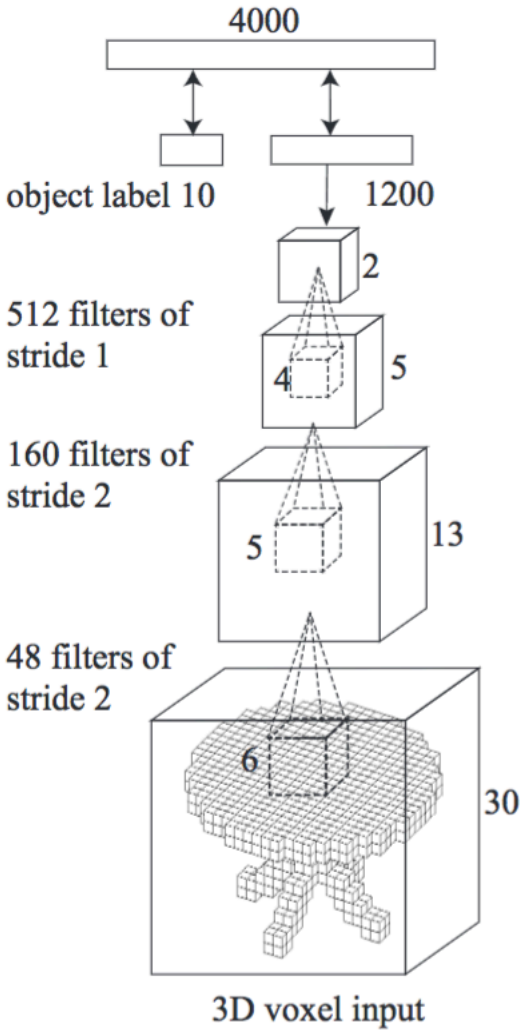


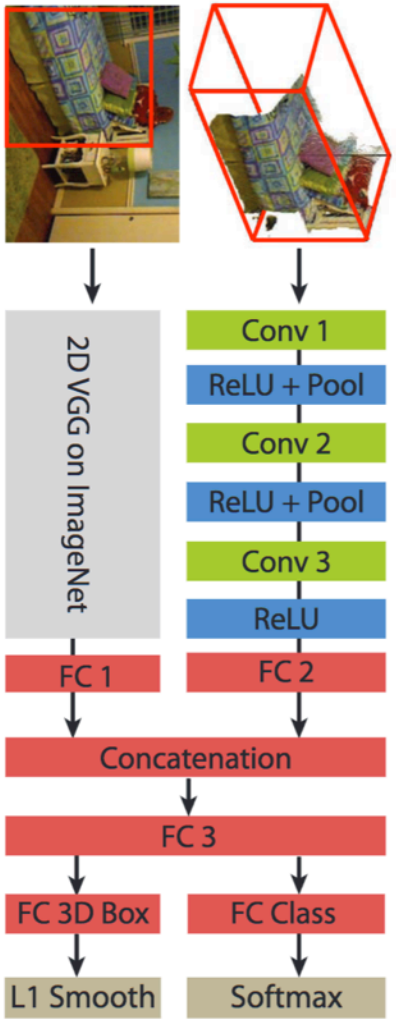
3D Recognition

3D ShapeNet, SUN RGB-D, Deep Sliding Window



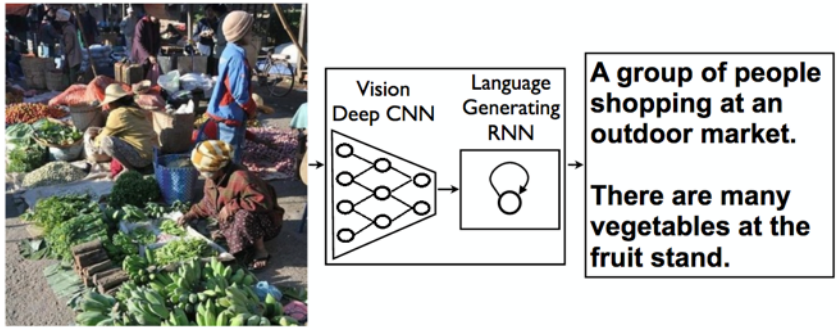


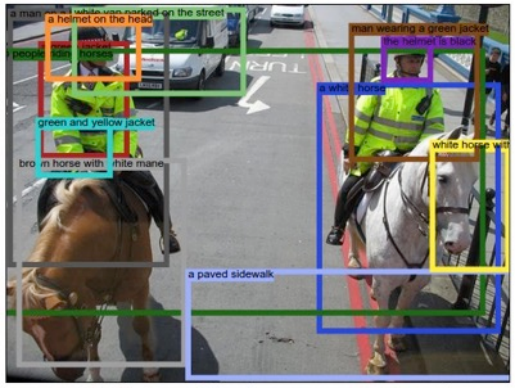




Image/Video Captioning

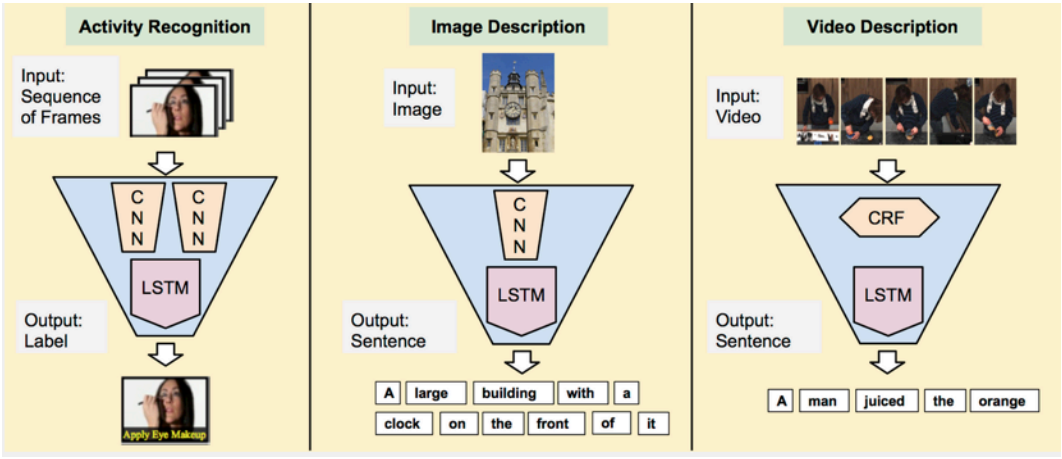
Show and Tell [Vinyals+, CVPR15] DenseCap [Johnson+, CVPR16]





a green jacket. a white horse. a man on a horse. two people riding horses. man wearing a green jacket. the helmet is black. brown horse with white mane. white van parked on the street. a paved sidewalk. green and yellow jacket. a helmet on the head. white horse with white face.

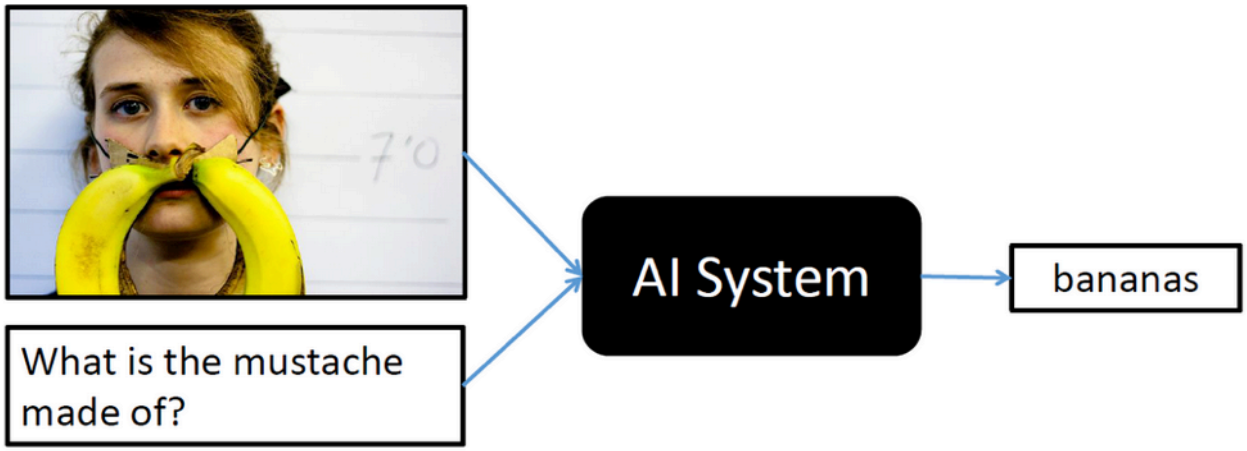
Long-term RNN [Donahue+, CVPR2015]





Visual Question Answering

Competition at CVPR 2016



12:15 PM - 12:30 PM



Challenge Winner Talk (Real Images)
Akira Fukui, Dong Huk Park, Daylen Yang, Anna Rohrbach, Trevor Darrell and Marcus Rohrbach
[slides]

11:40 AM - 11:55 AM

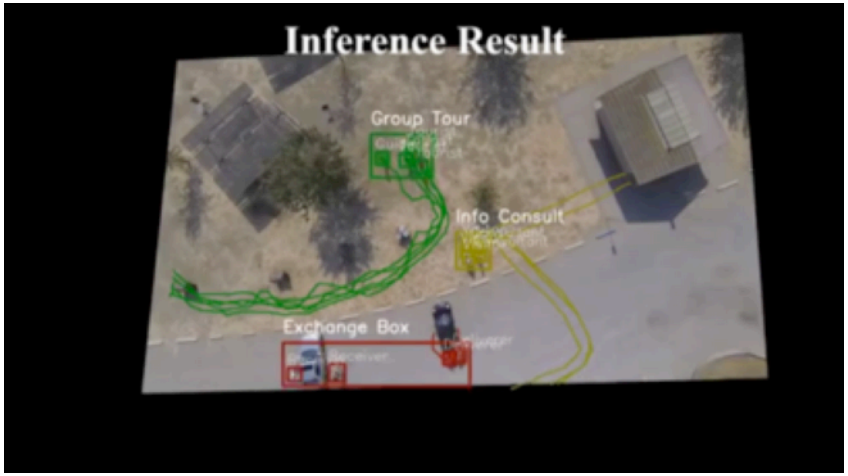


Challenge Winner Talk (Abstract Scenes)
Andrew Shin*, Kuniaki Saito*, Yoshitaka Ushiku and Tatsuya Harada
[slides]

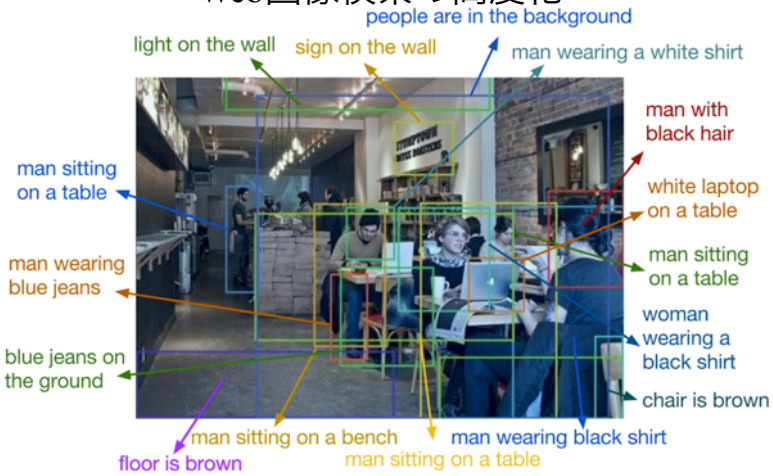
アプリケーション

ドローン，自動運転，Web，監視，一人称カメラ．．．

UAV [Shu+, CVPR15]
ドローンからのイベント認識



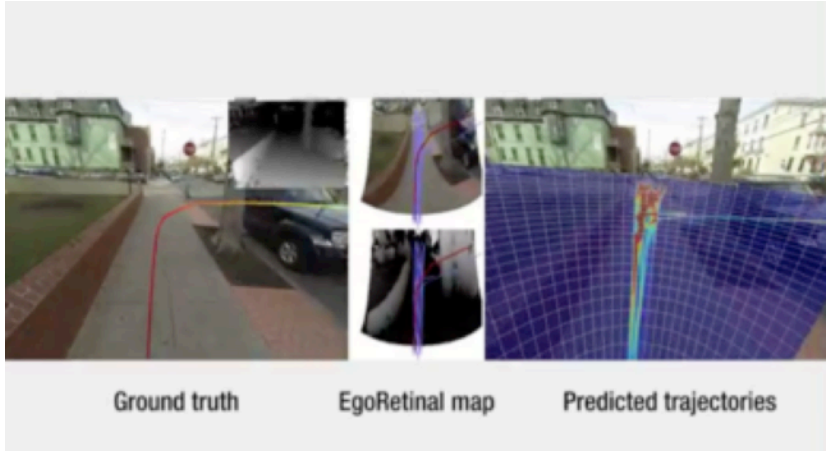
DenseCap [Johnson+, CVPR16]
Web画像検索の高度化



MegaFace [Kemelmacher-Shlizerman+, CVPR16]
地球規模(69万人)での顔認証問題への挑戦



Egocentric Future Localization [Park+, CVPR16]
一人称ビジョンによる道案内・経路推薦

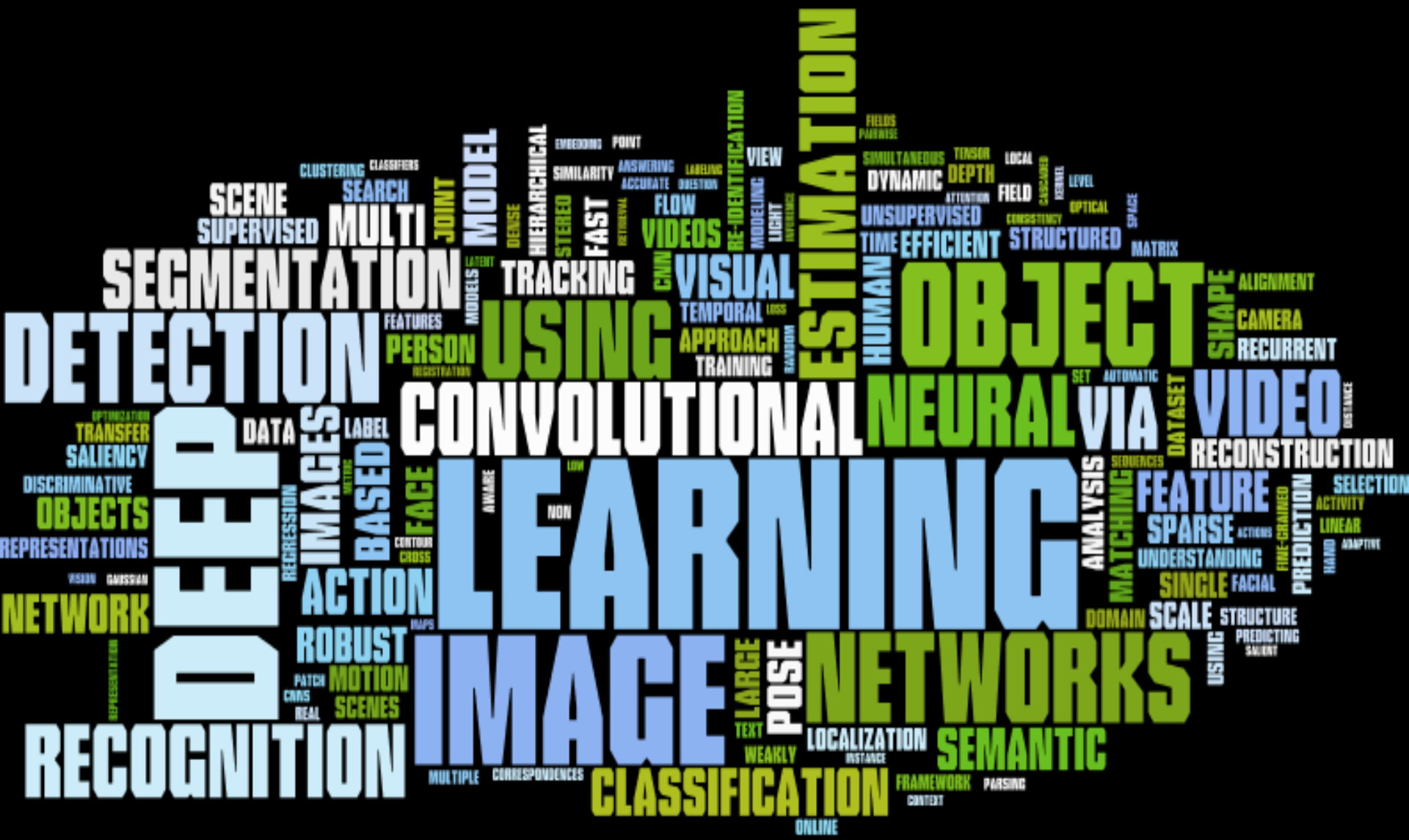


CV分野(特に認識)の動向

@CVPR2015


STEREO
HASHING
MODELING
RECONSTRUCTION
STRUCTURED
DESCRIPTOR
DEEP
DYNAMIC
DISCOVERING
DESCRIPTORS
ACCURATE
INFERENCE
LOCAL
ESTIMATION
MANIFOLD
ADAPTIVE
MATRIX
INSTANCE
SEARCH
CODING
SIMILARITY
SINGLE
DEPTH
IMAGE
OBJECTS
MULTIPLE
ALGORITHM
PARSING
JOINT
DETECTION
BINARY
MODELS
MOTION
VISUAL
PRIOR
BLUR
GLOBAL
BEYOND
FAST
RETRIEVAL
OPTIMIZATION
PREDICTION
STRUCTURE
DEFORMABLE
SURFACE
SEGMENTATION
MODEL
LEARNING
FINE-GRAINED
LOCALIZATION
VIDEOS
SELECTION
RE-IDENTIFICATION
CLASSIFICATION
CONVEX
FACIAL
PART
APPEARANCE
ATTRIBUTES
NEW
DOMAIN
LOW-RANK
SCENE
UNDERSTANDING
MATCHING
BAYESIAN
HUNING
GAUSSIAN
RECURRENT
LIGHT
ROBUST
IMAGES
RGB-D
RECOGNITION
FACE
OBJECT
CONVOLUTIONAL
RIEMANNIAN
FRAMEWORK
CONTEXT
HUMAN
ANALYSIS
SHAPE
ACTION
LABEL
ACTIVE
WILD
MACHINE
FLOW
VIA
USING
NETWORK
LARGE-SCALE
REAL-TIME
MATERIAL
METHOD
HIERARCHICAL
GRAPH
TEXTURE
FEATURES
REGRESSION
PHOTOMETRIC
SPARSE
REGULARIZATION
DATA
PERSON
CLUSTERING
SEMANTIC
NETWORKS
POSE
TRACKING
SALIENCY
PROPOSALS
WEAKLY
APPLICATION
LEARNED
KERNEL
TRANSFER
GENERALIZED
MULTI-LABEL
REGISTRATION
APPROACH
BASED
NEURAL
VIDEO
REPRESENTATION
VISION
SUBSPACE
DEBLURRING
UNSUPERVISED
PEDESTRIAN
ALIGNMENT
FEATURE
DATASET
UNCONSTRAINED
FUSION
METRIC
EFFICIENT
MONOCULAR
CAMERA
CORRESPONDENCE
RANDOM
REPRESENTATIONS
CORRELATION
SCENES
SUPERVISED
ADAPTATION

@CVPR2016



LEARNING
CONVOLUTIONAL
DETECTION
SEGMENTATION
RECOGNITION
IMAGE
NETWORKS
OBJECT
ESTIMATION
USING
NEURAL
VIA
VIDEO
DEEP
ACTION
ROBUST
MOTION
SCENES
NETWORK
OBJECTS
REPRESENTATIONS
DISCRIMINATIVE
SALIENCY
TRANSFER
DATA
IMAGES
BASED
FACE
SCENE
SUPERVISED
MULTI
SEARCH
CLUSTERING
CLASSIFIERS
JOINT
MODEL
HIERARCHICAL
STEREO
FAST
TRACKING
VISUAL
VIDEOS
FLOW
RE-IDENTIFICATION
VIEW
SIMILARITY
ANSWERING
ACCURATE
LABELING
QUESTION
MODELING
LIGHT
DENSE
POINT
EMBEDDING
FIELDS
PAIRWISE
SIMULTANEOUS
TENSOR
LOCAL
DYNAMIC
DEPTH
FIELD
LEVEL
OPTICAL
UNSUPERVISED
TIME
EFFICIENT
STRUCTURED
MATRIX
HUMAN
SHAPE
ALIGNMENT
CAMERA
RECURRENT
TEMPORAL
APPROACH
TRAINING
FEATURES
PERSON
MODELS
LATENT
CNN
RANDOM
SET
AUTOMATIC
DATASET
RECONSTRUCTION
SELECTION
ACTIVITY
LINEAR
ADAPTIVE
HAND
PREDICTION
FINE-GRAINED
ANALYSIS
MATCHING
SEQUENCES
FEATURE
SPARSE
ACTIONS
UNDERSTANDING
SINGLE
FACIAL
DOMAIN
SCALE
STRUCTURE
PREDICTING
SALIENT
USING
LARGE
POSE
TEXT
WEAKLY
LOCALIZATION
INSTANCE
SEMANTIC
FRAMEWORK
PARSING
CONTEXT
ONLINE
MULTIPLE
CORRESPONDENCES
CLASSIFICATION
PATCH
CNNS
REAL
REPRESENTATION
VISION
GAUSSIAN
REGRESSION
CONTOUR
CROSS
AWARE
NON
LOW
MAPS

CVPRで見る -2015と2016-

Deep Learning

- 2015: 約4割の人が使っている
- 2016: 80~85%に増加
- CNNのみならず，RNNの使用が目立った
  - 画像説明文/ 質問回答/ 画像解析/ 時系列解析など文脈を考慮するモデル

激戦区

- 画像認識 (CNN構築)
- 物体検出
- セマンティックセグメンテーション
- シーン認識

CVPRで見る -2015と2016-

次のフロンティア領域 – ボリュームデータや画像や物理量の取り扱いについて考察

- 時系列画像認識
  - 行動認識/ イベント認識/ / 動画解析/ 一人称ビジョン
- 3次元
  - 物体認識/ ポイントクラウド/ シーン解析/ 再構築
- フォトグラフィ
  - ライトフィールド
  - カメラアレイ
  - 物理量の変換

CVPRで見る -2015と2016-

蓄積があると有利

- 画像説明文/ 質問回答
- 画像生成/ モーフィング
- Style Transfer

どんな論文が通っているか？

戦略と難易度

- State-of-the-artを目指す (難易度: ★★★★★)
  - Top-1しか評価されない上にその後更新される
- 面白い数理理論や新しいモデルを提唱 (難易度: ★★★★☆)
  - 熟練が必要な上に実験が大変
- 意味のある評価を行っている (難易度: ★★★☆☆)
  - 愚直に実装・実験・考察を行う
- 問題設定を作成 (難易度: ★★★☆☆)
  - 一見難しいが動向を把握していれば比較的容易
- データセットを作成 (難易度: ★★★☆☆)
  - 一見難しいが頑張れば大丈夫

どんな論文が通っているか？

戦略と難易度

- State-of-the-artを目指す (難易度: ★★★★★)
  - Top-1しか評価されない上にその後更新される
- 面白い数理理論や新しいモデルを提唱 (難易度: ★★★★☆)
  - 熟練が必要な上に実験が大変
- 意味のある評価を行っている (難易度: ★★★☆☆)
  - 愚直に実装・実験・考察を行う
- 問題設定を作成※ (難易度: ★★★☆☆)
  - 一見難しいが動向を把握していれば比較的容易
- データセットを作成※ (難易度: ★★★☆☆)
  - 一見難しいが頑張れば大丈夫

※ 本プロジェクトで主に狙っている部分

ここまで1,000本超の論文を読んで

「優れた論文」は「優れた問題」を提供

- 追従型の研究よりもコンセプト重視の論文
- 次のトレンドを作るためのトライを繰り返す

洗練された問題を作る

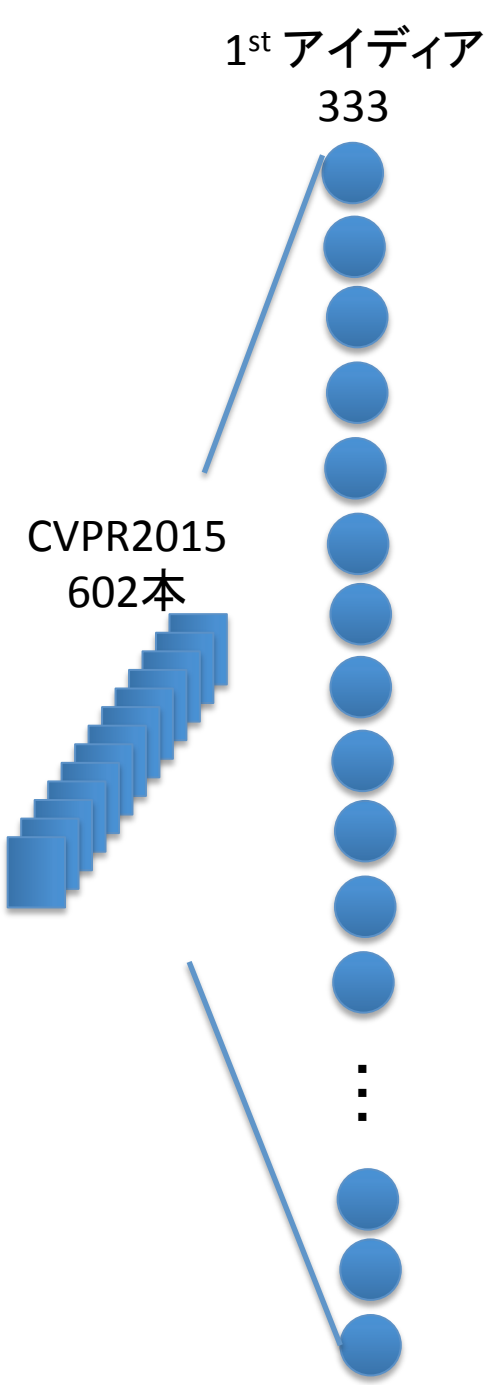
論文を「書く」

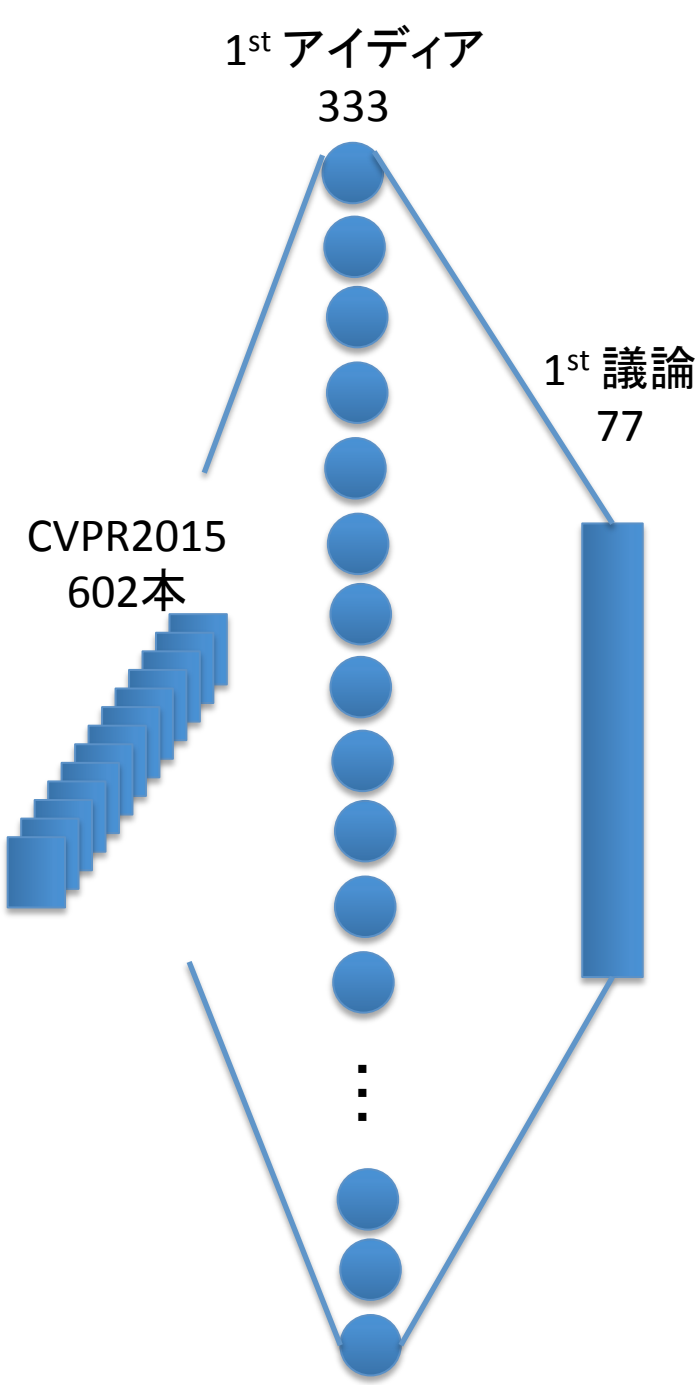
新しい問題設定を作る！

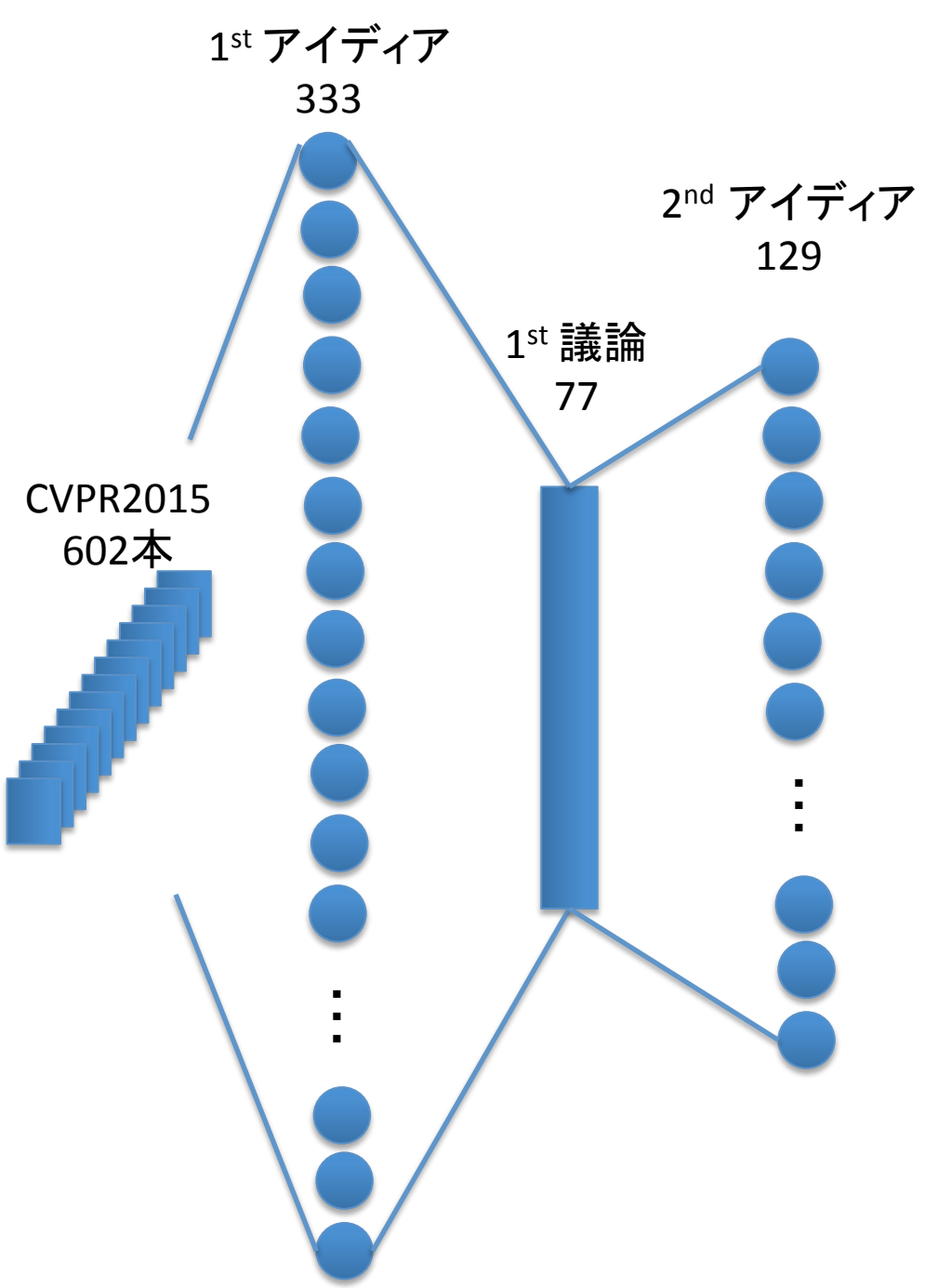
- 論文を網羅的に調査して歴史やトレンドを把握
- 膨大な知識から「洗練された」問題設定を世に送り出す
- アイディア考案・議論・実装の繰り返し

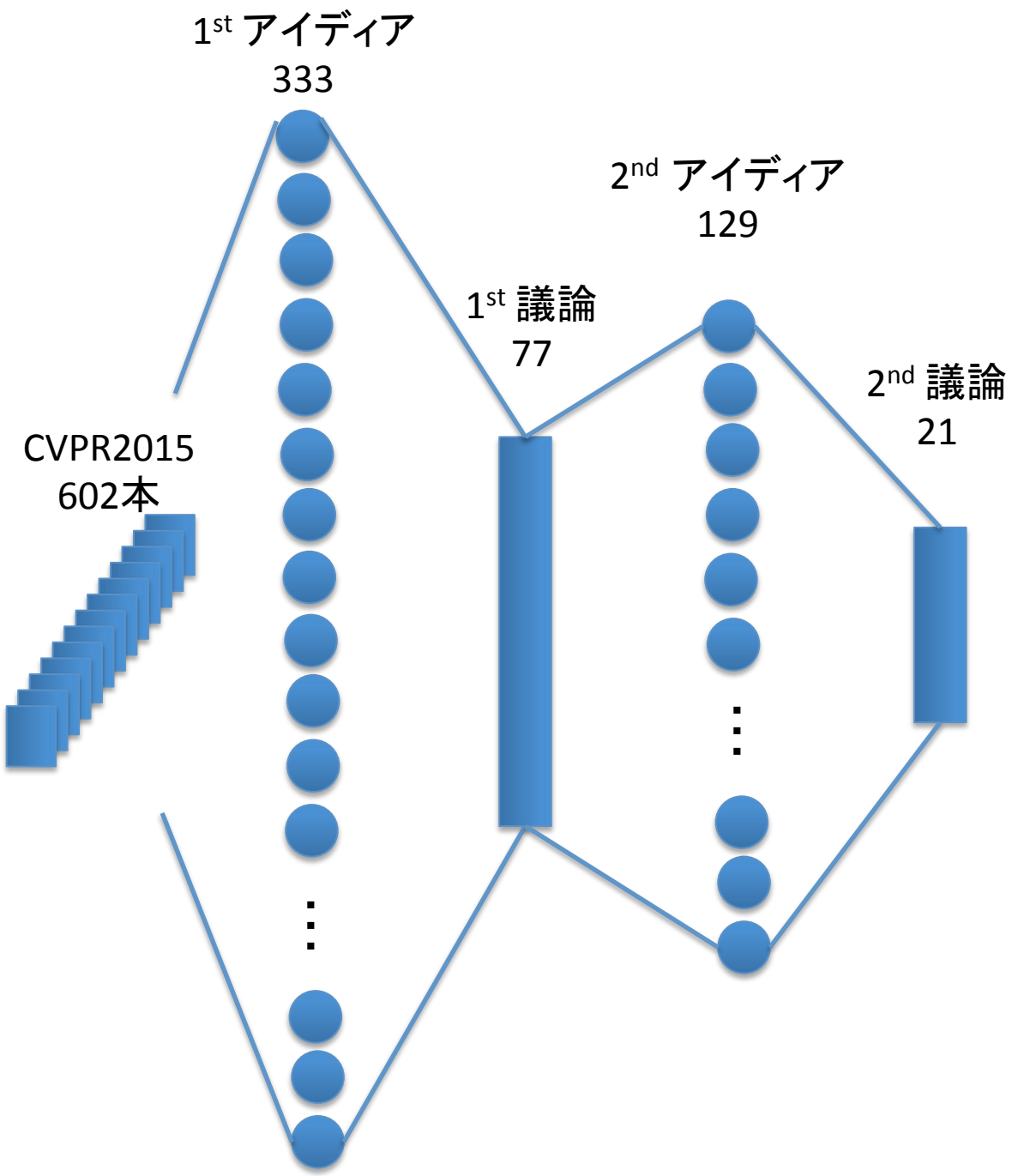
CVPR2015
602本

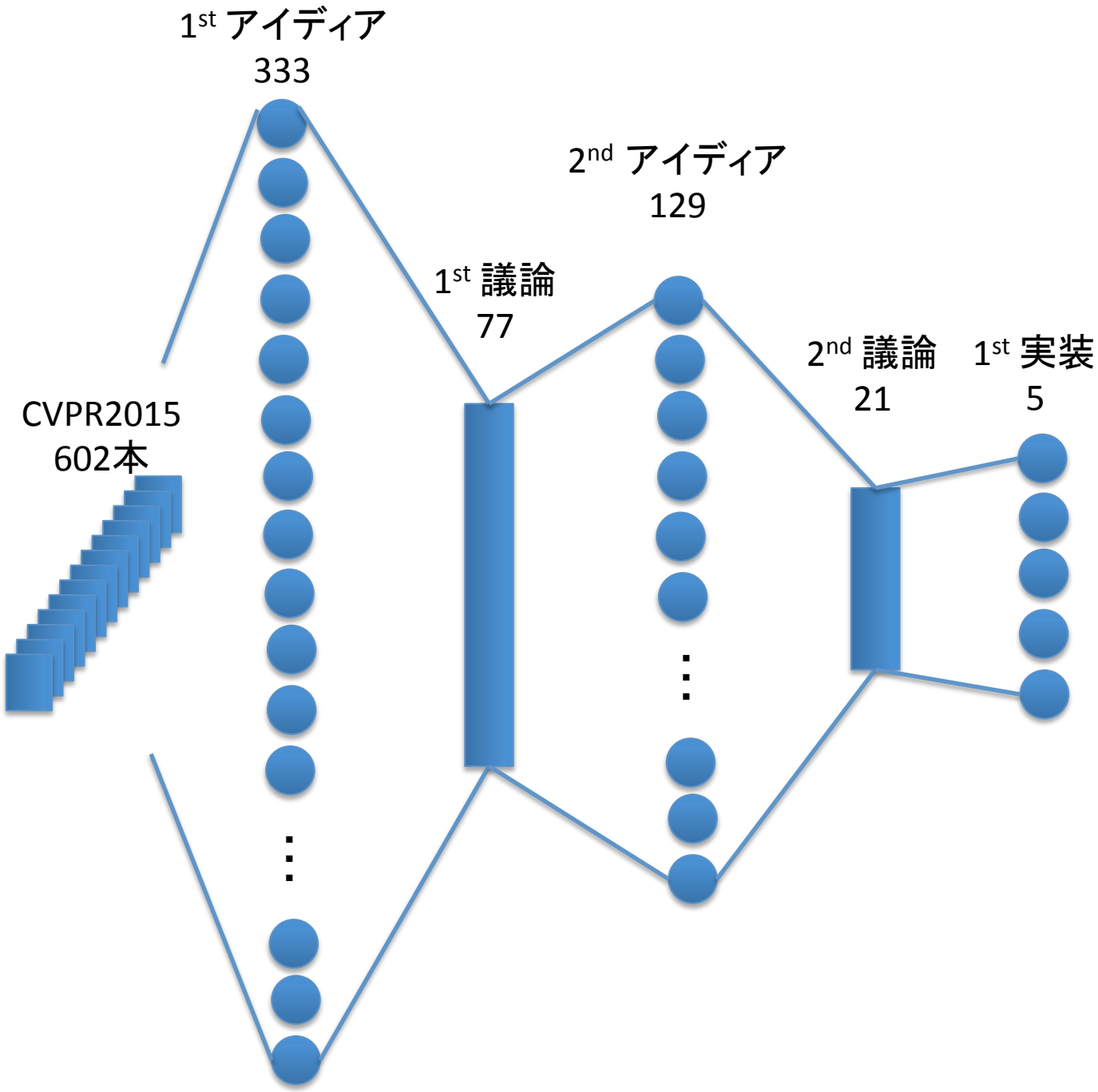


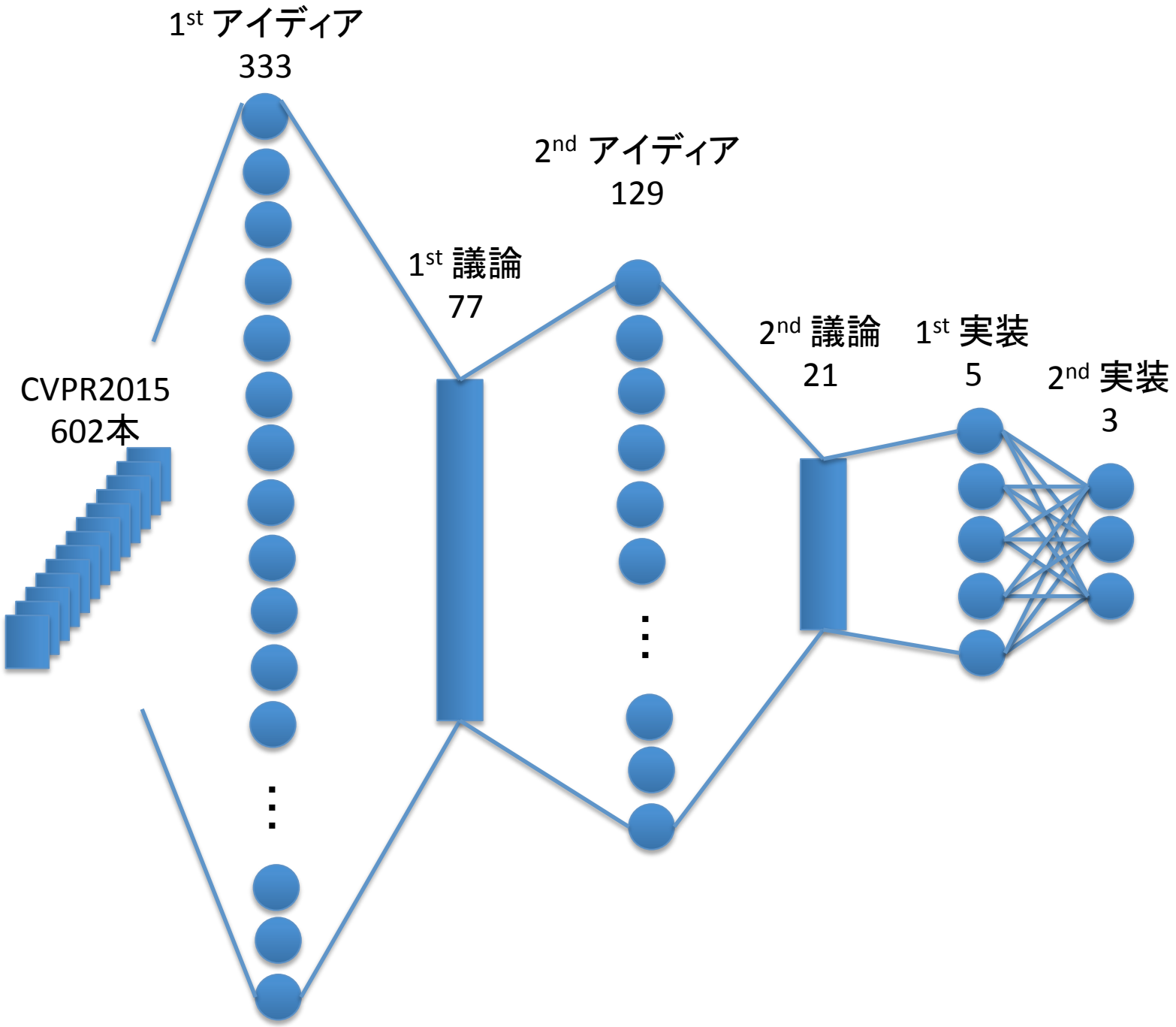

1st アイディア
333
CVPR2015
602本

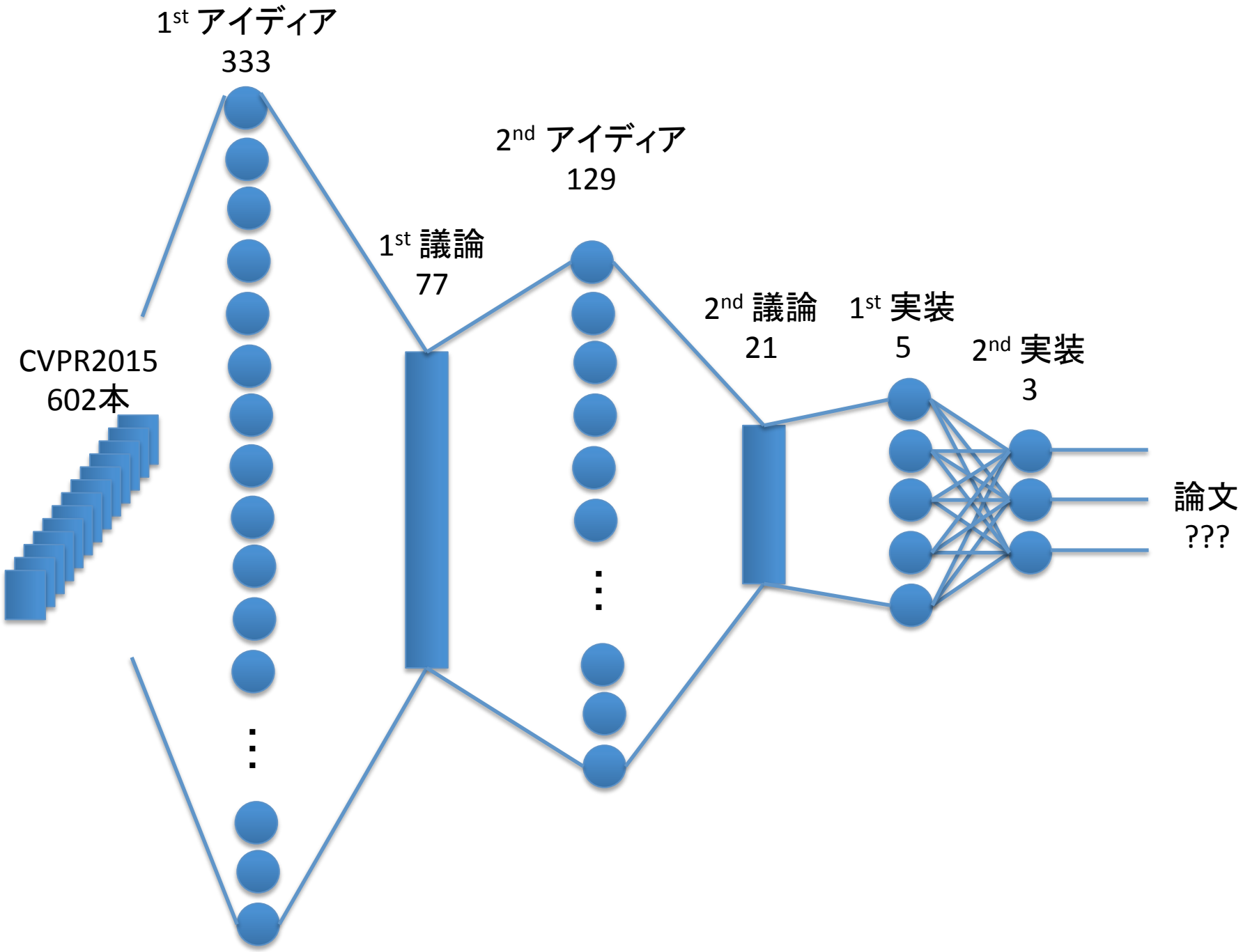

1st アイディア
333
1st 議論
77
CVPR2015
602本

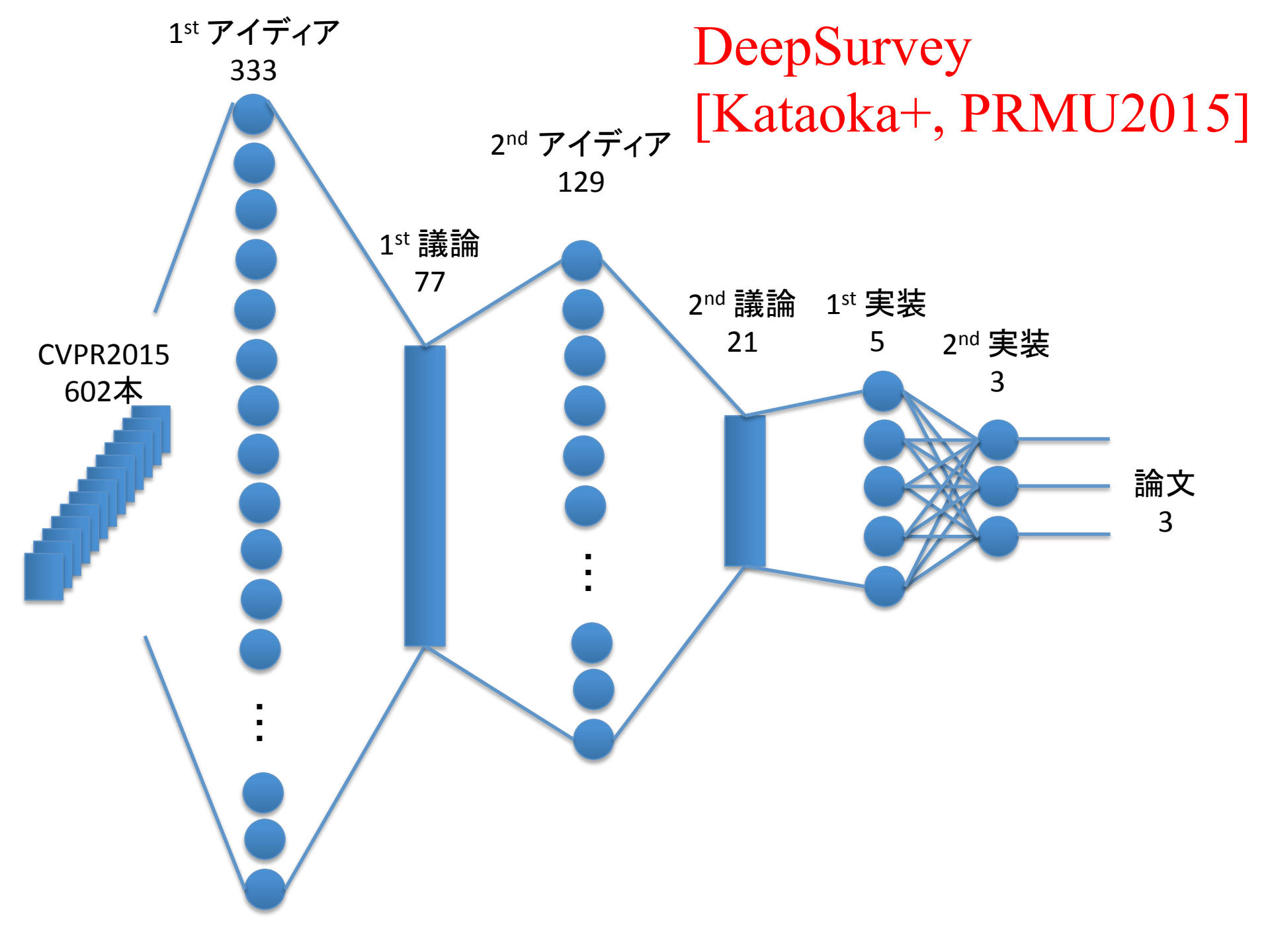

1st アイディア
333
2nd アイディア
129
1st 議論
77
CVPR2015
602本


1st アイディア
333
2nd アイディア
129
1st 議論
77
2nd 議論
21
CVPR2015
602本


1st アイディア
333
2nd アイディア
129
1st 議論
77
2nd 議論
21
1st 実装
5
CVPR2015
602本

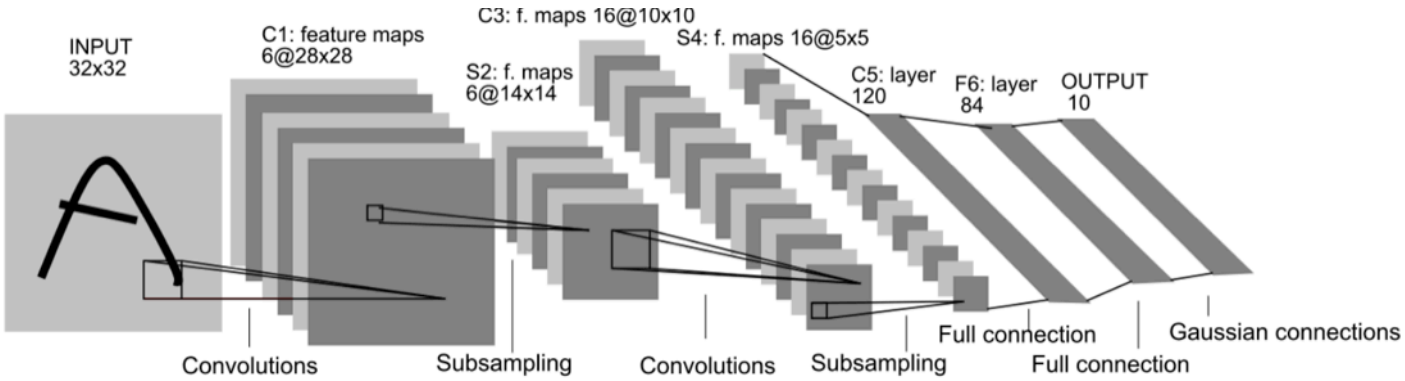

1st アイディア
333
2nd アイディア
129
1st 議論
77
2nd 議論
21
1st 実装
5
2nd 実装
3
CVPR2015
602本

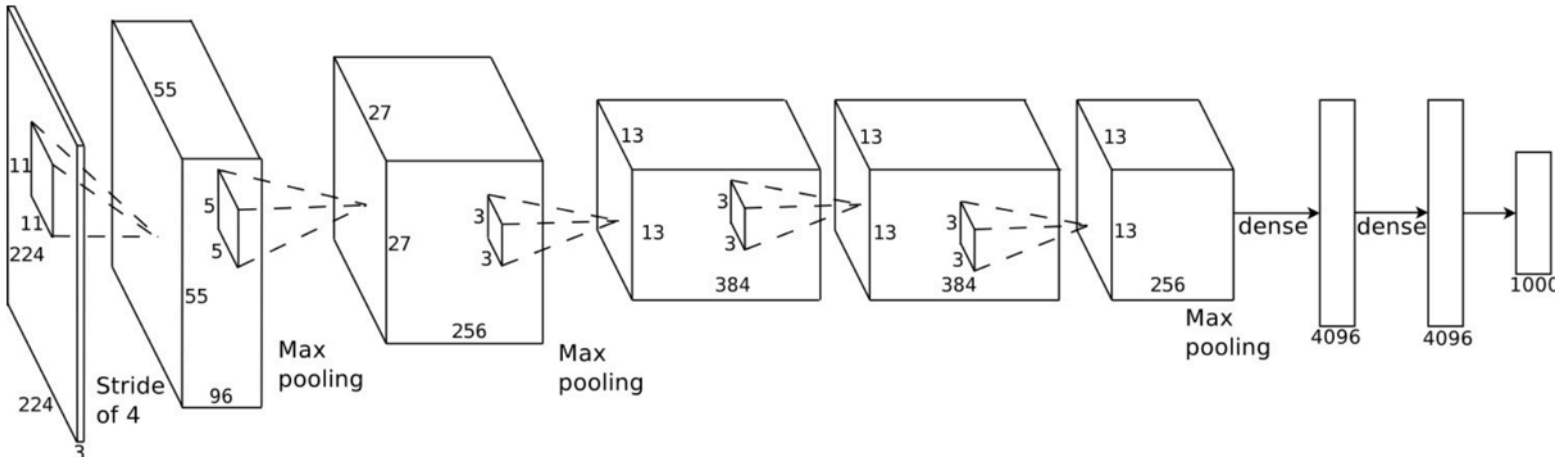

1st アイディア
333
2nd アイディア
129
1st 議論
77
2nd 議論
21
1st 実装
5
2nd 実装
3
CVPR2015
602本
論文
???


DeepSurvey
[Kataoka+, PRMU2015]
1st アイディア
333
2nd アイディア
129
1st 議論
77
2nd 議論
21
1st 実装
5
2nd 実装
3
CVPR2015
602本
論文
3

従来の深層学習と比較して．．．

- 「自らがニューロンの一部となる」という画期的な手法
- 「出力が研究成果という形式で出てくる」という画期的な手法
- 「一人だけではなくてみんなの力を合わせる」という画期的な手法
- 「学習によりニューロン自体が成長する」という画期的な手法





DeepSurvey 現在までの実績

【2016】

■ Hirokatsu Kataoka, Yudai Miyashita, Tomoaki Yamabe, Soma Shirakabe, Shin'ichi Sato, Hironori Hoshino, Ryo Kato, Kaori Abe, Takaaki Imanari, Naomichi Kobayashi, Shinichiro Morita, Akio Nakamura, "cvpaper.challenge in 2015 - A review of CVPR2015 and DeepSurvey", arXiv pre-print 1605.08247, 2016.

■ Hirokatsu Kataoka, Soma Shirakabe, Yudai Miyashita, Akio Nakamura, Kenji Iwata, Yutaka Satoh, "Semantic Change Detection with Hypermaps", arXiv preprint arXiv:1604.07513, 2016.

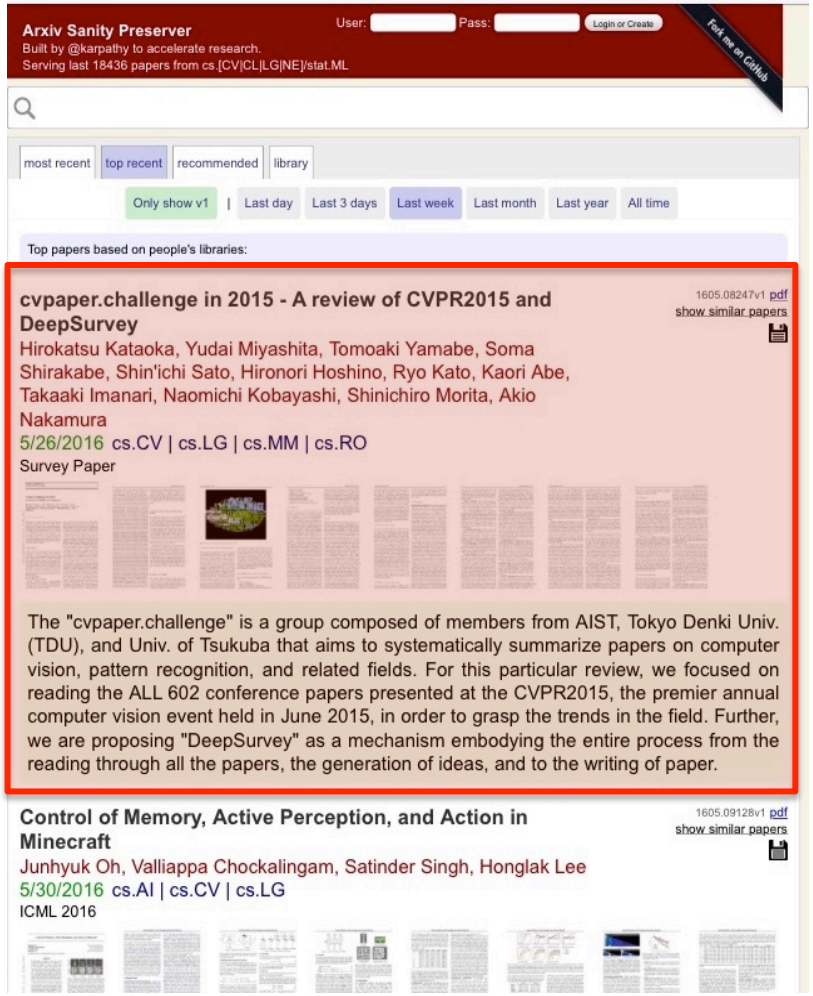
【2015】

■ 片岡裕雄, 宮下侑大, 山辺智晃, 白壁奏馬, 佐藤晋一, 星野浩範, 加藤遼, 阿部香織, 今成隆了, 小林直道, 森田慎一郎, 中村明生, "cvpaper.challenge in CVPR2015 -CVPR2015 のまとめ-", パターン認識・メディア理解研究会（PRMU）, 2015.

■ Tomoaki Yamabe, Yudai Miyashita, Shin'ichi Sato, Yudai Yamamoto, Akio Nakamura, Hirokatsu Kataoka, "What is an Effective Feature for a Detection Problem? -Feature Evaluation in Multiple Scenes-", IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2015.

arXivにて週間 世界第一位を達成！

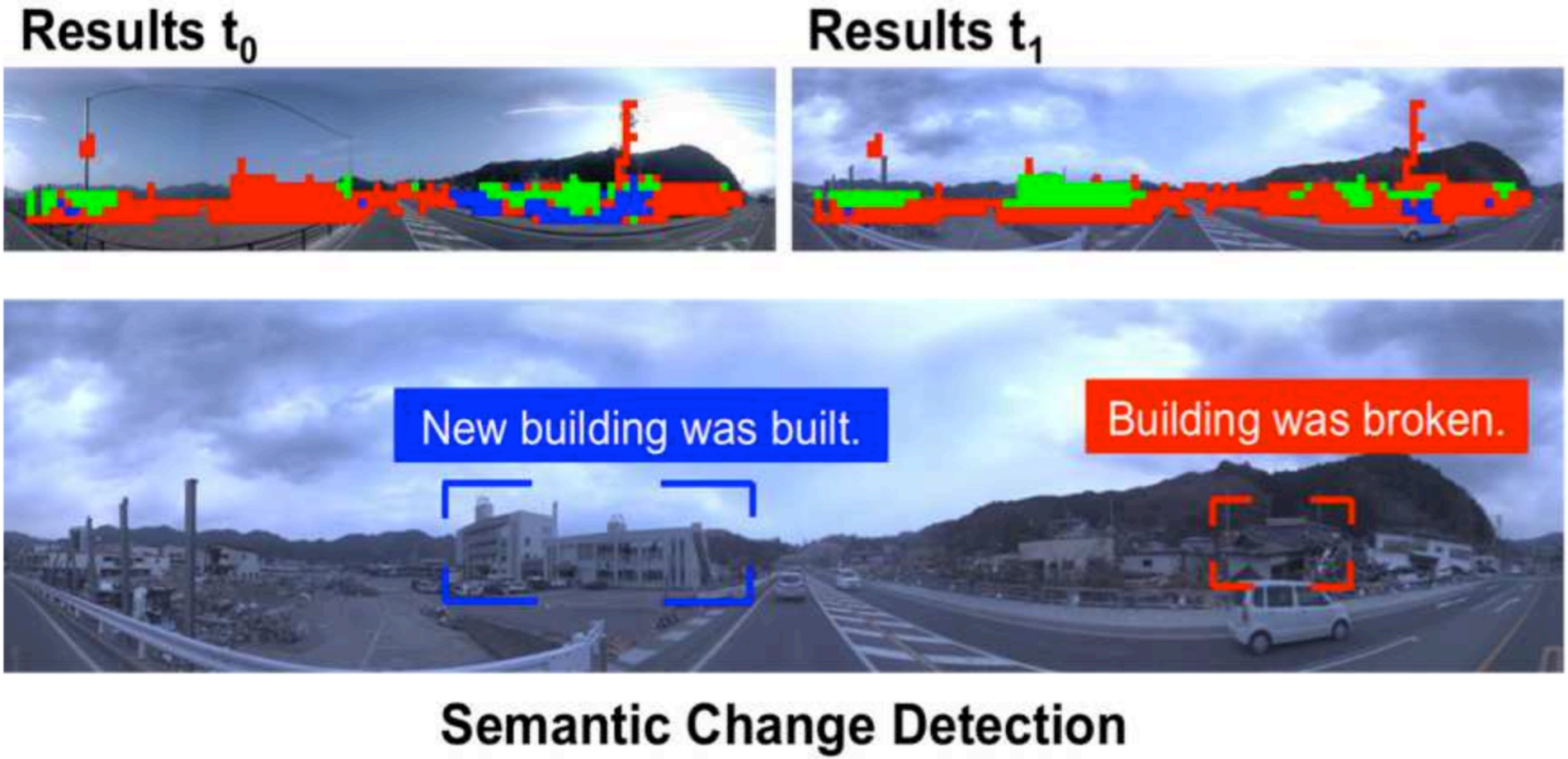
- 取り組みが世界的に話題 (参考文献数も)
- CambridgeやStanfordからも問い合わせあり



related fields. As the first step of this endeavor, we undertook to read all the 602 papers accepted during the CVPR2015 [1,2,3,4,5,6,7,8,9,10,11,12,13,14, 15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32, 33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50, 51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68, 69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86, 87,88,89,90,91,92,93,94,95,96,97,98,99,100,101,102,103, 104,105,106,107,108,109,110,111,112,113,114,115,116, 117,118,119,120,121,122,123,124,125,126,127,128,129, 130,131,132,133,134,135,136,137,138,139,140,141,142, 143,144,145,146,147,148,149,150,151,152,153,154,155, 156,157,158,159,160,161,162,163,164,165,166,167,168, 169,170,171,172,173,174,175,176,177,178,179,180,181, 182,183,184,185,186,187,188,189,190,191,192,193,194, 195,196,197,198,199,200,201,202,203,204,205,206,207, 208,209,210,211,212,213,214,215,216,217,218,219,220, 221,222,223,224,225,226,227,228,229,230,231,232,233, 234,235,236,237,238,239,240,241,242,243,244,245,246, 247,248,249,250,251,252,253,254,255,256,257,258,259, 260,261,262,263,264,265,266,267,268,269,270,271,272, 273,274,275,276,277,278,279,280,281,282,283,284,285, 286,287,288,289,290,291,292,293,294,295,296,297,298, 299,300,301,302,303,304,305,306,307,308,309,310,311, 312,313,314,315,316,317,318,319,320,321,322,323,324, 325,326,327,328,329,330,331,332,333,334,335,336,337, 338,339,340,341,342,343,344,345,346,347,348,349,350, 351,352,353,354,355,356,357,358,359,360,361,362,363, 364,365,366,367,368,369,370,371,372,373,374,375,376, 377,378,379,380,381,382,383,384,385,386,387,388,389, 390,391,392,393,394,395,396,397,398,399,400,401,402, 403,404,405,406,407,408,409,410,411,412,413,414,415, 416,417,418,419,420,421,422,423,424,425,426,427,428, 429,430,431,432,433,434,435,436,437,438,439,440,441, 442,443,444,445,446,447,448,449,450,451,452,453,454, 455,456,457,458,459,460,461,462,463,464,465,466,467, 468,469,470,471,472,473,474,475,476,477,478,479,480, 481,482,483,484,485,486,487,488,489,490,491,492,493, 494,495,496,497,498,499,500,501,502,503,504,505,506, 507,508,509,510,511,512,513,514,515,516,517,518,519, 520,521,522,523,524,525,526,527,528,529,530,531,532, 533,534,535,536,537,538,539,540,541,542,543,544,545, 546,547,548,549,550,551,552,553,554,555,556,557,558, 559,560,561,562,563,564,565,566,567,568,569,570,571, 572,573,574,575,576,577,578,579,580,581,582,583,584, 585,586,587,588,589,590,591,592,593,594,595,596,597, 598,599,600,601,602]. This review summarizes all the papers of CVPR2015 we read as the first project of

Semantic Change Detection

[Kataoka+, arXiv16]



入力のイメージ

点 (1論文)



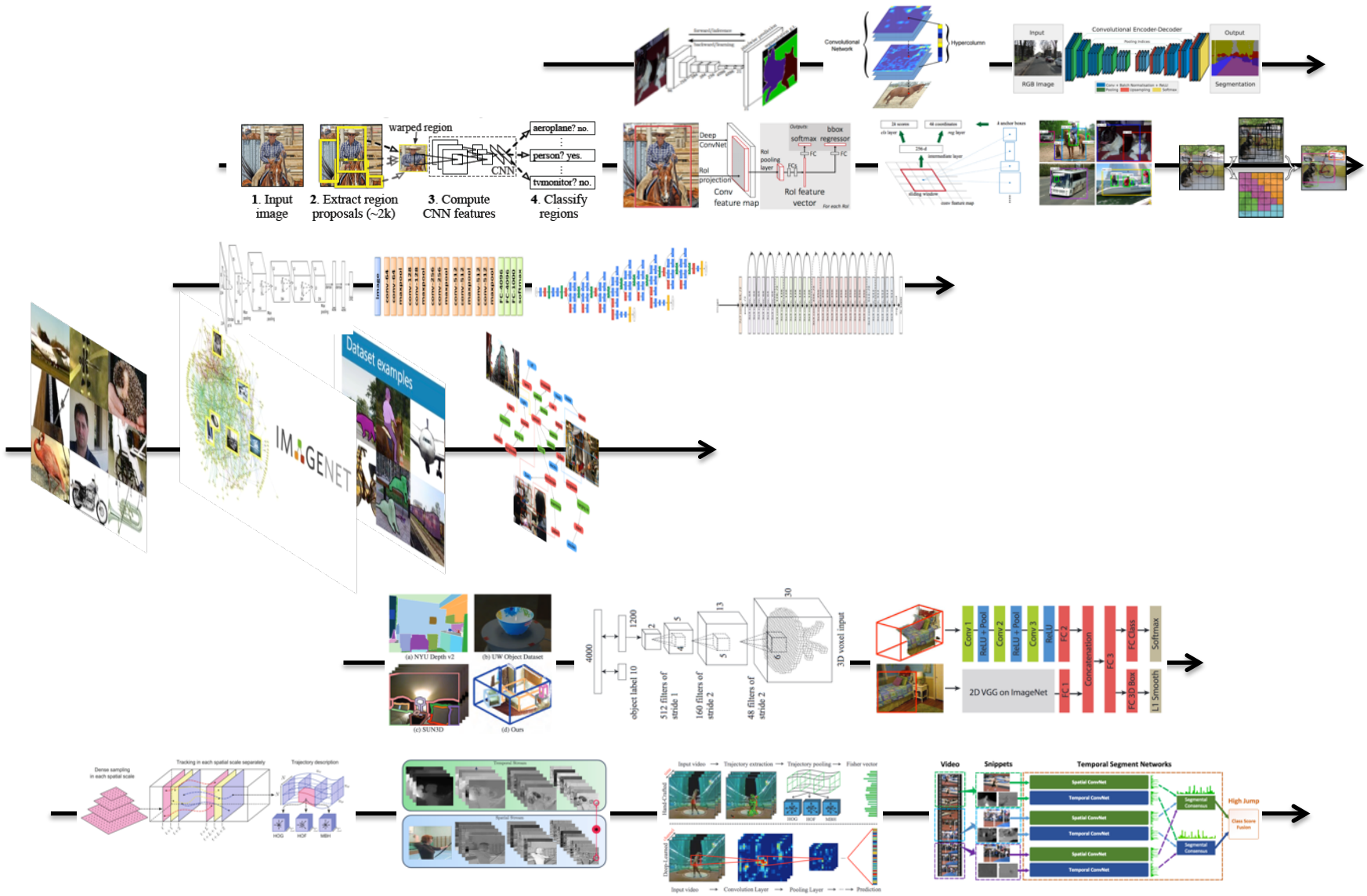
入力のイメージ

点 (1論文) => 線 (1分野)



入力のイメージ

点 (1論文) => 線 (1分野) => 面 (多分野)



良い出力のためには？

読む論文数は多いほうが良い

- 点 (Pixel), 線 (Line)より面 (Image)の入力

キーとなる論文の理解は重要

- 局所的に良い勾配を持ち，領域は大きいほうが良い

ブレストとアイディア整理の数は多いほうが良い

- 畳み込み，プーリング数を増やす

実施期間は重要

- 畳み込むタイミングは非常に重要 (ニューロンは忙しい)

実装はグループごとに密に行う

- 全結合は複数のストリームに分ける

To be continued in 2016…

2016年の目標は「1,000本読破」と「上位会議への投稿」

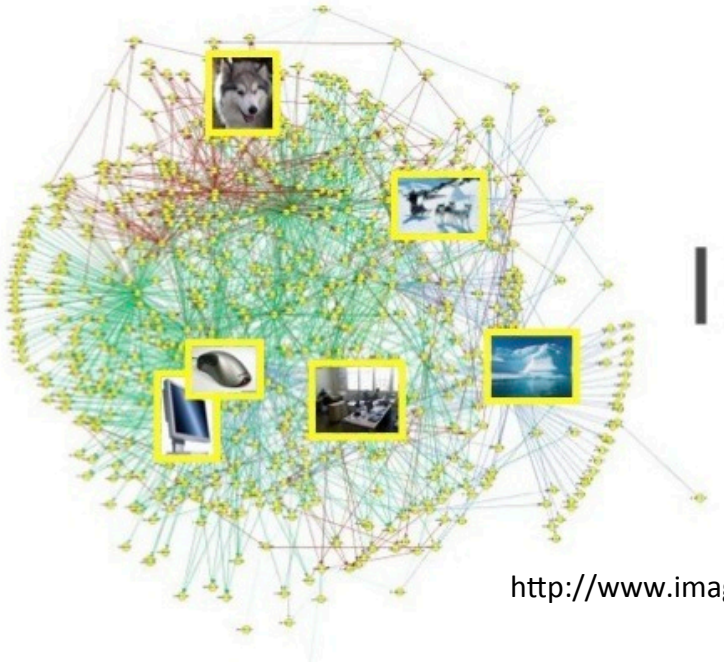
- 考案したアイディアを実現可能レベルに昇華
- 現在，まさに議論と実装を練り上げている最中

今後の取り組み

今後の方針

問題設定と同時にデータセットを作る

- 問題と同時にデータセットを自分で作る
- 覚悟する！
- データの作成が新規性のひとつとして論文に書ける



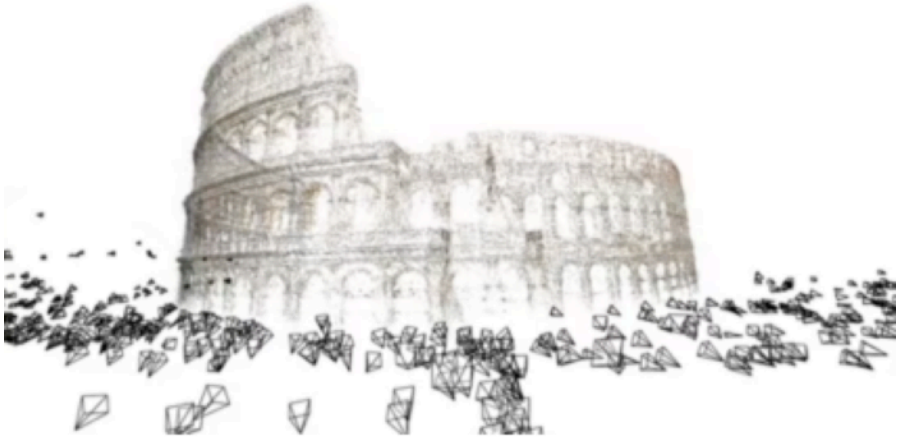
IMAGENET

http://www.image-net.org/

今後の方針

新規性をうまくアプリケーションとして見せる

- 昔のBuilding Rome in a Day (ICCV2009)のイメージ
- 2016年だとFace2Faceなど
- 少しの差分でもその違いで見える未来を（わかりやすく）説明
- 分かりやすく，というところが重要



https://www.youtube.com/watch?v=kxtQqYLRaSQ

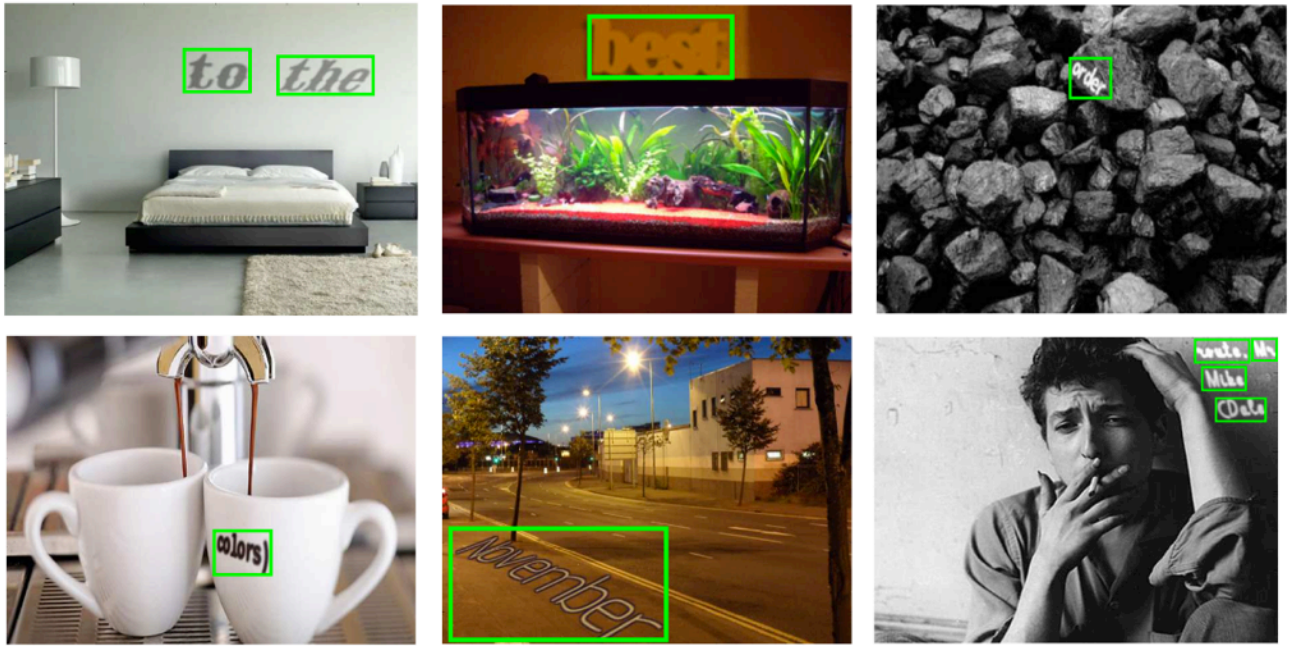


https://www.youtube.com/watch?v=ohmajJTcpNk

今後の方針

ロストアイディアを復活できないか？

- （一時的に）消されたアイディアという意味
- もちろん，今風にアレンジ
- 可能性があるにも関わらず，消えたアイディアという意味
- e.g. ニューラルネット，文字検出



次世代のアイディア？

自ら学習する機械

- Picking Robot (Google Brain)
  - カメラをアームに取り付けて80万回のトライ&エラー
- AlphaGo (Google DeepMind)
  - 自分と自分を6000万回対局

自分で学習データを生成し学習

次世代のアイディア？

時間軸・対象数・評価軸の飛躍

– 産総研大西さんのMIRU2016 DS論文より

画像数以上の価値があるデータ

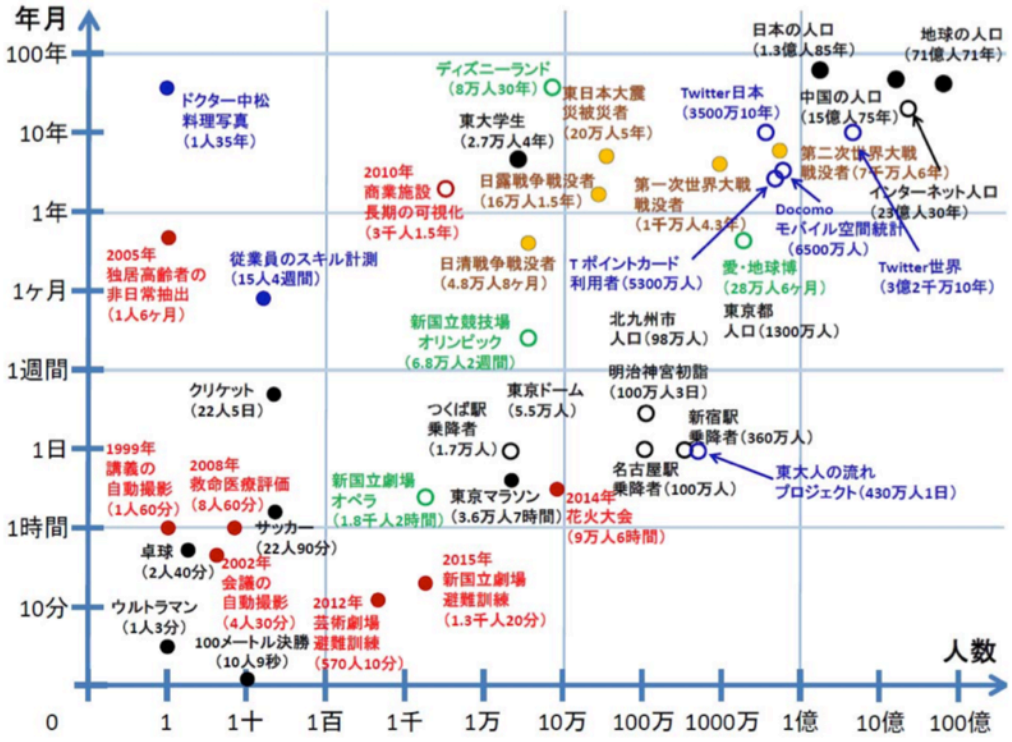


図 3　人数ー期間ダイアグラムとこれまでの研究のマッピング

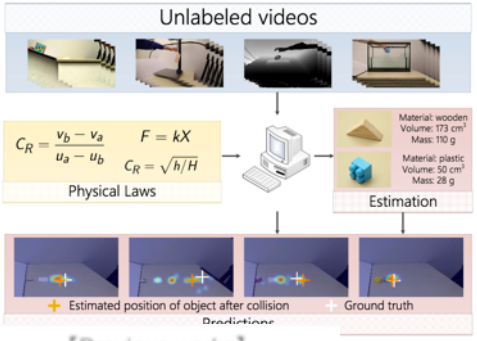
次世代のアイディア？

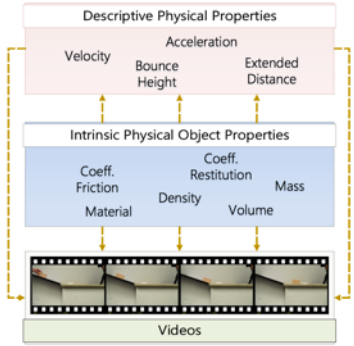
画像に映らないものを推定：物理量の変換，潜在的な知識

- Visually indicated sounds [Owens+, CVPR2016]
- Physics 101 [Wu+, BMVC2016]
- Transitional Action Recognition [Kataoka+, BMVC2016]



https://www.youtube.com/watch?v=JpZUZ9ZDECE







https://www.youtube.com/watch?v=m1AmdIYGaBY

未来の技術？

10~20年後を考えよう！

× いきなり新しい設定を考える

◎ 今の問題の劇的な進展が「未来」

歴史を尊重し新規問題を考えることこそが哲学 (Philosophy)に昇華され,
絶えず未来に繋がる技術へ